THESIS

# A VERY CONFUSING PROBLEM: INTERPRETING KEYNESIAN WEIGHT

Submitted by

Josh Brekel

Department of Philosophy

In partial fulfillment of the requirements

For the Degree of Master of Arts

Colorado State University

Fort Collins, Colorado

Fall 2022

Master's Committee:

Advisor: Jeff Kasser

Kenneth Shockley
Ben Prytherch

Copyright by Josh Brekel 2022

All Rights Reserved

ABSTRACT

A VERY CONFUSING PROBLEM: INTERPRETING KEYNESIAN WEIGHT

Initially outlined by John Maynard Keynes in 1921, *Keynesian weight* is a measure intended to characterize evidence independently of probability. As a concept that is often immersed in confusion, Keynesian weight requires thorough philosophical explication prior to any sort of legitimate use in decision-making, legal proceedings, or scientific inquiry. In this thesis, I attempt to explicate Keynesian weight by arguing in favor of Jochen Runde's *relative interpretation* of Keynesian weight. The aim of Chapter 1 is to introduce the basic idea of Keynesian weight. In Chapter 2, I demonstrate that Keynes's initial analysis of Keynesian weight creates an interpretative puzzle—two viable interpretations of Keynesian weight exist. Chapter 3 aims to solve the interpretative puzzle by consideration of how the interpretations of Keynesian weight respond to I.J. Good's criticism of Keynesian weight. Ultimately, I argue that Good's criticism demonstrates that the best interpretation of Keynesian weight is the relative interpretation.

## ACKNOWLEDGEMENTS

Graduate school is extremely difficult. I am sure plenty of people will tell you that. For most, philosophy graduate school means long hours of studying, reading, grading, writing papers, low pay, and stressing out. Although my experience involved all of those things, I loved (almost) every second of it. It blows my mind that I was paid to learn and (attempt to) teach others.

However, graduate school would not have been nearly as enjoyable without the wonderful people I am lucky enough to have in my life. Honestly, verbal or written expression of gratitude will never be enough to thank everyone who assisted me in making this dream into a reality. Maybe beer, coffee, or ice cream could suffice…But, I suppose written thanks is the best place to start. My whole-hearted gratitude goes out to everyone listed below, as well as the others who I (almost assuredly) forgot to mention. Thank you all!

**Jeff Kasser** – for showing me what philosophy is, for allowing me to constantly bug you these last six years, and for seeing potential in me despite my constant doubts. Stepping into PHIL 120 fundamentally changed my life (for the better). I don't think words can express how grateful I am to be your student and friend.

**Ben Prytherch** and **Ken Shockley** – for being so helpful and gracious with their time and comments while serving on my thesis committee. You both are excellent teachers, advisors, and stalwart members of the CSU community.

**Boston Lee** – for letting me join you in the philosophy of statistics independent study. The idea for this thesis grew out of those discussions.

Those who suffered through early drafts: **Javier Marquez**, **Blake Harris**, **Boston Lee**, **Caleb Hall**, and **Brandi Grauberger**.

All of the graduate students who I learned with at CSU. You were so welcoming, kind, and thought-provoking.

All of the teachers at CSU who took the time to get to know me, push me to be better, and provide an example for me to follow. I am particularly thankful to: **Jeff Kasser**, **Matt MacKenzie**, **Luke Petach**, **Domenica Romagni**, **Ken Shockley**, **Ben Prytherch**, **Jesse Burkhardt**, **Katie McShane**, **Moti Gorin**, **Beth Tropman**, **Dave Williams**, **Sean Waters**, **Cher Li**, **Robert Keller**, and **Harvey Cutler**.

**Shawn Brady** – for showing me what teaching philosophy can be.

**Kendall DeJonge** – for taking a chance on me.

The Colorado Avalanche (2022 Stanley Cup Champions) – for giving me something besides philosophy to fixate on.

All of my friends for their support over the years. I am particularly grateful to: **Javier Marquez**, **Bryan Dunlop**, **Bryce Retzlaff**, **Sophie Deam**, **John Kirchner**, **Savanna Johnson**, **Ethan Markey**, **Jocelyn Selan**, **Reilly Terranova**, and **Sergio Gordillo**.

**The Graubergers** – for welcoming me and treating me like one of your own.

**Matt** and **Bobbie**, **Kyle** and **Callie** – it would be difficult for someone to find better siblings. Matt taught me how to be an unwavering friend. Kyle taught me how to apply a blue-collar worth ethic to intellectual and creative endeavors. I would not be the person I am today without you all.

**Grandma Shirley** – for pushing me to "write the damn thing already." Well, I wrote the damn thing, Grandma. Now read the damn thing already. Thanks for encouraging me to take my education into my own hands.

**Mom** and **Dad** – for…well…everything. What can I say? You both worked so hard to provide me with every conceivable opportunity, teach me how to apply the skills you gifted me, and show how to overcome the hurdles that life puts in the way. You gave examples of how to approach difficulties with grace, give thanks, and remain optimistic. You are my earliest role models and I am so lucky to have you both in my life.

My dog **Mikko** – I realize he can't read (yet), but he still deserves some thanks. I guess I'll buy him a new ball or something.

Last and certainly not least – **Brandi**. Without you, I don't know where my life would be, but I know it wouldn't be as good. We helped each other claw happiness back into our lives and grow to become better people. I am so thankful for your love, support, kindness, eyerolls, honesty, and laughter. I am glad we have each other to lean on. I can't imagine a better person to travel through life with. Thank you for everything. I love you!

TABLE OF CONTENTS

# CHAPTER 1: AN INTRODUCTION TO KEYNESIAN WEIGHT

*"A little reflection will probably convince the reader that this is a very confusing problem." John Maynard Keynes,* A Treatise on Probability*, 85.*

## §1.0 INTRODUCTION

John Maynard Keynes altered economic orthodoxy enough to earn himself the title "inventor of macroeconomics."[1] For that reason, Keynes is typically remembered for his revolutionary work in economics—people rarely think of Keynes as a philosopher or statistician. However, Keynes's 1921 *A Treatise on Probability* (hereafter *TP*) outlines a thorough epistemic and statistical methodology. In *TP*, Keynes describes a *logical* approach to the probability calculus, which is an approach that makes the tie between statistics and philosophy quite strong. Thanks to the work done in *TP*, Keynes possesses some claim to fame as an epistemologist, statistician, and philosopher in addition to his usual economics honorifics.

Unfortunately, Keynes's logical approach to probability faces potentially insurmountable problems. Since the outset, Keynes faced criticisms for his unusual idea that probability arises from a rational assessment of the relation between the premises and conclusion of an argument.[2] Those criticisms, which began in earnest with Frank Ramsey's review of *TP*, only gained steam with the rise of the *personalist* or *subjective* approach to Bayesian probability.[3] According to the traditional division of probability interpretations,

---

[1] Wapshott, *Keynes Hayek*, 121, 196; Skidelsky, *Keynes: A Very Short Introduction*, 123.

[2] Faulkner, Feduzi, McCann, Runde, "Knight, Keynes After 100 Years," 857, 858.

[3] See Ramsey, "Truth and Probability."

the Keynesian logical approach to probability is an *objective* Bayesian approach, in stark contrast to subjective Bayesianism. On any Bayesian interpretation to probability, probabilities represent degrees of belief.[4] While the Keynesian approach to probability required Bayesian statisticians to attempt to determine *rational* degrees of belief free from personal bias, subjective Bayesianism allows statisticians to think of probability simply as degrees of belief, with only a few restrictions on how one arrives at those degrees of belief.[5] Another key difference between the Keynesian approach to probability and subjective Bayesianism is that Keynesian probabilities are sometimes wholly incomparable. These two idiosyncrasies, along with a few other problems, led to a large-scale abandonment of the Keynesian approach to probability. In the meantime, Ramsey, Bruno De Finetti, and Leonard J. Savage pioneered subjective Bayesianism by constructing a thorough apparatus for using people's behavior to interpret and quantify degrees of belief.[6] The abandonment of the Keynesian approach on one hand and the thorough quantitative explication of subjectivism on the other hand led to subjective Bayesianism becoming the orthodox form of Bayesianism.

Despite the drawbacks of Keynesian probability and the rise of a competing form of Bayesianism, Keynes's *TP* is far from a theoretical wasteland. Even after 100 years, aspects of the *TP* continue to generate interest. One particularly intriguing concept in Keynes's *TP*

[4] Curd, Cover, and Pincock, *Philosophy of Science*, 620. Jane Friedman gives a succinct account of degrees of belief by describing them as "doxastic attitudes whose strengths can be measured with real numbers in the unit interval, and which are normatively bound by the axioms of the probability calculus" ("Inquiry and Belief," 307).

[5] Curd, Cover, and Pincock, *Philosophy of Science*, 620.

[6] Curd, Cover, and Pincock, 620.

comes via his idea of the *weight of arguments*. At the most rudimentary level, weight of arguments (which I will refer to as *Keynesian weight* or simply *weight* in what follows) is meant to be a way of quantifying evidence independently of a probability statement. Like probability, Keynesian weight can be used to get more insight into your evidential situation. At its most basic, then, Keynesian weight is potentially another useful tool for statisticians, scientists, and decision-makers to add to their theoretical toolboxes.

Keynes began his discussion on weight by saying things such as, "The question to be raised in this chapter is somewhat novel; after much consideration I remain uncertain as to how much importance to attach to it."[7] In most of Keynes's discussion of weight, he hesitates about the relevance of the concept. Despite Keynes's hesitation, others have attempted to apply Keynesian weight to various issues. For example, Dale A. Nance proposes that optimizing Keynesian weight is a judicial role that falls under the responsibility of the courts.[8] On Nance's account, consideration of Keynesian weight gives courts an improved ability to determine when there is enough evidence available to render a verdict. Potential applications of Keynesian weight to other issues, such as diagnostic medicine, monetary policy, and decision theory more broadly, also exist. Over the course of this thesis, my examples tend to focus on issues in the philosophy of science and scientific inference. The details of how best to apply Keynesian weight to any subject are still contested; but generally speaking, whenever probability is considered, Keynesian weight can be considered as well.[9]

---

[7] Keynes, *A Treatise on Probability*, 78.

[8] Nance, "The Weights of Evidence," 278.

[9] I am grateful to Ken Shockley for his help with this section.

Regrettably, Keynesian weight faces substantial challenges prior to any sort of legitimate application. Due to the inherent difficulty of the concept, the debate surrounding Keynesian weight tends to generate more heat than light. When it comes to weight, concepts get conflated and interpretations of concepts tend to become muddled together. In this thesis, my most basic goal is to shed more light on Keynesian weight while attempting to avoid conflation and confusion. Although this thesis cannot solve every problem faced by Keynesian weight, I think it can help illuminate the way forward for solving many of those challenges. Illuminating Keynesian weight is not an easy task, and I do not expect to perfectly meet my aim.

To briefly summarize what is to come, I will say that this is a thesis largely focused on the challenge of how best to interpret Keynesian weight. In my mind, the best interpretation of Keynesian weight is an interpretation that sufficiently adheres to Keynes's own use of weight while also putting the concept in position to be useful in contemporary circumstances. For any interpretation of weight to make sense, quite a bit of set-up work will be necessary. Accordingly, the rest of this chapter focuses on continuing to introduce the concept of Keynesian weight. To that end, I first motivate the concept with a discussion of Karl Popper's Paradox of Ideal Evidence. Once the motivation for Keynesian weight is in place, I will explain various proposed applications of Keynesian weight while also discussing how weight relates to other similar concepts. Finally, I will conclude this first chapter with an analysis of two common conflations that arise in discussions involving Keynesian weight. The two conflations discussed at the end of this chapter will help us see how to find the best available interpretation of Keynesian weight, and they will provide a better understanding of the scope of the thesis as a whole.

Before digging in, it will also be helpful to provide an outline of my overall argument. As I said above, the overarching goal is to shed light on Keynesian weight by determining the best way to interpret the concept. I hope to attain my ultimate goal through consideration of two major interpretations of Keynesian weight. The purpose of Chapter 2 is to introduce and analyze those two interpretations. The upshot of that chapter is that the textual evidence from Keynes does not clearly favor one interpretation of the concept over another, which leads to an interpretative puzzle. In Chapter 3, I present I.J. Good's critique of Keynesian weight with the hope that Good's argument provides a way to approach the interpretative puzzle independently of textual evidence. In the second half of Chapter 3, I argue that a careful reading of Good's critique teaches us that the best interpretation of Keynesian weight is Jochen Runde's *relative* weight interpretation. In order to get to that point, we must first see why Keynesian weight matters.

## *§1.1 A Brief Technical Excursion*

I believe in plain speaking. Unfortunately, clear communication often requires technical terms loaded with philosophical baggage. I try to make my usage of such terms clear when they are introduced, but I recognize that background knowledge differs from reader to reader. For that reason, this section provides some preliminary definitions of two key terms, namely *probability* and *odds*. While some rudimentary understanding of these two terms is essential to my whole thesis, the details of this particular section only become directly crucial to my argument in section two of Chapter 3 (see §3.2.1). Thus, I think that readers who are familiar with probability and odds are justified in skipping this section for now, if they so please. With that in order, let's start our technical excursion into probability and odds.

In the introduction, we saw some disagreement about the nature of probability rear its head. Although specific battlelines are often shrouded in mystery, the debate over the fundamental nature of probability is well-known among philosophers, statisticians, and many scientists. Cutting through the noise of the debate, we should focus on the fact that the debate mostly focuses on the metaphysics, or fundamental nature, of probability itself. What things in the world, if any, does probability represent? How do we measure probabilities? To use fancy philosophical language, what is the ontological status of probability? These are the types of questions central to the well-known debate over probability.

Briefly setting aside tricky issues regarding the metaphysics of probability, I *think* it is fair to say that a probability (in the statistical sense) represents the proportion of the occurrence of some specific type of event to the total number of events in the same general class of events. It might help to connect this basic formulation of probability to the classic example of drawing marbles from an urn. If we say that drawing marbles from a specific urn is the general class of events, then the statistical probability of drawing a white marble from that urn is the number of possible white-marble-drawings (the specific type of event) over the total number of possible marble-drawings from the same urn (the general class of events). In other words, the probability of drawing a white marble from the urn is the number of white marbles in the urn divided by the total number of marbles in the urn. Once the process of drawing marbles from the urn begins, the statistical probability is then used to define the expected number of white marbles drawn when randomly sampling marbles from the urn.[10] Accordingly, a statistical (or numerical) probability that obeys Kolmogorov's axiomatization of probability is a number greater-than-or-equal-to zero and lesser-than-or-equal-to one.[11, 12] In other words, probability is a number that is either zero, one, or some number between zero and one.

Like probability, statistical odds is a ratio of two numbers. Statistical odds display the proportion of favorable outcomes to unfavorable outcomes.[13] Notice the structure of these ratios—probability and odds utilize the exact same number in their numerators. The difference is that the denominator of a probability is equivalent to the sum of the numerator and the denominator of an analogous odds. In other words, probability and odds are two different ways of displaying the same information about an event. The table below provides examples of odds and probability representing the same basic information in various ways.

---

[10] Thanks to Ben Prytherch for providing his helpful expertise and insight into the phrasing of this section.

[11] Hájek, "Interpretations of Probability."

[12] I limit myself to statistical/numerical probabilities here because (like every issue with respect to probability) things are contentious. Although many contemporary theorists think all probabilities are numerical, there is a long history of thinkers (including Keynes) who thought probabilities could legitimately be non-numerical. For more, see: Schum, *Evidential Foundations of Probabilistic Reasoning*, Chapter 2; and Fioretti, "Non-Numerical Probabilities Before Keynes." Entirely different sorts of probabilities also exist. For example, L. Jonathan Cohen develops *Baconian* probabilities as opposed to the traditional *Pascalian* treatment of probabilities. See Cohen, "Twelve Questions," 276-8 for a brief introduction. I am grateful to Jeff Kasser for reminding me of this detail.

[13] Technically speaking, this is a simplification. Statistical odds show the prospects of the occurrence of a specific type of event, and they do so by giving the ratio of that type of event over all *other* possible outcomes in the same general class of events. For what it is worth, the easiest example of odds for me to think about involves sports. The odds of a team winning their next game might be set at the ratio of their previous wins over their previous losses. Notice that this use of 'odds' is different from what might be referred to as 'betting odds', which represent a wager someone is willing to accept. Betting odds provide another way to represent the same basic information, but my focus is on statistical odds.

*Table 1.1: Various examples of probabilities and odds.*

| **Total Events** | **Type 1 Events** | **Non-Type 1 Events** | **Probability of Type 1 Event (fractional representation)** | **Probability of Type 1 Event (numerical representation)** | **Odds of Type 1 Event (fractional representation)** | **Odds of Type 1 Event (numerical representation)** |
|---|---|---|---|---|---|---|
| 100 | 92 | 8 | $\frac{92}{100}$ | 0.92 | $\frac{92}{8}$ | 11.5 to 1 |
| 96 | 50 | 46 | $\frac{50}{96}$ | 0.52083 | $\frac{50}{46}$ | 1.08696 to 1 |
| 2,022 | 26 | 1996 | $\frac{26}{2022}$ | 0.01286 | $\frac{26}{1996}$ | 0.01302 to 1 |
| Total | Favorable | Unfavorable | $\frac{Favorable}{Total}$ | Favorable ÷ Total | $\frac{Favorable}{Unfavorabe}$ | (Favorable ÷ Unfavorable) to 1 |

Since they use the same information, it is straightforward to derive a probability from an odds or to derive an odds from a probability. All that is required is thinking about the ratios involved and then either adding the numerator to the denominator and setting that sum as the new denominator (if one is moving from odds to probability), or subtracting the numerator from the denominator and setting that difference to the new denominator (if one is moving from probability to odds). For example, say we are presented with a probability of 6/10 and we want to derive the odds that is analogous to this probability. Recall that odds is the ratio of favorable cases to unfavorable cases. Since a probability is provided to us, we know the number of favorable cases (6) and the total number of cases (10), where the total number of cases is simply the sum of favorable and unfavorable cases. To determine the odds of the event, we need to find the number of unfavorable cases. We can find the number of unfavorable cases by subtracting the number of favorable cases from the total number of cases. In this case, we subtract 6 from 10, which is 4. Thus, the ratio 6/4 is the odds that is analogous to a probability of 6/10. Due to the structural similarities between statistical probability and statistical odds, they are often discussed as a pair. What we determine about the fundamental nature of probability usually translates to odds and vice versa.[14]

The bounds of probability imply that when a mathematical probability is stated, the probability inherently provides contextual information. The metaphysical interpretation of probability taken determines the details of the contextual information that accompanies a given probability. For an orthodox subjective Bayesian, an event with a 0 probability of occurring implies that the subject who stated the probability totally lacks belief in favor of the event's occurrence. In contrast, a frequentist interprets a 0 probability of an event happening as an indication that the event is not present in the frequencies (either past or hypothetical) relevant to the event happening. Of course, these are not the only interpretations of probability available, but they seem to currently be the most popular interpretations. Notice that Bayesians and frequentists agree that probabilities closer to zero are (for lack of a better phrase) less probable than probabilities closer to one.[15] Thus, although their interpretations of the probabilities differ, a Bayesian and a frequentist agree on the basic mathematics of probability. Regardless of the preferred interpretation, probability lies on a scale between zero and one, and the probability's location on the scale depends on the numbers in the ratio that constitutes the probability.

---

[14] I say "usually" because probability and odds convey the same information in different ways, so if we determine something fundamental about *the way in which* probability conveys information, that determination will not translate to odds. If, however, we determine something fundamental about the information itself, then that determination will apply to both probability and odds.

[15] If we want to use common English and avoid technical definitions, then we can say that a Bayesian and a frequentist should agree that probabilities closer to zero have a "lower chance of occurring" than probabilities closer to one. But, note that some frequentists may reject the sentence as I have it worded in the text above. If a frequentist builds their theoretic framework around historical frequencies rather than hypothetical frequencies, then they might find my sentence to be problematic.

Odds also come with contextual information, although that information is placed on a different scale from the scale used with probabilities. As we saw, an odds is the ratio of favorable outcomes to unfavorable outcomes. By providing a ratio of two numbers, odds communicate more contextual information than the information provided by a simple count of favorable outcomes, for example. Moreover, the mathematical definition of odds implies that the odds of an event occurring are bounded between zero and infinity. Perhaps a more specific mathematical example is in order. Consider an evidential situation that consists in two outcomes of an event occurring and nine outcomes of the event not occurring. In this example, the odds of the event occurring are 2/9 or 0.222 to 1. The more the odds are against something, the closer the odds move to zero. In cases that feature an equal number of favorable and unfavorable outcomes, the odds equal exactly 1 to 1.[16] In cases involving odds in favor of an event occurring, the odds are greater than one. For example, consider a case in which there are 100 favorable outcomes and 10 unfavorable outcomes. In this case, the odds are 100/10 which is equal to 10 to 1.

To summarize our technical excursion: we learned that probability and odds convey the same basic information, Bayesians and frequentists agree on the fundamental mathematics of probability, and probability and odds inherently provide scales that allow you to immediately contextualize the information they convey. As you can probably tell, I am trying my best to avoid letting the nitty-gritty details of the metaphysical nature of probability get in the way of the main points of this thesis. Nevertheless, completely avoiding the debate is impossible. Ultimately, I try to couch my thesis in a loosely subjective

[16] For example, modify the above case so that the odds are 9/9, which is equal to a probability of 9/18 or 0.5.

Bayesian probability framework, but I do not think that that decision should decrease the relevance of this work to someone who prefers a different interpretation of probability. I think what I say below is worthwhile—even when played in a different key.

## §1.2 Keynesian Weight Basics

In this section, I hope to roughly sketch the preliminary details of the concept I refer to as *Keynesian weight*. We will see numerous challenges to this general outline of Keynesian weight, but it is helpful to get some features of the concept front and center prior to moving forward.

To see how Keynesian weight works, let's first consider Karl Popper's Paradox of Ideal Evidence.[17] Although Popper hoped that the Paradox of Ideal Evidence would demonstrate problems with the Keynesian approach (and other broadly Bayesian approaches) to probability, it happens to provide a helpful introduction to Keynesian weight. In the "paradox", Popper asks us to consider the probability that a coin will turn up heads on the next toss. At the start of this thought experiment, we lack any knowledge about the coin other than the fact that it looks normal. Due to the normal appearance of the coin, we assume that there is no reason to suspect unfair bias in the coin.[18] As a result, we judge the probability that the coin will land on heads on the next toss to be 0.5. Then, Popper asks us to suppose that we find statistical evidence about the coin.[19] The statistical evidence shows that the coin underwent a million tosses in the past, approximately half of

[17] The paradox arises in the article "A Third Note on Degree of Corroboration or Confirmation," which can be found on 406-419 of Popper's *The Logic of Scientific Discovery*.

[18] Popper, "A Third Note," 407.

[19] Popper, 407.

which resulted in heads while the other half resulted in tails.[20] The statistical evidence is meant to be ideally favorable to the hypothesis that the coin is fair, so we assume that the evidence is trustworthy and that it gives us no reason to suspect any bias in the coin. Given this new evidence, our updated probability estimate of the coin landing heads on the next toss is 0.5. For Popper, the paradoxical part of this result is that our probability judgement fails to incorporate the *ideally favorable evidence* in any clear way.[21] In fact, the probability remains unchanged between the two stages of the example, even though the first probability incorporated quite little evidence and the second probability utilized a more robust set of evidence.

This is the point at which Keynesian weight enters the picture. Keynesian weight gives us a way to represent the change in the amount of evidence between the two stages of the example. Based on Keynes's outline of weight, the second probability in the "paradox" is *weightier* than the first probability. That is to say that the second probability judgement possesses more Keynesian weight than the first probability judgement because the second probability judgement uses more evidence to yield its result. Keynes gives a good summary of this feature of weight when he says that holding all else equal, "It seems plain that there is some sense in which a probability founded upon more evidence is superior to one founded upon less."[22] For Keynes, weight is the concept that conveys the comparative superiority of probabilities based on more substantial evidence.

---

[20] Popper, 407.

[21] Popper, 408. See also: Adler, *Belief's Own Ethics*, 252.

[22] Keynes, qtd. in O'Donnell, *Keynes: Philosophy, Economics, and Politics*, 67.

Consideration of potential applications of Keynesian weight further clarifies the nature of the concept. One obvious application of weight is to the question of when it is rational to stop acquiring information "in forming a probability judgment before making a decision."[23] Following Alberto Feduzi, I refer to the attempt to answer that question as *the stopping problem*. The stopping problem centers around the need for a rational principle for deciding when a probability judgement is "good enough" to use to make a decision. Popper's "paradox" demonstrates that probability alone is not enough to solve the stopping problem.[24] For instance, two probability judgements, A and B, can both possess a probability of 0.8, even if judgement A is based on a sample size of 10 while judgement B is based on a sample size of 100. But Keynesian weight arises independently of probabilities. Consequently, weight holds the potential to provide more leverage on the stopping problem, and our initial analysis of Keynesian weight should make it clear how that process might work. Keynesian weight is designed to convey the information about the amount of evidence utilized in a probability judgement. By basing standards for rational stopping around Keynesian weight, we may be better suited to solve the stopping problem than we would be if we tried to rely on probability alone.[25]

Another way to highlight the central features of Keynesian weight is to distinguish between what James Joyce calls "the balance of the evidence" and "the weight of the

[23] Feduzi, "Keynes's Conception of Weight of Evidence," 339.

[24] In "Probability, Anti-Resilience, and the Weight of Expectation," David Hamer argues that there is a probabilistic motivation for considering new evidence, and he provides a computer model of the relationship he describes. Hamer's argument may signal a way for probability to overcome the stopping problem.

[25] Feduzi, "On Keynes's Conception of Weight of Evidence", 345-8, features an account of how one might use weight to try to solve the stopping problem.

evidence."[26] In order to bring this distinction out, allow me to modify Popper's coin example.[27] Suppose that instead of being presented with a single coin, you are presented with Coin A and Coin B as well as trustworthy statistical evidence about each coin. In the case of Coin A, the statistical evidence indicates that the coin has undergone millions of tosses, exactly 50% of which landed on heads and exactly 50% of which landed on tails. As a result, the probability of the next toss of Coin A landing on heads is 0.5. In the case of Coin B, the statistical evidence indicates that the coin only underwent ten tosses since its production, all of which landed on heads. The resulting probability of Coin B landing on heads on the next toss is 1.0. Now we are in a position to see the distinction between balance and weight of evidence. In the case of Coin A, the evidence is perfectly balanced and there is a massive quantity of evidence showing that balance. Since it is exactly midway between 0 and 1, the probability of the next toss of Coin A resulting in heads displays the perfect balance of the evidence regarding Coin A. Since the evidence is perfectly balanced between two outcomes, the evidence fails to provide a decisive expectation about the result of the next toss of the coin. That said, the evidence about Coin A is quite weighty because there is so much of it. In the case of Coin B, the available evidence is not weighty, nor is it balanced between the possible outcomes of the next toss of Coin B. Instead, Coin B possesses a smaller evidential base, but the evidence plainly indicates that you should expect the next toss to be heads. As Joyce says, "The intuition here is that any body of

---

[26] Joyce, "How Probabilities Reflect Evidence," 158.

[27] For the purpose of this example, I assume a frequency conception of probability in addition to what Hans Reichenbach called "the straight rule." For more on the straight rule, see: Glymour and Eberhardt, "Hans Reichenbach."

evidence has both a kind of valence and a size."[28] Probability tells us the balance of the evidence, which makes probability an excellent tool for capturing the valence of the body of evidence. Keynesian weight is a measure meant to capture the size of the body evidence, rather than the balance of the evidence. As Keynes says, "The preceding paragraphs will have made it clear that the weighing of the amount of evidence is quite a separate process from the balancing of the evidence for and against."[29]

At this point, those familiar with statistics might assume that Keynesian weight is essentially the same thing as probable error.[30] If you made that assumption, then you are in good philosophical company. In 1878, Charles S. Peirce presented Keynesian-weight-like considerations as a central feature of his argument in favor of a frequentist interpretation of probability as opposed to a conceptualist (i.e. Bayesian) interpretation.[31] In that argument, Peirce recognized that as the amount of evidence relevant to a hypothesis increases, the probable error of that hypothesis decreases. Consequently, Peirce argues that probable error captures the essential functions that Keynesian weight is meant to capture—a decrease in probable error signifies an increase in weight. As Keynes later says, "The connection between probable error and weight, such as it is, is due to the fact that in scientific problems a large probable error is not uncommonly due to a great lack of

---

[28] Joyce "How Probabilities Reflect Evidence," 158.

[29] Keynes, *A Treatise on Probability*, 81.

[30] Or, perhaps you have standard error in mind. Probable error was the more popular measure at the time Keynes and Peirce wrote, but standard error is generally preferred now. While technically different (standard error is much more precise), standard error and probable error are functionally equivalent for the purposes of this thesis. While I will talk in terms of probable error in order to remain on Keynesian turf, note that the arguments about probable error should apply to standard error as well.

[31] Peirce, "The Probability of Induction," in *The Essential Peirce*, 160. See also: Kasser, "Two Conceptions of Weight," 645.

evidence, and that as the available evidence increases there is a tendency for the probable error to diminish."[32] However, Keynes argued that there can be instances in which Keynesian weight increases while probable error also increases, which indicates that the connection between probable error and weight sometimes fails to hold.[33] To demonstrate that point, Keynes gives a mathematical example in which new evidence eliminates some possibilities while also rendering previously unlikely possibilities more likely.[34] Keynes's mathematical example makes it plain to see how to sever the connection between weight and probable error. It can also be helpful to consider less-mathematically focused cases.[35] For instance, imagine you obtain new evidence relevant to your hypothesis, and the new evidence decreases the overall plausibility of your hypothesis in comparison to other hypotheses.[36] Thanks to the new evidence, your hypothesis becomes weightier. That said, the new evidence shows you more ways in which the hypothesis can be mistaken, which means that the probable error of the hypothesis also increases. As a result, probable error does not *necessarily* decrease with increases in Keynesian weight. Keynesian weight and probable error attempt to capture related, albeit distinct, phenomena.

Probability captures the balance of evidence nicely, but the Joyce example shows us that the Keynesian weight of evidence is distinct from the balance of the evidence. As such, some argue that the distinction between Keynesian weight and probability parallels F.H.

---

[32] Keynes, *A Treatise on Probability*, 82.

[33] Keynes, *TP*, 83.

[34] Keynes, *TP*, 83.

[35] By introducing a case that is less mathematically focused, I likely deviate from the technical definition of probable error. However, I think that the intuitive idea behind probable error remains present.

[36] Thanks to Ben Prytherch for his help with the wording in this sentence.

Knight's distinction between risk and uncertainty.[37] According to Knight, situations of risk are situations in which a probability can be attached to an inference. In contrast, situations of uncertainty are cases in which numerical probabilities cannot be obtained. For those who argue that weight indicates a commonality between Keynes and Knight, the thought is typically that Keynes's probability is akin to Knight's notion of risk, while Keynesian weight is akin to Knight's notion of uncertainty. However, Faulkner, Feduzi, McCann, and Runde maintain that while Keynes does in fact have a risk-uncertainty distinction analogous to Knight's distinction, Keynes's risk-uncertainty distinction actually comes via the difference between numerical and non-numerical probabilities.[38] Consequently, we should resist the urge to squeeze the distinction between probability and weight into Keynes's own risk-uncertainty distinction. The issue about numerical versus non-numerical probabilities, as well as Knight's risk-uncertainty distinction, ultimately focuses on how best to communicate probabilities. In contrast, Keynesian weight captures how robust our evidence is. Accordingly, there is at best a loose connection between Keynesian weight and the issue of how to communicate probabilities.[39] To summarize: it is wrong to think that the distinction between weight and probability is the Keynesian analogue to Knight's risk-uncertainty distinction.

[37] On page 5 of the otherwise excellent article "Keynes, Knight, and Fundamental Uncertainty," Dimand seems to make this argument by maintaining that uncertainty is characterized as the direct inverse of Keynesian weight. Of course, much depends on what sense of *uncertainty* Dimand had in mind

[38] Faulkner, Feduzi, McCann, Runde, "Knight, Keynes, After 100 Years," 863-4. Similarly, O'Donnell argues that the distinction is between known (which can sometimes be numerical) and unknown (which can only become numerical once known) probabilities. See O'Donnell, "Keynes and Knight," 1131.

[39] On page 864 of "Knight, Keynes, After 100 Years," Faulkner, Feduzi, McCann, and Runde point out that Knight comes up with a distinction that parallels Keynes's distinction between weight and probability.

However, some nuance begins to enter the frame when we consider the relationship between weight and uncertainty. Recognize that another primary function of Keynesian weight is to provide some account of the confidence we should feel in an inference.[40] In fact, this function of weight explains why the size of the evidential basis plays such a central role in determining Keynesian weight. Return to the cases of Coins A and B from above. Although the balance of Coin A's evidential base indicates that Coin A does not favor either heads or tails, the vast size of the evidential base should generate a great deal of confidence regarding our judgements about Coin A. It would take an extremely long string of heads (or tails) for the evidential balance of Coin A to cause a rational agent to expect heads (or tails) on the next toss. The large evidential basis of Coin A should cause us to expect an equal chance of heads or tails on the next toss of the coin, and an equal chance on the toss after that, and so on. In contrast, the balance of the evidence regarding Coin B should cause us to expect a particular result for the next toss, but our judgements about Coin B can be easily upended by future tosses. It is entirely possible that our current evidence regarding Coin B is misleading, and that Coin B is in fact just as fair as Coin A. Although the current sample of tosses indicates that Coin B possesses some bias towards heads, it is possible that the next ten tosses of the coin will land on tails and remove any claim to bias. Accordingly, we should feel less confident in our judgements regarding Coin B than we feel about our judgements regarding Coin A.

[40] In this sentence, I use the phrase "should feel" because my focus is on normative claims within the realm of epistemology. Whenever I talk about feelings of confidence in this work, I aim to describe what a rationally idealized agent should experience due to changes in their evidential circumstances. Taking a normative focus is somewhat different from Keynes's approach. Keynes tended to oscillate between descriptions of rational agents (in his more philosophical works) and descriptions of real people (in his economic works).

As Keynesian weight increases, then, our confidence should also increase.[41] The distinction between probable error and Keynesian weight shows that this tendency may not always prevail, though. Recall the case in which both weight and probable error increased due to evidence that raised the probability of competing hypotheses. In that case, it is plausible to say that although the new evidence increased Keynesian weight, the new evidence should decrease our confidence. This brings out one of the core complexities of Keynesian weight. When we obtain evidence that shows us more ways in which we might be wrong, that evidence is often surprising. The surprise may lead to a diminished confidence placed in our working hypothesis. However, once the initial surprise of the evidence fades, we might say that our overall confidence ultimately increased as a result of the increase in Keynesian weight. After all, the surprising evidence painted a clearer picture of the overall evidential situation. After the surprising evidence is obtained, we can better anticipate the range of evidence we might receive in the future. In such an event, perhaps the confidence placed in our initial hypothesis should diminish, but the recognition of our clearer evidential standing regarding the inquiry at hand should generate more confidence overall.

At this point, it is worth noticing all the different concepts intertwined with Keynesian weight. The previous paragraph shows that Keynesian weight is closely related

[41] I use the word *confidence* as a replacement for what Keynes calls "certainty". Keynes's terminology is often ambiguous, but I tend to interpret him as referring to a subjective feeling of certainty. The relationship between Keynesian weight and subjective feelings of certainty (or confidence) is a one way to relate weight to works throughout the history of philosophy. For instance, one might see foundations for Keynesian weight in Plato's *Meno*, Sextus Empiricus's *Outlines of Pyrrhonism*, Jean Buridan's theory of evidentness, Descartes's *Principles* and *Meditations*, Hume's *Enquiry*, and of course Peirce's *Illustrations in the Logic of Science*. These are just some prominent examples that come to mind. If you would like to explore these connections more, I suggest starting with Loeb, "Sextus, Descartes, Hume, and Peirce."

to confidence and surprise, but that the relationship between weight and these concepts is fuzzy at best. Moreover, we have seen that Keynesian weight is independent of probability and probable error, which means that any given evidential set might yield a probability measure, a measure of probable error, as well as some measure of weight. In considerations of weight, one needs to disentangle which concepts do which work. It is no wonder why Keynes said, "A little reflection will probably convince the reader that this is a very confusing problem."[42]

## §1.3 Higher-Order Evidence and Keynesian Weight

We already saw one potential use of Keynesian weight—as a potential solution to the stopping problem. Attempting to solve the stopping problem is only one of the many uses Keynesian weight has been put to in the last 100 years. Keynes himself also tied weight to the idea of liquidity preference, which is a vital measure Keynesian economics. Others connected Keynesian weight to the stability of belief, epistemic resilience, the philosophy of law, as well as the epistemology of disagreement. In this section, I hope to illuminate the connection between Keynesian weight and higher-order evidence. By analyzing that connection, we will recognize additional features of the concept. A thorough understanding of the main features of Keynesian weight will prove beneficial when we turn to our discussion of interpretations.

In recent years, higher-order evidence became a hot topic in epistemology. Put simply, higher-order evidence is *evidence about your evidence*. Quite a bit of recent work considers how higher-order evidence should affect epistemic considerations. For example,

[42] Keynes, *A Treatise on Probability*, 85.

the epistemology of disagreement centers around how disagreement with an epistemic peer should affect our doxastic lives. Although peer disagreement over an issue is not evidence directly about the issue itself, the existence of peer disagreement can serve as higher-order evidence about your current epistemic position. A multitude of interesting questions regarding the epistemic status of peer disagreement exist, but unfortunately we cannot venture too far in that direction here. I bring up disagreement simply because in my view, the existence of peer disagreement about a judgement and the Keynesian weight of a probability judgement both serve as varieties of higher-order evidence about those respective judgements. Just as the existence of peer disagreement regarding X might cause one to doubt their belief in X, a low Keynesian weight of probability judgement Y might cause one to doubt Y. In neither case does the higher-order evidence seem to directly tell you much about the truth of the judgement at hand. The mere existence of peer disagreement over X need not indicate whether X is true or false, and the low Keynesian weight of Y does not show that Y is mistaken. Neither of these species of higher-order evidence tell us what to believe about X and Y.

In other words, unlike evidence, Keynesian weight "has no *direction*, which refers to the function of evidence or reasons to raise or lower the credibility of a proposition."[43] Jonathan Adler argues that it is fallacious to think that Keynesian weight (and higher-order evidence generally) possesses directional bearing on belief. Adler calls this form of

[43] Adler, *Belief's Own Ethics*, 251. Notice that Adler's quote here is probably not true in cases where you learn that a proposition possesses a Keynesian weight close to zero. If you learn that a proposition entirely lacks evidential support, then that should probably alter the credibility of the proposition. In other words, there can be cases in which Keynesian weight possesses some directionality in the sense mentioned by Adler. For that reason, it is charitable to think that Adler does not use the word "credibility" in a literal manner in this quote.

fallacious thinking the *directionality of weight fallacy*.[44] The clearest example of the directionality of weight fallacy given by Adler comes via John Stuart Mill's argument about fallibility and free speech.[45] According to Mill, any suppression of speech is an admission of infallibility because such a suppression involves deciding that the suppressed speech is certainly false.[46] Mill argues that since we are not infallible, we should not suppress speech. In the case of Mill's argument, the directionality of weight fallacy involves the role played by fallibility in our reasoning. Instead of affecting *what* we should believe, Adler views fallibility as a type of higher-order evidence that affects confidence in beliefs. Adler argues that we often misconstrue our fallibility as negative evidence running contrary to our current belief.[47] In fact, fallibility lacks direction about what to believe.[48] Just as fallibility can show itself in cases when we make our beliefs too strong, it can also show itself when we render our belief in a proposition too low.[49] In this case, the fallacious thought comes from the assumption that fallibility implies that one will be wrong in a certain direction. For Adler, fallibility lacks that kind of directionality. When we are fallible about something, that implies that our judgement can be wrong "in either direction" as it were. So, contrary to Mill, Adler thinks that recognition of fallibility fails as a reason against the suppression of speech.

---

[44] Adler, 251.

[45] Adler, 255.

[46] Mill, "On Liberty," 85.

[47] Adler, *Belief's Own Ethics*, 256.

[48] Adler, 256.

[49] Adler, 256.

Adler clearly thought that Keynesian weight is analogous to fallibility in Mill's argument, which explains the naming of the directionality of weight fallacy. For Adler, it is a mistake to think that considerations of weight can legitimately generate some directional effect on the basic probability judgement.[50] Adler uses the directionality of weight fallacy to argue that Popper's Paradox of Ideal Evidence is not a paradox at all, but rather a fallacious yearning for higher-order evidence (in this case, Keynesian weight) to alter probability judgements.[51] Regardless of how successful Adler's critique of Popper is, the directionality of weight fallacy exposes many of the ties between Keynesian weight and higher-order evidence. Consequently, when we begin to consider interpretations of Keynesian weight, we must keep in mind that Keynesian weight should retain ties to both the lack of directionality of evidence on one hand and higher-order evidence generally on the other.[52] If the tie between Keynesian weight and higher-order evidence becomes severed, then it will not be clear whether we are still dealing with Keynesian weight. That said, I acknowledge that the interaction between the directionality of evidence and Keynesian weight needs more investigation. At least some examples in the disagreement literature suggest that at some evidential threshold, increases in higher-order evidence do generate directional bearing on belief.[53] Alas, such issues cannot concern us here.

[50] Adler, 251.

[51] Adler, 252.

[52] On 640 of "Two Conceptions of Weight," Kasser also asserts that Keynesian weight lacks directionality.

[53] See Kelly, "Peer Disagreement and Higher Order Evidence," 137 for one potential case.

## §1.4 Two More Key Points Regarding Keynesian Weight—Common Conflations

Before going any further, I will clarify two distinct (albeit related) points centered around Keynesian weight. At first glance, this section may appear like an unneeded appendix. However, seeing the subtlety of these points will be crucial to attaining a better understanding of the overall argument and avoiding quagmires of confusion in the chapters to come.

### §1.4.1 Weight of Argument vs. Weight of Evidence

First, we should recognize that there is potential for some ambiguity in discussions of Keynesian weight generated, at least in part, by the fact that the basic idea motivating weight is applicable to frameworks that utilize non-Keynesian conceptions of probability. The idiosyncrasy of the Keynesian approach to probability shines through when Keynes introduces weight. Keynes calls the concept "weight of arguments", a name that makes explicit reference to his argumentatively based approach to probability. Despite the idiosyncrasy of the name, one can read Chapter 6 of *A Treatise on Probability*, grasp the basic idea that motivates Keynesian weight, and then attempt to apply that basic idea to a different probability framework. Once the concept of weight gets separated from the rest of the Keynesian probability apparatus, it becomes natural to use the term *weight of evidence* in place of *weight of argument*. The translation of *weight of argument* into *weight of evidence* to better fit a different probability framework is not inherently problematic, provided that we do not lose valuable conceptual nuance in the translation. Nonetheless, Rod O'Donnell argues that we do lose subtleties from Keynes's discussion of weight of argument when we try to pull the concept out of its original context.[54] In fact, O'Donnell claims that much of the literature on Keynesian weight conflates the original concept described by Keynes with *weight of evidence*, which might be an altogether different concept than the one Keynes had in mind.[55] It is worth considering O'Donnell's view on this potential conflation, if only to delineate our exact focus.

For Keynes, a probability attaches to a logical argument as a whole rather than to the individual propositions of an argument.[56] As such, Keynes treated all probabilities as conditional on the arguments that they attach to. Thus, it makes no sense to speak of the Keynesian probability of an argument's conclusion separated from the premises of that argument. Like probability, Keynes saw the concept of weight to be an expression of the relation between the premises (or data/evidence) of an argument to the conclusion (or hypothesis/main claim) of an argument.[57] Just as it does not make sense to speak of the probability of a proposition divorced from an argument on the Keynesian framework, it likewise does not make sense to speak about the weight of some evidence divorced from the argument to which the evidence pertains. In other words, Keynes intended for weight to measure the evidence relevant to an entire argument rather than the evidence of a single proposition. O'Donnell's conflation argument signals the existence of subtle differences between what we might call the "weight of argument" and what one might call "weight of evidence." While *weight of argument* describes a probability-independent expression of the

[54] O'Donnell, *Keynes: Philosophy, Economics, and Politics*, 69.

[55] O'Donnell, 69.

[56] Faulkner, Feduzi, McCann, Runde, "Knight, Keynes, After 100 Years," 858.

[57] O'Donnell, *Keynes: Philosophy, Economics, and Politics*, 69.

relation between the premises and conclusion of an argument, *weight of evidence* refers to some general expression of the evidence at play in our probabilistic reasoning.

Despite the apparent difference between weight of argument and weight of evidence, it is not clear that abandoning the idea of *weight of argument* causes any real harm for our present purposes. Ultimately, the basic idea that motivates most interpretations of Keynesian weight can be separated from Keynes's general approach to probability, which largely fell out of favor. Remember that the primary question driving my thesis is this: Which interpretation of the Keynesian weight concept is it preferable for us to adopt? In order to answer that question, I will need to thoroughly explicate Keynesian weight. My explication will not center around textual evidence alone, but rather build in considerations of what we want out of a concept like Keynesian weight. That said, O'Donnell's worry about conflation should still be noted. It is a tangential issue to this thesis, but the potential conflation between weight of argument and weight of evidence implies that translating Keynesian weight to a different probability framework might cause us to lose valuable insight into what *Keynes* actually took the concept to mean. If O'Donnell's conflation argument is right, perhaps none of the interpretations discussed in this thesis are what Keynes himself meant by weight. However, whatever I say in what follows should still be useful for subjective Bayesians who wish to add a concept like Keynesian weight in their epistemological toolkit.

In reaction to O'Donnell's conflation argument, a brief note on approaches to the history of philosophy might be in order, if only to contextualize the upshot of my argument. Here, we should ask ourselves what we mean when we say that an interpretation is the "best interpretation" of a concept to take. I do not think there is a one-size-fits-all answer to that question. When determining which interpretation of Keynesian weight makes the most sense, we might be tempted to rest on our laurels once we determine which account Keynes himself would endorse (if such a task is even possible). However, I am not convinced that we should simply settle for the interpretation Keynes would prefer. Keynes's frequent remarks about the confusing and novel character of weight indicate that, at least through the publishing of *TP* in 1921, he remained perplexed about weight. Furthermore, it is possible that Keynes wrote on weight without recognizing the various ways of construing the concept. For these reasons, I think Keynesian weight is an issue ripe for a more revisionist approach. Keynes famously said, "*In the long run* we are all dead."[58] Keynes's personal long-run is now, and as such, I am not presently interested in what Keynes the living-breathing-man thought. Instead, I take interest in what Keynes can teach us about our contemporary approaches to analysis of evidence. So, in addition to textual consistency with *A Treatise on Probability*, we should also consider which interpretation of Keynesian weight makes the concept contribute most to contemporary literature. Not only is this step crucial to presenting Keynesian weight in the most charitable light, but it also puts Keynesian weight in a better position to illuminate any potential shortcomings in our current approaches to epistemology and scientific methodology. Accordingly, understand that in the chapters to come, I am trying to strike a balance between faithful consideration of the text and mining Keynes's work for conceptual resources.[59]

---

[58] Keynes, *A Tract on Monetary Reform*, 80.

[59] Perhaps the approach to the history of philosophy used in this thesis is best described as a version of Jonathan Bennett's collegial approach. See Garber, "What's Philosophical About the History of Philosophy?" for an excellent analysis of various approaches to the history of philosophy.

### §1.4.2 Two Types of Peircean Weight

Above, I mentioned that Keynesian weight can be separated from the Keynesian approach to probability. When that occurs, most people find it natural to begin calling the concept "evidential weight" or "weight of evidence." Such a change in terminology leads us to our second nuanced dispute regarding Keynesian weight, which focuses on how to use the metaphor of weight of evidence. In this dispute, it is crucial to recognize that Keynesian weight often gets conflated with the concept that I.J. Good called "weight of evidence." As we shall see in Chapter 3, Good argues that Keynesian weight is unworthy of being referred to as *weight of evidence*. The particular details of Good's critique of Keynesian weight will come into view later. For now, let us focus on the different ways in which Good and Keynes use the weight metaphor.

It did not become clear that Keynes and Good discussed wholly distinct concepts until fairly recently. In "Two Conceptions of Weight of Evidence in Peirce's *Illustrations of the Logic of Science*," Jeff Kasser shows that Charles S. Peirce anticipated both Keynesian weight as well as Good's measure of weight of evidence.[60] Kasser demonstrates how Peirce outlined both of these types of weight as early as 1878, and Kasser illustrates ways in which the various conceptions of weight of evidence became muddled in the literature since Peirce. To better illuminate the distinction between these concepts, Kasser refers to Keynesian weight as a version of *gross weight of evidence*, and (following Isaac Levi) he characterizes Good-Turing weight as a measure of *net weight of evidence*. Building on Kasser's terminology, I consider gross weight and net weight to be two types of *Peircean*

[60] In Section 1 of "The Weights of Evidence," Dale Nance anticipates Kasser's distinction between types of weight.

*weight*. Recognition of the differences between the two types of Peircean weight is necessary to understanding Keynesian weight.

As we saw above, Keynesian weight is traditionally construed as a function of the amount of evidence in question, which explains the name of *gross weight*. However, Kasser shows that Peirce did not use the weight of evidence metaphor to describe gross weight of evidence.[61] Instead, Peirce reserved the weight metaphor for descriptions of net weight of evidence.[62] Like gross weight of evidence, net weight of evidence measures the tendency of a probability to be altered by new evidence.[63] That said, these two types of Peircean weight operate in quite different ways. According to Kasser, net weight of evidence describes any measure that aggregates evidence in order to capture "the extent to which evidence favors a hypothesis."[64] In this way, net weight of evidence possesses directional bearing on belief.[65] A popular measure of net weight of evidence is the logarithm of the odds (also known as 'log-odds'), which is a measure developed by Peirce and then later independently formulated by Alan Turing and Good. As Peirce and later Good demonstrate, a Bayesian statistician can use log-odds to quantify the impact that a "piece" of evidence should have on belief.[66] So instead of indicating the amount of evidence utilized in an inference (like gross weight does), net weight of evidence demonstrates an aggregated balance of the

---

[61] Kasser, "Two Conceptions of Weight," 639.

[62] Kasser, 637, 644.

[63] Kasser, 644.

[64] Kasser, 639.

[65] Kasser, 640.

[66] Kasser, 638.

evidence. For Good, weighing evidence involves putting evidence on a mathematical scale and using that scale to measure the extent to which that evidence (dis)favors a hypothesis.

To avoid confusion, we must clarify whether we are talking about gross weight or net weight when we use the phrase "weight of evidence." Failure to do so only results in misunderstanding. In this thesis, my focus is on Keynesian weight, which is historically the paradigmatic instance of gross weight. However, both types of Peircean weight play a role in the chapters to come. For instance, Chapter 2 introduces Jochen Runde's *relative weight interpretation* of Keynesian weight, which runs the risk of transforming Keynesian weight into a form of net weight of evidence. In Chapter 3, I attempt to use Runde's relative weight interpretation in response to Good's critique of Keynesian weight. Good's critique will familiarize us with net weight of evidence. Ultimately, I think a careful reading of Good's critique demonstrates the way in which Keynesian weight remains a measure of gross weight, even on Runde's relative weight interpretation of Keynesian weight. By remaining a form of gross weight, the relative weight interpretation of Keynesian weight becomes the best interpretation of Keynesian weight available. Be that as it may, this chapter shows that Keynesian weight is shrouded in confusion. Thus, we must remain cautious in the chapters to come.

# CHAPTER 2: AN INTERPRETATIVE PUZZLE

*"It is reasonable, therefore, to be guided to a considerable degree by the facts about which we feel somewhat confident, even though they may be less decisively relevant to the issue than other facts about which our knowledge is vague and scanty." John Maynard Keynes,* The General Theory of Employment, Interest, and Money*, 148.*

## §2.0 Our Main Focus: Interpretations of Keynes

Hopefully Chapter 1 made it obvious why Keynesian weight is a very confusing problem. Chapter 2 represents a change in pace. In some ways, the present chapter retreads works already completed by Jochen Runde, Brian Weatherson, and James Joyce; those authors worked to illuminate the different ways in which Keynes presents weight.[67] However, it is crucial to outline the basic interpretations of Keynesian weight because they allow us to gain a clearer picture of the concept. Furthermore, many commentators on Keynesian weight fail to notice (or care) that Keynes often describes weight in vastly different ways.[68] Broadly speaking, two main interpretations of Keynesian weight exist in the literature.

First, Keynes's introduction of weight seems to describe a concept in which every new "piece" of evidence adds to the absolute total of evidential weight.[69] If that is all there

---

[67] See: Runde, "Keynesian Uncertainty and the Weight of Arguments," Weatherson, "Keynes, Uncertainty and Interest Rates," and Joyce, "How Probabilities Reflect Evidence."

[68] As one example, Michael Emmett Brady appears to run conceptions of weight together on page 361 of "Keynes's Theoretical Approach" when he says that a variable for weight "represents the amount of evidence or information or completeness of the information or the body of knowledge upon which the probability calculations are being based." Additionally, Karl Popper conflates Keynesian weight with Good-Turing weight of evidence, and Popper focuses solely on weight being a measure of the amount of evidence (see "Third Note," 406). Others only mention one interpretation of Keynesian weight—for example: Fox and Tversky, "Ambiguity Aversion," 585. In Chapter 3, we will see both Teddy Seidenfeld and I.J. Good focus only on one interpretation of Keynesian weight.

[69] Keynes, *A Treatise on Probability*, 78.

is to Keynesian weight, then weight appears to simply be a measure that increases monotonically with the accrual of new evidence.[70] This is the way Keynes's concept has traditionally been read, and it is how I introduced Keynesian weight in Chapter 1. For our purposes, we can call this interpretation of Keynesian weight *monotonic weight* or the *monotonic interpretation*. On the same page of Keynes's initial description of weight, he also alludes to a description of weight that is a ratio.[71] Here, weight seems to be the ratio of relevant knowledge to relevant ignorance or the ratio of relevant knowledge to total relevant information.[72] I follow Alberto Feduzi in calling this second conception of Keynesian weight *relative weight* or the *relative interpretation*.[73] In certain cases, relative weight functions quite differently from monotonic weight. To understand why I prefer a relative interpretation of Keynesian weight, we must first see how a relative interpretation can be considered viable. That is the task to which I turn next.

## *§2.1 The Basic Puzzle and Some Desiderata for Interpretations*

We are now in a position to appreciate the interpretative puzzle that confronts anyone analyzing Keynesian weight. Stated briefly, the puzzle comes down to the fact that there are at least two plausible ways to interpret Keynes's remarks about weight in *A Treatise on Probability* (*TP*) and it is up to us to determine which interpretation of Keynesian weight makes the most sense. In order to find a solution to the interpretative

---

[70] Runde, "Keynesian Uncertainty and the Weight of Arguments," 282; Feduzi, "Keynes's Conception of Weight of Evidence," 341.

[71] Keynes, *A Treatise on Probability*, 78.

[72] Runde, "Keynesian Uncertainty and the Weight of Arguments," 280-1.

[73] Feduzi, "Keynes's Conception of Weight of Evidence," 343-5.

puzzle, a list of desiderata with which to assess the interpretations will prove helpful. As such, here is a list of basic desiderata from the preferred interpretation of Keynesian weight:

1. Textual support;
2. Maintains the distinction between Keynesian weight and probable error;
3. Retains some believable (albeit nuanced) ties to rational feelings of confidence;
4. Maintains Keynesian weight's status as the paradigmatic species of gross weight of evidence, which requires the interpretation to:
    a. Preserve the connection to higher-order evidence via a lack of directional bearing on belief,
    b. And sustain the distinction between the balance of the evidence and the weight of the evidence.

The first desideratum is fairly self-explanatory. Any interpretation worth its salt will feature some degree of textual support. Additionally, Chapter 1 should make the second desideratum clear. One great achievement of Chapter 6 of *TP* comes via Keynes's recognition of the divergence between probable error and weight. Keynes provides a strong argument that shows the ways in which probable error and weight can theoretically come apart, even if that rarely occurs in our actual scientific practices. Moreover, Keynesian weight appears broad enough to apply outside of strictly statistical endeavors, while probable error is strictly a statistical concept. Thus, even if probable error and Keynesian weight turn out to be identical in statistical inference, they still cannot be identical concepts.[74] Accordingly, the interpretation of weight we accept should avoid causing weight and probable error to collapse in on one another.

The third desideratum might be less clear than the first two, so some explanation is in order. Keynes refers to weight in works other than *TP*. In Chapter 12 of *The General Theory of Employment, Interest, and Money*, Keynes clarifies that he maintains a distinction between something being *very improbable* and being *very uncertain* by making a footnote reference to Chapter 6 of *A Treatise on Probability*.[75] Keynes places the footnote at the end of this sentence: "It would be foolish, in forming our expectations, to attach great weight to matters which are very uncertain."[76] He continues by saying, "It is reasonable, therefore, to be guided to a considerable degree by the facts about which we feel somewhat confident, even though they may be less decisively relevant to the issue than other facts about which our knowledge is vague and scanty."[77] By using the word *weight* and then referring to the chapter on Keynesian weight, Keynes gives us an opening for figuring out what exactly he meant by the concept. Here, he says we should not place a large amount of weight on uncertain matters. According to Rod O'Donnell, Keynes's original terminology for weight replaced the word *weight* with the word *value*, where value is used in a somewhat economic sense.[78] With that in mind, we can safely rephrase Keynes's point as, "We should

---

[74] Thanks to Ben Prytherch for helping me see this point.

[75] Keynes also mentions the connection between weight and Chapter 12 of *General Theory* in a later letter. The relevant portion of that letter can be found on page 135 of Runde's "Keynesian Uncertainty and Liquidity Preference."

[76] Keynes, *The General Theory*, 148.

[77] Keynes, *The General Theory*, 148.

[78] O'Donnell, *Keynes: Philosophy, Economics, and Politics*, 69.

not attach great value to matters which are uncertain." In the next sentence, Keynes seems to say that we should instead value (or place great weight on) the facts about which we are confident, even if those facts are less relevant to the issue at hand.[79]

In these sentences, I see Keynes doing two things. First, he attempts to makes explicit that when we are uncertain about something, we should possess low Keynesian weight regarding that thing. Perhaps I am incorrect to think that Keynes intends to refer specifically to *Keynesian weight* in the quote from *The General Theory* above. After all, the way in which the above quote uses Keynesian weight places some strain on the relationship between weight and probable error that we saw in Chapter 1. Thus, we might wonder whether Keynes's usage aligns with the technical definition of Keynesian weight. However, I think that Keynes's reference back to the explication of weight in *TP* indicates that he did in fact have Keynesian weight in mind when writing the quote above. Furthermore, recall that Chapter 1 assumes a monotonic interpretation of Keynesian weight, but Keynes himself seemed to frequently depart from a monotonic interpretation of weight. With these considerations in mind, I think it is fair to assume that Keynes's intended referent in the above quote from *The General Theory* is Keynesian weight, and that it is the task of interpreters to reconcile that usage of weight with what Keynes says about weight in *TP*.

In the quote above, Keynes does something else of note. By saying that we should be guided by the facts that we are more confident in, the above quote from *The General Theory*

[79] It is common for Keynes to use the word 'weight' in this way. Another notable use can be found on page 2 of *A Treatise on Probability*. Cohen asserts that Keynes frequently uses 'weight' more loosely than his technical definition; see Cohen, "Twelve Questions," 272.

ties Keynesian weight to confidence in an interesting way.[80] Since Keynes mentions that uncertainty should lead to low weight, it seems to follow that the confidence Keynes talks about should be warranted due to some higher degree of certainty.[81] Since those facts possess a high degree of certainty, we should place more Keynesian weight in those facts; they are the valuable facts to consider when being guided by evidence. Of course, the connections between weight, confidence, and Keynesian (un)certainty are often difficult to comprehend.[82] In our assessment of interpretations of Keynesian weight, however, we must remember that there should be some sort of plausible connection between these concepts. If an interpretation of Keynesian weight completely severs the tie between weight and confidence, then the interpretation probably is too distanced from the concept that Keynes had in mind.

Several sections of Chapter 1 provide the basis for the fourth desideratum. As we saw in Sections 1.2 and 1.3, Keynesian weight is independent from probability, which means that Keynesian weight fails to tell us what to believe. Instead of possessing directional bearing on belief in that way, Keynesian weight measures the amount of evidence that gets incorporated into a probability statement. Similarly, the independence between Keynesian weight and probability indicates that there is a distinction between the Keynesian weight of the evidence and the balance of the evidence. While probability (as

---

[80] Rod O'Donnell, Sheila Dow, and Marco Crocco respectively argue that Keynesian weight is closely tied to confidence. See: O'Donnell, *Keynes: Philosophy, Economics, and Politics*, 73; Dow, "Keynes on Knowledge, Expectations," 115; Crocco, "Degrees of Uncertainty in Keynes," 13.

[81] By 'certainty', Keynes seems to have something like confidence in mind. Recall footnotes 37 and 38 above.

[82] Part of the difficulty stems from the fact that Keynes and contemporary commentators on Keynes use the word '(un)certainty' in several different ways.

well as net weight of evidence) provides a good measure of the balance of evidence, Keynesian weight performs a different function. We can summarize these two points, namely Keynesian weight's lack of directional bearing on belief and the distinction between the balance of the evidence and the weight of the evidence, by saying that Keynesian weight is the paradigmatic example of gross weight of evidence. The distinction between the two types of Peircean weight provides a good rule of thumb to use when considering interpretations of Keynesian weight. If an interpretation of Keynesian weight transforms the concept into the other type of Peircean weight (i.e. net weight of evidence), then that interpretation likely moves too far from the definition of Keynesian weight. Such a shift indicates a change of subject rather than a change of interpretation.

To see how the two interpretations of Keynesian weight fare with respect to these desiderata, we must start with the text itself. For that reason, the sections below begin with selections of textual evidence in support of each respective interpretation. One thing to note about these selections is that sometimes Keynes seems to switch between conceptions on the same page or even within the same paragraph. Once the selections of text are listed, I begin to flesh out the details of each respective interpretation.

## §2.2 Monotonic Keynesian Weight

### §2.2.1 Textual Evidence

> As the relevant evidence at our disposal increases, the magnitude of the probability of the argument may either decrease or increase, according as the new knowledge strengthens the unfavourable or the favourable evidence; but something seems to have increased in either case,—we have a more substantial basis upon which to rest our conclusion. I express this by saying that an accession of new evidence increases the *weight* of an argument. New evidence will sometimes decrease the probability of an argument, but it will always increase its 'weight'. (Keynes, *TP*, 78)

> Starting, therefore, with minimum weight, corresponding to *à priori* probability, the evidential weight of an argument rises, though its probability may either rise or fall, with every accession of relevant evidence. (Keynes, *TP*, 79)
>
> If we are to be able to treat 'weight' and 'relevance' as correlative terms, we must regard evidence as relevant, part of which is favourable and part unfavourable, even if, taken as a whole, it leaves the probability unchanged. With this definition, to say that a new piece of evidence is 'relevant' is the same thing as to say that it increases the 'weight' of the argument. (Keynes, *TP*, 79)
>
> The weight, to speak metaphorically, measures the sum of the favourable and unfavourable evidence, the probability measures the difference. (Keynes, *TP*, 85)

### §2.2.2 Explanation of Monotonic Weight

What I call the *monotonic interpretation* is the interpretation that dominates the literature on Keynesian weight, especially when it comes to the initial treatment of the concept. The early domination of the monotonic interpretation makes sense in part because this reading of Keynes is the most straightforward. On this interpretation, Keynesian weight ultimately comes down to the sheer amount of evidence that goes into a probability judgement. As Keynes explains, regardless of how new evidence affects a probability, we seem to have a "weightier" judgement on our hands once our judgement includes that new evidence. I call this reading *monotonic* because on it, new relevant evidence *only* ever increases Keynesian weight. Given this interpretation, the accrual of new evidence increases Keynesian weight. As such, the only way Keynesian weight can decrease is by somehow losing access to evidence. To draw an analogy to our everyday (or physical) notion of weight, the monotonic interpretation of Keynesian weight treats the accrual of evidence similarly to the addition of mass while holding the force of gravity constant—additional mass only ever increases weight.

With the monotonic interpretation of Keynesian weight, we get a clear picture of what exactly the concept does. To see this picture, consider the following example:

> Shirley has a credence of 0.8 that some event, E, will happen one year and one day hence. However, Shirley's credence is newly formed, and based upon only a handful of relevant cases. Over the next year, Shirley gathers more relevant evidence and updates her credence based upon that new evidence. After 365 days of daily evidence-gathering and updating, Shirley's credence fluctuated somewhat but ultimately settled at 0.8.

It should be clear that Shirley's updated credence possesses a much higher monotonic Keynesian weight than her initial credence possessed. Although the addition of relevant evidence made her credence fluctuate a bit, each addition of evidence increased her Keynesian weight.

One nice thing about the monotonic interpretation is that it clearly maintains the distinction between the balance of the evidence on one hand (which Keynes thinks is shown via probability) and the amount of evidence on the other hand. Since it only ever increases, monotonic Keynesian weight remains independent of the balance of evidence. Monotonic Keynesian weight also serves as evidence about your probability judgement. As we saw in Chapter 1, two probabilities can be equal even if they utilize extremely different amounts of evidence. By measuring the amount of evidence used by a probability, monotonic Keynesian weight preserves the connection between weight and higher-order evidence. In summary, monotonic Keynesian weight clearly enjoys some textual support and fares well insofar as it maintains Keynesian weight's status as the paradigmatic case of gross weight of evidence.

## §2.3 Relative Keynesian Weight

### §2.3.1 Textual Evidence

> The question to be raised in this chapter is somewhat novel; after much consideration I remain uncertain as to how much importance to attach to it. The magnitude of the probability of an argument, in the sense discussed in Chapter III., depends upon a balance between what may be termed the favourable and the unfavourable evidence; a new piece of evidence which leaves this balance

> unchanged, also leaves the probability of the argument unchanged. But it seems that there may be another respect in which some kind of quantitative comparison between arguments is possible. This comparison turns upon a balance, not between the favourable and the unfavourable evidence, but between the absolute amounts of relevant knowledge and of relevant ignorance respectively. (Keynes, *TP*, 78)

> The phenomenon of 'weight' can be described from the point of view of other theories of probability than that which is adopted here. If we follow certain German logicians in regarding probability as being based on the disjunctive judgment, we may say that the weight is increased when the number of alternatives is reduced, although the ratio of the number of favourable to the number of unfavourable alternatives may not have been disturbed; or, to adopt the phraseology of another German school, we may say that the weight of the probability is increased, as the field of possibility is contracted. (Keynes, *TP*, 85)

> In the present connection the question comes to this—if two probabilities are equal in degree, ought we, in choosing our course of action, to prefer that one which is based on a greater body of knowledge?...The question appears to me to be highly perplexing, and it is difficult to say much that is useful about it. But the degree of completeness of the information upon which a probability is based does seem to be relevant, as well as the actual magnitude of the probability, in making practical decisions. (Keynes, *TP*, 357-8)

> If, for one alternative, the available information is necessarily small, that does not seem to be a consideration which ought to be left out of account altogether. (Keynes, *TP*, 358)

### §2.3.2 Explanation of Relative Weight

As Jochen Runde points out, there are two related (albeit different in presentation) non-monotonic interpretations of Keynesian weight. Since these interpretations are closely related ways of displaying the same basic information, I consider them to be two sub-conceptions of relative weight. Taken together, these non-monotonic measures of Keynesian weight give us a robust picture of the relative weight interpretation. In this section, I will discuss these representations of relative weight as distinct. Unless noted otherwise, however, the rest of the thesis will treat relative weight as the disjunction of these two representations of relative weight.

We see the first version of relative weight in the first quote above, which is Keynes's introduction of weight in *A Treatise on Probability*. In that quote, Keynes points out that probability is not the only tool for comparing evidence, which signals the distinction between weight and probability. Notice that Keynes explicitly says that this other tool for evidential comparisons (which I take to be weight given the context of the quote), is similar to probability insofar as it is a balance of information. Probability consists in a balance between favorable and unfavorable evidence. But this new balance Keynes suggests is a balance of relevant knowledge and relevant ignorance. As such, Keynes's introduction of weight alludes to a ratio in which total evidence in our possession (or knowledge) forms the numerator and the total evidence we lack (or ignorance) forms the denominator.[83] One key thing to notice about this potential measure of weight is the way in which it parallels an odds measure. As we saw in §1.1, statistical odds display the proportion of favorable outcomes to unfavorable outcomes. This version of weight displays the proportion of relevant knowledge to relevant ignorance, making it similar in structure to statistical odds. Call this version of relative weight *relative weight$_{odds}$*.

Later in *A Treatise on Probability*, Keynes turns to potential practical applications of weight. In those sections, Keynes gives a slightly different version of relative weight. There, he says that weight captures the "completeness of the information."[84] For simplicity's sake, call this measure *relative weight$_{comp}$*. Runde explains that relative weight$_{\mathrm{comp}}$ is best thought of as a proportion (or a balance), but it is different from the ratio used to represent relative weight$_{\mathrm{odds}}$.[85] While the ratio given at the beginning of Chapter 6 parallels an odds measure,

---

83 See Dow, "Uncertainty: A Diagrammatic Treatment," page 12 for more on Keynes's use of the word 'ignorance'.

84 See the fourth quote from §2.3.1.

85 Runde, "Keynesian Uncertainty and the Weight of Arguments," 280-1.

the proportion given in the completeness of information quote parallels a probability measure. Recall from §1.1 that a probability (in the statistical sense) consists of the proportion of favorable outcomes over the total number of outcomes. To describe the completeness of information, relative weight$_{comp}$ gives the ratio of relevant knowledge over the total amount of relevant information.[86] Furthermore, we can think of the total amount of relevant information as the summation of relative ignorance and relative knowledge.[87] So just as one can easily transform an odds measure into a probability measure, it is straightforward to transform relative weight$_{odds}$ into relative weight$_{comp}$ (or vice versa). It is for this reason that I see relative weight$_{odds}$ and relative weight$_{comp}$ as two slightly different ways of construing a relative weight interpretation of Keynesian weight.

No matter the presentation, relative weight features relevant knowledge in the numerator of its ratio, and in this discussion, relevant knowledge is equivalent to the amount of evidence at our disposal. In this way, monotonic weight is built into relative weight. However, one cannot reduce relative weight to an absolute measurement of the amount of evidence involved in a judgement. Instead, relative weight involves consideration of the amount of relevant evidence already at our disposal, and then some estimation of the amount of other relevant evidence outside of our purview. As Ekaterina Svetlova says, "Importantly, the concept of weight allows a meaningful discussion of

[86] Those who are interested in philosophical pedagogy as well as those with an interest in the overlap between the work of Bertrand Russell and Keynes may notice some similarities between this version of relative Keynesian weight and the self/not-self distinction that Russell gives in Chapter 15 of *The Problems of Philosophy*.

[87] Runde, "Keynesian Uncertainty and the Weight of Arguments," 281.

*gradations of ignorance* and their dynamics."[88] The estimation of ignorance is admittedly the most difficult feature of relative weight because it asks us to quantify (or at least consider in a comparative, ordinal-ranking way), our *ignorance*.[89] Ignorance can be represented in several different ways. Feduzi seems to suggest utilizing ungrounded, second-order subjective probabilities to quantify ignorance.[90] If we use that method for quantifying our ignorance, then relative weight consists in a precise guess at how much of the relevant evidence our current probability judgement uses. On the other hand, we might try to use a little more epistemic caution when representing ignorance. A more cautious representation of relative Keynesian weight might come via the use of interval (or imprecise) probabilities as well as uninformed prior probabilities.[91] Either way, some methodologies for quantifying ignorance exist, so this quirky aspect of the relative interpretation need not count against the interpretation at this stage.

One feature unique to the relative weight interpretation is the possibility of new relevant evidence *decreasing* Keynesian weight. Relative Keynesian weight can decrease if

---

[88] Svetlova, "Relevance of Knight, Keynes, and Shackle," 999.

[89] As Runde says, "The question is whether it is possible to talk sensibly of knowing something about our ignorance, or, to be more precise, of knowing something about changes of the extent of our ignorance, on some or other proposition" ("Keynesian Uncertainty and the Weight of Arguments," 282). See also Svetlova, "Relevance of Knight, Keynes, and Shackle," 996-1002.

[90] Feduzi, "Keynes's Conception of Weight of Evidence," 345-7.

[91] I am tempted to think estimations of ignorance should be constrained in much the same way Wesley Salmon argues prior probability estimates should be "tempered" and constrained by facts (see Salmon, "Rationality and Objectivity," 525-33). However, I recognize that others may not share my temptation. The degree to which one thinks that estimations of ignorance should be constrained probably depends on how much "subjectivity" one is comfortable adding to their statistical apparatus. On a related note, Michael Emmett Brady argues that many scholars fail to read Chapters 15-17 of *TP* carefully enough to understand how Keynes himself hoped to apply weight. According to Brady, Keynes specified an "inexact numerical approximation approach" for application of the concept of weight (363). See Brady, "Keynes's Theoretical Approach," 360-3.

the new evidence shows us new ways in which we might be mistaken.[92] New relevant evidence might decrease your relative weight by showing how you have underestimated the extent of your ignorance on a topic. In such an event, the new evidence increases your level of knowledge. However, the new evidence still leads to an overall decrease in relative weight by increasing the estimate of relevant ignorance by a larger amount than the associated increase in relevant knowledge. To drive this point home, let us consider an example in which relative weight decreases with an increase in relevant evidence. Brian Weatherson provides one such example.[93] Weatherson describes a scenario in which we are playing a simple type of poker that involves betting on who has the best hand. As Weatherson says, "Before the bets start, I can work out the chance that some other player, say Monica, has a straight."[94] However, once betting starts, the information Monica provides via facial expressions, gestures, and vocal tones may decrease the relative weight I put on the chance that Monica has a straight, while keeping my probability estimate of

---

[92] On page 1003 of "Relevance of Knight, Keynes, and Shackle," Svetlova argues that we can also learn without gaining new information, provided that we reexamine the information we possess and discover new connections. This indicates that relative weight could also decrease via reexamination of an evidential situation, even if no new evidence is gained. As Svetlova explains, orthodox Bayesian methodology does not seem to allow for changes in beliefs without gaining new information. Thus, adding relative weight to the Bayesian apparatus would help Bayesians overcome this potential downside of their approach.

[93] Weatherson, "Keynes, Uncertainty and Interest Rates," 51-2. Weatherson's example is meant to demonstrate the difference between precise and imprecise credences, which implies that it is intended to focus on the risk-uncertainty distinction. However, the example seems capable of fitting our needs. Numerous other examples in which new evidence actually decreases relative weight also exist. For instance, Michael Lewis's book *The Big Short* describes many scenarios in which new evidence about the stock market in the leadup to the 2008 financial crash decreased the "weight" interested parties put on their credences regarding the stability of the stock market. Another recent real-life example might be James Comey reopening the investigation into Hillary Clinton in the leadup to the 2016 U.S. presidential election, which likely led to a decrease in the relative weight of most election forecast models.

[94] Weatherson, "Keynes, Uncertainty and Interest Rates," 52.

Monica's straight constant.[95] Monica's bets provide additional knowledge that gets incorporated into the relative Keynesian weight I possess about the proposition that she has a straight. Nonetheless, the way in which Monica bets creates doubt, which increases my estimation of my ignorance. Since my estimation of my ignorance increases at a faster pace than my knowledge about Monica's hand, Monica's betting leads to a net decrease in the weight I put on my estimation about her hand being a straight. Regardless of whether relative weight is represented in the odds-parallel form (relative $\text{weight}_{\text{odds}}$) or in the probability-parallel form (relative $\text{weight}_{\text{comp}}$), Keynesian weight decreases when the ratio's denominator grows at a faster rate than the growth in the ratio's numerator. The upshot is that unlike a monotonic interpretation, relative Keynesian weight allows new evidence to decrease weight.

It is also possible to see key events in the history of science as instances in which new evidence decreased relative Keynesian weight. For example, imagine the potential evidential situation faced by a 17th century astronomer who learns of Galileo's telescopic discoveries. Suppose the astronomer is a relatively unbiased observer who is well-trained in Aristotelian physics, Ptolemaic astronomy, and Copernican astronomy. It is rational for such an astronomer to be confident in Aristotelian physics, especially because that system formed the foundation of her entire scientific framework. But it is also rational for her to be confident in the results learned through Galileo's telescopic evidence. Of course, the problem is that the telescopic evidence is incompatible with large swaths of Aristotelian

[95] Weatherson, 51. You might doubt that the probability estimate remains constant, but the example can be set up to guarantee that the overall probability estimate does not change. For instance, perhaps Monica's betting affects my probability estimate such that it fluctuates throughout the hand but ends up equivalent to the initial estimate.

physics. Thus, we can imagine our astronomer being thrown into a state of doubt by the new evidence from Galileo's telescope. Although the evidence from the telescope increased her knowledge, it did so in a way that greatly increased the astronomer's estimation of her ignorance. Accordingly, the new evidence decreased her relative Keynesian weight.[96]

## §2.4 Summary of Interpretations

With the multitude of interpretations floating around, this chapter is admittedly somewhat complex. For that reason, I will now briefly summarize the ground covered so far and create a table of the pertinent interpretations of Keynesian weight.

At the outset of this chapter, we saw that there are two plausible ways to interpret the concept of Keynesian weight. First, there is the monotonic interpretation. On the monotonic view of weight, new relevant evidence necessarily increases Keynesian weight. According to the monotonic view, Keynesian weight is equivalent to the total amount of relevant evidence in our possession. As we saw, the monotonic interpretation of Keynes is generally the one accepted in the literature on weight. Despite being the most popular in the literature, the textual evidence on weight does not decisively count in favor of the monotonic interpretation.

Second, there is the relative weight interpretation. Relative weight comes in two flavors, which are in rows two and three of the table below. First, there is the

[96] Note that this example need not correspond to the evidential situation of any actual person in history—it should simply paint a plausible picture of what the evidential situation of such an astronomer might believably be. To make the changes in Keynesian weight sharper, we can even imagine a possible world in which, simultaneously with Galileo's telescopic evidence, some theoretical advance in another area of Aristotelian physics occurs and that advance makes Aristotelian physics more robust. In that case, a large portion of the change in the astronomer's evidential situation will be reflected via weight rather than via probability. A different paper would consider whether decreasing Keynesian weight is a symptom of a Kuhnian paradigm shift.

representation of relative weight that parallels an odds ratio, which I have called relative weight$_{\text{odds}}$. An equivalent representation of relative weight$_{\text{odds}}$ is relative weight$_{\text{comp}}$. Relative weight$_{\text{comp}}$ has the same numerator as relative weight$_{\text{odds}}$, but the denominator of relative weight$_{\text{comp}}$ is the summation of the numerator and denominator of relative weight$_{\text{odds}}$. Consequently, these representations of relative weight are merely different ways of displaying the same information. The unique aspect of the relative weight interpretation is that it allows new relevant evidence to result in a decrease of overall Keynesian weight. Although only a few recent commentators on Keynes have adopted the relative weight interpretation, the available textual evidence indicates that it is a plausible view of Keynesian weight.

*Table 2.1: List of names, suggested notations, and textual support of the considered interpretations of Keynesian weight. In the notation column, K is short for 'Knowledge' and I is short for 'Ignorance'. Suggested notation is based on Runde's "Keynesian Uncertainty and the Weight of Arguments," 280-1.*

| **Interpretations of Keynesian Weight** | | |
|---|---|---|
| **Name** | **Suggested Notation** | **Textual Support (pages)** |
| Monotonic Weight | $K$ | *TP*: 78, 79, 85 |
| Relative Weight$_{\text{odds}}$ | $\frac{K}{I}$ | *TP*: 78, 85 |
| Relative Weight$_{\text{comp}}$ | $\frac{K}{I+K}$ | *TP*: 85, 357-8, 358 |

## §2.5 Revisiting the Interpretative Puzzle and Desiderata

At this stage, the tension between the monotonic and relative interpretations of Keynesian weight might remain cloudy. Monotonic weight measures absolute knowledge. As such, monotonic weight is identical to the numerator of relative weight. That relationship between the two interpretations implies that all cases of monotonic weight can be explained in terms of relative weight, making relative weight into an instance of conceptual expansion. In fact, the relationship between monotonic weight and relative weight actually gives a fairly clean case of (a philosophical version of) logical positivist reductive explanation.[97] As a result of this close relationship between the two interpretations, the interpretations do not result in any meaningful differences in a large swath of cases. However, the interpretations of Keynesian weight result in meaningful differences in some interesting cases, such as those in which relative Keynesian weight decreases in response to new evidence. As we saw, the monotonic interpretation dictates that new evidence only ever results in increases to Keynesian weight. Despite initial impressions otherwise, the two interpretations sometimes diverge.

So, the two interpretations of Keynesian weight lead to meaningful differences. What creates an interpretative puzzle? Put simply, Keynes's failure to realize that his comments on weight generate distinct interpretations of the concept creates the puzzle. It is the job of commentators on Keynes's work to determine which version of weight is closer to what he took the concept to be, provided that there is an answer to that sort of question (maybe Keynes had no clear idea about what he meant). Most of the literature that touches on Keynesian weight seems to fixate on the quotes from Keynes that mention weight increasing with increases in new relevant evidence. Even though many commentators mention that Keynesian weight is meant to give a measure of the completeness of our information, the monotonic weight interpretation still dominates the formal treatment of the concept.[98] Often, the relative weight interpretation fails to be

[97] For more on reductive explanation, see Ernest Nagel, "Issues in the Logic of Reductive Explanations."

[98] While an excellent article overall, Nance's "The Weights of Evidence," provides an exemplary instance of mentioning the completeness of information aspect of weight without incorporating a relative weight

mentioned, even when an author regards Keynesian weight as a measure of completeness of evidence. Regardless of the failure of commentators to recognize these distinct conceptions of weight given by Keynes, we should now be in a position to recognize that the textual evidence does not decisively tell us which interpretation Keynes himself preferred. Keynes wavers between interpretations on the same page, and he seems to frequently make references in which either interpretation can plausibly be the intended referent. That leaves us with at least two distinct and plausible interpretations of Keynesian weight and no authoritative source on which interpretation should be adopted. As such, we have an interpretative puzzle on our hands.

The existence of an interpretative puzzle demonstrates that textual evidence fails to tell us which interpretation of Keynesian weight to prefer. Luckily, we already saw other desiderata to consider when choosing the preferred interpretation. Remember that for any interpretation of Keynesian weight to make sense, the interpretation must have a plausible relation to the issues that Keynes brings to bear on the concept. Disconformity with what Keynes says about weight generally should signal that we have ceased to deal with something worth calling Keynesian weight. As such, I will spend the next two paragraphs outlining impressions of how the two interpretations fare with respect to our desiderata for interpretations of Keynesian weight.

Following textual support, the second item thing to be desired from an interpretation of Keynesian weight is an interpretation that maintains the distinction between weight and probable error. Luckily, I do not think either of the proposed

interpretation into the formal treatment of weight. Chapter 4 of O'Donnell's *Keynes: Philosophy, Economics, and Politics* provides another instance of an assumption of a monotonic interpretation.

interpretations of Keynesian weight obviously creates an identity between weight and probable error. That said, I am tempted to give a slight edge to the relative weight interpretation when it comes to maintaining an intuitive understanding of the distinction between weight and probable error. From what I can tell, many of the instances of divergence between weight and probable error are precisely those instances in which one is forced to reevaluate their estimation of their relevant ignorance. Sometimes, new evidence indicates that the field of inquiry potentially contains what Nassim Nicholas Taleb calls *black swans* — high impact events with low predictability.[99] In these cases, perhaps the new evidence gives us a clue regarding more possibilities, which increases our probable error. However, the new evidence itself (evidence that directly alters our estimation of our ignorance) can simultaneously increase our Keynesian weight. In other words, although probable error increases, the new evidence renders what is to come less surprising. Since most of the effects of the evidence in such a case seem tied to the estimation of ignorance, I am willing to say that relative Keynesian weight might do a better job at explaining the difference between weight and probable error. However, I am not wholly convinced of this point. Probable error maintains a precise technical definition, and I am not sure that monotonic weight performs poorly at explaining all cases in which probable error and Keynesian weight simultaneously increase or decrease. For that reason,

[99] Taleb, *The Black Swan: The Impact of the Highly Improbable*, xxii. Taleb frequently talks about black swans possessing low probability (i.e. being outliers), but he also frequently describes them as events with low predictability. I am not convinced that a low probability event is the same as an event that possesses low predictability. A great deal depends on what exactly we mean by "predictable" and "probable." Despite these complications, I think black swans are helpful for analyzing the sorts of situations described in this paragraph. In fact, Runde argues that thorough investigation of black swans requires an apparatus capable of handling shades of what Taleb calls "unknowledge" (i.e. ignorance). Relative Keynesian weight provides just that sort of apparatus. See Runde, "Dissecting the Black Swan," particularly pages 495-500, for more information.

I think both interpretations fare roughly the same on our second desideratum, with the chance that relative Keynesian weight performs slightly better.

The third desideratum given above focuses on the relationship between Keynesian weight, certainty, and confidence. This desideratum is the point at which relative Keynesian weight and monotonic Keynesian weight truly begin to separate. We saw above that by definition, monotonic Keynesian weight cannot decrease. Additional evidence only increases monotonic Keynesian weight. However, it is obvious that confidence can decrease. For example, think of a case of disagreement with an epistemic peer. Although the existence of disagreement does not affect whether your belief is true, the discovery of disagreement with someone you consider to be your peer will probably cause you to doubt your judgement. That doubt signifies decreased confidence. Insofar as the purpose of Keynesian weight is to measure (even if that measure is rough) our feelings of confidence, monotonic Keynesian weight provides too blunt of a tool. By allowing Keynesian weight to decrease, relative Keynesian weight creates a closer connection to feelings of confidence. For that reason, relative Keynesian weight performs better when it comes to our third desideratum.

Nonetheless, the interpretative puzzle remains intact. Is relative Keynesian weight still a form of gross weight of evidence, or is it a form of net weight of evidence? Recall that like gross weight of evidence, net weight of evidence gives us an idea of a probability's tendency to be altered by new evidence.[100] However, net weight operates differently than gross weight. Net weight of evidence describes any measure that aggregates evidence in

[100] Kasser, "Two Conceptions of Weight," 644.

order to show the extent to which evidence favors a hypothesis.[101] By representing our evidential state via a ratio of knowledge and ignorance (or knowledge and information) relative Keynesian weight seems to move closer to being a measure of net weight of evidence rather than gross weight of evidence. Consider that a high relative Keynesian weight indicates that the amount of relevant information we possess (i.e. knowledge) is greater than the amount of relevant information that we lack (i.e. ignorance). This might seem to suggest that a high relative weight indicates—via an aggregation of our evidence—that the evidence greatly supports the hypothesis, regardless of the content of the hypothesis. At this point, it is tempting to slip into thinking that relative weight provides direct support to particular hypotheses. Unfortunately for relative weight, Keynes obviously did not intend for his concept of weight to indicate how evidence favors or disfavors a hypothesis. For instance, Keynes says, "The preceding paragraphs will have made it clear that the weighing of the amount of evidence is quite a separate process from the balancing of the evidence for and against."[102] As such, the relative weight interpretation runs the risk of transforming Keynesian weight into the wrong type of Peircean weight, namely net weight. On the other hand, the monotonic interpretation clearly makes Keynesian weight into a form of gross weight. After all, the paradigmatic example of gross weight of evidence is monotonic weight. In Chapter 3, we will discover why it is wrong to think that relative Keynesian weight is a form of net weight of evidence. At this point

[101] Kasser, 639.

[102] Keynes, *A Treatise on Probability*, 81.

however, monotonic weight appears to perform better than the relative weight interpretation when it comes to our fourth desideratum.

In summary, neither textual support nor the relationship to probable error decisively favors either of our plausible interpretations of Keynesian weight. Furthermore, while relative weight performs much better than monotonic weight with respect to the connection to confidence, relative Keynesian weight runs the risk of transforming Keynesian weight into the wrong type of Peircean weight. Thus, even after consideration of the desiderata for Keynesian weight, an interpretative puzzle remains. For that reason, it seems likely that we will need to resort to reasons independent of Keynes's writing on the topic of weight in order to determine which interpretative lens we should use when dealing with Keynesian weight. In Chapter 3, I show how relative weight and monotonic weight handle I.J. Good's critique of Keynesian weight. Perhaps a Keynesian can live with the drawbacks of one of these interpretations if that interpretation overcomes a popular critique of Keynesian weight.

# CHAPTER 3: A GOOD LESSON

*"No concept is fundamental if only statisticians use it." - I. J. Good, "Weight of Evidence: A Brief Survey", 250.*

## §3.0 Chapter Overview

In this chapter, I turn my attention to a prominent critique of Keynesian weight. The critique that I focus on comes from Irving John Good. Good is best known for his work alongside Alan Turing during the Allied code-breaking efforts at Bletchley Park in World War II. Amidst their efforts to break the Nazi Enigma Code, Turing and Good co-developed a measure that they called *weight of evidence*.[103] In his later explication and defense of weight of evidence, Good critiqued Keynesian weight as a competing version of weight of evidence. I will now explain both the context and the substance of Good's critique of Keynesian weight. Once the critique has been explained, I will then demonstrate how a defender of Keynes might use a relative weight interpretation in an attempt to circumvent Good's main points of criticism. Ultimately, I argue that careful analysis of Good's critique teaches us a lesson about the relative interpretation of Keynesian weight. Namely, we learn that relative Keynesian weight is indeed a type of gross weight of evidence, which helps us solve Chapter 2's interpretative puzzle in favor of the relative weight interpretation.

## §3.1 Good's Critique

### §3.1.1 Context: Good-Turing Weight

In 1985, Good provided a survey and review of weight of evidence. In that survey, Good's primary aim was to summarize the conception of weight of evidence that he

[103] Hodges, *Alan Turing: The Enigma*, 247-9.

developed in conjunction with Turing during their code-breaking efforts. As an aside, I try to refer to Good's preferred version of weight of evidence as *Good-Turing weight*. Good's survey of Good-Turing weight includes responses from statistician Herman Rubin as well as philosopher Teddy Seidenfeld. Good concludes the survey with his reply to those responses. Although the initial survey makes no mention of Keynesian weight, Keynesian weight plays a key role in Seidenfeld's response to Good. Consequently, the version of Good's critique of Keynesian weight that I focus on is nested within a defense and promotion of Good's own concept of weight. Since Good offers Good-Turing weight as a competitor to Keynesian weight, it will be helpful to begin with an overview of Good-Turing weight. Recognize that I do not intend to give a thorough analysis of Good-Turing weight; for that, the reader should turn to Good's survey itself. Instead, I aim to summarize enough of the Good-Turing conception of weight of evidence in order to contrast Good-Turing weight with Keynesian weight.[104]

Good begins the survey with analysis of the metaphor "weight of evidence." He considers the ways in which juries, detectives, and doctors "weigh their evidence" when making decisions.[105] At the outset, it becomes clear that Good takes weight of evidence to

---

[104] One thing to recognize off the bat is that while Turing seems to be the British originator of the mathematics behind Good-Turing weight of evidence, Good is the one who proposes the metaphor of "weight of evidence." As such, both Good and Turing have legitimate claim to being the eponyms of this version of weight. However, Kasser points out that Peirce arrived at Good-Turing weight of evidence close to 60 years prior to the work by Turing and Good. One of Turing's innovations was to name the units of weight of evidence, but there is evidence that Peirce thought about units in a similar way to Turing. Turing called the units of weight of evidence *bans*, which refers to the town of Banbury. Banbury is where the sheets of paper used in Turing's procedure for calculating bans at Bletchley Park were printed. Although a ban is the base unit, Good and Turing most commonly spoke in terms of *decibans* (i.e. 1/10th of a ban). See Kasser, "Two Conceptions of Weight," 637-9 for details. The history of Good-Turing weight is an excellent example of Stigler's law of eponymy (Chapter 14, *Statistics on the Table*).

[105] Good, "Weight of Evidence," 249.

be closely related to the imagery of using a double-pan balance scale. Like using a double-pan balance to compare the weights of objects, Good thinks weight of evidence should show us how the evidence in favor of a hypothesis balances against the evidence contrary to that hypothesis. Good takes the balancing-scale metaphor so seriously that throughout the survey, he makes frequent reference to Themis (the Greek goddess of justice), who "is usually represented as carrying a pair of scales."[106] Good makes it abundantly clear that he takes the process of weighing evidence to be fundamental to human life. He even admits that he takes the concept of weight of evidence to be at least as important as the concept of probability.[107] Good's preferred conception of weight of evidence involves some intricate mathematical and statistical theorizing, but his overall aim is to provide a measure that aligns with our common-sense notion of weighing evidence prior to making a decision. As he says:

> I believe that the basic concepts of probability and weight of evidence should be the same for all rational people and should not depend on whether you are a statistician. There should be a unity of rational thought applying, for example, to statistics, science, law, and politics. This assumption will set the tone of my survey. No concept is fundamental if only statisticians use it. (Good, "Weight of Evidence: A Brief Survey, 249-50)

Since Good takes the double-pan balance weighing process to be so fundamental to human life, he thinks concepts that fail to make sense of that particular metaphor are not worthy of the title *weight of evidence*. Any concept that deviates from this metaphor in a substantial manner is too revisionary to count as weighing evidence.

[106] Good, "Weight of Evidence," 249.

[107] Good, "Weight of Evidence," 249.

Notice what the balancing metaphor tells us about Good-Turing weight. Both Keynesian weight and Good-Turing weight are forms of Peircean weight. While Keynes focuses on gross weight of evidence, the references to balancing evidence show that Good focuses on net weight of evidence.[108] Good thinks that weight of evidence ought to provide a representation of the way in which evidence supports (or fails to support) a hypothesis.[109] As such, Good-Turing weight of evidence is meant to capture what happens to our belief when we say things such as, "That evidence tipped the scales in favor of the defendant."

Now that we understand the metaphor Good-Turing weight is meant to capture, we can briefly consider the mathematics of Good-Turing weight of evidence. Put in the simplest possible terms, Good-Turing weight is the logarithm of the odds (or *logodds*) of a hypothesis. To see how Good-Turing weight of evidence works, we need to unpack what that means.

First, recognize that since Good-Turing weight refers to the odds of *a hypothesis*, it assumes the Bayesian metaphysics of probability and odds. It takes more than a little stretching to make the frequentist metaphysics of probability align with the idea of *a hypothesis* possessing a probability between zero and one. To understand this point, it helps to first realize that frequentism about hypotheses is different from frequentism about events. Frequentists usually take probabilities (and thereby odds) to correspond to real-world frequencies. In the case of events, we can repeatedly sample real-world frequencies

[108] Kasser, "Two Conceptions of Weight," 637-9.

[109] Kasser, 640.

to find the probability of the event. However, a hypothesis cannot be repeatedly sampled. A hypothesis is either true of our world or it is false of our world. In the former case, the frequentist probability of the hypothesis being true is one, and in the latter case, the frequentist probability of the hypothesis being true is zero. To report a probability between these extremes would indicate that the frequentist, to quote Peirce, believes that "universes were as plenty as blackberries."[110] For the traditional frequentist, there is no frequentist probability of a single case that falls between zero and one. In contrast, since Bayesians think probabilities reflect degrees of belief, they can easily talk about meaningful probabilities of single hypotheses. The Bayesian need not check the world for an entity that corresponds to probability or odds.[111] Thus, for the Good-Turing conception of weight of evidence to be a conception of any interest whatsoever, it must treat probability as a measure of degrees of belief. In other words, Good-Turing weight inherently utilizes some of the orthodox Bayesian metaphysics of probability and odds.

Second, we should understand why Good-Turing weight features a logarithmic transformation of the odds of the hypothesis. Without getting bogged down in the mathematical details (which involve consideration of the quotient rule for logarithms), recognize that taking the logarithm of odds provides a mathematical picture of our evidence that is strikingly similar to a double-pan balance. In this case, the center of the scale is the number 0, while the pans of the scale span the infinite number line on either side of zero. To see how Good-Turing weight works on this scale, consider the odds for a

---

[110] Peirce, "The Probability of Induction," in *The Essential Peirce*, 165.

[111] Technically, they might need some correspondence to betting behavior in order to utilize meaningful probabilities, but that is quite different from the need to find a real-world frequency.

hypothesis and the corresponding odds of the negation of the hypothesis. Let's start by assuming that the odds for the hypothesis and the odds for its negation are equivalent at a ratio of 50/50. In this case, the logarithm of the odds for the hypothesis is zero, which means that the evidence is perfectly balanced on our scale. Now, assume that the odds for the hypothesis are 70/30. Because the odds for the negation of the hypothesis are simply the reciprocal of the odds of the hypothesis, the odds for the negation of the hypothesis are 30/70. In this case, the logarithm of the odds for the hypothesis is 0.3678, while the logarithm of the odds for the negation of the hypothesis is -0.3678. By taking the logarithm of the odds of the hypothesis, we arrive at a scale for balancing the evidence relevant to the hypothesis. The scale centers itself at zero, which represents perfectly balanced evidence. Negative log-odds values indicate evidence against the hypothesis while positive values indicate evidence in support of the hypothesis.

The mathematical apparatus outlined in the previous paragraph is the fundamental feature of Good-Turing weight of evidence. By seeing how the evidence regarding a hypothesis falls on this scale, we can see how the evidence (dis)favors our hypothesis. The log-odds weighs evidence in a way that parallels the process of using a balance scale to compare the weights of objects. We have now seen that Good-Turing weight of evidence is a conception of net weight of evidence, it is fundamentally Bayesian, and it features a mathematical apparatus that parallels a double-pan balancing scale.[112] In the next section, I summarize Seidenfeld's response to Good's survey as a way to introduce Good's critique of Keynesian weight.

---

[112] The reader may also notice that Good-Turing weight measures what we saw Joyce call "the balance of evidence."

### §3.1.2 GOOD'S CRITIQUE OF KEYNESIAN WEIGHT

Seidenfeld begins his response to Good by drawing a humorous comparison between philosophy and horseradish. Paraphrasing a character in *Dr. Zhivago*, Seidenfeld says that philosophy

> is good if taken in small amounts in combination with other things. But it is not good in large amounts by itself. The risk with philosophy, as with horseradish, is the temptation to use ever stronger concentrations to maintain the sensation of that first taste. Soon you are serving up pure horseradish! (Seidenfeld in Good, "Weight of Evidence," 264. ).

Seidenfeld's discussion of horseradish illuminates his worry about Good's weight of evidence. For Seidenfeld, Good's explication and application of weight of evidence risks being pure philosophical horseradish rather than the clarification of a useful tool for decision-making. Seidenfeld's main question for Good comes down to how to use weight of evidence in a decision. Like many, Seidenfeld does not see a straightforward application of Good-Turing weight of evidence.[113] Seidenfeld analyzes two potential applications of weight of evidence, neither of which he considers fruitful uses of the concept.

In the first potential application of Good-Turing weight discussed by Seidenfeld, Seidenfeld introduces Keynesian weight in an analysis of the stopping problem.[114] Seidenfeld says, "For Keynes, *weight of evidence* cannot be defined by probability as he sees weight monotonically increasing with increasing evidence."[115] From this quote, we see that Seidenfeld maintained a monotonic interpretation of Keynesian weight, which will prove

---

[113] Seidenfeld in Good, "Weight of Evidence," 265.

[114] See §1.2 for a refresher on the stopping problem.

[115] Seidenfeld, 265. While Seidenfeld talks about Keynes's conception of "weight of evidence", §1.4.1 showed that Keynes never used that specific phrasing. As such, it is not clear that Keynes would see Keynesian weight as a competitor to Good-Turing weight.

crucial to our analysis of Good's critique. Seidenfeld continues by noting the way in which Keynes considered tying weight to the stopping problem, but that Keynes eventually concluded that the application of weight of evidence is a "very confusing problem."[116] Seidenfeld then summarizes Good's own contributions to the stopping problem, but points out that "the general theory of optimal stopping is tangential to Good's concept of *weight*."[117] The intended upshot is that weight of evidence (of any variety) fails to substantially figure into solutions to the stopping problem.

Good couches his critique of Keynes in his reply to Seidenfeld. I now quote Good's treatment of Keynes at length:

> Keynes's definition of weights of arguments, in which he puts all the weight in one scale, whether they are positive or negative, is like interpreting weight of evidence as the weight of the documents on which they are printed. I think, if not horseradish, it is at least a crummy concept in comparison with the explicatum of weight of evidence that I support. Keynes himself said of his discussion (1921, p. 71) "...after much consideration I remain uncertain as to how much importance to attach to it. The magnitude of the probability of an argument…depends upon a *balance* between what may be termed the favourable and unfavourable evidence…". In other words he clearly recognizes that Themis is right to use both scales. It is a standard English expression that the weight of evidence favours such and such. Of course this refers to the *balance* of the evidence, *not* to the sum of all the pieces irrespective of their signs.
>
> If you must have a quantitative interpretation of Keynes's "weight of arguments", just compute the weights of evidence in my sense for each "piece" of evidence and add their absolute values. This then is yet another application of my explicatum, to give a somewhat quantitative interpretation to the crummy one. But Keynes's discussion of this matter is purely qualitative. (Good, "Weight of Evidence: A Brief Survey," 267-8)

Like Keynes, Good often wrote with a pithy and pejorative pen. Good's critique of Keynesian weight borders on polemical, which indicates that he was at least a little

[116] Seidenfeld, 265.

[117] Seidenfeld, 265.

perturbed to see Good-Turing weight compared to Keynesian weight. In the next section, I analyze Good's critique of Keynesian weight with the aim of reconstructing the substance of the criticisms while simultaneously removing (or at least explaining) the polemics.

### §3.1.3 Reconstructing Good's Critique

The most obvious problem that Good seemed to have with Keynesian weight is the lack of a scale for the evidence, where a scale is taken to be something that allows comparisons between "pieces" of evidence. Based on the last paragraph in the above quote from Good, it seems as though Good interpreted measuring total accumulated evidence as the only function of Keynesian weight. Consequently, Good operates with a monotonic interpretation of Keynesian weight in mind. Recall that on the monotonic interpretation, Keynesian weight increases with any additions to the evidence involved in the probability judgement, regardless of how the additional evidence affects the probability judgement.[118] Accordingly, monotonic Keynesian weight only ever provides a single, absolute value for the weight of a probability judgement. Since Keynesian weight is meant to measure the gross amount of evidence incorporated into a probability judgement, using a single number makes sense. However, Good's comments indicate that he did not see Keynesian weight as a useful statistical measure.

To understand why, let us think about what we might mean when we say that a concept features a scale. We might say that the minimum qualification needed for something to be considered a scale is the capacity to permit comparisons between two

[118] To the best of my knowledge, Keynes does not say much about how much weight increases with increases in evidence. However, Brady seems to maintain that Keynes gave a fully fleshed-out apparatus for the functioning of weight. See Brady, "Keynes's Theoretical Approach" for more details.

entities, whether the entities are objects, sets of objects, events, sets of events, or even something more abstract such as hypotheses. Building on that basic definition, we might say that a scale should also allow us to quantify some quality of an entity in order to assess differences in that quality between entities. This is the sense in which an engineer's ruler, a thermometer, or even the Richter scale is a scale.

At first pass, we might think that Good criticizes Keynesian weight because Good thinks that Keynesian weight fails to amount to any sort of scale whatsoever. However, when translated into a non-Keynesian approach to probability, monotonic Keynesian weight should meet the initial, simple definition of a scale.[119] Furthermore, notice that insofar as we can reach agreement about the method of quantifying the Keynesian weight of hypotheses, monotonic Keynesian weight plausibly meets the more stringent definition of a scale. For example, let us assume that details regarding units of monotonic Keynesian weight have been worked out and that we have all agreed that one *Maynard* (however we might define that) is the proper unit for measuring Keynesian weight. Perhaps one judgement possesses a Keynesian weight of 15 Maynards while a second judgement weighs 30 Maynards. So long as we can coherently say that the second judgement weighs 15 Maynards more than the first judgement weighs, we can also say that monotonic Keynesian weight provides a scale for quantifying differences. To be fair, when Good describes Keynes's treatment of weight as "purely qualitative," he alludes to the difficulty of how to solve the unit problem and quantify monotonic Keynesian weight. How do you quantify the

[119] Keynes argues that just as it is sometimes impossible to compare Keynesian probabilities, it will sometimes be impossible to compare Keynesian weights. See O'Donnell, *Keynes: Philosophy, Economics, and Politics*, 71. I interpret this as a holdover from the Keynesian probability apparatus.

monotonic Keynesian weight of two wildly different types of evidence? Answering that question is a difficult problem without a clear answer as far as I can tell. However, we should recognize that Good's critique is not centered solely around the unit problem or the potential failure of Keynesian weight to include any scale whatsoever.

For Good, the problem with Keynesian weight is that it puts all the evidence, whether that evidence is in favor or against a hypothesis, on *the same scale*. As he says in an earlier work:

> The expression [weight of evidence] was used by J.M. Keynes (1921, p. 71) in a less satisfactory sense, to apply to the total bulk of evidence whether any part of it supports or undermines a hypothesis, almost as if he had the weight of the documents in mind. (Good, *Good Thinking*, 160)

Even if we determine units that quantify Keynesian weight, Keynesian weight still fails to differentiate between evidence that supports a hypothesis and evidence that runs contrary to a hypothesis. In Good's mind, weighing the evidence should be a process that puts the evidence in favor of the judgement on one side of a figurative double-pan balance scale, and then puts the evidence against the judgement in the other "pan" of the balance scale. Such a process signals how well the available evidence supports a hypothesis.

Monotonic Keynesian weight fails to provide a weighing apparatus like the one Good describes. The process of weighing evidence provided by the monotonic interpretation of Keynesian weight places all the available evidence relevant to a hypothesis together in order to see how *much* evidence went into the judgement. A major problem with the monotonic interpretation of Keynesian weight is figuring out what the relevant contrast class is supposed to be. That is to say that on the monotonic interpretation of Keynesian weight, it will be difficult to say whether a hypothesis, taken by itself, possesses a substantial Keynesian weight. Of course, when we compare two hypotheses, we may (sometimes) be able to determine which hypothesis possesses a higher monotonic Keynesian weight.[120] It should become even easier to understand the weight of a hypothesis if we develop units for monotonic weight, like the Maynards mentioned above. But the point is that units alone do not provide the context we need when we seek to understand the "weightiness" of a hypothesis—some experience of the units in action is required to understand the meaning of the units.

The difficulties pertaining to the determination of the relevant contrast class make it challenging to see what a measure of monotonic Keynesian weight tells us about the way evidence supports a hypothesis. Does this monotonic Keynesian weight indicate that the hypothesis is well supported? That question remains unanswered (unless there is a competing hypothesis with a Keynesian weight that we can compare to the Keynesian weight of the first hypothesis). As a result, monotonic Keynesian weight fails to provide a contrast class to use when determining how well-supported a hypothesis is. This is the reason that Good would have considered the statement, "This hypothesis is based upon 2,000 'pieces' of evidence," as useful as "This hypothesis is printed on 200 pieces of paper." Because it does not implicitly feature a contrast class, Keynesian weight lacks a scale that corresponds to the idea of weighing evidence on a double-pan balance.

At this point, it is important to recognize how deeply Good's criticism cuts. Good seems to get at a more fundamental problem regarding what we should expect out of our statistical and theoretical concepts. Recall Good's insistence on weight of evidence being

[120] This assumes either that the two hypotheses have the same general type of evidence or that we have developed some methodology for comparing the Keynesian weights of different types of evidence. Even then, though, making such comparisons is tricky. Keynesian weight seems like Keynesian probability in that comparisons are not always possible. Thanks to Jeff Kasser for reminding me of this point.

"the same for all rational people" and not dependent on statistical training.[121] Good took the process of weighing evidence to be fundamental to human inquiry. After all, the Greek goddess of justice and fairness used a scale to weigh evidence! The problem with Keynesian weight is not just that it lacks use-value, but rather that it deviates too much from our normal understanding of weighing evidence. In Good's mind, weighing evidence *must* describe a process that measures the balance of the evidence, and he presents several real-life examples to support that interpretation of weighing evidence. For Good, Keynes's notion is simply too distanced from our use of the phrase from which it derives its name.

By using a crude methodology to put all evidence on the same scale, Keynesian weight struggles to capture our everyday intuitions regarding weighing evidence. By failing to capture those intuitions, it becomes unclear how to apply Keynesian weight, which makes the concept lose appeal. To put the final nail in the Keynesian-weight coffin, Good describes how to use Good-Turing weight of evidence to quantify monotonic Keynesian weight. So, Keynesian weight lacks a tie to our understanding of weighing evidence and it can easily be superseded by Good-Turing weight of evidence. On this view, Keynesian weight appears under-motivated at best and downright misleading at worst. No wonder Good called Keynesian weight crummy.

## *§3.2 Relative Keynesian Weight Response*

Both Good and Seidenfeld discussed Keynesian weight with the monotonic interpretation in mind. Nonetheless, Chapter 2 showed that the monotonic interpretation of Keynesian weight is not the only game in town. Do Good's criticisms affect both

[121] Good, "Weight of Evidence," 249-50.

interpretations of Keynesian weight in the same way? In this section, I will argue that the answer to that question is "No." More specifically, I think that a relative weight interpretation avoids the most straightforward reading of Good's criticisms because relative weight features the general type of scale that Good seeks in a conception of evidential weight. However, analysis of the scale implicit to relative Keynesian weight indicates that Keynesian weight is meant to be utilized in a different way than the scale used in Good-Turing weight. Accordingly, applying the relative Keynesian weight interpretation to Good's critique leads to a deeper understanding of that critique. Before getting to that deeper reading, let us see how exactly a relative weight interpretation allows an advocate of Keynesian weight to respond to the most straightforward reading of Good.

### *§3.2.1 The Scale of Relative Weight*

Above, we saw that the process of weighing that Good wants out of a conception of weight utilizes a scale that allows one to easily judge how evidence supports a hypothesis. Such a scale seems to require a contrast class for making judgements of evidential support. Here, I argue that a relative weight interpretation of Keynesian weight provides a scale in this sense. Recall that the two formulations of relative Keynesian weight considered in Chapter 2 parallel probability and odds measures respectively (see Table 2.1 for a quick refresher). Thus, in order to understand why relative Keynesian weight inherently involves a scale, we will be well-served to first briefly return to the mathematical nature of probability and odds (see §1.1 above). In the next paragraphs, I will explain how the mathematical nature of probability and odds shows that probability and odds both feature a scale in the sense Good uses the word.

In §1.1, we saw that almost everyone agrees about the mathematical nature of probability and odds, even if they disagree about the interpretations of these terms. The agreement over the mathematics of probability may seem inconsequential, but it is crucial to understanding how probability (and consequently odds) inherently provide a scale for comparisons. On its own, a number conveys little information. This is the problem with the monotonic interpretation of Keynesian weight—the measure simply does not provide the contextual information necessary to make the concept worthwhile. If I walked up to you and said, "My hypothesis has a monotonic Keynesian weight of 15," then you would not know what that means unless I gave you information about the Keynesian weight of other similar hypotheses. We saw above that provision of units for measuring monotonic Keynesian weight would ease this problem, but units do not entirely overcome the problem.

In contrast to monotonic Keynesian weight, a first-order probability inherently communicates information that allows us to contextualize its implications. If I were to walk up to you and say, "This event has a probability of occurring of 0.99," then despite lingering questions about the relevant interpretation of that probability, you would know that I take the event to be far more probable than not. Since a probability is definitionally bound between 0 and 1, it inherently involves a contrast class with which we can glean information. Furthermore, positive and negative evidence do not alter probabilities in exactly the same way.[122] When we add more favorable outcomes to our evidential set, the

[122] 'Pieces' of positive and negative evidence might alter a probability by the same magnitude, but the direction of the changes in probability will be opposite. Accordingly, positive and negative evidence might have a mirrored effect on a probability, but they will not alter the probability in exactly the same way.

probability increases. When we add more unfavorable outcomes to our evidential set, the probability decreases. At the extremes (i.e. close to 0 or 1), these changes in probability become smaller and smaller. However, the important thing to recognize is that a probability decreases when unfavorable evidence is added to our evidential set, and it increases when favorable evidence is added to our evidential set. Contrast that mechanism with monotonic Keynesian weight, which Good criticizes because it treats all evidence identically.

In §1.1, we similarly saw how odds inherently convey information with which we can contextualize their meaning, although that information is placed on a different scale from the one used with probabilities. When someone says, "The odds of this hypothesis being true are 50/10," the very definition of odds allows you to immediately know that the odds point in favor of the truth of the hypothesis. Accordingly, the mathematical treatment of odds and probabilities implies that they inherently feature a scale ripe for making comparative judgements.

With the mathematical treatment of probability and odds in place, we can now see why relative Keynesian weight features the sort of scale that Good seeks out of weight of evidence. Let us start by considering the conception of relative Keynesian weight that parallels odds, which I called *relative weight$_{odds}$*. As we saw in Chapter 2, relative weight$_{\text{odds}}$ is the ratio of relevant knowledge over relevant ignorance. The mathematical operation of relative weight$_{\text{odds}}$ is just like the mathematics of odds.[123] In cases in which relevant

[123] It is tempting to try to perform a logarithmic transformation of relative weight$_{\text{odds}}$ in order to make it more comparable to Good-Turing weight. We might try to take the logarithm of relative weight$_{\text{odds}}$ to try to give it a scale that ranges from negative infinity to positive infinity and centers around zero. However, it is difficult to see what such a scale would tell us in the case of relative Keynesian weight. As we will soon see, Good's critique will show us that relative weight is fundamentally a conception of gross weight of evidence. By giving

ignorance far exceeds relevant knowledge, relative $\text{weight}_{\text{odds}}$ will be less than 1 and approach 0. For example, imagine a scenario in which I find (what I know to be) a coin, and I flip it twice. Suppose that I have never seen a coin like it before and I have no idea where it may have come from. Assume I lack knowledge about the date and location of the production of the coin, the materials involved in making the coin, the intended use of the coin, and so on. In this scenario, my previous two flips of the coin and my knowledge that it is a coin exhaust all of my relevant knowledge about the next flip.[124] In this case, I possess little relevant knowledge regarding the coin in comparison to the vast amount of relevant ignorance regarding the coin. Consequently, the relative $\text{weight}_{\text{odds}}$ of my probability judgement of the result of the next flip of the coin will be less than one. We can imagine another case that reverses the amounts of relevant knowledge and relevant ignorance. Perhaps I was present when the coin was made with an extremely precise machine (which I invented) for making biased coins. In that case, my relevant knowledge about the coin would far exceed my relevant ignorance regarding the coin. Although I will not know the result of every flip (provided the coin is not totally biased), the relative $\text{weight}_{\text{odds}}$ of my probability judgement about the next flip of the coin will be far greater than one. Since relative $\text{weight}_{\text{odds}}$ is a ratio of relevant knowledge over relevant ignorance, it inherently features a useful scale for understanding the way in which evidence supports a probability judgement.

---

it a scale similar to the scale suited to a leading conception of net weight of evidence, we seem to be packing too much into our conception of Keynesian weight. I leave this point for the reader to cogitate upon.

[124] Perhaps the knowledge of the location of the next flip in space and time should be added to our consideration of relevant knowledge. Defining 'relevance' is a tricky issue, but there are several subjective Bayesian solutions to that problem (including use of log-odds!). See Fitelson, "A Bayesian Account" for more information.

In Chapter 2, we saw how relative weight$_{comp}$ parallels a probability measure. Recall that a probability measure is a ratio of favorable outcomes over all possible outcomes. In the case of relative weight$_{comp}$, the numerator of the ratio is relevant knowledge while the denominator of the ratio is all relevant information. Thus, relative weight$_{comp}$ is a ratio that communicates the amount of relevant knowledge possessed in comparison to the set of all relevant information pertaining to the judgement. Since it is a ratio, relative weight$_{comp}$ is a measure that comes with a contrast class inherently built into it. Just like a probability, relative weight$_{comp}$ is bound between zero and one. The zero case indicates that we take ourselves to completely lack relevant knowledge regarding the probability judgement, and the case in which relative weight$_{comp}$ equals one indicates that we think that we possess all of the information relevant to the probability judgement. Relative weight$_{comp}$ might provide an even more intuitive scale for Keynesian weight (than relative weight$_{odds}$) since we can think of relative weight$_{comp}$ as an estimate of the completeness of our relevant information. Regardless, it is clear that relative weight$_{comp}$ features a scale similar to the scale provided by mathematical probabilities.

Both relative weight$_{odds}$ and relative weight$_{comp}$ allow new evidence to decrease Keynesian weight. Allowing new evidence to decrease Keynesian weight is a crucial feature of any relative weight interpretation of Keynesian weight. By allowing evidence to lower Keynesian weight, the relative weight interpretation of Keynesian weight sidesteps Good's critique about Keynesian weight putting all evidence, regardless of sign, on the same scale.

### *§3.2.2 A Deeper Reading of Good's Critique*

A relative weight interpretation fails to get Keynesian weight all of the way out of the woods because Good's critique is more subtle than it first appears. Suppose we are testing a hypothesis that features a somewhat high initial measure of relative Keynesian weight. In the course of testing that hypothesis, assume we receive some unfavorable but not wholly surprising evidence. Although the evidence we received in this example runs contrary to our hypothesis, it fails to give us reason to doubt our estimation of our ignorance. For that reason, the new evidence increases relative Keynesian weight, despite rendering our hypothesis less likely. Moreover, although the evidence is unfavorable to our hypothesis, the evidence still increases relative Keynesian weight by the same degree favorable evidence of the same variety would increase it. Consequently, Good might maintain that even on the relative weight interpretation, Keynesian weight still forces us to put all evidence, whether positive or negative, on the same scale. Accordingly, a relative weight interpretation of Keynesian weigh still ultimately falls victim to Good's critique, at least in some cases.

Although the deeper reading of Good's critique shows that a relative weight interpretation of Keynesian weight falls victim to Good's complaint, the deeper reading simultaneously solves our interpretative puzzle from Chapter 2. The subtlety in Good's critique brings out a crucial difference between the analyses of Good and Keynes. I think that the fundamental difference between Keynesian weight and Good-Turing weight comes from the fact that these measures operate on different "levels" of analysis. Consider some evidence that runs contrary to the truth of our hypothesis. As competent Bayesians, we can clearly see that the evidence decreases the probability placed on the truth of our hypothesis. According to Good's critique and discussions of weight of evidence, such contrary evidence should decrease the weight of evidence in favor of our hypothesis. The Good-Turing conception of weight of evidence operates in this fashion; a change in probability necessitates a change in Good-Turing weight. Our deeper reading of Good's critique indicates that Keynesian weight fails to operate in this manner, regardless of the interpretation taken. For Keynes, the probability of the hypothesis is (by definition) independent of the weight of the hypothesis. Thus, the contrary evidence actually increases the Keynesian weight placed on the hypothesis, provided that the evidence fails to alter our understanding of our ignorance of the situation.

I want to drive home the difference in levels of analysis between Keynesian weight and Good-Turing weight with one more toy example. Suppose we have two competing, exclusive and exhaustive hypotheses. Call them *A* and *not-A*. Suppose that the Bayesian probability in favor of the truth of hypothesis A is 0.99 and consequently the Bayesian probability in favor of the truth of *not-A* is 0.01. Both hypotheses are based on the same 100 cases of evidence, and all the evidence is of the same general type. The probabilities indicate that *A* is much more likely to be true than hypothesis *not-A*. This fact is represented nicely by the Good-Turing weight of evidence of these hypotheses. In the case of *A*, the Good-Turing weight of evidence is 1.99564 bans (which is Turing's unit for measuring Good-Turing weight). In the case of hypothesis *not-A*, the Good-Turing weight of evidence is -1.99564 bans. Good's fundamental critique seems to come down to the fact that Keynesian weight does not do a good job of representing this difference between hypotheses such as *A* and *not-A*. Regardless of the interpretation of Keynesian weight taken, the Keynesian weight of hypotheses *A* and *not-A* will be the same because both hypotheses are based on the same exact evidence. Keynesian weight fails to convey the balance of the evidence in the evidential set. Thus, Keynesian weight fails to tell us which hypothesis we should believe to be true.

Despite this failure, Keynesian weight is not crummy. Rather than telling us which particular hypothesis to believe in, Keynesian weight tells us something about the inferential waters in which we tread. When we say that the Bayesian probability of the truth of hypothesis *A* equals 0.99, we know that this means that our current evidence indicates that *A* is more likely to be true than *not-A*. However, that probability says quite little about the nature of the evidential set that the probability is based on.

Is this the type of inquiry in which we are likely to attain the truth? Or is this the type of inquiry in which truth currently escapes our grasp, regardless of what our best hypotheses say? Is this the type of inquiry in which we are likely to gain a high degree of warranted confidence? Are we confident in our current hypotheses? Should we even hold beliefs in this sort of inquiry? It seems to me that Keynesian weight is meant to answer these sorts of questions. By making reference to our current evidence regarding a hypothesis and our best estimate of our ignorance of the remaining relevant information out there, relative Keynesian weight gives us an approximate idea of how to answer these questions. Unlike Good-Turing weight, Keynesian weight is not meant to tell us how our evidence favors a hypothesis. Instead, Keynesian weight tells us how much stock, worth, or weight to put into *any* of our hypotheses in the specified area of inquiry. The way in which relative weight accomplishes that task is different from the way in which monotonic weight tries to accomplish the task. By utilizing a scale that parallels probability or odds, relative weight attempts to tell us how much evidence is at our disposal in a way that also provides context as to how much more relevant information might be out there. Monotonic weight provides no such context.

I think the crucial difference between Keynesian weight and Good-Turing weight comes down to the questions that each conception of weight attempts to answer. Like a Bayesian probability, Good-Turing weight tells us something about our degree of belief in the truth of the hypothesis. In contrast, Keynesian weight gives us a picture of the overall amount of relevant evidence upon which the hypothesis is based. As such, it is more natural to read Keynesian weight as a measure of our confidence in our judgements about a hypothesis, regardless of what those judgements happen to be. By applying Good's critique to relative Keynesian weight, we learn that regardless of the interpretation selected, Keynesian weight serves a different inferential role than the role served by Good-Turing weight.

### *§3.2.3 The Solution to the Interpretative Puzzle*

The difference in levels of analysis taken by Good-Turing and Keynesian weight teaches us something interesting about Good's critique of Keynesian weight. Insofar as Good criticizes the fact that Keynesian weight always treats all evidence alike, his critique fails. There is a plausible interpretation of Keynesian weight (i.e., relative weight) that can avoid putting all evidence on a single scale. On the relative interpretation of Keynesian weight, certain types of evidence, namely evidence that increases our estimation of our ignorance, decrease Keynesian weight. The relative weight interpretation allows defenders of Keynesian weight to sidestep Good's critique in many cases.

However, if Good's critique centers around the failure of Keynesian weight to indicate how evidence favors or disfavors a hypothesis (which seems to be the case), then even Keynesians who take a relative weight interpretation fall subject to Good's critique. Thorough consideration of Good's critique shows that relative Keynesian weight fails to tell us anything about the balance of the evidence relevant to a hypothesis. As a result, Good shows that relative weight still operates on a different level of analysis than the level of analysis taken by Good-Turing weight. While Good-Turing weight provides an analysis of how evidence favors or disfavors a hypothesis, relative Keynesian weight provides a picture of the amount of evidence our hypothesis is based on.

Consequently, Good's critique teaches us something about relative Keynesian weight. Despite the fact that relative weight can decrease, Good's critique shows us that relative weight is truly a measure of what Kasser called "gross weight of evidence." By showing us that the relative interpretation of Keynesian weight still renders Keynesian weight to be a measure of gross weight, Good helps us solve the interpretative puzzle surrounding Keynesian weight. When initially introduced, the relative weight interpretation of Keynesian weight appears far too revisionary. It can seem like a subject change rather than a plausible interpretation of the text. Then, we start to see some textual evidence in support of the relative weight interpretation, and it begins to appear less revisionary. But a nagging question remains: is relative weight actually the concept Keynes had in mind when he introduced weight? For it to be so, we need to see the ways in which relative weight can account for the text that seems to speak in favor of a monotonic interpretation of Keynesian weight. I hope that Chapter 2 showed that the text is not decisive toward either interpretation of Keynesian weight. But we still might wonder whether relative weight counts as a measure of gross weight of evidence rather than net weight of evidence. Our assessment of the desiderata at the end of Chapter 2 showed that the interpretative puzzle seems to boil down to whether relative weight is the proper type of Peircean weight. No conception of Keynesian weight will be acceptable if it transforms

Keynesian weight into something other than a gross weight measure because such a transformation would indicate a shift in concepts rather than an alternative interpretation of a concept. Moreover, the textual evidence in favor of the monotonic interpretation of Keynesian weight shows that Keynesian weight must be a measure of gross weight in some fashion.

Since Good's critique applies to relative Keynesian weight, Good's critique shows us that relative weight is still a measure of gross weight of evidence. Like other measures of gross weight of evidence, relative Keynesian weight operates on a different level of analysis than the level of analysis taken by net weight of evidence. Even though relative weight features a scale similar to the scale of probability or odds, it is not a measure of the balance of evidence in a probabilistic inference. By illuminating the fact that relative weight remains a measure of gross weight, Good gives us grounds for solving the interpretative puzzle in favor of relative Keynesian weight.

Now, it is worth briefly considering how fatal Good's critique actually is if it is fundamentally a criticism of Keynesian weight's failure to be a measure of net weight of evidence. In my view, if Good is critiquing Keynesian weight simply because it is a gross weight measure, then his criticism seems thin. Good does not explain why we should limit our focus to net weight. Instead, Good seems to hold a basic preference for the information provided by net weight rather than the information provided by gross weight. But a preference is not an argument. Both types of weight provide useful information. Additionally, since gross weight and net weight of evidence are meant to fulfill different tasks, they should not be treated as direct competitors. In fact, it is plausible to use conceptions of gross weight and net weight of evidence in conjunction with one another. To that end, we can use the Good-Turing log-odds conception of net weight of evidence as a way to show how decisively our evidence points in (dis)favor of a hypothesis. Then, we can also provide a measure of relative Keynesian weight in an attempt to demonstrate how convincing our evidence is based on our best estimations of how much evidence is out there. An approach ecumenical to both conceptions of weight exists.

## §3.3 CONCLUSION

In Chapter 1, I tried to untangle the web of issues tangentially related to Keynesian weight in order to limit the focus of the thesis. To accomplish that task, I introduced the two types of Peircean weight, which Kasser calls "gross weight of evidence" and "net weight of evidence." We saw that the traditional and most-straightforward interpretation of Keynesian weight makes Keynesian weight out to be a paradigmatic type of gross weight of evidence. In Chapter 2, I outlined the monotonic and relative interpretations of Keynesian weight as well as the interpretative puzzle generated by these two interpretations. In that chapter, I attempted to show that the less-traditional relative weight interpretation of Keynes possesses strong textual support, although that textual support fails to decisively resolve the interpretative puzzle. Furthermore, the relative weight interpretation of Keynes runs the risk of turning Keynesian weight into a measure of net weight of evidence. In Chapter 3, I applied the relative weight interpretation to I.J. Good's critique of Keynesian weight in order to discover whether Good's critique gives us leverage on the interpretative puzzle. We discovered that because relative Keynesian weight utilizes scales that parallel probability and odds, relative weight is able to answer Good's critique of Keynesian weight when we take a straightforward reading of that critique. The scales of relative Keynesian weight allow certain kinds of evidence to decrease weight, which means that relative Keynesian weight does not treat all evidence exactly alike.

That said, we also discovered that even when we use a relative interpretation, Keynesian weight and Good-Turing weight operate on different levels of analysis. We can see this when we consider a case of using these two concepts to weigh a hypothesis and its negation. In this case, the Keynesian weights are equivalent, while the Good-Turing weights possess opposite signs. This test case shows us that relative weight is still a form of gross weight of evidence rather than a conception of net weight of evidence. Accordingly, a deeper reading of Good's critique shows us that the relative weight interpretation of Keynesian weight avoids collapsing into a form of net weight. In other words, a relative weight interpretation preserves the distinction between the two types of Peircean weight. Thus, Keynesian weight is not in direct competition with Good-Turing weight of evidence. By illuminating the ways in which relative Keynesian weight differs from his preferred conception of net weight of evidence, Good's critique of Keynes gives us grounds for solving the interpretative puzzle in favor of relative weight of evidence. Despite traditional interpretations to the contrary, Keynesian weight is relative weight.

# BIBLIOGRAPHY


Adler, Jonathan E. *Belief's Own Ethics*. Cambridge, MA: The MIT Press, 2002.

Brady, Michael Emmett. "J.M. Keynes's Theoretical Approach to Decision-Making Under Conditions of Risk and Uncertainty." *The British Journal for the Philosophy of Science*, 44, no. 2, (1993): 357-76. doi: 10.1093/bjps/44.2.357.

Cohen, L. Jonathan. "Twelve Questions About Keynes's Concept of Weight." *British Journal for the Philosophy of Science*, 37, no. 3 (1986): 263-78. doi: 10.1093/bjps/37.3.263.

Crocco, Marco. "The Concept of Degrees of Uncertainty in Keynes, Shackle, and Davidson." *Nova Economia*, Economics Department, Universidade Federal de Minas Gerais (Brazil), 12 (2002): 11-28.

Curd, Martin, J.A. Cover, and Christopher Pincock. *Philosophy of Science*, Second Edition. New York: W.W. Norton & Company, Inc., 2013.

Dimand, Robert W. "Keynes, Knight, and Fundamental Uncertainty: A Double Centenary 1921-2021." *Review of Political Economy* (2021). doi: 10.1080/09538259.2021.1924470.

Dow, Sheila. "Keynes on Knowledge, Expectations, and Rationality." In *Rethinking Expectations: The Way Forward for Macroeconomics*, edited by Roman Frydman and Edmund S. Phelps, 112-130. Princeton: Princeton University Press, 2013.

Dow, Sheila. "Uncertainty: A Diagrammatic Treatment." *Economics: The Open-Access, Open-Assessment E-Journal*, 10, no. 3 (2016): 1-25. doi: 10.5018/economics-ejournal.ja.2016-3.

Faulkner, Phil, Alberto Feduzi, C.R. McCann Jr., and Jochen Runde. "F.H. Knight's Risk, Uncertainty, and Profit and J.M. Keynes' Treatise on Probability After 100 Years." *Cambridge Journal of Economics*, 45, no. 5 (2021): 857-82. doi: 10.1093/cje/beab035.

Feduzi, Alberto. "On Keynes's Conception of the Weight of Evidence." *Journal of Economic Behavior & Organization*, 76 (2010): 338-51.

Fioretti, Guido. "Von Kries and the Other 'German Logicians': Non-Numerical Probabilities Before Keynes." *Economics and Philosophy*, 17 (2001): 245-73.

Fitelson, Branden. "A Bayesian Account of Independent Evidence With Applications." *Philosophy of Science*, 68, no. 3, Supplement: Proceedings of the 2000 Biennial Meeting of the Philosophy of Science Association. Part I: Contributed Papers (Sept. 2001): S123-S140. http://www.jstor.org/stable/3080940.

Fox, Craig R. and Amos Tversky. “Ambiguity Aversion and Comparative Ignorance.” *The Quarterly Journal of Economics*, 110, no. 3 (1995): 585–603. doi: 10.2307/2946693.

Friedman, Jane. “Inquiry and Belief.” *Nous*, 53, no. 2 (2019): 296-315. doi: 10.1111/nous.12222.

Garber, Daniel. “What’s Philosophical About the History of Philosophy?” In *Analytic Philosophy and History of Philosophy*, edited by Tom Sorell and G.A.J. Rogers, 129-46. Oxford: Oxford University Press, 2005.

Glymour, Clark and Frederick Eberhardt. “Hans Reichenbach.” In *The Stanford Encyclopedia of Philosophy*, (Spring 2022 edition), edited by Edward N. Zalta. https://plato.stanford.edu/archives/spr2022/entries/reichenbach/.

Good, I. J. “Weight of Evidence: A Brief Survey.” *Bayesian Statistics*, 2 (1985): 249-70. https://www.cs.tufts.edu/~nr/cs257/archive/jack-good/weight-of-evidence.pdf.

Good, Irving J. *Good Thinking: The Foundations of Probability and Its Applications*. Mineola, NY: Dover Publications Inc., 1983.

Hájek, Alan. “Interpretations of Probability.” In *The Stanford Encyclopedia of Philosophy*, (Fall 2019 edition), edited by Edward N. Zalta. https://plato.stanford.edu/archives/fall2019/entries/probability-interpret/.

Hamer, David. “Probability, Anti-Resilience, and the Weight of Expectation.” *Law, Probability and Risk*, 11 (2012): 135-58. doi: 10.1093/lpr/mgs004.

Hodges, Andrew. *Alan Turing: The Enigma*. Princeton: Princeton University Press, 1983; London: Vintage, 2014.

Joyce, James M. “How Probabilities Reflect Evidence.” *Philosophical Perspectives*, 19, no. 1 (2005): 153-78. doi: 10.1111/j.1520- 8583.2005.00058.x.

Kasser, Jeff. “Two Conceptions of Weight of Evidence in Peirce’s *Illustrations of the Logic of Science*.” *Erkenntnis*, 81, no. 3 (2016): 629-48. doi: 10.1007/s10670-015-9759-5.

Kelly, Thomas. “Peer Disagreement and Higher-Order Evidence.” In *Disagreement*, edited by Richard Feldman and Ted A. Warfield, 111-74. Oxford: Oxford University Press, 2010.

Keynes, John Maynard. *A Tract on Monetary Reform*. London: 1924; Project Gutenberg, 2021. https://www.gutenberg.org/ebooks/65278.

Keynes, John Maynard. *A Treatise on Probability*. London: 1921; Project Gutenberg, 2014. https://www.gutenberg.org/ebooks/32625.

Keynes, John Maynard. *The General Theory of Employment, Interest, and Money*. New York: First Harvest/Harcourt Inc, 1964.

Lewis, Michael. *The Big Short: Inside the Doomsday Machine*. New York: W.W. Norton & Company, Inc., 2011.

Loeb, Louis. “Sextus, Descartes, Hume, and Peirce: On Securing Settled Doxastic States.” *Nous*, 32, no. 2 (1998): 205-30. doi: 10.1111/0029-4624.00097.

Mill, John Stuart. “On Liberty.” In *Utilitarianism, On Liberty, Considerations on Representative Government, Remarks on Bentham’s Philosophy*, edited by Geraint Williams, 69-186. London: Everyman’s Library, 1993.

Nagel, Ernest. “Issues in the Logic of Reductive Explanations.” In *Philosophy of Science: The Central Issues* (Second Edition), edited by Martin Curd, J.A. Cover, and Christopher Pincock, 911-26. New York: W.W. Norton & Company, Inc., 2013.

Nance, Dale A. “The Weights of Evidence.” *Episteme*, 5, no. 3 (2008): 267-81. doi: 10.3366/E1742360008000385.

O’Donnell, Rod M. “Keynes and Knight: Risk-Uncertainty Distinctions, Priority, Coherence, and Change.” *Cambridge Journal of Economics*, 45, no. 5 (2021): 1127-1144. doi: 10.1093/cje/beab034.

O’Donnell, Rod M. *Keynes: Philosophy, Economics, and Politics: The Philosophical Foundations of Keynes’s Thought and Their Influence on His Economics and Politics*. New York: St. Martin’s Press, 1989.

Peirce, Charles Sanders. “The Probability of Induction.” In *The Essential Peirce: Selected Philosophical Writings, Volume 1 (1867-1893)*, edited by Nathan Houser and Christian Kloesel, 155-70. Bloomington: Indiana University Press, 1992.

Popper, Karl. “A Third Note on Degree of Corroboration or Confirmation.” In *The Logic of Scientific Discovery*, 10th (revised) impression, 406-19. London: Hutchinson & Co., 1980.

Ramsey, Frank P. “Truth and Probability.” In *The Foundations of Mathematics and other Logical Essays*, edited by R.B. Braithwaite, 156-98. London: Kegan, Paul, Trench, Trubner & Co., New York: Harcourt, Brace and Company, 1931. 1999 electronic edition URL: http://fitelson.org/probability/ramsey.pdf.

Runde, Jochen. “Dissecting the Black Swan.” *Critical Review*, 21, no. 4 (2009): 491-505. doi: 10.1080/08913810903441427.

Runde, Jochen. “Keynesian Uncertainty and Liquidity Preference.” *Cambridge Journal of Economics*, 18, no. 2 (1994): 124-44. http://www.jstor.org/stable/24231904.

Runde, Jochen. “Keynesian Uncertainty and the Weight of Arguments.” *Economics and Philosophy*, 6, no. 2 (1990): 275-92. doi: 10.1017/S0266267100001255.

Russell, Bertrand. *The Problems of Philosophy*. London: 1912; Project Gutenberg, 2004. https://www.gutenberg.org/ebooks/5827.

Salmon, Wesley. "Rationality and Objectivity in Science or Tom Kuhn Meets Tom Bayes." In *Philosophy of Science: The Central Issues* (Second Edition), edited by Martin Curd, J.A. Cover, and Christopher Pincock, 518-49. New York: W.W. Norton & Company, Inc., 2013.

Schum, David A. *The Evidential Foundations of Probabilistic Reasoning*. New York: J. Wiley, 1994.

Seidenfeld, T. Reply to Good's Survey of Weight of Evidence. In "Weight of Evidence: A Brief Survey," by I.J. Good, in *Bayesian Statistics*, 2 (1985): 264-6. https://www.cs.tufts.edu/~nr/cs257/archive/jack-good/weight-of-evidence.pdf.

Skidelsky, Robert. *Keynes: A Very Short Introduction*. Oxford: Oxford University Press, 2010.

Stigler, Stephen M. *Statistics on the Table: The History of Statistical Concepts and Methods*. Cambridge, MA: Harvard University Press, 1999.

Svetlova, Ekaterina. "On the Relevance of Knight, Keynes, and Shackle for Unawareness Research." *Cambridge Journal of Economics*, 45, no. 5 (2021): 989-1007. doi: 10.1093/cje/beab033.

Taleb, Nassim Nicholas. *The Black Swan: The Impact of the Highly Improbable* (Second Edition). New York: Random House Trade Paperbacks, 2010.

Wapshott, Nicholas. *Keynes Hayek: The Clash That Defined Modern Economics*. New York: W.W. Norton & Company, Inc., 2011.

Weatherson, Brian. "Keynes, Uncertainty and Interest Rates." *Cambridge Journal of Economics*, 26 (2002): 47-62.