Journal of Philosophical Logic (2007) 36: 97–121
DOI: 10.1007/s10992-006-9030-2

MARTIN SMITH

CETERIS PARIBUS CONDITIONALS AND COMPARATIVE NORMALCY

Received 27 October 2005

ABSTRACT. Our understanding of subjunctive conditionals has been greatly enhanced through the use of possible world semantics and, more precisely, by the idea that they involve variably strict quantification over possible worlds. I propose to extend this treatment to *ceteris paribus conditionals* – that is, conditionals that incorporate a ceteris paribus or 'other things being equal' clause. Although such conditionals are commonly invoked in scientific theorising, they traditionally arouse suspicion and apprehensiveness amongst philosophers. By treating ceteris paribus conditionals as a species of variably strict conditional I hope to shed new light upon their content and their logic.

KEY WORDS: ceteris paribus conditional, comparative normalcy, possible world semantics, subjunctive conditional

1. INTRODUCTION

By a 'ceteris paribus conditional' I shall mean a subjunctive conditional in which the relationship between antecedent and consequent is mediated by a qualifying ceteris paribus or 'other things being equal' clause. Any conditional of the form 'If __ were the case, then ceteris paribus ... would be the case' counts as a ceteris paribus conditional. Many law-statements in the sciences have the form of ceteris paribus conditionals, so defined.

Consider, for instance, the principle of natural selection. As formulated by Elliott Sober (1984, pp. 27) the principle of natural selection states the following: If (a) the organisms in a population possessing trait T were better able to survive and reproduce than organisms possessing trait T' and (b) T and T' are heritable traits, then (c) the proportion of organisms in the population with trait T would increase.

The evolution of actual populations will not always conform to this principle, however. There are various possible sources of interference. The genes controlling trait T could, for instance, mutate or they could be linked to genes controlling maladaptive traits. Furthermore, random genetic drift can be a significant factor, particularly in small populations. It would be quite possible for the antecedent of the above conditional to be satisfied without the consequent being satisfied. Thus, as Sober remarks, ‘a ceteris paribus clause needs to be added here’ (Sober, 1984, pp. 27–28).

Ceteris paribus conditionals are not peculiar to scientific theorising. In more colloquial settings, however, we sometimes prefer to use qualifying phrases such as ‘normally,’ ‘ordinarily,’ ‘typically,’ ‘as a rule’ and the like rather than the more formal ‘other things being equal’ or ‘ceteris paribus.’ In my view these colloquial hedging clauses can be perfectly well substituted for ceteris paribus clauses without any change in content. Indeed, most of my examples will be drawn not from science, but the mundane.

In this paper I offer a semantic analysis of ceteris paribus conditionals. I shall employ the framework of possible world semantics. That is, I shall offer an account of how the truth value, at a particular possible world, of a ceteris paribus conditional is determined by the truth values, at various possible worlds, of its antecedent and its consequent.

It is widely believed that ceteris paribus clauses are either (1) shorthand for an explicit list of background provisos or (2) catch-alls that render a conditional logically or vacuously true (see, for instance, Shiffer, 1987, pp. 287, 1991; Hempel, 1988; Earman and Roberts, 1999; Earman et al., 2002). I accept neither alternative. I believe that ceteris paribus clauses are needed precisely when no explicit list of background provisos is available – as is arguably the case with the principle of natural selection (see Pietroski and Rey, 1995, pp. 87). I also believe that ceteris paribus conditionals can express perfectly substantial claims about the world and are governed by a non-trivial logic.

The (1)/(2) dilemma partly stems from the idea that a ceteris paribus clause works by further qualifying or *strengthening* the antecedent of a conditional. I propose that a ceteris paribus clause be viewed not as part of the *antecedent* of a conditional but, rather, as part of the *conditional operator* itself. In this paper, I shall treat the presence of a ceteris paribus clause as a kind of logical or grammatical feature – like the presence of a subjunctive verb – that is correlated with a distinctive sort of conditional operator.

Although I do not regard ceteris paribus conditionals as vacuously true, I do not expect my semantic analysis to necessarily reassure those who do – at least not by itself. Those who believe that ceteris paribus conditionals are devoid of substantial content may take a similar attitude toward the very parameter that my analysis will exploit. They are free to do so. However, I think it is a parameter that is here to stay whether or not we decide to use it in the semantic analysis of ceteris paribus conditionals. It is the relation of *comparative normalcy*.

We are quite comfortable assenting to things like ‘it is *more normal* for a human to have 46 chromosomes than 47’ and ‘it is *more normal* for me to have breakfast at home than at work.’ It is this comparative normalcy relation that will be invoked in my proposed semantic analysis of ceteris paribus conditionals. Naturally, it will be the comparative normalcy of entire *possible worlds*, and not limited states of affairs, that will be at issue. The idea that possible worlds or states of affairs might be ordered to reflect a relation of comparative normalcy is not new – though it is not exactly commonplace either. The idea has been explored in connection with conditional logics for defeasible reasoning (Delgrande, 1987; Boutilier, 1994; Boutilier and Becher, 1995), in connection with subjunctive conditionals (Gundersen, 2004) and in connection with conditional analyses of causation (Menzies, 2004).

Consider the conditional 'If it had not rained today, then we would have gone to the cricket.' David Lewis (1973a, 1973b) has suggested that a bare subjunctive conditional such as this means something like: In any possible world in which the weather is fine and which resembles the actual world as much as the weather being fine permits it to, we go to the cricket. Alternately, the most similar worlds in which it is not raining and we go to the cricket are *more similar* than the most similar worlds in which it is not raining and we do not go to the cricket.

Consider the conditional 'If it had not rained today, then other things being equal we would have gone to the cricket.' I suggest that a ceteris paribus conditional such as this means something like: In any possible world in which the weather is fine and which is as *normal*, from the perspective of the actual world, as the weather being fine permits it to be, we go to the cricket. Alternately, the most normal worlds in which it is not raining and we go to the cricket are *more normal* than the most normal worlds in which it is not raining and we don't go to the cricket. A ceteris paribus clause indicates that the worlds relative to which a subjunctive conditional is evaluated are to be selected on the basis of their *normalcy* and not their *similarity*.

A ceteris paribus conditional and its bare subjunctive counterpart may well diverge in truth value. Suppose that our only viable means of transportation to the cricket is the train and that, unbeknownst to me, the trains are not running as a result of a serious mechanical fault. In this case the bare subjunctive conditional 'If it had not rained today we would have gone to the cricket' will be false. Since the trains are not running at the actual world, the trains are not running at all those fine weather worlds that resemble the actual world as much as the fine weather permits them to. Plausibly, at some of these worlds at least, we are unable to go to the cricket.

In contrast, the ceteris paribus conditional 'If it had not rained today, then other things being equal we would have gone to the cricket' could still be *true*. Even though the trains are not running at the actual world, those fine weather worlds in which the trains *are* running are arguably more normal than those fine weather worlds in which the trains are not. At the actual world, it is more normal for the trains to run than for all the trains to be stopped. Therefore, the trains are running at all those fine weather worlds that are as normal, from the perspective of the actual world, as the fine weather permits them to be. Plausibly, at all these worlds, we go to the cricket. While the no-rain, no-train worlds outrank the no-rain, train worlds with respect to comparative similarity, the opposite is true when comparative normalcy is our measure. This is what accounts for the divergence in truth value between the bare subjunctive conditional and its ceteris paribus counterpart.

In terms of comparative world normalcy, Sober's version of the principle of natural selection has the following truth condition: The most normal worlds in which (a) organisms possessing trait T are better able to survive and reproduce than organisms possessing trait T', (b) T and T' are heritable traits and (c) the proportion of organisms possessing trait T increases, are *more normal* than the most normal worlds in which (a) organisms possessing trait T are better able to survive and reproduce than organisms possessing trait T', (b) T and T' are heritable traits and (d) the proportion of organisms possessing trait T decreases or stagnates.

One might complain, at this point, that judgments of comparative normalcy are *context sensitive* – that is, that the content of at least some comparative normalcy judgments will be sensitive to features of the context of utterance. This is doubtless true – but it is all for the better, since ceteris paribus conditionals are *also* context sensitive. If my analysis is on the right track, then the contextual factors to which ceteris paribus conditionals are responsive are the very same as the contextual factors to which comparative normalcy judgments are responsive. It is the job of a semantic analysis to *respect* context sensitivity and not to eradicate it.

2. Subjunctive Conditionals and Comparative Similarity

Lewis, in providing his semantic analysis of subjunctive conditionals, supplements the standard possible world semantic toolkit with the notion of *comparative world similarity*. He elucidates this relation graphically. Let w be the index world to which all other worlds are being compared. If we visualise possible worlds arranged in space with proximity serving as a metaphor for similarity, then we can imagine a series of concentric spheres radiating out from w – each representing a class of worlds that resemble w equally. This is the mathematician's sense of sphere – a locus of points (in three-dimensional space) equidistant from a single point. Now think of a sphere as solid rather than hollow – containing not just the worlds that compose its surface, but also those that fall inside it. This is how Lewis intends that the term be used. For Lewis the system of spheres associated with an index world can be well-ordered by inclusion or 'size' – a constraint he terms *nesting*.

According to Lewis, the smallest nonempty sphere in a system will be a singleton set containing the index world. Presumably no world resembles the index world more closely or as closely as it resembles itself. The largest sphere in a system will be the set containing all the worlds accessible from the index world. This is intended to be the very same accessibility relation that governs the quantificational range of necessity and possibility operators. The worlds accessible from a world *w* are those worlds that are possible from the perspective of *w*. Say that a sphere around a world *w permits* a sentence φ just in case it contains φ-worlds.

Lewis introduces two conditional operators □→ and ◇→ to be read, respectively, as 'If __ were the case, then ... would be the case' and 'If __ were the case, then ... might be the case' (see Lewis, 1973b, pp. 1–2) Equipped with the notion of comparative similarity, the truth conditions for the two types of subjunctive conditional can be given as follows:

φ □→ ψ is true at a world *w* iff there exists a φ-permitting sphere of similarity around *w* in which all the φ-worlds are also ψ-worlds.
φ ◇→ ψ is true at a world *w* iff in all φ-permitting spheres of similarity around *w*, there exists a (φ ∧ ψ)-world.

Lewis allows for one exception: Despite the fact that its truth condition has the form of an existential quantification over spheres, φ □→ ψ should be deemed vacuously true at *w* if φ is necessarily false – that is, if there is *no* φ-permitting sphere around *w*. This allows us to maintain that subjunctive conditionals are logically weaker than *strict* conditionals. That is, that (φ → ψ) → (φ □→ ψ). A strict conditional φ → ψ, of course, is true at a world just in case *all* accessible φ-worlds are ψ-worlds.

Subjunctive conditionals, like strict conditionals, involve quantification over possible worlds. However, while strict conditionals are associated with a single sphere of accessibility governing their quantificational range, subjunctive conditionals are associated with a set of spheres of accessibility, each governing one *possible* quantificational range. Subjunctive conditionals might, then, be described as *variably strict* (see Lewis, 1973a, 1973b, pp. 13–19). A subjunctive conditional will quantify over as large a sphere as it needs to in order to accommodate the truth of its antecedent (assuming this can be done).

According to Lewis, sphere systems must focus upon, or radiate from, the world to which they are assigned – that is {*w*} is guaranteed to be the smallest nonempty sphere around *w*. Lewis terms this constraint *centering*.

(It is technically convenient to include the empty set amongst the spheres around an index world. Evidently, this makes no difference to the truth conditions for subjunctive conditionals). Some of Lewis' commentators have expressed misgivings about the strength of the centering constraint (Bowie, 1979; Nozick, 1981, pp. 176, 690, 681; Gundersen, 2004). Amongst other things, it lands us with some unusual logical consequences. With centering in place, for instance, a subjunctive conditional with a true antecedent and a true consequent is guaranteed to be true.

In a *weakly* centered system of spheres, while the index world remains an element of the smallest nonempty sphere around itself, it is no longer the *sole* element. Further worlds are permitted to infiltrate. If we supplant centering with weak centering, a subjunctive conditional $\varphi \;\Box\!\!\rightarrow \psi$ with a true antecedent and consequent need not be true – rather it will retain some modal strength. This conditional will be true at a $(\varphi \wedge \psi)$-world just in case all of the φ-worlds in the smallest or innermost nonempty sphere of similarity around the world are also ψ-worlds. I take no stand here on whether the stronger or the weaker centering constraint is preferable. I raise this issue simply as a way of finessing the transition to my own semantics for ceteris paribus conditionals.

3. From Imperative Conditionals to Ceteris Paribus Conditionals

In a centered system of spheres, the index world enjoys a uniquely privileged position as the sole member of the smallest nonempty sphere. In a weakly centered system of spheres, the index world shares this particular privilege with a range of further worlds. Clearly, there is a third possibility. In a *decentered* system of spheres, the index world enjoys *no* privileges whatsoever. In a decentered system, the index world does not even feature as an element of the smallest nonempty sphere. While out of place in the modelling of bare subjunctive conditionals, Lewis suggested that such systems might find a home in the semantic analysis of certain other kinds of conditional – he chose *imperative conditionals* as his example.

When evaluating a conditional such as 'If Jesse robbed the bank, then he ought to return the loot and confess' we are still interested in the properties of possible worlds in which Jesse robs the bank. But, rather than selecting these worlds on the basis of the extent to which they resemble the actual world, we would be better served by examining the degree to which they exemplify some moral ideal. The content of this conditional, according to Lewis, is that the *best* worlds in which Jesse robs the bank are worlds in which he then confesses and returns the loot. Similar suggestions have been made by Bengt Hansson (1969) and Bas Van Fraassen (1972).

If a sphere system is intended to represent the comparative goodness of possible worlds from the perspective of the index world, then there is no reason to expect that the system will be centered or even weakly centered. Indeed, these conditions will only be met on the proviso that the index world estimates itself to be perfect. Ceteris paribus conditionals and imperative conditionals appear to have a good deal in common. Both, I think, implicate some implicit *idealisation* of the world at which they are to be evaluated. However, in the case of a ceteris paribus conditional, this is not a romantic idealisation so much as a *simplifying* one – a practice in which we, surely, just as often indulge.

When evaluating a conditional such as ‘If the supply of oil were to decrease while demand remains constant, then ceteris paribus the price of oil would rise’ we are interested in the properties of certain possible worlds in which the supply of oil decreases while the demand remains constant. What the ceteris paribus clause does, I have suggested, is to signal that these worlds are to be selected on the basis of their *normalcy* or simplicity rather than their similarity to the actual world. The content of the conditional is that the most normal worlds in which the supply of oil decreases while the demand remains constant are worlds in which the price increases.

It would be interesting to compare the notion of a normal world with the idea of an *idealised model* of a phenomenon, as used across the breadth of sciences.[1] In the same way that normal worlds can be ranked according to their degree of normalcy, certain kinds of models can be ranked according to their degree of idealisation. Various models of projectile motion provide one good example (see Shaffer, 2001, pp. 39–41; Arthur and Fenster, 1969, chap. 7). Furthermore, the relationship between ceteris paribus conditionals and the goings-on in idealised models has been noted (Cartwright, 1983, 1999; Pietroski and Rey, 1995; Menzies, 2004). I won’t pursue this comparison further here. Sticking with a widespread custom in deontic logic, I will say nothing precise about the significance of the world orderings to which I appeal.

I introduce two ceteris paribus conditional operators ■→ and ♦→ to be read, respectively, as ‘If__ were the case, then ceteris paribus ... would be the case’ and ‘If __ were the case, then ceteris paribus ... might be the case.’ The truth conditions for the two types of ceteris paribus conditional can be given as follows:

φ ■→ ψ is true at a world *w* iff there exists a φ-permitting sphere of normalcy associated with *w* in which all the φ-worlds are also ψ-worlds.
φ ♦→ ψ is true at a world *w* iff in all φ-permitting spheres of normalcy associated with *w* there exists a $(\varphi \wedge \psi)$-world.

Once again, we need to add the proviso that $\varphi \blacksquare\!\!\rightarrow \psi$ should be deemed vacuously true and $\varphi \blacklozenge\!\!\rightarrow \psi$ vacuously false at a world *w* in case there is no φ-permitting sphere of normalcy associated with *w*.

The failure of weak centering means that a number of notable inference patterns will fail to preserve truth in the logic of ceteris paribus conditionals. The following notable patterns, all of which are perfectly valid in the logic of subjunctive conditionals, are fallacious in the logic of ceteris paribus conditionals:

$$\frac{\begin{array}{l}\varphi\\ \varphi\blacksquare\rightarrow\psi\end{array}}{\psi}\ \text{Modus Ponens}\qquad \frac{\begin{array}{l}\sim\psi\\ \varphi\blacksquare\rightarrow\psi\end{array}}{\sim\varphi}\ \text{Modus Tollens}\qquad \frac{\begin{array}{l}\varphi\\ \sim\psi\end{array}}{\sim(\varphi\blacksquare\rightarrow\psi)}\ \text{Refutation}$$

The following notable patterns, however, remain valid:

$$\frac{\begin{array}{l}\varphi\blacksquare\rightarrow\psi\\ (\psi\wedge\varphi)\blacksquare\rightarrow\chi\end{array}}{\psi\blacksquare\rightarrow\chi}\ \text{Restricted Hypothetical Syllogism}\qquad \frac{\begin{array}{l}\varphi\blacksquare\rightarrow\psi\\ \psi\rightarrow\chi\end{array}}{\varphi\blacksquare\rightarrow\chi}\ \text{Weakening the Consequent}$$

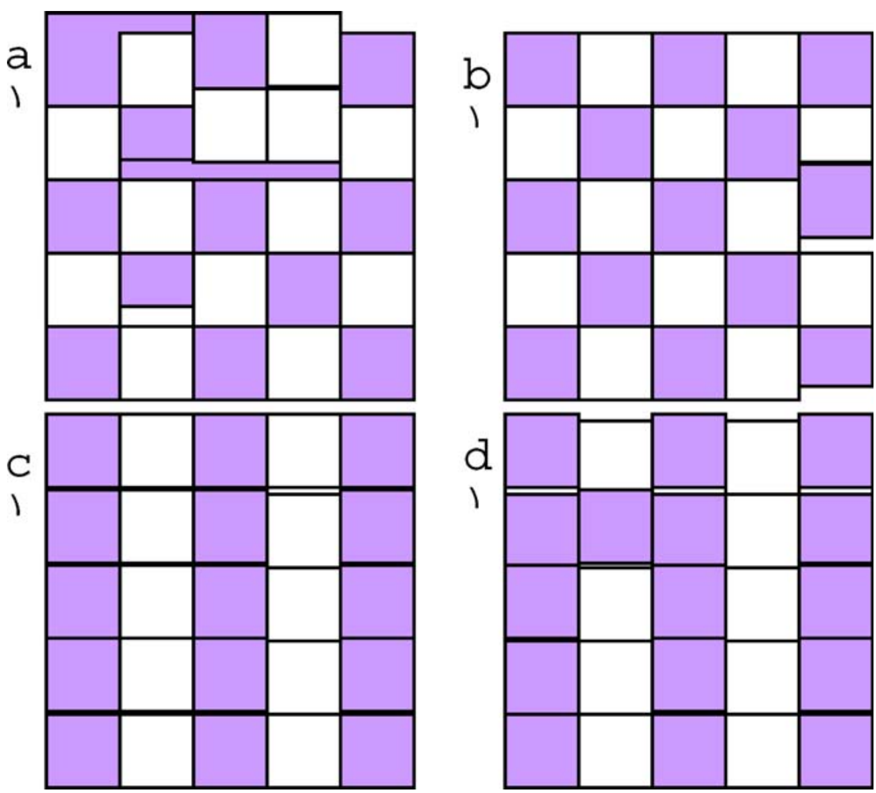
The logic of ceteris paribus conditionals will be explored in some detail in the final section.

Not only will world rankings based upon comparative goodness fail, in general, to center upon the index world, they will fail, in general, to center upon any *single* world. That is to say, the smallest nonempty sphere in a comparative goodness ordering will typically contain more than just one world. After all, why should a single possible world outshine all others? Intuitively, there are innumerable facts about any given world that simply do not bear upon its goodness. Similarly, one would presume that there are countless minutiae about most possible worlds that do not affect their normalcy in any way. The worlds that comprise the smallest nonempty sphere in a comparative normalcy ordering will differ in just these irrelevant respects. Comparative normalcy orderings, then, will typically center upon a non-singleton class of possible worlds.

One might think that the correct ranking of worlds with respect to comparative goodness is unchanged from the perspective of any given world – that is, every world gets assigned the one true comparative goodness ordering. Conventional wisdom regarding the fact/value gap would tend to support this supposition. Plausibly, what *ought* to be the case does not vary as a function of what *is* the case. However, if one held, say, a divine command theory of goodness, according to which what is good is simply that which is dictated to be good by an appropriate deity, then one may well take a different view.

For a divine command theorist, the correct ranking of worlds with respect to comparative goodness would indeed vary from world to world reflecting the incumbent deity's preferences. (Those worlds without an appropriate deity would, presumably, receive an empty or nihilistic sphere system.) The best worlds, from the standpoint of *w*, will be those worlds at which the preferences of the deity incumbent at *w* are all satisfied (as far as this is possible). There will be one comparative goodness ordering corresponding to every possible set of divine preferences. Similarly, the correct ranking of worlds with respect to comparative *normalcy* will, intuitively, vary from world to world – but with explanatory generalisations or tendencies taking over the role of divine preferences. The most normal worlds, from the standpoint of *w*, will be those worlds at which the generalisations or tendencies that play an important explanatory role at the actual world are all *exceptionless* (as far as this is possible).

Consider the following four figures:



If we were to order these figures with respect to their similarity to (a), then plausibly we would obtain the following ranking (from most similar to least similar): (a), (b), (d), (c). ((b) has 24 squares in common with (a), (d) has 17 and (c) has 16). If, on the other hand, we were to order these figures with respect to their faithfulness to the *pattern* suggested by (a) (their 'comparative normalcy' from the perspective of (a)), then plausibly we would obtain the following ranking (from most faithful to least faithful): (b), (a), (d), (c). We can form a defeasible generalisation about the distribution of squares in (a) that happens to describe, *without exception*, the distribution of squares in (b).

Similarly, if we were to order the figures with respect to their similarity to (d) we would obtain the ordering (d), (c), (a), (b) and if we were to order them with respect to their faithfulness to the pattern established by (d) we would obtain the ordering (c), (d), (a), (b). While each of (a), (b), (c) and (d) is associated with its own unique comparative similarity ordering, both (a) and (b) share a comparative normalcy ordering as do (c) and (d). They serve to 'establish' the same recognisable pattern. The standards of comparative normalcy imposed by these figures are not, then, absolute. They are, however, *less relative* or *more robust* than the standards of comparative similarity.

When dealing with objects more complex than these figures, it is possible to generalise about the *exceptions* to generalisations. Consider, for instance, the generalisations 'Most peacocks are blue' and 'Most albino peacocks are white.' These two generalisations can, of course, be made simultaneously exceptionless – namely, by emptying the extension of albino peacocks. However, the second generalisation will still leave its mark upon a comparative world normalcy ordering. It is as if a deity were to issue the following commands: 'All peacocks are to be blue' and 'All albino peacocks are to be white.' The second command is not superfluous. The worlds that most faithfully satisfy the deity's preferences will be worlds in which all peacocks are blue and there are no albino peacocks at all. However, with respect to worlds that *do* contain albino peacocks, those in which all such peacocks are white will better satisfy the deity's preferences than those in which they are not. So it is when normalcy rather than divine preference is our measure.

The idea that standards of normalcy are world relative is related to a view expressed in Section I – namely, that ceteris paribus conditionals can be used to express substantial claims about the world. If standards of normalcy were indifferent to hypothetical variation in the nature of the world then, given my semantic analysis, so too would the truth of ceteris paribus conditionals be indifferent to hypothetical variation in the nature of the world. In this case, ceteris paribus conditionals could never be *contingent*.

Craig Boutilier (1994) proposes a possible world semantics for what he terms 'normative' conditionals – that is, conditionals of the form 'If __, then normally ...'. He suggests that a normative conditional should be deemed true at a possible world *w* iff for every possible world at which the antecedent is true and the consequent false, there is a more normal possible world at which (1) antecedent and consequent are both true and (2) at all possible worlds that are more normal still, if the antecedent is true, then the consequent is true (Boutilier, 1994, pp. 103). Boutilier's semantic apparatus is more austere than mine, consisting of a class of possible worlds and a single ordering relation upon that class intended to represent a relation of comparative normalcy.[2] As such, his analysis depends for its success upon the assumption that standards of normalcy are not themselves world relative.

It is interesting to note that Boutilier does not appear to regard normative conditionals as expressing substantial claims about the world (Boutilier, 1994, pp. 96, 110–116). For Boutilier, normative conditionals serve primarily as a vehicle for expressing a kind of *expectation* or *reasoning preference* – a default or defeasible willingness to accept the consequent given the antecedent. These assumptions could help to motivate a semantic framework in which standards of normalcy are world absolute – but they strike me as dubious. When I say 'If I were to drop this glass, then normally it would break,' I am telling you something (contingent) about the properties *of the glass* – not just something about my reasoning preferences. Furthermore, although there is some relation between the endorsement of a conditional such as this and a default willingness to draw certain inferences, I am not inclined to think that the relationship is straightforward. Having said this, I should point out that Boutilier's primary concern is to provide a semantic analysis of *defeasible inference rules* and that his theory can be assessed in this light quite independently of the issues I have raised here.

One must be careful not to place unrealistic demands upon a semantic analysis. The purpose of my comparative normalcy semantics for ceteris paribus conditionals is not to help us determine the truth value of particular conditionals – at least not on its own. No one would expect a comparative goodness semantics for imperative conditionals to provide genuine moral guidance – and the situation with my analysis is no different. A comparative goodness semantics for imperative conditionals, it might be said, is no substitute for a genuine theory of *what goodness is*. By the same token, a comparative normalcy semantics for ceteris paribus conditionals is no substitute for a genuine theory of *what normalcy is* – and, aside from a few suggestive remarks, this is a matter upon which I have remained silent.[3]

Much the same point can be made regarding the standard possible world semantics for necessity and possibility operators (see Kripke, 1980, pp. 19, footnote 18). Possible world semantics does not tell us what is possible and what isn't. But it is, of course, often used to clarify just what is at stake in such disputes. Similarly, my proposed comparative world normalcy semantics may prove useful in clarifying precisely what is at stake in disputes over particular ceteris paribus conditionals.

Furthermore, although my comparative world normalcy semantics cannot settle disputes over particular ceteris paribus conditionals, it can settle matters of *logic* – at least, in combination with a little reflection upon the formal features of comparative normalcy. My proposed semantics has the potential to *explain* and *motivate*, in a unified way, the logical principles to which ceteris paribus conditionals are intuitively subject. As such, it offers an alternative to piecemeal theorising about the logic of ceteris paribus conditionals. This is perhaps the most significant benefit of a semantic analysis. In the concluding section, I provide technical details of the logic to which ceteris paribus conditionals, and various related operators, are subject and show just how this logic emerges from my analysis.

4. The Logic of Ceteris Paribus Conditionals

I shall begin this final section by setting things up in a purely formal fashion. The formal language **L** to be interpreted includes in its vocabulary countably many sentence letters (A, B, C ...), the sentential constants $\top$ and $\bot$ and punctuation. It also contains the truth functional operators ($\sim$, $\wedge$, $\vee$, $\supset$, $\equiv$), the special operators (■→, ♦→, →, □, ◇) and four more sentential operators to be introduced shortly. The sentences of **L** are built up using the standard recursive clauses for operators and punctuation. The metalanguage, in which the truth conditions for **L** sentences are described, is an extensional, first order language. For ease, I will use the same symbols for the truth functional operators of this language. I will use lower case Greek letters as metalinguistic sentence variables.

Language **L** is interpreted relative to a nonempty set of indices **I** and a function $S(x)$ assigning a system of spheres to each $x \in$ **I**. A sphere is simply a set of indices and a system of such spheres is simply a set of such sets meeting certain structural conditions. I define an associated interpretation function $I_x(y)$ which maps pairs of elements of **I** and sentences of **L** into the set {0, 1}. An interpretation $\mathcal{J}$, then, is a triple <**I**, $S(x)$, $I_x(y)$>. For all interpretations, $I_x(\top) = 1$ and $I_x(\bot) = 0$ for any index x. If $I_i(\varphi) = 1$, we can say that sentence φ holds at i under interpretation $\mathcal{J}$. A sentence φ is *valid* under an interpretation $\mathcal{J}$ just in case it holds at every index – that is, just in case $\forall x \in \mathbf{I}$, $I_x(\varphi) = 1$. A sentence φ might be described as *semantically valid* or *valid* simpliciter just in case it is valid under all permissible interpretations.

The function $I_x(y)$ will meet the following conditions:

$$I_i(\sim \varphi) = 1 \textit{ iff } \sim (I_i(\varphi) = 1)$$
$$I_i(\varphi \wedge \psi) = 1 \textit{ iff } (I_i(\varphi) = 1) \wedge (I_i(\psi) = 1)$$
$$I_i(\varphi \vee \psi) = 1 \textit{ iff } (I_i(\varphi) = 1) \vee (I_i(\psi) = 1)$$
$$I_i(\varphi \supset \psi) = 1 \textit{ iff } (I_i(\varphi) = 1) \supset (I_i(\psi) = 1)$$
$$I_i(\varphi \equiv \psi) = 1 \textit{ iff } (I_i(\varphi) = 1) \equiv (I_i(\psi) = 1)$$
$$I_i(\Box\varphi) = 1 \textit{ iff } \forall S \in S(i),\ \forall x \in S,\ I_x(\varphi) = 1$$
$$I_i(\Diamond\varphi) = 1 \textit{ iff } \exists S \in S(i),\ \exists x \in S,\ I_x(\varphi) = 1$$
$$I_i(\varphi \rightarrow \psi) = 1 \textit{ iff } \forall S \in S(i),\ \forall x \in S,\ I_x(\varphi \supset \psi) = 1$$
$$I_i(\varphi \blacksquare\!\rightarrow \psi) = 1 \textit{ iff } = \exists S \in S(i),\ \big(\exists x \in S,\ I_x(\varphi) = 1 \wedge \forall y \in S,\ I_y(\varphi \supset \psi) = 1\big)$$
$$\vee(\sim \exists S \in S(i),\ \exists x \in S,\ I_x(\varphi) = 1)$$
$$I_i(\varphi \blacklozenge\!\rightarrow \psi) = 1 \textit{ iff } \forall S \in S(i),\ \big(\exists x \in S,\ I_x(\varphi) = 1 \supset \exists y \in \mathrm{S},\ I_y(\varphi \wedge \psi) = 1\big)$$
$$\wedge(\exists S \in S(i),\ \exists \mathrm{x} \in S,\ I_x(\varphi) = 1)$$

The additional disjunct added to the truth conditions for φ ■→ ψ and the additional conjunct added to the truth conditions for φ ♦→ ψ capture Lewis's provisos about impossible antecedents.

This purely formal characterisation obviously leaves it quite neutral just what **I**, $I_x(y)$ and $S(x)$ are to represent. Different thoughts about this give rise to different interpretations of **L**. On my *intended interpretation*, **I** is the totality of possible worlds, $I_i(\varphi) = 1$ *iff* φ is true at world i and $S(x)$ is the function assigning to each world the system of spheres of normalcy associated with that world. That is, every $S \in S(i)$ represents the set of possible worlds that satisfies some given standard of normalcy from the perspective of world i. Given this intended interpretation, I have argued that ■→ and ♦→ can be read as ceteris paribus conditionals.

We can now introduce a pair of 'inner' modal operators (■,♦) to serve as the counterparts of the 'outer' modal operators (□,◇):

$$I_i(\blacksquare\varphi) = 1 \textit{ iff } \exists S \in S(i),\ \forall x \in S,\ I_x(\varphi) = 1$$
$$I_i(\blacklozenge\varphi) = 1 \textit{ iff } \forall S \in S(i),\ (S = \Lambda \vee \exists x \in S,\ I_x(\varphi) = 1)$$

While □φ holds at an index *i* just in case φ holds throughout all spheres assigned to *i*, ■φ holds at an index *i* iff φ holds throughout *some* sphere assigned to *i*. Given my intended interpretation, ■ can be read as '__ would normally be the case' and ◆ as '__ could normally be the case' or 'It would not be abnormal for __ to be the case.' Call these *normalcy* operators.

We can also introduce two binary sentential operators <■ and ≤■ as follows:

$$I_i\big(\varphi <\blacksquare\psi\big) = 1 \textit{ iff } \exists S \in S(i),\ \big(\exists x \in S,\ I_x(\varphi) = 1 \wedge \forall y \in S, \sim \big(I_y(\psi) = 1\big)\big)$$

$$I_i\big(\varphi \leq \blacksquare\psi\big) = 1 \textit{ iff } \forall S \in S(i),\ \big(\exists x \in S,\ I_x(\psi) = 1 \supset \exists y \in S,\ I_y(\varphi) = 1\big)$$

φ <■ ψ holds at an index *i* iff there is a sphere *S* in *S*(*i*) such that *S* permits φ but not ψ. φ ≤■ ψ holds at an index *i* iff for all spheres *S* in *S*(*i*), if *S* permits ψ then *S* permits φ. Given my intended interpretation, ≤■ can be read as '__ would be no less normal than ...' and <■ can be read as '__ would be more normal than ...'. Call these *comparative normalcy* operators.

Both the normalcy and comparative normalcy operators can be defined in terms of ceteris paribus conditionals (and possibility) by exploiting the following equivalences:

$$\blacksquare\varphi \equiv \big(T\blacksquare \rightarrow \varphi\big)$$

$$\blacklozenge\varphi \equiv (T\blacklozenge \rightarrow \varphi)$$

$$\big(\varphi \leq \blacksquare\psi\big) \equiv ((\varphi \vee \psi)\blacklozenge \rightarrow \varphi) \vee \sim \Diamond\psi$$

$$\big(\varphi <\blacksquare\psi\big) \equiv \Big((\varphi \vee \psi)\blacksquare \rightarrow \sim \psi\Big) \wedge \Diamond\varphi$$

We can define possibility as follows:

$$\Diamond\varphi \equiv \varphi\blacklozenge \rightarrow \varphi$$

Ceteris paribus conditionals can also be defined in terms of the comparative normalcy operators (and possibility) by exploiting the equivalences:

$$\big(\varphi\blacksquare \rightarrow \psi\big) \equiv \Big((\varphi \wedge \psi) <\blacksquare (\varphi \wedge \sim \psi)\Big) \vee \sim \Diamond\varphi$$

$$\big(\varphi\blacklozenge \rightarrow \psi\big) \equiv \Big((\varphi \wedge \psi) \leq \blacksquare (\varphi \wedge \sim \psi)\Big) \wedge \Diamond\varphi$$

We can also define possibility as follows:

$$\Diamond\varphi \equiv \varphi <\blacksquare\bot$$

The claim that if φ were the case, then ceteris paribus ψ would be the case is equivalent to the claim that it would be more normal for φ and ψ to both be true than for φ to be true and ψ false (or φ is impossible).' The claim that if φ were the case, then ceteris paribus ψ might be the case is equivalent to the claim that it would be no less normal for φ and ψ to be true than for φ to be true and ψ to be false (and φ is possible).'

There are three compulsory conditions that a function $S(x)$ taking indices to sets of sets of indices must meet in order to qualify as a sphere system assignment function. The first condition – *nesting*, requires that $\forall x \in \mathbf{I}, \forall X, Y \in S(x), X \subseteq Y \vee Y \subseteq X$. The second and third are closure under *unions* and (nonempty) *intersections*, respectively: $\forall x \in \mathbf{I}, \forall \mathbf{X} \subseteq S(x), \cup\mathbf{X} \in S(x))$ and $\forall x \in \mathbf{I}, \forall \mathbf{X} \subseteq S(x), (\mathbf{X} \neq \Lambda \supset \cap\mathbf{X} \in S(x))$.

These compulsory conditions furnish us with two axiom schemata:

$$(1)\ ((\varphi \leq\blacksquare\ \psi) \wedge (\psi \leq\blacksquare\ \chi)) \supset (\varphi \leq\blacksquare\ \chi)$$

$$(2)\ (\varphi \leq\blacksquare\ \psi) \vee (\psi \leq\blacksquare\ \varphi)$$

The content of these principles is most perspicuous when they are expressed in terms of comparative normalcy – they jointly express the fact that comparative normalcy is a weak ordering.

The following table lists a series of optional constraints that might be placed upon a sphere assignment function. Each constraint, save for the first, is listed by Lewis (1973b, pp. 121). Constraints are listed along with the formal semantic postulates that capture them and the characteristic axiom schemata validated when those postulates are implemented:

Constraint	**Postulate**	**Axiom Schema**
Triviality	$\forall x \in \mathbf{I},\ S(x) = \Lambda$	$\bot \leq\blacksquare\top$
Significance	$\forall x \in \mathbf{I},\ S(x) \neq \Lambda$	$\top <\blacksquare\bot$
Total reflexivity	$\forall x \in \mathbf{I},\ x \in \cup S(x)$	$\Box\varphi \supset \varphi$
Weak centering	$\forall x \in \mathbf{I},\ x \in \cap(S(x) - \Lambda)$	$\blacksquare\varphi \supset \varphi$
Centering	$\forall x \in \mathbf{I},\ \{x\} \in S(x)$	$\blacklozenge\varphi \supset \varphi$
Local uniformity	$\forall x \in \mathbf{I},\ \forall y \in \cup S(x),\ \cup S(x) = \cup S(y)$ }	{ $\Diamond\varphi \supset \Box\Diamond\varphi$
Uniformity	$\forall x, y \in \mathbf{I},\ \cup S(x) = \cup S(y)$ }	{ $\Box\varphi \supset \Box\Box\varphi$
Local absoluteness	$\forall x \in \mathbf{I},\ \forall y \in \cup S(y),\ S(x) = S(y)$ }	{ $\varphi \leq\blacksquare\psi \supset \Box(\varphi \leq\blacksquare\psi)$
Absoluteness	$\forall x, y \in \mathbf{I},\ S(x) = S(y)$ }	{ $\varphi <\blacksquare\psi \supset \Box(\varphi <\blacksquare\psi)$

Axioms linked by brackets are *both* yielded by *either* of the corresponding postulates. The conditions listed here are not logically independent. Given the compulsory constraints upon $S(x)$, centering implies weak centering, which implies total reflexivity which, in turn, implies significance. Triviality implies absoluteness which implies local absoluteness and uniformity, and uniformity implies local uniformity.

I include the first constraint – triviality – since this is the constraint to which sceptics who regard ceteris paribus conditionals as vacuously true must subscribe – provided, at any rate, that they accept my semantic analysis. Given my intended interpretation, one can easily confirm that, under this condition, any sentence of the form φ ■→ ψ will be true at all possible worlds and any sentence of the form φ ♦→ ψ will be true at none. One need not probe the resultant logic too deeply to unearth principles that are intuitively bizarre. I think it is quite appropriate to use the counterintuitive nature of the principles as part of a case against scepticism about ceteris paribus conditionals. However, as I've stated, I'm not concerned to refute the sceptics here.

On my intended interpretation of **I** and $S(x)$, $\Box$ and $\Diamond$ can be read as metaphysical necessity and possibility, respectively. If we suppose, for ease, that the logic governing possibility and necessity is the modal logic **S5**, it follows immediately that $S(x)$ should be subject to the total reflexivity and local uniformity constraints. The accessibility relation **R** governing the quantificational range of the necessity and possibility operators is defined as follows: $\forall x, y\ x\mathbf{R}y$ iff $\exists S \in S(x), y \in S$. The necessity and possibility operators quantify, at a world w, over those worlds that appear in some sphere of normalcy assigned to w. If **R** is to be reflexive – $\forall x\ x\mathbf{R}x$ – then $S(x)$ must be constrained by total reflexivity. If **R** is to be, in addition, transitive and symmetric – $\forall x, y, z\ ((x\mathbf{R}y \wedge y\mathbf{R}z) \supset x\mathbf{R}z)$ and $\forall x, y\ (x\mathbf{R}y \supset y\mathbf{R}x)$ – then $S(x)$ must also be constrained by local uniformity.

The weak centering constraint upon $S(x)$ corresponds, given my intended interpretation, to the presumption that every world estimates itself to be maximally normal, while the full centering constraint upon $S(x)$ corresponds to the presumption that every world estimates itself to be *uniquely* maximally normal. As I have argued, both of these presumptions are untenable. Thus, $S(x)$ should be free from both centering and weak centering. It is intuitively correct that $\blacksquare\varphi \supset \varphi$ should fail to be semantically valid. The claim that φ would normally be the case does not logically imply that φ is in fact the case.

Given my intended interpretation, the absoluteness constraint upon $S(x)$ corresponds to the presumption that every possible world imposes the very same standards of normalcy, while the local absoluteness constraint corresponds to the presumption that every world that is *possible* from the perspective of world w imposes the very same standards of normalcy as w. As I have argued, these presumptions are also untenable and both constraints should be relaxed in the case of $S(x)$. It is intuitively correct that $(\varphi \leq_{\blacksquare} \psi) \supset \Box(\varphi \leq_{\blacksquare} \psi)$ and $(\varphi <_{\blacksquare} \psi) \supset \Box(\varphi <_{\blacksquare} \psi)$ should not be valid. The claim that φ would be at least as normal as ψ does not logically imply that it is *necessarily* the case that φ would be at least as normal as ψ. The claim that φ would be more normal than ψ does not logically imply that it is necessarily the case that φ would be more normal than ψ.

Lewis's list of semantic restrictions is, of course, far from exhaustive. Consider the following constraint that I shall term *robustness*: $\forall x, y \in \mathbf{I}$, $(\forall S \in S(x), S = \Lambda \vee y \in S) \supset S(y) = S(x)$. This constraint upon $S(x)$, weaker than both local absoluteness and centering, corresponds on my intended interpretation to the presumption that, if a world i is estimated to be maximally normal by a world j – if i is a member of every nonempty sphere of normalcy associated with j – then i and j will *share* the same standards of normalcy. That is, if a possible world perfectly exemplifies any world's ideals of normalcy, then it will share these ideals and, thus, perfectly exemplify its own ideals of normalcy. Robustness does some justice to the intuition, explored briefly in the previous section, that standards of normalcy, while not entirely aloof from the contingent nature of the world, are insensitive with respect to certain contingent differences.

Two principles are jointly characteristic of robustness – namely, $(\varphi \leq\blacksquare \psi) \equiv \blacksquare(\varphi \leq\blacksquare \psi)$ and $(\varphi <\blacksquare \psi) \equiv \blacksquare(\varphi <\blacksquare \psi)$. Robustness will also serve to validate the equivalences $(\varphi \blacksquare\!\!\rightarrow \psi) \equiv \blacksquare(\varphi \blacksquare\!\!\rightarrow \psi)$ and $(\varphi \blacklozenge\!\!\rightarrow \psi) \equiv \blacksquare(\varphi \blacklozenge\!\!\rightarrow \psi)$. These principles seem intuitively plausible. The claim that φ would be at least as normal as ψ does seem logically equivalent to the claim that *normally* φ would be at least as normal as ψ. Similarly, the claim that φ would be more normal than ψ seems logically equivalent to the claim that normally φ would be more normal than ψ. In both of these cases, the additional normalcy clause seems redundant. My suggestion, then, is that the logic governing ceteris paribus conditionals is the logic generated by a system assignment function constrained by total reflexivity, local uniformity and robustness.

Given this logic for ceteris paribus conditionals, we can derive a modal logic for the normalcy operators $(\blacksquare, \blacklozenge)$. The accessibility relation $\mathbf{R}'$ governing the quantificational range of the normalcy operators is defined as follows: $\forall x, y\ x\mathbf{R}'y$ iff $\forall S \in S(x)\ (S = \Lambda \vee y \in S)$. That is, the normalcy operators quantify, at a world w, over the worlds that appear in all nonempty spheres of normalcy assigned to w. By total reflexivity $\mathbf{R}'$ is serial – $\forall w\ \exists x\ w\mathbf{R}'x$ – and by robustness $\mathbf{R}'$ is both transitive – $\forall w, x, y\ (w\mathbf{R}'x \wedge x\mathbf{R}'y) \supset w\mathbf{R}'y$ – and Euclidean – $\forall w, x, y\ (w\mathbf{R}'x \wedge w\mathbf{R}'y) \supset x\mathbf{R}y$. By the failure of weak centering, $\mathbf{R}'$ is irreflexive – $\sim\forall w\ w\mathbf{R}'w$ – and by robustness and the failure of weak centering $\mathbf{R}'$ is asymmetric – $\sim (\forall w, x\ (w\mathbf{R}'x \supset x\mathbf{R}'w))$. The distribution principle $\mathbf{K}$: $\blacksquare(\varphi \supset \psi) \supset (\blacksquare\varphi \supset \blacksquare\psi)$, holds for a completely unconstrained $\mathbf{R}'$. A serial $\mathbf{R}'$ validates **D**: $\blacksquare\varphi \supset \blacklozenge\varphi$, a transitive $\mathbf{R}'$ validates **4**: $\blacksquare\varphi \supset \blacksquare\blacksquare\varphi$ and a Euclidean $\mathbf{R}'$ validates **5**: $\blacklozenge\varphi \supset \blacksquare\blacklozenge\varphi$.

It is worth pointing out that a Euclidean $\mathbf{R}'$ will also be *shift reflexive* – $\forall w, x\ (w\mathbf{R}x \supset x\mathbf{R}x)$. A serial and Euclidean $\mathbf{R}'$ will also be *dense* – $\forall w, x\ (w\mathbf{R}'x \supset \exists y\ (w\mathbf{R}'y \wedge y\mathbf{R}'x))$. Shift reflexivity validates the principle $\blacksquare$**T**: $\blacksquare(\blacksquare\varphi \supset \varphi)$, and density validates the principle **C4**: $\blacksquare\blacksquare\varphi \supset \blacksquare\varphi$.

The modal logic governing the normalcy operators will have as axioms all of the truth functional tautologies, all instances of the definition $\blacklozenge\varphi \equiv \sim\blacksquare\sim\varphi$ and all instances of the following schemata:

$$(K)\ \blacksquare(\varphi \supset \psi) \supset \left(\blacksquare\varphi \supset \blacksquare\psi\right)$$
$$(D) \blacksquare\varphi \supset \blacklozenge\varphi$$
$$(4) \blacksquare\varphi \supset \blacksquare\blacksquare\varphi$$
$$(5) \blacklozenge\varphi \supset \blacksquare\blacklozenge\varphi$$

It will have two inference rules:

(1) $\Rightarrow\varphi, \Rightarrow\varphi \supset \psi$
$\Rightarrow \psi$
Modus Ponens

(2) $\Rightarrow\varphi$
$\Rightarrow \blacksquare\varphi$
Normic Necessitation

This is the modal logic **KD45** – a logic that has been investigated as a possible deontic and doxastic logic.

I conclude by offering an axiomatisation for the full logic of ceteris paribus conditionals and all related operators. The logic will include, as axioms, all of the truth functional tautologies, schemata defining the operators $\blacksquare$, $\blacklozenge$, $\Box$, $\Diamond$, $<\blacksquare$, $\blacksquare\!\rightarrow$, $\blacklozenge\!\rightarrow$ and $\rightarrow$ in terms of the operator $\leq\blacksquare$ (along with the truth functional operators), and all instances of the following schemata:

(1) $\left((\varphi \leq \blacksquare\psi) \wedge (\psi \leq \blacksquare\chi)\right) \supset (\varphi \leq \blacksquare\chi)$
(2) $(\varphi \leq \blacksquare\psi) \vee (\psi \leq \blacksquare\varphi)$
(3) $\Box\varphi \supset \varphi$
(4) $\Diamond\varphi \supset \Box\Diamond\varphi$
(5) $\Box\varphi \supset \Box\Box\varphi$
(6) $(\varphi \leq \blacksquare\psi) \equiv \blacksquare(\varphi \leq \blacksquare\psi)$
(7) $(\varphi < \blacksquare\psi) \equiv \blacksquare(\varphi < \blacksquare\psi)$

There will be two rules of inference:

(1) $\Rightarrow\varphi$, $\Rightarrow \varphi \supset \psi$
$\Rightarrow \psi$
Modus Ponens

(2) for any $n \geq 1$
$\Rightarrow \varphi \supset (\psi_1 \vee \ldots \vee \psi_n)$
$\Rightarrow (\psi_1 \leq\blacksquare\ \varphi) \vee \ldots \vee (\psi_n \leq\blacksquare\ \varphi)$
Rule for Comparative Normalcy

Call this logic **VTRU**. Following Lewis, **V** indicates a variably strict conditional logic and **T**, **R** and **U** represent total reflexivity, robustness and local uniformity, respectively.

Demonstrating that **VTRU** is sound with respect to the proposed semantics is relatively straightforward. As can be easily verified, all instances of schemata (1)–(7) are valid under any interpretation with the total reflexivity, local uniformity and robustness constraints (along with nesting and closure under unions and intersections). Further, modus ponens and the rule for comparative normalcy clearly preserve validity – they will never take us from valid to non-valid sentences. Therefore, there are no theorems provable in **VTRU** that are not validated by the proposed semantics. **VTRU** is consistent – it does not prove $\perp$ as a theorem. Completeness is proved in the Appendix.

I have suggested that ceteris paribus conditionals should be understood as a distinct species of variably strict conditional alongside subjunctive conditionals, imperative conditionals and others. The content of a ceteris paribus conditional is essentially this: The most normal worlds in which antecedent and consequent jointly hold are more normal than the most normal worlds in which the antecedent holds and the consequent fails. I have suggested further that **VTRU** is the logical system that emerges from this conception of the content of ceteris paribus conditionals, given a little careful reflection upon the relation of comparative normalcy.

ACKNOWLEDGMENTS

Thanks go to Martin Davies, Andy Egan, Peter Roeper and an anonymous referee for this journal, all of whom provided valuable comments and criticism. Thanks also to audiences at the Australian National University and the 2005 Australasian Association of Logic conference.

APPENDIX

*Appendix: Completeness of **VTRU***

The following technique for proving completeness is a variant of the technique devised by Lewis (1973b, pp. 124–130) which draws, in turn, upon the work of Lemmon and Scott (see Lemmon, 1977, Sections 2 and 3) and Makinson (1966). We begin by constructing a canonical interpretation of language **L**. Call a set of sentences Σ of **L**, *consistent* just in case it does not allow us to prove $\perp$ in the logic **VTRU**. Say that a set of sentences Σ is *consistent with* a sentence φ just in case $\Sigma \cup \{\varphi\}$ is consistent. Say that a set of sentences is *maximally consistent* just in case it is consistent but not consistent with any sentence that is not already contained in it.

VTRU satisfies Lindenbaum's Lemma. That is, any consistent set of sentences can be extended into a maximally consistent set.

Proof. The countably many sentences of **L** can be numbered and ordered. Call this sequence φ_1, φ_2, φ_3 .... For any consistent set of sentences Σ_0, let $\Sigma_{n+1} = \Sigma_n \cup \{\varphi_n\}$ if Σ_n and φ_n are consistent and let $\Sigma_{n+1} = \Sigma_n$ otherwise. Every set in this sequence is consistent. Let Σ_∞ be the union of all sets in this sequence. Σ_∞ includes Σ_0. Σ_∞ is consistent. If not, some finite subset of Σ_∞ must be inconsistent. But every finite subset of Σ_∞ is included in some Σ_n contradicting the consistency of each Σ_n. Σ_∞ is maximally consistent. If not, then it must be consistent with some sentence φ_n that is not included in it. If Σ_∞ is consistent with φ_n, then Σ_n must be consistent with φ_n, in which case Σ_∞ will include φ_n. □

We construct the canonical interpretation of **L** as follows: **I** is the set of all maximally consistent sets of sentences of **L** and $I_i(\varphi) = 1$ iff $\varphi \in i$. That is, a sentence φ holds at an index i, under the canonical interpretation just in case φ is a member of i. All and only the theorems of **VTRU** will be valid under the canonical interpretation of **L**.

Proof. An index that did not contain a theorem of **VTRU** would be inconsistent with that theorem and, hence, inconsistent simpliciter, contrary to stipulation. Thus, all theorems of **VTRU** are valid under the canonical interpretation of **L**. If φ is not a theorem of **VTRU** then $\sim\varphi$ is consistent, in which case $\{\sim\varphi\}$ can be extended to a maximally consistent set – call it i. Index i does not contain sentence φ. Thus, no non-theorems of **VTRU** are valid under the canonical interpretation of **L**. □

Provided we can show that this canonical interpretation of **L** *is* a genuine interpretation meeting all requisite conditions, completeness will follow at once. First, on the canonical interpretation, **I** is nonempty.

Proof. The set of theorems of **VTRU** is a consistent set of sentences and can be expanded into a maximally consistent set which will belong to **I**. □

Second, the canonical interpretation provides the correct truth conditions for the sentential constants and truth functional operators.

Proof. $\top$ is a theorem of **VTRU**. Hence, $\top$ is a member of every maximally consistent set of sentences and, thus, true at every index under the canonical interpretation. $\perp$ is not a member of any maximally consistent sets of sentences. Thus, it is not true at any index under the canonical interpretation. $I_i(\sim\varphi) = 1$ iff $\sim(I_i(\varphi) = 1)$ holds, since a maximal consistent set must, for any φ, contain either but not both of φ and

$\sim\varphi$. If a maximally consistent set contained neither φ nor $\sim\varphi$ it would be inconsistent with both and thus inconsistent with $(\varphi \vee \sim\varphi)$ – a theorem of **VTRU** – in which case it would be inconsistent simpliciter. $I_i(\varphi \wedge \psi) = 1$ iff $(I_i(\varphi) = 1) \wedge (I_i(\psi) = 1)$ is true, since a maximally consistent set of sentences must contain every sentence that it implies in **VTRU**. If it did not contain a sentence it implied, it would be inconsistent with a sentence it implied and thus inconsistent simpliciter. These results can easily be extended to the other truth functional operators. □

The canonical interpretation of the function $S(x)$ is constructed as follows: Call a set of sentences Σ *characteristic* of an index i just in case, (1) if a sentence $\sim\Diamond\varphi$ is a member of i then φ is not a member of Σ and (2) if a sentence φ is a member of Σ and $\psi \leq\blacksquare\, \varphi$ is a member of i then ψ is a member Σ. These two conditions will not clash – that is, a sentence will never be included in Σ in accordance with (2) and excluded from Σ in accordance with (1). This is ensured by the fact that $(\Diamond\varphi \wedge (\psi \leq\blacksquare\, \varphi)) \supset \Diamond\psi$ is a theorem of **VTRU** and, thus, a member of every index i.

Call a set of sentences *saturated* just in case, for every sentence φ of **L**, it contains either φ or $\sim\varphi$ or both. If a set of sentences Σ is characteristic for an index i then, provided it is nonempty, it will be saturated.

Proof. Assume a set of sentences Σ characteristic for index i is nonempty and not saturated. Since Σ is nonempty it must contain a sentence – call it ψ. If neither φ nor $\sim\varphi$ are members of Σ, then neither $(\varphi \leq\blacksquare\, \psi)$ nor $(\sim\varphi \leq\blacksquare\, \psi))$ can be members of i. In this case, i must be inconsistent with a theorem of **VTRU**. $\psi \supset (\varphi \vee \sim\varphi)$ is a truth functional tautology which implies, by the rule for comparative normalcy, $(\varphi \leq\blacksquare\, \psi) \vee (\sim\varphi \leq\blacksquare\, \psi)$. As a result, characteristic sets of sentences, provided they are nonempty, must be saturated. □

Characteristic sets of sentences are closed under single premise consequence in **VTRU**. That is, if a set of sentences Σ, characteristic for an index i, contains a sentence φ and $\varphi \supset \psi$ is a theorem of **VTRU** then Σ will also contain sentence ψ.

Proof. If $\varphi \supset \psi$ is a theorem of **VTRU** then, by the rule for comparative normalcy, $\psi \leq\blacksquare\, \varphi$ is a theorem of **VTRU** and, thus, a member of i. If φ is a member of Σ and $\psi \leq\blacksquare\, \varphi$ is a member of i, then ψ is a member of Σ. □

Under the canonical interpretation, a *sphere* associated with an index i is a set of maximally consistent subsets of a characteristic set of i. $S(x)$ assigns to an index x the set of spheres, so defined, that are associated with it.

Under the canonical interpretation, the set of spheres assigned to an index by the function $S(x)$ is nested.

Proof. The sets that are characteristic for an index i are nested. If not, then there are two sets Σ and Π characteristic of an index i, such that for two sentences φ and ψ, $\varphi \in \Sigma$, $\varphi \notin \Pi$, $\psi \in \Pi$ and $\psi \notin \Sigma$. It follows from this that neither $\varphi \leq\blacksquare\, \psi$ nor $\psi \leq\blacksquare\, \varphi$ will be members of i. In this case i must be inconsistent with the **VTRU** theorem $(\varphi \leq\blacksquare\, \psi) \vee (\psi \leq\blacksquare\, \varphi)$ and hence inconsistent simpliciter. If a set of sentences Σ is a subset of a set of sentences Π, then the set of maximally consistent subsets of Σ will be a subset of the set of maximally consistent subsets of Π. If the sets characteristic of an index i are nested, then so too are the spheres associated with i. Therefore the set of spheres assigned to an index i is nested. □

Under the canonical interpretation, the set of spheres assigned to an index i by the function $S(x)$ is closed under unions and nonempty intersections.

Proof. The union Σ of any set of sets characteristic of an index i must itself be characteristic of i. If not, then either (a) Σ must contain a sentence φ such that ∼◇φ is a member of i or (b) Σ must contain a sentence ψ but not contain a sentence φ even though φ ≤■ ψ is a member of i. If (a) then a characteristic set of i must contain a sentence φ such that ∼◇φ is a member of i. If (b) then a characteristic set of i must contain a sentence ψ but not contain a sentence φ even though φ ≤■ ψ is a member of i. Both are impossible. If a set Σ characteristic for i is equal to the union of a set ξ of sets characteristic for i, then the set of maximally consistent subsets of Σ is equal to the union of the set of sets of maximally consistent subsets of members of ξ. The intersection Σ of any nonempty set of sets characteristic of an index i is itself characteristic of i. If not, then either (a) Σ must contain a sentence φ such that ∼◇φ is a member of i or (b) Σ must contain a sentence ψ but not a sentence φ even though φ ≤■ ψ is a member of i. If (a), then a characteristic set of i must contain a sentence φ such that ∼◇φ is a member of i. If (b), then a characteristic set of i must contain a sentence ψ but not contain a sentence φ even though φ ≤■ ψ is a member of i. Both are impossible. If a set Σ characteristic for i is equal to the intersection of a set ξ of sets characteristic for i then the set of maximally consistent subsets of Σ is equal to the intersection of the set of sets of maximally consistent subsets of members of ξ.□

The canonical interpretation gives the correct truth conditions for the operator ≤■.

Proof. Every sphere in $S(i)$ that contains a ψ-index contains a φ-index iff for every set of sentences Σ characteristic for i, Σ contains ψ only if Σ contains φ. Suppose φ ≤■ ψ is a member of i. It follows immediately that any set that is characteristic for i will contain φ if it contains ψ and any sphere will contain a φ-index if it contains a ψ-index. Suppose φ ≤■ ψ is not a member of i. Since ∼◇ψ ⊃ (φ ≤■ ψ) is a theorem of **VTRU**, ◇ψ must be a member of i. Consider the set Σ that contains all and only those sentences χ such that χ ≤■ ψ is a member of i. Σ will not contain any sentence χ such that ∼◇χ is a member of i, since ◇ψ is a member of i and (◇ψ ∧ (χ ≤■ ψ)) ⊃ ◇χ is a theorem of **VTRU**. Since all instances of ((λ ≤■ χ) ∧ (χ ≤■ ψ)) ⊃ (λ ≤■ ψ) are theorems of **VTRU** and, thus, members of Σ, Σ will, then, be characteristic for i. Σ does not contain sentence φ. Therefore, there is a set of sentences Σ characteristic for i, such that Σ contains ψ but not φ and thus a sphere that contains a ψ-index but not a φ-index. □

Under the canonical interpretation, the function $S(x)$ is subject to total reflexivity.

Proof. The set of all sentences φ such that ◇φ is a member of i is characteristic of i. Call this set Σ. Given that □φ ⊃ φ and the contraposed principle φ ⊃ ◇φ are theorems of **VTRU**, every sentence that is a member of i must be a member of Σ. It follows that i itself is one of the maximally consistent subsets of Σ. Therefore, i is a member of one of the members of $S(i)$.□

Under the canonical interpretation, the function $S(x)$ is subject to local uniformity.

Proof. Consider two indices i and j such that $j \in \cup S(i)$. Suppose that there is a sentence φ such that ◇φ is a member of i and ∼◇φ is a member of j. Given that ◇φ ⊃ □◇φ and the consequence ◇φ ⊃ ∼◇∼◇φ are theorems of **VTRU**, it follows that ∼◇φ is not a member of any set characteristic for i. It follows from this that j is not a member of $\cup S(i)$, contrary to stipulation. In this case, if ◇φ is a member of i then ◇φ is a member of j. Suppose that there is a sentence φ such that ◇φ is a member of j and ∼◇φ is a member of i. Given that □φ ⊃ □□φ and the consequence ∼◇φ ⊃ ∼◇◇φ are theorems of **VTRU** it follows that ◇φ is not a member of any set characteristic for i. It follows from this that

j is not a member of $\cup S(i)$, contrary to stipulation. In this case, if $\Diamond\varphi$ is a member of j then $\Diamond\varphi$ is a member of i. The set of all sentences φ such that $\Diamond\varphi$ is a member of i is the largest set that is characteristic of i, and the set of all sentences φ such that $\Diamond\varphi$ is a member of j is the largest set that is characteristic of j. Since these sets are equal and the members of $S(i)$ and $S(j)$ are nested, it follows that $\cup S(i) = \cup S(j)$. □

Under the canonical interpretation, the function $S(x)$ is subject to robustness.

Proof. Consider two indices i and j such that $\forall S \in S(i), (S = \Lambda \vee j \in S)$. If j is a member of every nonempty member of $S(i)$ then j must be a maximally consistent subset of every nonempty set characteristic for i. Consider a sentence φ such that $\varphi \leq\blacksquare\ \psi$ is a member of i for any sentence ψ. Sentence φ must be a member of every nonempty set that is characteristic for i. Given that $(\varphi \leq\blacksquare\ \psi) \equiv \blacksquare(\varphi \leq\blacksquare\ \psi)$ is a theorem of **VTRU**, if a sentence $\varphi \leq\blacksquare\ \psi$ is a member of i then $(\varphi \leq\blacksquare\ \psi) \leq\blacksquare\ \chi$ will be a member of i for any χ. Therefore, if a sentence $\varphi \leq\blacksquare\ \psi$ is a member of i, $\varphi \leq\blacksquare\ \psi$ must be a member of j. It follows that a set that is characteristic for j will be characteristic for i. Given that $(\varphi <\blacksquare\ \psi) \equiv \blacksquare(\varphi <\blacksquare\ \psi)$ is a theorem of **VTRU**, if a sentence $\sim(\varphi \leq\blacksquare\ \psi)$ is a member of i, $(\sim(\varphi \leq\blacksquare\ \psi)) \leq\blacksquare\ \chi$ will be a member of i for any χ. Therefore, if a sentence $\sim(\varphi \leq\blacksquare\ \psi)$ is a member of i, $\sim(\varphi \leq\blacksquare\ \psi)$ must be a member of j. It follows that a set that is characteristic for i will be characteristic for j. In this case we have $S(i) = S(j)$ as required. □

NOTES

[1] The widespread use of idealisations in scientific explanation and prediction has been emphasised by Cartwright (1983, 1999) and Laymon (1985, 1989) amongst others. I would suggest that the use of idealisations is just as widespread throughout ‘folk’ explanation and prediction.

[2] Boutilier, in a sense, limits himself to the resources of standard possible world semantics. He introduces a reflexive, transitive accessibility relation upon the class of possible worlds, to be understood as a comparative normalcy ordering. That is, aRb means that b is at least as normal a world as a (Boutilier, 1995, pp. 96–97). He then introduces two modal operators – one that quantifies, at a given world, over all equally or more normal worlds (accessible worlds) and one that quantifies over all less normal worlds (inaccessible worlds) giving us, in effect, a bimodal logic in which both modal operators are governed by the same accessibility relation. Boutilier proceeds to define his normative conditional operator in terms of these two unary operators.

[3] A few more suggestive remarks: When it comes to normalcy, I am inclined to think that both frequentist accounts (Gundersen, 2004) and teleological accounts (Millikan, 1984, pp. 5, 33–34) are on the wrong track. I suggest that one fruitful, if somewhat elliptical, way to shed light upon the nature of normalcy is by investigating the utility of idealised models in prediction and explanation and, in particular, the conditions under which explanation and prediction can successfully proceed in the absence of complete theories.

REFERENCES

Arthur, W. and Fenster, S.: 1969, *Mechanics*, Holt, Rinehart and Winston, New York.
Boutilier, C.: 1994, Conditional logics of normality: A modal approach, *Artificial Intelligence* **68**, 87–154.

Boutilier, C. and Becher, V.: 1995, Abduction as belief revision, *Artificial Intelligence* **77**, 43–94.

Bowie, G.: 1979, The similarity approach to counterfactuals: Some problems, *Noûs* **13**, 477–498.

Cartwright, N.: 1983, *How the Laws of Physics Lie*, Clarendon, Oxford.

Cartwright, N.: 1999, *The Dappled World*, Cambridge University Press, Cambridge.

Delgrande, J.: 1987, A first order logic for prototypical properties, *Artificial Intelligence* **33**, 105–130.

Earman, J. and Roberts, J.: 1999, "Ceteris paribus" there is no problem of provisos, *Synthese* **118**, 439–478.

Earman, J., Roberts, J. and Smith, S.: 2002, Ceteris paribus lost, *Erkenntnis* **57**, 281–301.

Gundersen, L.: 2004, Outline of a new semantics for counterfactuals, *Pacific Philosophical Quarterly* **85**, 1–20.

Halpern, J. and Rabin, M.: 1987, A logic to reason about likelihood, *Artificial Intelligence* **32**, 379–405.

Hansson, B.: 1969, An analysis of some deontic logics, *Noûs* **3**, 373–398.

Hempel, C.: 1988, Provisos: A problem concerning the inferential function of scientific theories, *Erkenntnis* **28**, 147–164.

Krabbe, E.: 1978, Note on a completeness theorem in the theory of counterfactuals, *Journal of Philosophical Logic* **7**, 91–93.

Kripke, S.: 1980, *Naming and Necessity*, Basil Blackwell, Oxford.

Laymon, R.: 1985, Idealization and the testing of theories by experimentation, in P. Achinstein and O. Hannaway (eds.), *Observation, Experiment and Hypothesis in Modern Physical Science*, MIT, Cambridge, Massachusetts.

Laymon, R.: 1989, Cartwright and the lying laws of physics, *Journal of Philosophy* **86**, 353–372.

Lemmon, E.: 1977, *An Introduction to Modal Logic*, Blackwell, Oxford.

Lewis, D.: 1971, Completeness and decidability of three logics of counterfactual conditionals, *Theoria* **37**, 74–85.

Lewis, D.: 1973a, Counterfactuals and comparative possibility, *Journal of Philosophical Logic* **2**, 418–444.

Lewis, D.: 1973b, *Counterfactuals*, Blackwell, Oxford.

Lewis, D.: 1974, Semantic analyses for dyadic deontic logic, in S. Stenlund (ed.), *Logical Theory and Semantical Analysis*, Reidel, Dordrecht.

Lewis, D.: 1979, Counterfactual dependence and time's arrow, *Noûs* **13**, 455–476.

Makinson, D.: 1966, On some completeness theorems in modal logic, *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* **12**, 379–384.

McCarthy, J.: 1986, Application of circumscription to formalizing common sense knowledge, *Artificial Intelligence* **28**, 89–116.

Menzies, P.: 2004, Difference making in context, in J. Collins (ed.), *Causation and Counterfactuals*, MIT, Cambridge, Massachusetts.

Millikan, R.: 1984, *Language, Thought and Other Biological Categories*, MIT, Cambridge, Massachusetts.

Nozick, R.: 1981, *Philosophical Explanations*, Harvard University Press, Cambridge, Massachusetts.

Pietroski, P. and Rey, G.: 1995, When other things aren't equal: Saving "ceteris paribus" laws from vacuity, *British Journal for the Philosophy of Science* **46**, 81–110.

Shaffer, M.: 2001, Bayesian confirmation theories that incorporate idealizations, *Philosophy of Science* **68**, 36–52.

Shiffer, S.: 1987, *Remnants of Meaning*, MIT, Cambridge, Massachusetts.
Shiffer, S.: 1991, Ceteris paribus laws, *Mind* **100**, 1–18.
Sober, E.: 1984, *The Nature of Selection*, MIT, Cambridge, Massachusetts.
Van Fraassen, B.: 1972, The logic of conditional obligation, *Journal of Philosophical Logic* **1**, 417–438.

Department of Philosophy,
University of York,
Heslington, York, YO10 5DD,
UK
E-mail: ms567@york.ac.uk