Counterfactual Probability

Ginger Schultheis
Forthcoming in *The Journal of Philosophy*

1 Introduction

Stalnaker's Thesis about indicative conditionals is, roughly, that the probability one ought to assign to an indicative conditional equals the probability that one ought to assign to its consequent conditional on its antecedent. The thesis seems right. If you draw a card from a standard 52-card deck, how confident are you that the card is a diamond if it's a red card? To answer this, you calculate the proportion of red cards that are diamonds—that is, you calculate the probability of drawing a diamond conditional on drawing a red card.

Skyrms' Thesis about counterfactual conditionals is, roughly, that the probability that one ought to assign to a counterfactual equals one's rational expectation of the chance, at a relevant past time, of its consequent conditional on its antecedent.[1] This thesis also seems right. If you decide not to enter a 100-ticket lottery, how confident are you that you would have won had you bought a ticket? To answer this, you calculate the prior chance—that is, the chance just before your decision not to buy a ticket—of winning conditional on entering the lottery.

The central project of this article is to develop a new uniform theory of conditionals that allows us to derive a version of Skyrms' Thesis from a version of Stalnaker's Thesis, together with a *chance-deference norm* relating rational credence to beliefs about objective chance.[2]

I say a *version of* Stalnaker's Thesis because it is well known that Stalnaker's Thesis itself is subject to a series of *triviality results*. Assuming orthodox probability theory, it can be shown that, except in trivial cases, there is no way to

[1] The label 'counterfactual conditional' is misleading. Consider:

(1) If I caught the four o'clock train today, I would make it to the meeting by five.

An utterance of (1) suggests that the speaker *leaves open* the possibility that she will catch the four o'clock train. Thus, 'counterfactual conditional' hardly seems an apt label if it is to cover conditionals like (1). Some authors use the term 'subjunctive conditional'. But this label is also misleading. It suggests that the main grammatical difference between indicative conditionals and conditionals like (1) has to do with subjunctive mood. But that is not the case. In most languages, the primary grammatical difference between indicatives and conditionals like (1) is that the latter exhibit an extra layer of past tense morphology. I will continue to use the term 'counterfactual conditional' to refer to conditionals, like (1), that contain this extra layer of past tense morphology because the term is familiar and I know of no better label.

[2] For discussion of Skyrms' Thesis, see Skyrms (1980), Edgington (2008), Williams (2012), Moss (2013), Schwarz (2016), Schulz (2017), and Khoo (ms).

interpret the indicative conditional uniformly so that Stalnaker's Thesis holds universally, that is, for all rational probability functions. And I say a *version of* Skyrms' Thesis because, as I will show in §3 of this paper, that thesis also has unacceptable trivializing consequences given orthodox probability theory.

The paper opens in §2 with a discussion of Stalnaker's Thesis and, following van Fraassen (1976) and Bacon (2015), suggests an improved, *context-sensitive* version of the thesis, which I will call the *Local Thesis*. The rest of the paper breaks into two main parts. The first part (§3-§5) refutes Skyrms' Thesis and develops a context-sensitive replacement, the *Local Conditional Principal Principle*—a counterfactual analogue of David Lewis's *Principal Principle*.

The second part (§6-§8) begins by introducing a neo-Stalnakerian, uniform theory of conditionals. At a high level, my view says that all of the semantic differences between indicative conditionals and counterfactual conditionals boil down to differences in what is held fixed. When we evaluate an indicative conditional, we hold fixed all of our knowledge; when we evaluate a counterfactual, we hold fixed a contextually-determined subset of our knowledge. I show that this theory allows us to derive the Local Conditional Principal Principle from the Local Thesis and Lewis's Principal Principle. And although a full tenability result for the Local Conditional Principal Principle is beyond the scope of this paper, I will argue in the final section that there is good reason to be optimistic that the principle is indeed tenable within the Stalnakerian framework that I develop.

2 From Stalnaker's Thesis to The Local Thesis

Our eventual goal to derive a plausible, contextualist-friendly version of Skyrms' Thesis from a plausible, contextualist-friendly version of Stalnaker's Thesis and a plausible chance-deference norm. Here I introduce the contextualist-friendly version of Stalnaker's Thesis—the *Local Thesis*.

Suppose I'm a detective working on a murder case. I know it was either the butler or the gardener. My credence that it was the gardener, on the supposition that it wasn't the butler, is high. Correspondingly, I will be confident in (2).

(2) If the butler didn't do it, it was the gardener.

Take another case. Suppose I know that the four o'clock train arrives within an hour about 75% of the time. So I am 75% confident that John will make it by five, supposing he catches the four o'clock train. Correspondingly, I will be 75% confident in (3).

(3) If John catches the four o'clock train, John will be here by five.

Examples like these are easy to multiply, and the pattern of probability assignments is robust, leading many theorists to endorse some version of Stalnaker's Thesis. Where $A > B$ stands for the indicative conditional with antecedent A and consequent B, Stalnaker's Thesis is as follows:

> **Stalnaker's Thesis**
> For any rational credence function P such that $P(A) > 0$: $P(A > B) = P(B|A)$.

Stalnaker's Thesis, as I will understand it, is a normative thesis. It says that, *if you're rational*, then your credence in $A > B$ is equal to your credence in B conditional on A (whenever the conditional probability is defined). I take no stand on whether there are irrational subjects whose credences in indicative conditionals diverge from their conditional credences.

Despite its initial plausibility, Stalnaker's Thesis is false. David Lewis (1976) showed that Stalnaker's Thesis has trivializing consequences given just two standard assumptions: (1) that rational credence functions obey the laws of probability; and (2) that the set of rational credence functions is closed under conditionalization. Given (1) and (2), Stalnaker's Thesis entails that whenever you think that A and B are compatible, and that A and $\neg B$ are compatible, you are certain of B conditional on the indicative conditional $A > B$. This consequence is unacceptable. To illustrate, suppose that it's compatible with my beliefs that Milo is at a picnic and in a good mood, and compatible with my beliefs that he's at a picnic and in a bad mood. Stalnaker's Thesis predicts that I should be certain that Milo is in a good mood, conditional on *if Milo is at a picnic, he's in a good mood*. In other words, if I learn the conditional *if Milo is at a picnic, he's in a good mood*, then I should be certain that he's in a good mood. But that's absurd! For all I know, he's not at a picnic; for all I know, he's not in a good mood. So if we keep (1) and (2), we have no choice but to reject Stalnaker's Thesis.

Fortunately there are limited versions of Stalnaker's Thesis that capture its intuitive motivation but are not subject to the Lewisian triviality results. The one that I will be concerned with—the *Local Thesis*—is motivated by a contextualist theory of indicative conditionals. Before I state the thesis, let me say a few words to motivate contextualism independent of its connection to Stalnaker's Thesis and avoiding triviality. Contextualism about indicative conditionals is the view that what proposition is expressed by an utterance of an indicative conditional depends, in part, on a contextually-supplied body of information. Often that information is simply the speaker's knowledge. Other times it is some other body of information. For example, it may be the knowledge of some other individual or group. And sometimes the standards are more demanding than knowledge—such as being known *with certainty*. Other times they are less demanding. To allow for this variability, I refer to this contextually-supplied body of information simply as the information *associated with the context*.

Why accept contextualism? One argument comes from so-called *stand-off cases*. Consider:

> Sly Pete and Mr. Stone are playing poker on a Mississippi riverboat. It is now up to Pete to call or fold. My henchman Zach sees Stone's hand, which is good, and signals its content to Pete. My henchman Jack sees both hands and sees that Pete's hand is low, so that Stone's is the winning hand. At this point, the room is cleared. A few minutes later, Zack slips me a note which says, 'If Pete called, he won,' and Jack slips me a note which says 'If Pete called, he lost.' I know both notes come from my trusted henchmen but do not know which of them sent which note. I conclude Pete folded. (Gibbard 1981, p. 231)

According to the contextualist, Zack says something true when he writes:

(4) If Pete called, he won.

Likewise, Jack says something true when he writes:

(5) If Pete called, he lost.

Zach's conditional is true *relative to Zack's information*. Jack's conditional is true *relative to Jack's information*. Nevertheless, there is no information state—that is, no context—relative to which *both* conditionals are true.

If we're contextualists, Stalnaker's Thesis needs to be refined because it doesn't mention context.And, as we will see, these refinements are also sufficient for avoiding the triviality results. Specifically, we need to do two things. First, we need to add contextual parameters. Both the indicative conditional and the probability function need to be indexed to a context. Second, we must *coordinate* these two contextual parameters—the indicative conditional proposition on the left side of the equation must be indexed to the same context as the probability function on the right side. The notation is as follows. I write $A >_c B$ for the proposition expressed by the indicative conditional in a given context c. I write P_c for the probability function associated with c. (To simplify, I assume that P_c is result of conditioning a uniquely rational initial credence function P_o on the information associated with context c.) The Local Thesis is as follows:

The Local Thesis
$P_c(A >_c B) = P_c(B|A)$ whenever $P_c(A) > 0$.

Suppose, for a moment, that the information associated with a given context c is the speaker's knowledge. Then the Local Thesis says that the probability that the speaker in c assigns to the proposition expressed by the indicative, relative her information—*her* indicative conditional, as I will sometimes say—is equal to the probability that she assigns to B conditional on A. But importantly, it is silent about the probability that she assigns to propositions expressed by the indicative conditional in contexts other than her own. I won't get into the details, but as Bacon (2015) and others have shown, it is for precisely this reason that the Local Thesis is not subject to Lewisian triviality results. Indeed, it is not subject to any triviality results. Building on the work of van Fraassen (1976), Bacon (2015) has shown that the Local Thesis—or, more carefully, a thesis that is very close to the Local Thesis—is tenable within a possible-worlds semantics for indicatives based on Stalnaker's *selection semantics*. I return to these tenability results in §8.[3]

In the next few sections—sections §3-§5—I will set the Local Thesis to one side as I work up to my preferred formulation of Skyrms' Thesis—the *Local Conditional Principal Principle*. As we will see, that principle is similar in spirit to the Local Thesis. In this section, I hope to have set the basic groundwork for articulating a contextualist-friendly connection between chance and counterfactuals.

3 Triviality for Counterfactuals: Skyrms' Thesis

The last section concerned the relationship between indicative conditionals and probability. In this section, we turn to my primary concern in this paper—the relationship between counterfactuals and chance. I begin by motivating the most natural formulation of this connection—*Skyrms' Thesis*. Then I present a new argument showing that Skyrms' Thesis has unacceptable trivializing consequences.

Suppose that I decide not to flip a fair coin at noon. And suppose I know that the coin had a 50% chance of landing heads and a 50% chance of landing tails. How confident should I be in the counterfactual (6)?

(6) If the coin had been flipped at noon, it would have landed heads.

50% seems to be the only reasonable answer.

Now imagine that I don't know the coin is fair. I divide my credence evenly between two hypotheses about the chance of heads—that the chance of heads is 30% and that the chance of heads is 60%. How confident should I be in (6) in this case? A natural answer: $(50\% \times 30\%) + (50\% \times 60\%) = 45\%$. That is, my

[3] van Fraassen himself is not explicit about how contextualism figures in his tenability results. Nevertheless, it is natural to interpret his results within a contextualist framework. See Stefan Kaufman (2005, 2005, 2009) for important work in this tradition.

credence in (6) should be equal to my *expectation* of the conditional chance, just before noon, of the coin landing heads conditional on being flipped.

This data motivates Skyrms' Thesis—a general principle that ties rational credences in counterfactuals to rational expectations of prior chances.[4] To state the thesis, let t be a relevant past time; let $Ch_t(B|A) = x$ be the proposition that the chance, at t, of B conditional on A is equal to x; and let $A \mathrel{\Box\!\!\rightarrow} B$ stand for the counterfactual with antecedent A and consequent B. Then we have:

Skyrms' Thesis
For any rational P: $P(A \mathrel{\Box\!\!\rightarrow} B) = \sum_x x \times P(Ch_t(B|A) = x)$

Skyrms assumption that t is always a past time is not quite right.[5] This issue is not especially important for my purposes, and I will often speak as though the relevant time is in the past.

Like Stalnaker's Thesis, I take Skyrms' Thesis to be a normative thesis. As I will understand the thesis here, it says that *if you're rational*, then your credence in $A \mathrel{\Box\!\!\rightarrow} B$ is equal to your expectation of the chance, at the relevant time, of B conditional on A.[6]

Skyrms' Thesis, I argue, has unacceptable trivializing consequences. Given orthodox probability theory, it entails that if you give positive credence to $A \mathrel{\Box\!\!\rightarrow} B$, and positive credence to $\neg(A \mathrel{\Box\!\!\rightarrow} B)$, then you are sure that the conditional

[4]Note that Skyrms' himself formulates the thesis in terms of *propensities*. Following recent literature on Skyrms' Thesis, I replace talk of propensities with talk of objective chance. (See, among others, Williams (2012), Moss (2013), and Schulz (2017) for formulations of Skyrms' Thesis in terms of objective chances.) I will assume a broadly Lewisian account of chance.

[5]Take, for example:

(1) If I caught the four o'clock train today, I would make it to the meeting by five.

The probability that I assign to (1) is equal to the *present* chance of making the meeting conditional on catching the four o'clock train. In general, when we evaluate counterfactual conditionals whose antecedents concern events that will occur at some future time, we set our credence in the counterfactual to our expectation of the present chance of the consequent given the antecedent.

[6]There are known counterexamples to Skyrms' Thesis involving *counterlegals*—counterfactuals whose antecedents concern events that violate the laws of nature. A counterlegal may have positive probability even though its antecedent has chance zero, in which case the chance of the consequent conditional on the antecedent is undefined. To deal with cases like this, one option is to use *Popper functions*, which would allow conditional chances to be well-defined even if the conditioned proposition is chance zero. Another option is to treat Skyrms' Thesis as a special case of a more general thesis stated in terms of *hypothetical* probability functions. On this view, one's credence in a counterfactual is given by the probability of the consequent conditional on the antecedent, relative to a hypothetical probability distribution that assigns positive probability to the antecedent and is suitably related to one's actual probability distribution. If this hypothetical probability distribution matches the objective chances whenever the latter are defined, Skyrms' Thesis would come out as a special case of this more general norm. (See Edgington (2008) for discussion.) For the purposes of this paper, I do not need to take a stand on how to handle counterlegals and other counterfactuals with chance-zero antecedents, so I set these aside.

chance of B given A is equal to one or zero—you do not give positive credence to non-extreme hypotheses about the conditional chance of B given A.

Here is the triviality argument. First observe that Skyrms' Thesis entails (a) and (b) below (I omit time references for readability):

(a) For any rational P, if $P(A \boxright B) = 1$, then $P(Ch(B|A) = 1) = 1$

(b) For any rational P, if $P(A \boxright B) = 0$, then $P(Ch(B|A) = 1) = 0$

If you are certain that $A \boxright B$ is true, then you are certain that the prior chance of B conditional on A is one. If you are certain that the counterfactual is false, then you are certain that the prior chance of B conditional on A is zero.

Now consider any rational probability function P. As Lewis assumed in his triviality results for Stalnaker's Thesis, I assume that the class of rational probability functions is closed under conditionalization: if P is rational, and $P(A) > 0$, then $P(\cdot|A)$ is rational.

Suppose that $P(A \boxright B) > 0$ and $P(\neg(A \boxright B)) > 0$. Then (a) entails (c), and (b) entails (d):

(c) $P(Ch(B|A) = 1|A \boxright B) = 1$

(d) $P(Ch(B|A) = 1|\neg(A \boxright B)) = 0$

And, by the Law of Total Probability, we know (e):

(e) $P(Ch(B|A) = 1) = P(Ch(B|A) = 1|A \boxright B) \times P(A \boxright B) + P(Ch(B|A) = 1|\neg(A \boxright B)) \times P(\neg(A \boxright B))$

(c), (d), and (e) together give us:

(f) $P(Ch(B|A) = 1) = P(A \boxright B)$

And finally, by another application of Skyrms' Thesis, (f) entails (g):

(g) $P(Ch(B|A) = 1) = \sum_x x \cdot P(Ch(B|A) = x)$

(g) entails that you are sure that the chance of B given A is either one or zero: $P(0 < Ch(B|A) < 1) = 0$.[7] Skyrms' Thesis has allowed us to derive this conclusion from the assumption that you give positive credence to $A \boxright B$ and positive credence to $\neg(A \boxright B)$. This result is unacceptable.

[7]Suppose, for reductio, that for some y s.t. $0 < y < 1$, $P(Ch(B|A) = y) > 0$. We know that: $\sum_x x \cdot P(Ch(B|A) = x) \geq (y \cdot P(Ch(B|A) = y)) + (1 \cdot P(Ch(B|A) = 1))$. This entails: $\sum_x x \cdot P(Ch(B|A) = x) \geq (y \cdot P(Ch(B|A) = y)) + P(Ch(B|A) = 1)$. Since the first term is positive, it follows that: $\sum_x x \cdot P(Ch(B|A) = x) > P(Ch(B|A) = 1)$. And that contradicts (g). So we conclude that $P(Ch(B|A) = y) = 0$.

4 The Conditional Principal Principle

Skyrms' Thesis seemed plausible on first glance, but closer inspection revealed it to be untenable. Where do we go from here? To answer this question, I turn to the literature on *chance-deference norms*. Chance-deference norm are norms governing the relationship between our credences and our beliefs about objective chance. They tell us to defer to the chances when setting our opinions. Viewed abstractly, Skyrms' Thesis is also a kind of chance-deference norm; it tells us to defer to certain prior *conditional chances* when setting our credences in counterfactuals.[8] If we think of Skyrms' Thesis as a counterfactual chance-deference principle, it is natural to formulate Skyrms' Thesis in a way that mirrors our preferred formulation of the unconditional chance-deference norm. My starting point is David Lewis's *Principal Principle*. After introducing this principle, I propose a counterfactual version of it—the *Conditional Principal Principle*. (A context-sensitive version of this—the Local Conditional Principal Principle—will be my final proposal.)

Note that Lewis's Principal Principle is just one of many non-trivializing formulations of the norm to defer to objective chance. There are others, such as Ned Hall's *New Principle*.[9] I will not defend the Principal Principle over its rivals. I only wish to describe one plausible formulation of the norm to defer to objective chance, and to construct a counterfactual analogue of that principle. I am confident that we can formulate counterfactual analogues of other chance norms—a counterfactual version of Hall's New Principle, for example—but I leave this for future research.

Following Hall (2004) among others, I will state Lewis's Principal Principle in terms of an *ur-chance function*. Consider any world w. If w contains an earliest moment, then the ur-chance function of w—denoted ch_w—is a function that takes a proposition and returns its chance at the earliest moment of w. Later chance functions are defined in terms of ch_w as follows. Where $H_{t,w}$ is a com-

[8] I am not the first to draw an analogy between Skyrms' Thesis and chance-deference norms. See Schulz (2017) for extended discussion. Schulz endorses a counterfactual analogue of the Principal Principle that is similar to my Conditional Principal Principle. There are important differences between my principle and the one defended by Schulz, however. One important difference is that he distinguishes two chance functions: the chance function that figures in the Principal Principle itself (the physical chances) and the one that figures in the counterfactual analogue of the Principal Principle (the counterfactual chances). The two chance functions have different properties. I do not distinguish two chance functions—the chance function that figures in the Conditional Principal Principle is the very same chance function as the one that figures in the Principal Principle. (The reason for this is that I intend to derive the Conditional Principal Principle from the Principal Principle (and the Local Thesis)).

[9] See Hall (1994) and Hall (2004).

plete specification of the history at w up to the moment t, then $ch_w(\cdot|H_{t,w})$ is a function that takes a proposition and returns its chance, at t, in w. Let P_o be any reasonable initial credence function. Let 'π' be a rigid designator that picks out a particular ur-chance function. Let 'Ch' be a definite description for 'the initial chance, whatever it is'. Finally, let E be any total body of evidence that is compatible with the proposition $Ch = \pi$. With this notation, we state the Principal Principle as follows.

The Principal Principle
$P_o(A|E \wedge Ch = \pi) = \pi(A|E)$

Suppose that your total evidence is E. And suppose that you learn what the initial chance function is, which is to say that you learn $Ch = \pi$. Then, the Principal Principle says, you should adopt the opinions π would have were it given your evidence E. To use Ned Hall's metaphor, this version of the Principal Principle tells us to treat chance as an *analyst expert*. We defer to the initial objective chance function not because it has evidence than we don't have—it doesn't have any evidence at all—but because we think it's especially good at *evaluating* evidence. Upon learning what the initial chance function is, we feed it our evidence, and then defer to its *conditional opinions*—the opinions it would have, if it knew everything we know.[10]

I suggest that we replace Skyrms' Thesis with a counterfactual version of the Principal Principle, which I will call *Conditional Principal Principle*. Now, my statement of the Principal Principle tells us to give the chances all of our evidence and then align our credences with the objective chances conditional on our total evidence. Clearly, this won't work in the counterfactual case. Suppose I know that I did not strike the match at noon. In that case, the chance, at noon, that the match would light conditional on my striking it *and all of my evidence* is undefined. But, we may suppose, my credence in (7) is close to one.

(7) If I had struck the match at noon, it would have lit.

So we can't give the initial chances all of our evidence, as we did with the Principal Principle. But what body of evidence should we use instead? An immediate answer that won't work: give the chances all of our evidence *minus* our evidence

[10] You might wonder how could this principle be useful for ordinary subjects. It tells us what to do if we learn what the entire initial chance function is like, but we're never in that situation. What we do learn are facts about the chances of *specific* propositions—the proposition that the chance that a certain die will land on an even number is 50%, for example. However, this concern fails to appreciate the strength of my formulation of the Principal Principle when combined with the laws of probability and our definition of chance. In fact, one show that the formulation of the Principal Principle in the main text entails Lewis's more familiar formulation. See Pettigrew (2016), chapter 9.

that the antecedent is false. This won't work because in any ordinary context in which I assert (7), I also know that the match did not light. The prior chance that the match would light conditional on being struck and *this* piece of knowledge of mine is equal to zero. But again, we may suppose that my credence in (7) is close to one.

Here is what I suggest. Instead of giving the chances all of our evidence, we give them the subset of our evidence *that we hold fixed when we evaluate the counterfactual*. Let me take a moment to explain just what this subset is because it will be very important in what follows.

When we evaluate a counterfactual, we imagine a hypothetical scenario in which the antecedent is true and ask ourselves whether the consequent is also true in that scenario. To do this, we temporarily release some of our knowledge—our knowledge of the antecedent's falsity, among other things. But we don't release all of our knowledge, as philosophers have long observed. We hold much of what we know fixed. Take, for instance, Adams' famous example:

(8) If Oswald hadn't shot Kennedy, someone else would have.

When we evaluate (8), we tend to hold fixed our knowledge of how things went before the assassination—that Oswald acted alone, that he was not part of a conspiracy, and so forth. We clearly do not hold fixed all of our knowledge of how things went *after* the assassination—that the papers reported that Kennedy was shot, that his funeral took place, or that Johnson assumed the presidency in 1963.

My *Conditional Principal Principle* says that your credence in the counterfactual $A \mathrel{\Box\!\!\rightarrow} B$, upon learning what the ur-chance function is, should be equal to the ur-chance of B conditional on A *and the evidence you're holding fixed*. To state the principle, we use the notation that we introduced to state the Principal Principle. We let P_o be any reasonable initial credence function. We let $Ch = \pi$ be the proposition that the ur-chance function is identical to π. And we let E be any total body of evidence that is compatible with the proposition $Ch = \pi$. We introduce one new piece of notation: E^- will be the set of worlds consistent with all of the information that is held fixed when evaluating the counterfactual—a strict superset of E. The Conditional Principal Principle is as follows.

The Conditional Principal Principle
$P_o(A \mathrel{\Box\!\!\rightarrow} B|E \wedge Ch = \pi) = \pi(B|A \wedge E^-)$

Suppose you learn that the initial chance function is π. Then the Conditional Principal Principle says: if you're rational, then your credence in the counterfactual $A \mathrel{\Box\!\!\rightarrow} B$ equals the credence that π would have in B given A *if* π *were given all of the information you are holding fixed*.

Think about it this way. The information that you hold fixed is the information that you judge relevant to determining whether B would have been true if A had been true. In the example of Kennedy's assassination, for instance, you hold fixed what you know about the events leading up to Oswald pulling the trigger, as well as your general knowledge about presidential assassinations, among other things. The initial chance function π should have this information if it is to determine how likely it is that someone else shoots Kennedy supposing Oswald doesn't. You don't hold fixed that Oswald shot Kennedy, and that, as a result, nobody else did. Intuitively, this information is irrelevant to what would have happened if Oswald *hadn't* shot Kennedy. So the initial chance function π has no use for this information.

The Conditional Principal Principle says that, upon learning that the initial chance function is π, give π all of the information that you hold fixed when evaluating (8), *if Oswald hadn't shot Kennedy, someone else would have*. Then ask π: Given this information, how likely do you think it is that someone else shoots Kennedy supposing Oswald doesn't? If you're rational, the answer to this question is the credence that you assign to (8).

Before we move on, a brief methodological remark. In this essay, I will not provide a novel theory of what is held fixed (relative to a context of utterance). We have a rough-and-ready understanding of the notion, and I believe that is enough for my main goal in this paper—to formulate a plausible, context-sensitive version of Skyrms' Thesis, and to show how it can be derived from the Principal Principle and the Local Thesis. If the reader would benefit from a more concrete account, I propose that we adopt the causal independence account presented by Dorothy Edgington (2004). On this view, the set of worlds consistent with what we hold fixed is, roughly, the set of worlds consistent with all of the facts about history before the antecedent time, as well as the facts concerning the time between antecedent and consequent that are *causally independent* of the antecedent.

5 The Local Conditional Principal Principle

We're on the right track. We have a counterfactual version of the the Principal Principle. Still, the principle is not quite right as it stands. Counterfactuals, it is widely agreed, are context sensitive—which proposition is expressed by an utterance of a counterfactual conditional depends, in part, on the conversational context in which the utterance occurs. But the Conditional Principal Principle does not mention context. So it needs refinement. And, as we will see, the necessary refinements are also sufficient for avoiding the triviality result in §3, as well as a recent triviality due to Williams (2012).

To layer context-sensitivity on top of the Conditional Principal Principle, we need to do two things, both of which will be familiar from when we layered context-sensitivity on top of Stalnaker's Thesis. First, we need to add contextual parameters. Both the counterfactual conditional and the information that is held fixed need to be indexed to a context. Second, we must *coordinate* these two contextual parameters, just as we saw with the Local Thesis—the counterfactual conditional proposition must be indexed to the same context as the information that is held fixed.

Note that what's held fixed doesn't depend purely on context, but also on the antecedent of the conditional. As we've seen, when we evaluate a counterfactual whose antecedent concerns a particular period of time, we hold fixed a broad range of facts about history before that time, but not after. Consider an example from Dorr (2016). Suppose John has had breakfast every day this year. You say:

(9) If John had forgotten to have breakfast on Tuesday, that would have been the first time this year.

To evaluate (9), I hold fixed history before Tuesday—that John had breakfast on Monday, that he had breakfast on Sunday, and so forth. But plainly I do not hold fixed that he had breakfast on Tuesday. Now imagine that you had said (10) instead of (9):

(10) If John had forgotten to have breakfast on Wednesday, that would have been the first time this year.

In that case, I *would* have held fixed that John had breakfast on Tuesday, and I would have assented to (10).

I will write $E_c^-(A)$ to refer to the information that is held fixed in context c when we are evaluating a counterfactual with antecedent A. I will write $A \mathrel{\Box\!\!\to}_c B$ for the proposition expressed by the counterfactual in context c. I will continue to assume that P_o is the uniquely rational initial credence function. With this notation, the Local Conditional Principal Principle is as follows.

The Local Conditional Principal Principle

$P_o(A \mathrel{\Box\!\!\to}_c B|E_c \wedge Ch = \pi) = \pi(B|A \wedge E_c^-(A))$

To illustrate, suppose that you are the speaker of a certain context c. The Local Conditional Principle says that, if you're rational, then upon learning that the ur-chance function is π, the credence that you assign to your counterfactual—the proposition expressed by the counterfactual, relative to your context c—is equal to the chance, relative to π, of B conditional on A *and all of the information held fixed in c, relative to antecedent A*.

Thanks to this contextual coordination, the Local Conditional Principal Principle is not subject to the triviality result presented in §3. To see this, recall that Skyrms' Thesis implies (a) and (b) below (I omit time references for readability):

(a) For any rational P, if $P(A \mathrel{\Box\!\!\rightarrow} B) = 1$, then $P(Ch(B|A) = 1) = 1$

(b) For any rational P, if $P(A \mathrel{\Box\!\!\rightarrow} B) = 0$, then $P(Ch(B|A) = 1) = 0$

Taken together, (a) and (b) imply that, if you give positive credence to $A \mathrel{\Box\!\!\rightarrow} B$ and positive credence to $\neg(A \mathrel{\Box\!\!\rightarrow} B)$, then $P(A \mathrel{\Box\!\!\rightarrow} B) = P(Ch(B|A) = 1$.

The Local Conditional Principle escapes the triviality result because it does not entail (a) and (b). Instead, it entails (a′) and (b′), where P^{E_c} is our (uniquely rational) initial probability function P_o conditioned on evidence E_c:

(a′) If $P^{E_c}(A \mathrel{\Box\!\!\rightarrow}_c B) = 1$, then $P^{E_c}(Ch(B|A \wedge E_c^-(A)) = 1) = 1$

(b′) If $P^{E_c}(A \mathrel{\Box\!\!\rightarrow}_c B) = 0$, then $P^{E_c}(Ch(B|A \wedge E_c^-(A)) = 1) = 0$

But if we have (a′) and (b′) in place of (a) and (b), we can block the next step of the argument. For (a′) and (b′) do not entail (c′) and (d′):

(c′) $P^{E_c}(Ch(B|A \wedge E_c^-(A)) = 1|A \mathrel{\Box\!\!\rightarrow}_c B) = 1$

(d′) $P^{E_c}(Ch(B|A \wedge E_c^-(A)) = 1|\neg(A \mathrel{\Box\!\!\rightarrow}_c B)) = 0$

And without (c′) and (d′) we can't complete the argument that we used to refute Skyrms' Thesis.[11]

[11] Let me briefly explain why. Notice that (a) and (b) have the following form:

1. For all rational P, if $P(X) = 1$, then $P(Y) = 1$.
2. For all rational P, if $P(X) = 0$, then $P(Y) = 0$.

Observe that if $P(X) > 0$ and $P(\neg X) > 0$, then (1) entails (3) and (2) entails (4):

3. $P(Y|X) = 1$
4. $P(Y|\neg X) = 0$.

Given the Law of Total Probability, (3) and (4) entail that $P(Y) = P(X)$. This is the argument I used to show that (a) and (b) entail that $P(A \mathrel{\Box\!\!\rightarrow} B) = P(Ch(B|A) = 1)$.

(1) says that there are two propositions X and Y, such that, for *any* rational probability function P, if $P(X) = 1$, then $P(Y) = 1$. It is clear that (a) has the same quantificational structure, with $A \mathrel{\Box\!\!\rightarrow} B$ taking the place of X and $Ch(B|A) = 1$ taking the place of Y. But notice that (a′) does not have the same form. It does not assert that there is any *single* proposition Y such that, for *all* rational probability functions P, if $P(A \mathrel{\Box\!\!\rightarrow} B) = 1$, then $P(Y) = 1$. Rather, the proposition that takes the place of Y is coordinated with the probability function. And likewise for (b′).

The Local Conditional Principle is also not subject to a recent triviality proof due to Williams (2012). (Note that my presentation of Williams' argument differs from his own presentation; Williams' original argument targets Skyrms' Thesis, but I am interested in exploring how a version of it might be used to refute the Conditional Principal Principle. Although the details differ, the basic strategies behind the arguments are the same.) Consider a rational subject in context c who has no evidence, and thus, is not holding any evidence fixed. The Principal Principle, applied to our subject in c, entails (a) below (where $A \mathrel{\Box\!\!\rightarrow}_c B$ is the proposition expressed by the counterfactual in c):

(a) $P_o(A \mathrel{\Box\!\!\rightarrow}_c B|Ch = \pi) = \pi(A \mathrel{\Box\!\!\rightarrow}_c B)$

Since nothing is being held fixed in c, the Local Conditional Principal Principle entails:

(b) $P_o(A \mathrel{\Box\!\!\rightarrow}_c B|Ch = \pi) = \pi(B|A)$

Notice that (a) and (b) together entail (c):

(c) $\pi(A \mathrel{\Box\!\!\rightarrow}_c B) = \pi(B|A)$

Now, as Williams observes, *that* equation looks a lot like Stalnaker's Thesis. The indicative conditional has been replaced with a counterfactual and the rational credence function with an objective chance function. But the Lewisian triviality results that refute Stalnaker's Thesis do not presuppose any particular interpretation of the conditional operator, nor do they depend on any particular interpretation of probability. Perhaps, then, we can use a version of Lewis's argument to refute the Conditional Principal Principle.

There are two critical lemmas in Lewis's argument, stated in terms of chance and counterfactuals below:

Lemma 1. $\pi(A \mathrel{\Box\!\!\rightarrow}_c B|B) = 1$

Lemma 2. $\pi(A \mathrel{\Box\!\!\rightarrow}_c B|\neg B) = 0$

If we can derive these two lemmas from the Local Conditional Principal Principle and the Principal Principle, then we can use Lewis's reasoning to derive the absurd conclusion that, if $\pi(B) > 0$ and $\pi(\neg B) > 0$, then $\pi(B|A) = \pi(B)$. That is, if the initial chance function π assigns positive probability to B, and positive probability to $\neg B$, then B is probabilistically independent of A, relative to

π.[12] Fortunately, Lemma 1 and Lemma 2 don't follow from the Local Conditional Principal and the Principal Principle. To see why not, consider how we might try to derive Lemma 1, following the argument in (a) to (c) above. The first step would be to obtain (a′) from the Principal Principle:

(a′) $P_o(A \mathrel{\Box\!\!\rightarrow}_c B|Ch = \pi \wedge B) = \pi(A \mathrel{\Box\!\!\rightarrow}_c B|B)$

The second step would be to obtain (b′) from the Local Conditional Principle (in a moment we'll see that this is the step that's blocked):

(b′) $P_o(A \mathrel{\Box\!\!\rightarrow}_c B|Ch = \pi \wedge B) = \pi(B|A \wedge B) = 1$

The third step would be to derive Lemma 1 from (a′) and (b′). The problem with this argument is that the Local Conditional Principal Principle does not entail (b′) because it requires $A \mathrel{\Box\!\!\rightarrow}_c B$ to be coordinated with the information that is held fixed in context c, relative to A. Thus, it requires the probability of $A \mathrel{\Box\!\!\rightarrow}_c B$, conditional on $Ch = \pi$, to be equal to $\pi(B|A \wedge B)$ *only if B is held fixed in context c*. But, by hypothesis, B is not held fixed in c—nothing is held fixed in c.

Before moving on, let's take a moment to draw some broader morals about my contextualist response to the triviality results for Skyrms' Thesis. Contextualist responses to Lewisian triviality results for Stalnaker's Thesis exploit the fact that indicative conditionals are information sensitive. When your information changes, the proposition that you express when you utter an indicative conditional changes. Recall that the Local Thesis says that the probability that you assign to *your* indicative conditional (the proposition expressed by the indicative in your context) must be equal to your conditional credence in its consequent given its antecedent. The Local Thesis escapes Lewisian triviality by enforcing coordination between the indicative conditional and the probability function. You might have thought that we can't use this strategy to avoid triviality results targeting Skyrms' Thesis; after all, it is standardly assumed that counterfactuals are not information sensitive. I have argued that this is a mistake. On my view, both indicative conditionals and counterfactuals are information sensitive, and what distinguishes them is the kind of information that they are sensitive to.

[12] Remember that (c) says that $\pi(A \mathrel{\Box\!\!\rightarrow}_c B) = \pi(B|A)$. So if we can show that $\pi(A \mathrel{\Box\!\!\rightarrow}_c B) = \pi(B)$, we can conclude that $\pi(B|A) = \pi(B)$. Here is the proof of $\pi(A \mathrel{\Box\!\!\rightarrow}_c B) = \pi(B)$ from Lemma 1 and Lemma 2.

$$\begin{aligned}\pi(A \mathrel{\Box\!\!\rightarrow}_c B) &= \pi(A \mathrel{\Box\!\!\rightarrow}_c B|B) \times \pi(B) + \pi(A \mathrel{\Box\!\!\rightarrow}_c B|\neg B) \times \pi(\neg B)\\ &= 1 \times \pi(B) + 0 \times \pi(\neg B)\\ &= \pi(B)\end{aligned}$$

The step from the first line to the second uses the Law of Total Probability. The step from the second to the third uses Lemma 1 and Lemma 2.

Like other contextualists about indicatives, I say that the content of an indicative conditional, relative to a given context, depends on what's known in that context. Likewise, the content of a counterfactual conditional also depends on a contextually-determined body of information—the information that is being held fixed in that context, relative to the antecedent of the counterfactual. When your information changes, what you're holding fixed changes, and so the content of the counterfactual changes, too. It is for this reason that the Local Conditional Principal Principle is not subject to the triviality results in §3 or to the triviality result due to Williams (2012).[13]

6 A Sketch of a Theory of Conditionals

My goal when I started this paper was to derive a plausible, contextualist-friendly version of Skyrms' Thesis from a plausible, contextualist-friendly version of Stalnaker's Thesis and a plausible chance-deference norm. We now have the first three ingredients. Our chance-deference norm is the Principal Principle. Our contextualist-friendly version of Stalnaker's Thesis is the Local Thesis. And our contextualist-friendly version of Skyrms' Thesis is the Local Conditional Principal Principle. Here I turn to the final ingredient—the theory of conditionals.

I develop a theory on which all of the semantic differences between indicative conditionals and counterfactuals boil down to differences in what is held fixed in the context in which we evaluate the conditional. Following Stalnaker and others, I say that when we evaluate indicative conditionals, we hold fixed all of our knowledge. And, as we've seen in previous sections, when we evaluate counterfactuals, we hold fixed a contextually-determined subset of our knowledge. Because indicative conditionals and counterfactual conditionals, on my view, differ only in what is held fixed, there is a systematic connection between the truth conditions for indicatives and the truth conditions for counterfactuals. Roughly, a counterfactual is true, relative to our present context, just in case the corresponding indicative conditional is true relative to the information we are holding fixed.

[13] Here I take a leaf out of the *suppositional* theorist's book. Suppositional theories of conditionals aim to give theories of conditionals that are based on probabilistic considerations. (See Adams (1966) and Edgington (1995; 2008), among others.) Some suppositional theorists—most notably, the theory of Edgington (2008)—maintain that indicative and counterfactual conditionals are to be distinguished by their characteristic mode of supposition. For indicatives, we have supposition in the *indicative* mode, and for counterfactuals we have supposition in the *subjunctive* mode. (See Joyce (1999) for discussion.) Although suppositional theorists tend to be *expressivists* about conditionals, we can separate this commitment to expressivism from the commitment to giving a theory of conditionals based on probabilistic considerations. Andrew Bacon (2015) has shown that, using the resources of contextualism, we can give a truth-conditional theory of conditionals that is based on probabilistic considerations. In this essay, I'm arguing that we can do the same for counterfactuals.

In this section, I show how to implement this idea within a Stalnakerian *selection semantics* framework for conditionals.

Stalnaker's theory is a uniform theory of conditionals. He states the truth conditions for conditionals in terms of a contextually-supplied *selection function* f. This is a function that takes a world w, and an antecedent A, and yields a world where A is true—the *selected A-world*, relative to w. Then Stalnaker says that, a conditional, whether indicative or subjunctive, is true at a world w just in case the selected antecedent-world, at w, is a consequent world.

To adopt a uniform theory of conditionals is not, of course, to say that indicatives and counterfactuals have the same meaning. They do not. Indicative conditionals are about epistemic possibilities; counterfactuals usually concern possibilities that are incompatible with our knowledge. Adams's famous minimal pair highlights the contrast:

(8) If Oswald didn't shoot Kennedy, someone else did.

(11) If Oswald hadn't shot Kennedy, someone else would have.

While (11) strikes us as a dubious claim about an alternative course of history, (8) looks straightforwardly true.

How do we account for these differences within a uniform theory? Stalnaker proposes that the selection function we use to evaluate indicative conditionals is subject to a special constraint: roughly, the selected antecedent-world must be an epistemically possible world. Here is a precise statement of the constraint on indicative selection functions.

> **Stalnaker's Constraint**
> If $A \cap E_c \neq \emptyset$, then if $w \in E_c, f_c(w, A) \in E_c$.

If A is compatible with the information associated with context c, then for any world w in E_c, the selected A-world, at w, is also in E_c. Counterfactuals, Stalnaker says, are not subject to this constraint; their selection functions may reach outside the set of epistemically possible worlds.

I am going to account for the differences between indicatives and counterfactuals in a different way. It is clear that we cannot uphold Stalnaker's Constraint, in its current form, for counterfactuals. One response is to dispense with the constraint altogether, as Stalnaker seems to suggest. But another response is to replace it with something else. On an abstract level, it is not hard to see what the replacement should be. For indicatives, Stalnaker's Constraint requires that the selected antecedent-world be one where everything we're holding fixed when we evaluate *the indicative* is true—a world where *everything* we know is true. For counterfactuals, the selected antecedent-world should be one where everything we're holding fixed when we evaluate the *counterfactual* is true—a world where *some* of what we know is true, the part we're holding fixed.

To implement this idea, I propose that a conditional, whether indicative or subjunctive, is evaluated relative to a *conditional information function*. This is a function s that takes an information state E and delivers a selection function that is *Stalnakerian relative to E*—a selection function that satisfies the constraints that Stalnaker imposes on indicative selection functions, relative to information state E. The constraint that matters for my purposes is a generalized version of Stalnaker's Constraint, *Generalized Stalnaker's Constraint* (I leave the others to a footnote):[14]

Generalized Stalnaker's Constraint
If $E \cap A \neq \emptyset$, then for all $w \in E$, $s(E)(w, A) \in E$.

The only difference between indicative and subjunctive conditionals, on my view, is that they supply different *arguments* to the conditional information function. For an indicative conditional, the argument to the conditional information function is E_c, the set of worlds compatible with everything we know. This gives us the following semantic entry, which is roughly equivalent to Stalnaker's own theory of indicative conditionals:

Indicative Selection Semantics
$[\![A > B]\!]^{c,w,s} = 1$ *iff* $s(E_c)(w, A) \in B$

This says: An indicative conditional is true at a world w, relative to a context c and conditional information function s, just in case s takes E_c—the information associated with c—to a selection function that takes w and the antecedent A to a world where the consequent B is true. The selection function is Stalnakerian relative to E_c so it satisfies the Generalized Stalnaker's Constraint relative to E_c. This means that we evaluate an indicative conditional by checking whether the consequent holds at an antecedent world that is compatible with everything we know.

For counterfactuals, the informational argument to the conditional informa-

[14]The other four constraints are:

Success. $s(E)(w, A) \in A$ if $A \neq \emptyset$.

Minimality. $s(E)(w, A) = w$ if $w \in A$

Absurd. Where γ is an absurd world that makes all sentences true, $s(E)(w, A) = \gamma$ if and only if $A = \emptyset$

CSO. If $s(E)(w, A) \in B$ and $s(E)(w, B) \in A$, then $s(E)(w, A) = s(E)(w, B)$.

tion function is the set of worlds consistent with what we are holding fixed, relative to the antecedent of the counterfactual.[15] (Remember that what we hold fixed for counterfactuals varies by antecedent.) The semantic entry is as follows:

Counterfactual Selection Semantics
$[\![A \mathrel{\Box\!\!\rightarrow} B]\!]^{c,w,s} = 1$ iff $s(E_c^-(A))(w, A) \in B$

This says: A counterfactual is true at a world w, relative to a context c and conditional information function s, just in case s takes $E_c^-(A)$—the information that we hold fixed, relative to antecedent A—to a selection function that takes w and the antecedent A to a world where the consequent B is true. The selection function is Stalnakerian relative to $E_c^-(A)$ so it satisfies the Generalized Stalnaker's Constraint relative to $E_c^-(A)$. This means that we evaluate a counterfactual conditional by checking whether the consequent holds at an antecedent world that is compatible with everything we're holding fixed.

On my theory, both indicatives and counterfactuals are governed by Generalized Stalnaker's Constraint, and there are no other differences between the selection functions that we use to interpret the two kinds of conditional. As a result, there is a close connection between the truth conditions for indicatives and the truth conditions for counterfactuals. To make this connection precise, let me introduce some notation. Consider a context c. $E_c^-(A)$ is, to repeat, the set of worlds compatible with everything we're holding fixed in c, relative to antecedent A. Let c^- be a hypothetical context in which our information is characterized by $E_c^-(A)$. In other words, $E_c^-(A) = E_{c^-}$. My theory predicts:[16]

$$[\![A \mathrel{\Box\!\!\rightarrow} B]\!]^{c,s} = [\![A > B]\!]^{c^-,s}$$

The proposition expressed by the counterfactual $A \mathrel{\Box\!\!\rightarrow} B$ relative to $\langle c, s\rangle$ is identical to the proposition expressed by the indicative $A > B$ relative to $\langle c^-, s\rangle$.

We have my neo-Stalnakerian uniform theory conditionals in place. I will close this section by giving three brief arguments for my uniform theory of conditionals, on which both indicatives and counterfactuals are subject to Generalized Stalnaker's Constraint, and there are no other differences between indicative and counterfactual selection functions.[17]

[15]Thanks to David Boylan for discussion.

[16]*Proof*. Suppose $A \mathrel{\Box\!\!\rightarrow} B$ is true relative $\langle c, w, s\rangle$. By the Counterfactual Selection Semantics, it follows that $s(E_c^-(A))(w, A) \in B$. Then, by the Indicative Selection Semantics and the definition of c^-, it follows that $A > B$ is true relative to $\langle c^-, w, s\rangle$.

Now suppose that $A > B$ is true relative to $\langle c^-, w, s\rangle$. By the Indicative Selection Semantics and the definition of c^-, it follows that that $s(E_c^-(A))(w, A) \in B$. Then, by the Counterfactual Selection Semantics, it follows that $A \mathrel{\Box\!\!\rightarrow} B$ is true relative $\langle c, w, s\rangle$.

[17]Thanks to Matt Mandelkern and Harvey Lederman for discussion.

First is an abductive argument based on the main claims of this paper. We have good reason to believe that some version of Skyrms' Thesis is true. I argue that we can derive this principle from the Local Thesis, the Principal Principle, and the unified semantics that I propose, on which both indicatives and subjunctives are subject to Generalized Stalnaker's Constraint. This gives us some reason to believe that the premises of that derivation are true. Since one of the premises is my uniform theory of conditionals, we have some reason to believe that this uniform theory is right.

A second, closely related argument concerns the fact that the probability one assigns to a counterfactual is often equal to the probability that one assigned to the corresponding indicative conditional at an earlier time. Suppose I know the Lakers are playing the Clippers in the NBA semi-finals, and that whoever wins that series will go on to the NBA finals and play the Celtics. Before the series starts I am confident in the indicative conditional (12).

(12) If the Lakers beat the Clippers, they will win the NBA championship.

The series between the Lakers and the Clippers concludes and the Clippers have won. I now endorse the counterfactual:

(13) If the Lakers had beat the Clippers, they would have won the NBA championship.

The probability I now assign to the counterfactual (13) at the conclusion of the series matches the probability I assigned to the indicative (12) at the start of the series. My theory easily accounts for this observation. If we assume that the information I hold fixed when evaluating (13) is identical to my total evidence at the earlier time when evaluating (12), then, on my theory, the proposition I am evaluating now *just is* the proposition I was evaluating then. And if these are just the same propositions, then of course I assign them equal probability.

A final argument concerns *presupposition*. Some contemporary research about presupposition starts from the idea that the presuppositions of a clause must be satisfied relative to their *local contexts*. The local context of an embedded clause is, very roughly, the information information that is *already available*—the information that we can draw on to evaluate the clause—in the course of processing the sentence. Schlenker (2009) develops an algorithm for calculating local contexts, which says, very roughly, that the local context for an embedded clause is the strongest proposition that you can add to that clause without changing the truth-value of the whole sentence at any world in the global context. Mandelkern and Romoli (2017) have shown that, given Stalnaker's semantics for indicatives, this algorithm rightly predicts that the local context for the antecedent of an indicative conditional is (using my notation) E_c, the set of worlds consistent with the information associated with the context. Stalnaker's Constraint plays a critical role in their argument. For, together with Stalnaker's other constraints on selection functions, Stalnaker's Constraint entails that, for any $w \in E_c$, and any A compatible with E_c, $f(w,A) = f(w, A \cap E_c)$. And once this constraint is in place, it is not hard to show that adding E_c to the antecedent won't change the truth-value of the conditional at worlds compatible with our information.

Now, the local context for the antecedent of a counterfactual clearly isn't E_c. Counterfactual antecedents are often inconsistent with what we know. So their local contexts can't contain all of our information. But they do seem to contain some of our information. Take an example from Heim (1992). You and I both know that Mary went to the party. I'm wondering whether she attended with her partner John. You're pretty sure John didn't attend, and you say:

(14) If John had attended too, I would have seen him.

An utterance of (14) is predicted to be felicitous only if the presupposition of its antecedent—that a salient individual attended the party—is satisfied relative to its local context. Hence, our theory of local contexts had better predict that the local context of the antecedent of (14) entails that a salient person—in this case, Mary—attended the party. In light of examples like (14), a natural hypothesis is that the local context of the antecedent of a counterfactual is the set of worlds consistent with what we're holding fixed.[18] Generalized Stalnaker's Constraint will play a central role in deriving this prediction, just as we saw with indicatives. For again, together with the other constraints, Stalnaker's Constraint entails that, for any $w \in E_c^-(A)$, and any A compatible with $E_c^-(A)$, $f(w,A) = f(w, A \cap E_c^-(A))$. Once this constraint is in place, adding $E_c^-(A)$ to the antecedent won't change the truth-value of the conditional at worlds compatible with our information.

7 Deriving the Local Conditional Principal Principle

Now that I have outlined my theory of conditionals, I am ready to show how we can use that theory to derive the Local Conditional Principal Principle from the Local Thesis and the Principal Principle.[19]

[18]See Heim (1992) who makes a similar suggestion.

[19]See Moss (2013) for a derivation of a version of Skyrms' Thesis for *future-directed* subjunctive conditionals from the Principal Principle. Moss does not show how to extend her argument to the case of past-directed subjunctive conditionals. But, as she notes, this does not mean that her argument has no implications for past subjunctives. If the proposition expressed by an earlier utterance of a future-directed subjunctive is the very same proposition as the proposition expressed by a current utterance of a past-directed subjunctive, then constraints on credences in

So that we have everything in front of us, here is the Local Conditional Principal Principle:

Local Conditional Principal Principle
$P_o(A \mathrel{\Box\!\!\rightarrow}_c B | E_c \wedge Ch = \pi) = \pi(B | A \wedge E_c^-(A))$

Remember that P_o is the uniquely rational initial credence function; E_c is the information associated with context c; and $E_c^-(A)$ is the information that is held fixed, relative to antecedent A.

My derivation of the Local Conditional Principal Principle will rely on three principles. We have already seen two of these principles—the Local Thesis and the Principal Principle, repeated below.

Local Thesis
$P_o(A >_c B | E_c) = P_o(B | A \wedge E_c)$

The Principal Principle
$P_o(A | E \wedge Ch = \pi) = \pi(A | E)$

The third is a principle we have not yet seen that concerns the relationship between E_c, our actual information, and $E_c^-(A)$, the information we hold fixed when we evaluate a counterfactual with antecedent A. That principle is:

Independence
$P_o(A \mathrel{\Box\!\!\rightarrow}_c B | E_c^-(A)) = P_o(A \mathrel{\Box\!\!\rightarrow}_c B | E_c)$

Independence says that the probability of $A \mathrel{\Box\!\!\rightarrow}_c B$ conditional on $E_c^-(A)$ (the evidence we hold fixed in context c) is equal to the probability of $A \mathrel{\Box\!\!\rightarrow}_c B$ conditional on E_c (our evidence in context c).

To see why this assumption is warranted, remember what $E_c^-(A)$ is supposed to represent. The information that you hold fixed when you evaluate a counterfactual with antecedent A is the information that you judge relevant to determining what would have happened if A had been true. The information that you do not hold fixed is information you do not judge relevant to determining what would have happened if A had been true. Recall the case of Kennedy's assassination. You hold fixed a broad range of facts about what happened before Oswald pulled the trigger—that Oswald acted alone, that he was not part of a conspiracy, and so forth. You do not hold fixed what happened after the assassination—that Oswald shot Kennedy, that nobody else shot Kennedy, or that Johnson assumed the presidency in 1963. Consider a rational subject who knows everything you're holding fixed, and nothing more—that is, a rational subject whose total evidence consists

future-directed subjunctives will entail constraints on credences in past-directed subjunctives. In many ways, then, the project of this essay is quite friendly to Moss's framework.

of everything you know about history before Oswald pulled the trigger, and nothing you know about history after. Independence says that the probability that this subject assigns to the counterfactual (8), *if Oswald hadn't shot Kennedy, nobody else would have*, equals the probability that you assign to the counterfactual. In other words, learning what you're not holding fixed—that is, learning that Oswald shot Kennedy, that nobody else shot Kennedy, that Johnson assumed the presidency in 1963, and so forth—should not change her view about the counterfactual (8). For if it did, then E_c (your total evidence) must know something that $E_c^-(A)$ (the evidence you're holding fixed) doesn't know and that you judge relevant to determining what would have happened if Oswald hadn't shot Kennedy. But in that case you should have been holding it fixed![20]

To make things concrete, I will show how the derivation of the Local Conditional Principal Principle works for a specific example. Suppose that on Monday at noon you are deciding whether to buy a lottery ticket. For simplicity, I will assume that you know $Ch = \pi$. Suppose that you decide not to purchase the ticket. Later you are evaluating the counterfactual:

(16) If you had bought the ticket, you would have lost.

Let c be the context in which you are evaluating (16). E_c is your current information and, where *Buy* is the proposition that you buy the ticket, $E_c^-(Buy)$ is the information you hold fixed when evaluating (16). Let c^- be your context at noon, just before deciding not to buy the ticket. I will assume that the information associated with c^- just is the information that you hold fixed in c—specifically, everything you knew before deciding not to buy the ticket and nothing you have learned since.

Our derivation begins with an instance of the Principal Principle (where *Win* is the proposition that you win the lottery).

[20] Independence places constraints on the relationship between E_c (your evidence) and $E_c^-(A)$ (what you hold fixed, relative to antecedent A). It might be helpful to look at a specific case in which Independence is satisfied. Often when you're evaluating a counterfactual $A \mathrel{\Box\!\!\rightarrow} B$, the relationship between E_c and $E_c^-(A)$ is the following: $E_c = E_c^-(A) \cap (\neg A \wedge \neg B)$. That is, your current knowledge is the result of intersecting what you hold fixed with the negation of the antecedent and the negation of the consequent. Take the coin case. I decide not to flip a fair coin. I am evaluating the counterfactual (6), *if the coin had been flipped, it would have landed heads*. I hold fixed all of my knowledge *except* for my knowledge of the fact that I did not flip the coin and that, as a result, it did not land heads. If this case, Independence says (where $E_c^-(Heads)$ is the set of worlds consistent with what I hold fixed):

(15) $P_o(Heads \mathrel{\Box\!\!\rightarrow}_c Flip | E_c^-(Heads)) = P_o(Heads \mathrel{\Box\!\!\rightarrow}_c Flip | E_c^-(Heads) \cap (\neg Flip \wedge \neg Heads))$

In the next section, I will show that that this instance of Independence holds in van Fraassen's *Stalnaker-Bernoulli* models—the models that van Fraassen (1976) and Bacon (2015) use to establish the tenability of the Local Thesis.

(a) $P_o(Win|Buy \wedge E_c^-(Buy)) = \pi(Win|Buy \wedge E_c^-(Buy))$

(a) says that your credence, at noon, that you win the lottery, conditional on buying the ticket, is equal to the initial chance of winning conditional on buying a ticket and your total evidence at noon. (This follows from the Principal Principle because we have stipulated that you know that the initial chance function is π.)

Next, observe that (b) follows from (a) and the Local Thesis:

(b) $P_o(Buy >_{c^-} Win|E_c^-(Buy)) = \pi(Win|Buy \wedge E_c^-(Buy))$

Your credence, at noon, in $Buy >_{c^-} Win$—the proposition expressed by the indicative conditional relative to your information at noon—is equal to the initial chance of winning conditional on buying a ticket and your total evidence at noon.

On page 19, I showed that the counterfactual $Buy \,\Box\!\!\rightarrow_c Win$ is the very same proposition as the indicative $Buy >_{c^-} Win$. This means that the probability of the former is always equal to the probability of the latter. Thus, (c) follows from (b):

(c) $P_o(Buy \,\Box\!\!\rightarrow_c Win|E_c^-(Buy)) = \pi(Win|Buy \wedge E_c^-(Buy))$

Your credence, at noon, in $Buy \,\Box\!\!\rightarrow_c Win$—the proposition expressed by the counterfactual relative to your *present* context, after deciding not to buy the lottery ticket—is equal to the initial chance of winning conditional on buying and your total evidence at noon.

Next, we appeal to Independence, which says that the probability of $Buy \,\Box\!\!\rightarrow_c Win$, conditional on $E_c^-(Buy)$—what you hold fixed—is equal to your credence $Buy \,\Box\!\!\rightarrow_c Win$, conditional on you E_c—your total evidence.[21] Applying Independence to (c) gives us:

(d) $P_o(Buy \,\Box\!\!\rightarrow_c Win|E_c) = \pi(Win|Buy \wedge E_c^-(Buy))$

We said that $E_c^-(Buy)$—your evidence at noon—entails $Ch = \pi$. Since E_c entails $E_c^-(Buy)$, it follows that E_c also entails $Ch = \pi$. Thus, (d) entails (e):

(e) $P_o(Buy \,\Box\!\!\rightarrow_c Win|E_c \wedge Ch = \pi) = \pi(B|A \wedge E_c^-(Buy))$

And (e) is an instance of the Local Conditional Principal Principle.

[21] One way to secure Independence in this example is be to assume that: $E_c = E_c^-(Buy) \cap \neg Buy$. (See §8 for explanation.) This assumption seems plausible given the setup of the case.

We have shown that under the assumption that you know $Ch = \pi$, the Local Conditional Principal Principle follows from the Principal Principle, the Local Thesis, and Independence.[22]

The details of the derivation are somewhat involved, so let me take a moment to walk through it in a more informal way. Suppose you are evaluating a counterfactual with antecedent A and consequent B. Consider a subject whose total evidence is the evidence that you hold fixed (and, we will assume, who knows all of the chance facts). By the Principal Principle, her credence in B given A equals the conditional chance of B given A, and by the Local Thesis, her credence in B given A equals her credence in her indicative conditional—that is, the proposition expressed by the indicative conditional, relative to her information. Thus it follows that her credence in her indicative is equal to the conditional chance of B given A.

Now, according to the uniform semantics for conditionals that I have proposed, the proposition expressed by the indicative, relative to her information, is equivalent to the proposition expressed by the corresponding counterfactual $A \mathrel{\Box\!\!\rightarrow} B$, relative to your context. This means that her credence in her indicative is equal to her credence in your counterfactual. But remember that she has all of the information that you have and that you judge relevant to evaluating the counterfactual. Thus, it stands to reason that *your* credence in your counterfac-

[22] Here's how it goes when $E_c^-(Buy)$ doesn't 'know' the chance of *Win* conditional on *Buy*. Let c_{ch} be the context that results from updating the information in c, E_c, with the proposition $Ch = \pi$. I will assume that what's held fixed in this new context is the intersection of what's held fixed in c and $Ch = \pi$. So we want to show that:

$$P_o(Buy \mathrel{\Box\!\!\rightarrow}_{c_{ch}} Win | E_c \wedge Ch = \pi) = \pi(Win | Buy \wedge (E_c^-(Buy) \wedge Ch = \pi))$$

We begin with an instance of the Principal Principle:

(a) $P_o(Win | Buy \wedge (E_c^-(Buy) \wedge Ch = \pi)) = \pi(Win | Buy \wedge E_c^-(Buy))$

Next, (a) entails (b):

(b) $P_o(Win | Buy \wedge (E_c^-(Buy) \wedge Ch = \pi)) = \pi(Win | Buy \wedge (E_c^-(Buy) \wedge Ch = \pi))$

(The reason that (a) entails (b) is that it is a consequence of the Principal Principle that the initial chance function knows that it is the initial chance function: $\pi(Ch = \pi) = 1$.)

Let c^- be any context such that the information associated with c^- is the information that is held fixed in c_{ch}. Then (b) and the Local Thesis entail:

(c) $P_o(Buy >_{c^-} Win | E_c^-(Buy) \wedge Ch = \pi) = \pi(Win | Buy \wedge (E_c^-(Buy) \wedge Ch = \pi))$

By the theory of conditionals outlined in §6, (c) entails (d):

(d) $P_o(Buy \mathrel{\Box\!\!\rightarrow}_{c_{ch}} Win | E_c^-(Buy) \wedge Ch = \pi) = \pi(Win | Buy \wedge (E_c^-(Buy) \wedge Ch = \pi))$

And finally, (d) and Independence entail:

(e) $P_o(Buy \mathrel{\Box\!\!\rightarrow}_{c_{ch}} Win | E_c \wedge Ch = \pi) = \pi(Win | Buy \wedge (E_c^-(Buy) \wedge Ch = \pi))$

tual should equal *her* credence in your counterfactual.

If your credence in your counterfactual is equal to her credence in your counterfactual, which, in turn, is equal to her credence in her indicative, then your credence in your counterfactual is equal to her credence in her indicative. And we have already seen that her credence in her indicative is equal to the conditional chance of *B* given *A*. So, putting everything together, it follows that your credence in your counterfactual is equal to the conditional chance of *B* given *A*, just as the Local Conditional Principal Principle requires.

8 Looking Forward: Tenability

Bas van Fraassen (1976) and Andrew Bacon (2015) have shown that the Local Thesis is *tenable* within a Stalnakerian semantic framework. There are non-trivial models in which the Local Thesis holds for the Stalnaker conditional. On an abstract level, it is not hard to see how to extend these results to establish the tenability of the Local Conditional Principle. That principle follows from the Principal Principle, the Local Thesis, and Independence. So if there are non-trivial models in which all three of these principles hold, then there are non-trivial models in which the Local Conditional Principal Principle holds—that is, the Local Conditional Principal Principle is tenable. (Note: I am only going to talk about *simple conditionals* in this section—conditionals with non-conditional antecedents and consequents.)

Here is a simplified overview of van Fraassen's *Stalnaker-Bernoulli* models.[23] We evaluate conditional sentences relative to *sequences* of worlds. The first world in the sequence represents all of the *non-conditional* propositions that are true at the sequence—that is, all of the facts that can be specified without mentioning conditionals. And the rest of the sequence represents the conditional facts. To construct a Stalnaker-Bernoulli model, we begin with a set of worlds $\mathcal{I}$, which I will take to be the set of worlds compatible with all of the non-conditional information in a given context. We define $O_{\mathcal{I}}$ as the set of all sequences of worlds in $\mathcal{I}$. For example, if $\mathcal{I} = \{w_1, w_2, w_3\}$, then:

$$O_{\mathcal{I}} = \{\langle w_1, w_2, w_3\rangle, \langle w_1, w_3, w_2\rangle, \langle w_2, w_1, w_3\rangle\langle w_2, w_3, w_1\rangle, \langle w_3, w_1, w_2\rangle, \langle w_3, w_2, w_1\rangle\}$$

A non-conditional sentence *A* is true at a sequence just in case the first world in that sequence is an *A*-world. A conditional $A > B$ is true at a sequence just in case the first *A*-world in the sequence is also a *B*-world. Suppose, for instance,

[23] I am heavily indebted to lecture notes from Justin Khoo and Paolo Santorio for this presentation. See Khoo and Santorio, 'Lecture Notes: Probabilities of Conditionals in Modal Semantics.'

that A is true at w_1 and w_2, but false at w_3, and that B is true at w_1, but false at w_2 and w_3. Then the conditional $A > B$ is true at $\langle w_1, w_2, w_3 \rangle$, $\langle w_1, w_3, w_2 \rangle$, and $\langle w_3, w_1, w_2 \rangle$, and false at the other three sequences.

To model the probabilities of conditionals, van Fraassen provides a recipe for taking us from a probability function P defined over $\mathcal{I}$ to a probability function P' defined over $O_{\mathcal{I}}$. He shows that the resulting probability function (1) *extends* P in the sense that $P'(A) = P(A)$ for all non-conditional A, and (2) for simple conditionals, the probability of the conditional is the corresponding conditional probability (whenever the conditional probability is defined). Although van Fraassen himself is not explicit about the role of context-sensitivity, there is a natural way of interpreting his results within a contextualist framework. If you feed the construction E_c and P_c—the information associated with c, and the probability function associated with c, respectively—the construction will output an interpretation of the conditional $A >_c B$ and an extended probability function P'_c such that Stalnaker's equation holds for $A >_c B$ relative to P'_c. This establishes the tenability of the Local Thesis for simple conditionals.

Because P' extends P, if we start with a probability function P that obeys the Principal Principle with respect to non-conditional sentences, then P' will also obey the Principal Principle with respect to non-conditional sentences. So there is no obstacle to upholding both the Principal Principle (with respect to non-conditional sentences) and the Local Thesis.

My derivation of the Conditional Principal Principle also relied on the principle of Independence, repeated below.

Independence
$P_o(A \mathrel{\Box\!\!\to}_c B | E_c^-(A)) = P_o(A \mathrel{\Box\!\!\to}_c B | E_c)$

Independence can be reformulated as a principle about indicative conditionals. Let c^- be a hypothetical context in which our information is characterized by $E_c^-(A)$. On my theory, the proposition expressed by the counterfactual, relative to context c, is identical to the proposition expressed by the indicative conditional, relative to context c^-. So Independence becomes:

Indicative Independence
$P_o(A >_{c^-} B | E_c^-(A)) = P_o(A >_{c^-} B | E_c)$

Stated as a principle about indicatives, it is easy to show that certain instances of this principle are consistent with the Local Thesis. Indeed, we can show that a special case of the principle is entailed by the Local Thesis, given a Stalnakerian semantics for the conditional. Let $E_c = E_c^-(A) \cap \neg A$, and assume that $P_o(\neg A | E_c^-(A))$. Then Independence says that the proposition expressed by the indicative conditional, relative to c^-, is probabilistically independent of the negation of its antecedent, relative to the information in c^-. And this fact is a well-known consequence of the Local Thesis, assuming Stalnaker's selection semantics for conditionals.

Of course, showing that Independence holds for this particular choice of E_c and $E_c^-(A)$ does not show much. That's because it's not normally the case that $E_c = E_c^-(A) \cap \neg A$. Suppose I know that neither you nor your partner went to the party last night, and that you often attend parties together. When I evaluate:

(17) If you had gone to the party, your partner would have gone to the party.

I don't hold fixed that you didn't go, but I also don't hold fixed that your partner didn't go. So we don't get to my present knowledge by intersecting what I hold fixed with the proposition that you did not go to the party—the negation of the counterfactual's antecedent. We get to my present knowledge by intersecting what I hold fixed with the proposition that *neither you nor your partner* attended the party—the conjunction of the negation of the antecedent and the negation of the consequent.

Often our current knowledge results from intersecting what we're holding fixed, relative to some proposition A, with some proposition Q that is *stronger* than $\neg A$. It would be good to show that, for any such Q, the conditional $A >_{c^-} B$ is probabilistically independent of Q relative to $P_o(\cdot|E_c^-(A))$. Formally, where Q is any proposition entailing $\neg A$ such that $P_o(Q|E_c^-(A)) > 0$:

$$P_o(A >_{c^-} B|E_c^-(A)) = P_o(A >_{c^-} B|E_c^-(A) \cap Q)$$

Stated in terms of counterfactuals, this becomes:

$$P_o(A \mathrel{\Box\!\!\to}_c B|E_c^-(A)) = P_o(A \mathrel{\Box\!\!\to}_c B|E_c^-(A) \cap Q)$$

This says: The probability of the proposition expressed by the counterfactual in c, conditional on what's held fixed in c, is equal to the probability of the proposition expressed by the counterfactual in c, conditional on the intersection of what's held fixed in c and Q, where Q is any proposition that entails $\neg A$ (and is assigned positive probability by $P_o(\cdot|E_c^-(A))$). In the appendix, I show that this fact holds in our simplified Stalnaker-Bernoulli models. This establishes the tenability of many plausible instances of Independence with respect to these models.

Let's look at a simple example. Suppose that $E_c^-(A) = \{w_1, w_2, w_3, w_4\}$. Suppose that A is true at w_1 and w_2, but false at w_3, and w_4. And suppose that B is true at w_1 and w_3, but false at w_2 and w_4. If we assume that each world has equal probability, then the probability $A >_{c^-} B$, relative to $E_c^-(A)$, is equal to the proportion of sequences of $O_{E_c^-(A)}$ whose first A-world is B-world. It is easy to verify that $O_{E_c^-(A)}$ contains 24 sequences and that 12 of these sequences are such that their first A-world is a B-world. So the probability of the conditional $A >_{c^-} B$, relative to $E_c^-(A)$, is 1/2.

Now consider a $\neg A$-entailing factual proposition—say, $\neg A \wedge \neg B$. This proposition is true at all and only the sequences in $O_{E_c^-(A)}$ whose first world is w_4:

$\langle w_4, w_1, w_2, w_3\rangle, \langle w_4, w_1, w_3, w_2\rangle, \langle w_4, w_2, w_1, w_3\rangle$
$\langle w_4, w_2, w_3, w_1\rangle, \langle w_4, w_3, w_1, w_2\rangle, \langle w_4, w_3, w_2, w_1\rangle$

There are six sequences in total beginning with w_4, three of which are such that their first A-world is a B-world. So the conditional is true at half of sequences beginning with w_4, which is to say that the probability of $A >_{c^-} B$, conditional on $\neg A \wedge \neg B$, is again 1/2. The conditional is probabilistically independent of $\neg A \wedge \neg B$, relative to $E_c^-(A)$. Formally we have shown that:

$$P_o(A >_{c^-} B|E_c^-(A)) = P_o(A >_{c^-} B|E_c^-(A) \cap (\neg A \wedge \neg B))$$

Stated in terms of counterfactuals, this says:

$$P_o(A \mathrel{\Box\!\!\rightarrow}_c B|E_c^-(A)) = P_o(A \mathrel{\Box\!\!\rightarrow}_c B|E_c^-(A) \cap (\neg A \wedge \neg B))$$

The same will be true of any factual Q entailing $\neg A$. Zoom in on the set of sequences in $O_{E_c^-(A)}$ that make Q true. The proportion of sequences in this new set whose first A-world is B-world will be equal to 1/2.

We have seen that there are Stalnaker-Bernoulli models—of the simplified variety that I have presented in this section—in which each of the Principal Principle, the Local Thesis, and Independence holds. In each of these, the Local Conditional Principal Principle holds, too. I leave a full tenability proof—one that dispenses with the simplifying assumptions I have made here—to future research.

9 Conclusion

The project of this article has been to sketch a neo-Stalnakerian, uniform theory of conditionals that allows us to derive a plausible, contextualist-friendly version of Skyrms' Thesis (the Local Conditional Principal Principle) from a plausible, contextualist-friendly version of Stalnaker's Thesis (the Local Thesis) and a plausible chance-deference norm (the Principal Principle). I close by outlining two questions for future research.

One question is about chance. I used the Principal Principle to derive a version of Skyrms' Thesis. But the Principal Principle is just one candidate chance-deference norm—one way of formalizing the claim that one's credences ought to be guided by objective chances. Other candidates are Hall's *New Principle* and Dorst's *Trust Principle*.[24] One area of future research involves determining whether we can derive counterfactual versions of these principles from the Local Thesis and their non-conditional counterparts.

Another question is about semantics. I've given a semantics for counterfactuals on which the meaning of a counterfactual is closely related to the meaning of an indicative conditional. A full defense of this theory would require showing how to derive this meaning compositionally. I am optimistic about the prospects of this project if one adopts a certain approach to the role of tense in counterfactuals. Let me conclude by saying something about the approach I favor. Plausibly, a 'would'-conditional is composed of a 'will'-conditional under a past tense operator. (For defense of this claim, see, for example, Ippolito (2013).) There are two main hypotheses about what this past tense operator does, one of which I take to be particularly promising: the *past-as-modal* view, on which the past tense is interpreted as a modal.[25] (The alternative approach is the *past-as-past* view, on which the past tense has its usual temporal meaning in counterfactuals. See Khoo (2015) for a defense of this approach.) Inspired by Schulz (2014), one understanding of the past-as-modal view says that the past tense shifts the information state relative to which we interpret the embedded indicative conditional. Specifically, the past tense operator shifts the information state from the set of worlds consistent with our actual information to the set of worlds consistent with what we hold fixed. The result is that a counterfactual is true, relative to our present context, just in case the corresponding indicative conditional is true relative to the information we're holding fixed when we evaluate the counterfactual. This, of course, is exactly what my uniform theory of conditionals predicts: the only difference between indicative conditionals and counterfactuals is the information that is held fixed when we evaluate the conditional.[26]

[24]See Kevin Dorst (2020). Note that Dorst's principle is formulated as a principle about deference to one's own (future) evidence, but Ben Levinstein (ms) advocates adopting Trust for deference to chance.

[25]See, for example, Iatridou (2000) and Schulz (2014) for defenses of this approach.

[26]Thanks to Zach Barnett, David Boylan, Fabrizio Cariani, Kevin Dorst, Branden Fitelson, Melissa Fusco, Simon Goldstein, Ben Holguín, Justin Khoo, Arc Kocurek, Harvey Lederman, Matt Mandelkern, Sarah Moss, Milo Phillips-Brown, Bernhard Salow, Paolo Santorio, Robbie Williams, and two anonymous referees for helpful feedback.

10 Appendix

Begin with some notation and terminology.

- Let X be any finite set of worlds.
- Let S be the set of all sequences of worlds in X.
- Let $A >_S B$ be the set of all sequences in S whose first A-world is a B-world.
- Let S_x be the set of sequences in S whose first element is x.
- For any $x \in X$, let X_{-x} be $X - \{x\}$. Let S_{-x} be the set of all sequences of worlds in X_{-x}.

We begin by showing:

> **Claim 1**. For any x such that $x \in X$ and $x \notin A$: $\frac{|A>_S B|}{|S|} = \frac{|A>_{S_x} B|}{|S_x|}$

We will show Claim 1 by proving two sub-claims that together entail Claim 1. Those claims are (1) and (2) below, for any x such that $x \in X$ and $x \notin A$:

1. $\frac{|A>_S B|}{|S|} = \frac{|A>_{S_{-x}} B|}{|S_{-x}|}$
2. $\frac{|A>_{S_{-x}} B|}{|S_{-x}|} = \frac{|A>_{S_x} B|}{|S_x|}$

Proof of (1).

Let $|X^{-x}| = n$. Then $|S_{-x}| = n!$ and $|S| = (n+1)!$. We know that for any sequence $o \in S_{-x}$, there are exactly $n+1$ sequences in S that preserve the order of the elements in o. Roughly, that is because, for any $o \in S_{-x}$, there are $n+1$ places where we can insert x: at the beginning of the sequence, after the first element, after the second element, and so forth. For example, consider $o = \langle w_1, w_2, w_3, \ldots, w_n \rangle$. There are $n+1$ sequences in S that preserve the order of the elements of o:

$\langle x, w_1, w_2, w_3, ... \rangle$

$\langle w_1, x, w_2, w_3, ... \rangle$

$\langle w_1, w_2, x, w_3, ... \rangle$

$\langle w_1, w_2, w_3, x ... \rangle$

And so forth. For each of these sequences, the first A-world in the sequence will be a B-world just in case the first A-world in o—the original sequence—is a B-world. So, for any $o \in A >_{S_{-x}} B$, there are exactly $n+1$ sequences in $A >_S B$. This means that we can reason as follows:

$$\begin{aligned}
\frac{|A >_S B|}{|S|} &= \\
\frac{|A >_{S_{-x}} B|(n+1)}{(n+1)!} &= \\
\frac{|A >_{S_{-x}} B|(n+1)}{(n+1)(n!)} &= \\
\frac{|A >_{S_{-x}} B|}{n!} &= \\
\frac{|A >_{S_{-x}} B|}{|S_{-x}|}
\end{aligned}$$

Proof of (2).

The sequences in S_x are the same as the sequences in S_{-x}, except that x is tacked on to the beginning of each. Let $f : < w_1, \ldots, w_n > \mapsto < x, w_1, \ldots, w_n >$. Then f is a bijection from S_{-x} to S_x as well as from $A >_{S_{-x}} B$ to $A >_{S_x} B$. So Claim 2 immediately follows: $\frac{|A >_{S_{-x}} B|}{|S_{-x}|} = \frac{|A >_{S_x} B|}{|S_x|}$

We have shown Claim 1. Next we want to show Claim 2 (where $|S_Q|$ is the set of sequences in S whose first world is a Q-world and $|A >_{S_Q} B|$ be the set of sequences in S_Q whose first A-world is a B-world):

> **Claim 2**. Where Q is any proposition that entails $\neg A$: $\frac{|A >_S B|}{|S|} = \frac{|A >_{S_Q} B|}{|S_Q|}$

Proof of Claim 2.

Let $Q = \{x_1, \ldots, x_n\}$. We know:

- $|S_Q| = |S_{x_1}| + \ldots + |S_{x_n}|$
- $|A >_{S_Q} B| = |A >_{S_{x_1}} B| + \ldots + |A > S_{x_n} B|$

Then we can reason as follows:

$$\frac{|A>_{S_Q} B|}{|S_Q|} =$$

$$\frac{|A>_{S_{x_1}} B| + \ldots + |A>_{S_{xn}} B|}{|S_{x_1}| + \ldots + |S_{xn}|} =$$

$$\frac{|A>_{S_{x_1}} B|(n)}{|S_{x_1}|(n)} =$$

$$\frac{|A>_{S_{x_1}} B|}{|S_{x_1}|} =$$

$$\frac{|A>_S B|}{|S|} =$$

The third line follows from the second because (a) $|S_{x_1}| = |S_{x_2}| = \ldots = |S_{x_n}|$ and (b) $|A >_{S_{x_1}} B| = |A >_{S_{x_2}} B| = \ldots = |A >_{S_{x_n}} B|$. And Claim 1 secures the inference from the fourth line to the fifth line.

11 References

1. Bacon, Andrew (2015). 'Stalnaker's thesis in context.' *Review of Symbolic Logic* 8, pp. 131 - 163.

2. Bradley, Richard (2000). 'A Preservation Condition for Conditionals.' *Analysis* 60, pp. 219-222.

3. Dorst, Kevin (2020). 'Evidence: A Guide for the Uncertain.' *Philosophy and Phenomenological Research* 100, pp. 586-632.

4. Edgington, Dorothy (2008). 'Counterfactuals.' *Proceedings of the Aristotelian Society* 108, pp. 1-21.

5. von Fintel, Kai (1998). 'The presupposition of subjunctive conditionals.' In Orin Percus & Uli Sauerland (eds.), *The interpretive tract*, Vol. 25, pp. 25-44. Cambridge, Massachusetts: MIT Working Papers in Linguistics.

6. van Fraassen, Bas C. (1976). 'Probabilities of conditionals.' In *Foundations of probability theory, statistical inference, and statistical theories of science*, Springer, pp. 261?308.

7. Gillies, Anthony (2009). 'On the truth conditions for If (but not quite only If).' *Philosophical Review* 118, 325 - 349.

8. Hall, Ned (1994). 'Correcting the Guide to Objective Chance.' *Mind* 103, pp. 504-18.

9. Hall, Ned (2004). 'Two Mistakes about Credence and Chance.' *Australasian Journal of Philosophy* 82, pp. 93-111

10. Harper, William L., Robert Stalnaker & Glenn Pearce (1981). *Ifs: Conditionals, Belief, Decision, Chance, and Time*. D. Reidel Publishing Company, Dordrecht.

11. Holguin, Ben (2020). 'Knowledge in the Face of Conspiracy Conditionals.' *Linguistics and Philosophy*.

12. Iatridou, Sabine (2000). 'The Grammatical Ingredients of Counterfactuality.' *Linguistic Inquiry* 31, pp. 231 - 270.

13. Ippolito, Michela (2013). 'Subjunctive Conditionals: A Linguistic Analysis.' *Linguistic Inquiry Monograph* (Series 65), Cambridge: MIT Press.

14. Kaufmann, Stefan (2004). 'Conditioning against the grain: Abduction and indicative conditionals.' *Journal of Philosophical Logic* 33, pp. 583-606.

15. Kaufmann, Stefan (2005). 'Conditionals predictions: A probabilistic account.' *Linguistics and Philosophy* 28, pp. 181-231.

16. Kaufmann, Stefan (2009). 'Conditionals right and left: Probabilities for the whole family.' *Journal of Philosophical Logic* 38, pp. 1-53.

17. Khoo, Justin (2015). 'On Indicative and Subjunctive Conditionals.' *Philosopher's Imprint* 15, pp. 1-40.

18. Khoo, Justin (2020). 'The meaning of *If*.' Ms.

19. Levinstein, Ben (2020). 'Accuracy, Deference, and Chance.' Ms.

20. Lewis, David (1976). 'Probabilities of conditionals and conditional probabilities.' *Philosophical Review* 85, pp. 297-315.

21. Lewis, David (1980). 'A Subjectivist's Guide to Objective Chance.' In Jeffrey (ed.) *Studies in Inductive Logic and Probability,* Vol. 2, pp. 263-93. Berkeley: University of California Press.

22. Lewis, David (1994). 'Humean Supervenience Debugged.' *Mind* 103, pp. 473-490.

23. Mandelkern, Matthew (2019). 'If p, then p!' Ms.

24. McGee, Vann (1989). 'Conditional probabilities and compounds of conditionals.' *Philosophical Review* 98, 485-541.

25. Moss, Sarah (2013). 'Subjunctive Credences and Semantic Humility.' *Philosophy and Phenomenological Research* 87, pp. 251 - 278

26. Schulz, Katrin (2014). 'Fake tense in conditional sentences: A modal approach.' *Natural Language Semantics* 22, pp. 117 - 144.

27. Schulz, Moritz (2017). *Counterfactuals and Probability*. Oxford: Oxford University Press.

28. Schwarz, Wolfgang (2016). 'Subjunctive Conditional Probability.' Forthcoming in *Journal of Philosophical Logic*.

29. Skyrms, Bryan (1980). 'The Prior Propensity Account of Subjunctive Conditionals.' In Harper et al. (1981).

30. Stalnaker, Robert (1968). ‘A Theory of Conditionals.’ In Nicholas Rescher (ed.) *Studies in Logical Theory*(American Philosophical Quarterly Monographs 2), Oxford: Blackwell. pp. 98-112.

31. Stalnaker, Robert (1975). ‘Indicative conditionals.’ *Philosophia* 5, pp. 269-286.

32. Stalnaker, Robert (2002). ‘Common ground.’ *Linguistics and Philosophy* 25, pp. 701-721.

33. Stalnaker, Robert and Richard Jeffrey (1994). ‘Conditionals as random variables.’ In Ellery Ells, Brian Skyrms & Ernest W. Adams (eds.) *Probability and Conditionals: Belief Revision and Rational Decision,* Cambridge: Cambridge University Press.

34. Williams, J. Robert G. (2012). ‘Counterfactual Triviality: A Lewis-Impossibility Argument for Counterfactuals.’ *Philosophy and Phenomenal Research* 85, pp. 648 - 670