Marco Hausmann
Redemptoris Mater Berlin
marco.hausmann1@gmx.de

Original Scientific Paper
UDC 123.1
123.2
Received: May 07, 2024
Revised: October 03, 2024
Accepted: October 11, 2024

HOW TO MAKE BETA BETTER

Abstract
In this paper, I appeal to a quantified counterfactual logic in order to argue that van Inwagen's new version of the Consequence Argument, unlike the original version of his argument, is logically valid. To this end, I build on an argument that has been developed by Alexander Pruss. Pruss has argued, on the basis of Lewis's semantics for counterfactual conditionals, that, given a suitable interpretation, David Widerker's version of the Consequence Argument is logically valid. In my paper, I extend Pruss' proposal and I argue, on the basis of a quantified counterfactual logic, that, given a suitable interpretation, not only Widerker's but also van Inwagen's version of the Consequence Argument is logically valid.

Keywords Free Will · Determinism · Incompatibilism · Consequence Argument · Rule Beta · Counterfactuals

1. Introduction

Arguments for incompatibilism are still viewed with an air of suspicion. As is well known, doubts about the validity of those arguments have been pervasive throughout the history of the debate. And, as recently turned out, not without reason: Thomas McKay and David Johnson (1996) have been able to show that one of the most promising arguments for incompatibilism, Peter van Inwagen's Consequence Argument, is indeed invalid – it rests on a fallacious rule of inference. In response, van Inwagen (2008, 2013, 2015, 2017a, 2017b, 2017c) suggested several ways to reinterpret and modify his original rule of inference. It is unclear, however, whether his suggestions lead to significant improvements. For, as in the case of his original version, van Inwagen had to appeal to intuitions in order to justify his revised rule of inference. And there appears to be little reason to trust his intuitions given that in the case of his original rule his intuitions proved to be off the mark.

The aim of my paper is to show that doubts about the validity of van Inwagen's new argument are ill-founded. I show that, once compatibilists and incompatibilists agree upon a few basic principles of quantified counterfactual logic, they may safely ignore questions about the validity of van Inwagen's argument, and they may safely focus on questions about the truth of van Inwagen's premises.

My paper proceeds as follows: In the first part of my paper, I recapitulate van Inwagen's original version of the argument, I present McKay and Johnson's case against van Inwagen's original version and I show why McKay and Johnson's case becomes harmless once we incorporate one of van Inwagen's several suggestions for a new version of the argument. In the second part of my paper, I discuss van Inwagen's new version of the argument. To this end, I build on an argument that has been developed by Alexander Pruss (2013) in support of David Widerker's version of the Consequence Argument. In his argument, in order to establish the validity of Widerker's version of the argument, Pruss appeals to David Lewis's semantics for counterfactual conditionals. Unlike Pruss, however, I appeal to principles of counterfactual *logic* (and not to principles of counterfactual *semantics*). This will pave the way for an argument in favor of the validity of van Inwagen's – and not only in favor of the validity of Widerker's – argument.

2. Peter van Inwagen's original version of the argument

To start with, it might be helpful to recapitulate the original version of the Consequence Argument. In order to argue for the incompatibility of free will and determinism, van Inwagen (1983, 93) introduced a new modal operator 'N' that is supposed to obey the following two inference rules ('N p' stands for 'p and no one has, or ever had, any choice about whether p'):

Rule α: $\Box\, p \vdash N\, p$
Rule β: $N\, p, N\, (p \rightarrow q) \vdash N\, q$

He then defined determinism as the conjunction of the following two theses:

> "For every instant of time, there is a proposition that expresses the state of the world at that instant;
> If p and q are any propositions that express the state of the world at some instants, then the conjunction of p with the laws of nature entails q" (1983, 65).

Now let 'p_0' be a proposition that expresses the state of the world in the remote past (at a time long before the existence of the first human being), let 'p' be an arbitrary truth, and let 'l' be the conjunction of all laws of nature. The argument for the incompatibility of free will and determinism then runs as follows:

(1) $\Box\, ((p_0\ \&\ l) \rightarrow p)$	follows from determinism
(2) $\Box\, ((p_0 \rightarrow (l \rightarrow p))$	follows from (1)
(3) $N\, ((p_0 \rightarrow (l \rightarrow p))$	follows from (2) and rule α

(4) N p_0	remise
(5) N ($l \rightarrow p$)	follows from (3), (4) and rule β
(6) N l	premise
(7) N p	follows from (5), (6) and rule β

The main idea of the argument seems to be the following: if determinism is true, any truth follows from the conjunction of ‘p_0’ and ‘l’. But nobody has a choice about whether ‘p_0’ is true, for nobody has a choice about whether a true proposition about the past is true. And nobody has a choice about whether ‘l’ is true, for nobody has a choice about whether a true proposition about the laws of nature is true. It follows that, if determinism is true, nobody has a choice about anything.

3. A counterexample and van Inwagen’s new version of the argument

Much of the discussion of the Consequence Argument has focused on the validity of beta.[1] In fact, van Inwagen had to revise his interpretation of ‘N’, given that McKay and Johnson (1996) were able to argue against beta by showing that alpha and beta together entail a principle of agglomeration, and by providing a counterexample against that principle. Roughly speaking, according to agglomeration the conjunction of two unavoidable propositions is itself unavoidable:

Principle of Agglomeration: $N\,p, N\,q \vdash N\,(p \,\&\, q)$

It is not difficult to see that agglomeration is derivable from alpha and beta:

(1) $\Box\,(p \rightarrow (q \rightarrow (p \,\&\, q)))$	logical truth
(2) N p	assumption
(3) N q	assumption
(4) N $(p \rightarrow (q \rightarrow (p \,\&\, q)))$	follows from (1) and rule α
(5) N $(q \rightarrow (p \,\&\, q))$	follows from (2), (4) and rule β
(6) N $(p \,\&\, q)$	follows from (3), (5) and rule β

Although agglomeration might seem plausible at first glance, McKay and Johnson were able to present the following counterexample to agglomeration:

> “Suppose that I do not toss a coin, but could have.
>
> p = the coin does not fall heads.
>
> q = the coin does not fall tails.

1 See, for example, Widerker (1987), O’Connor (1993), Fischer (1994), Kapitan (1996, 2002), McKay and Johnson (1996), Finch and Warfield (1998), Carlson (2000), Crisp and Warfield (2000), Huemer (2000), Blum (2000, 2003), Schnieder (2008), van Inwagen (2008, 2013, 2015, 2017a, 2017b, 2017c), Pruss (2013), Vihvelin (2013), Gustafsson (2017) and Hausmann (2018).

> Both premises of agglomeration are true, 'Np' and 'Nq'; no one can choose to falsify p (no one can choose to make the coin fall heads) and no one can choose to falsify q (no one can choose to make the coin fall tails). The conclusion, 'N (p & q)', is false, however. I could have chosen to make '(p & q)' false by choosing to toss the coin, so I had a choice about whether 'N (p & q)' is true" (1996, 115).

Hence, agglomeration is not valid. And given that agglomeration follows from alpha and beta and given further that alpha seems undeniably valid, it follows that beta is not valid. That is to say, in order to preserve the argument's validity, van Inwagen had to revise his definition of 'N'. Recently, he therefore came to suggest the following modified interpretation of his operator:

> "Say that a proposition *p* is untouchable just in the case that: *p* is true and no human being is or ever has been able to act in such a way that, if he or she did act that way, *p* might be or might have been false" (2013, 214; see also 2017c, 228, fn. 34).

Following a suggestion of Pruss, we might formally represent his proposal as follows (if we let 'Up' stand for 'it is an untouchable truth that p'):[2]

> (*Van Inwagen*) Up ↔ (p & ~ ∃*x* ∃*y* (x is able to do y & (x does y ◊→ ~ p)))[3]

Accordingly, what the argument now purports to show is that any truth is untouchable, if determinism is true.

As expected, given this revised interpretation of the operator, our counterexample against agglomeration fails. For our counterexample tries to show that (this is a translation of what it tries to show into van Inwagen's terms), although it is an untouchable truth that the coin does not fall heads and it is an untouchable truth that the coin does not fall tails, it is not an untouchable truth that the coin neither falls heads nor tails. Applying the revised interpretation of the operator, however, yields the following result: of course there is someone who is able to do something such that, if he did it, the coin *might* fall heads and, of course, there is someone who is able to do something such that, if he did it, the coin *might* fall tails (there is someone who is able to toss the coin). Hence, it is neither an untouchable truth that the coin does not fall heads, nor is it an untouchable truth that the coin does not fall tails. It follows that – given this revised interpretation of the operator – our alleged counterexample against agglomeration fails.

2 I follow David Lewis (1973, 2) in using '◊→' for the 'might' counterfactual, in using '□→' for the 'would' counterfactual, and in assuming that the following interdefinability holds: '(p ◊→ q) ↔ ~ (p □→ ~ q)'.

3 In what follows, 'x is able to do y' stands for 'x is or ever has been able to do y'.

4. The validity of the new version of the argument

In what follows I am going to argue that van Inwagen's revised interpretation of the operator leads to a promising version of the Consequence Argument. The main idea of van Inwagen's modified version of the argument (the version that results from presupposing his revised interpretation of the operator) seems essentially to be this: say that someone is able to falsify a proposition just in case that he is able to do something such that, if he did it, that proposition might be false. If determinism is true, any truth follows from the conjunction of 'p_0' and 'l'. But nobody is able to falsify 'p_0', for nobody is able to falsify a true proposition about the past. And nobody is able to falsify 'l', for nobody is able to falsify a true proposition about the laws of nature. It follows that, if determinism is true, nobody is able to falsify a truth.

This version of the Consequence Argument, as I shall show in a moment, has a desirable feature (a feature that other versions of the Consequence Argument unfortunately lack): it is logically valid. Worries might arise, though, given that the argument rests on a somewhat artificial or technical explication of being able to falsify a proposition. For it cannot be denied that this explication has some rather weird implications (or, at any rate, implications one would not expect given a natural reading of 'being able to falsify a proposition'). Just to mention one example: given this explication, anyone who does what he is able to do falsifies any falsehood. Take, for instance, Bart, the barber, who is not only able to shave himself, but who is in fact shaving himself. If it is true that Bart is shaving himself then it is certainly true with respect to any falsehood 'p' that:

(1) Bart shaves himself & ~ p

It follows, that it *might not* be the case that p, if Bart shaved himself:[4]

(2) Bart shaves himself ◊→ ~ p

It has been supposed, however, that Bart is able to shave himself:

(3) Bart is able to shave himself & (Bart shaves himself ◊→ ~ p)

Hence, given the explication of being able to falsify a proposition that has been supposed, Bart is able to falsify any falsehood. It follows that Bart is able to falsify the proposition that Caesar did not cross the Rubicon, the proposition that Columbus did not travel to America, or the proposition that 2+2=5. And this, of course, is not what one would expect given a natural reading of being able to falsify a proposition. For, intuitively, Bart has nothing to do with the falsity of either of these propositions.

4 It is easy to see that the involved inference rule is valid. For suppose 'p & q' is true. If 'p ◊→ q' would be false, 'p □→ ~ q' would be true. But then 'q & ~ q' would be true (because, as has been supposed, 'p' is true). Thus, the involved inference rule is valid.

Of course, from the fact that Bart is able to falsify any falsehood, it doesn't follow that Bart is able to falsify 'p_0' or 'l' (for 'p_0' and 'l' are by definition truths, not falsehoods). It is, nonetheless, a matter of controversy whether the introduction of technical notions (such as the notion of being able to falsify a proposition) threatens the purposes of van Inwagen's argument. On the one hand, David Lewis (without doubt one of the most important critics of van Inwagen's argument) notes that it "does not matter what 'could have rendered false' [= was able to falsify] means in ordinary language." For, as he goes on to explain, "van Inwagen introduced the phrase as a term of art. It does not even matter what meaning van Inwagen gave it. What matters is whether we can give it any meaning that would meet his needs" (1981, 119). On the other hand, Benjamin Schnieder points out that "the Consequence Argument owes much of its strikingness to the intuitive plausibility of its premises. An evaluation of the argument, not based on some technical notion of rendering false [= falsifying] which departs from our intuitive understanding is therefore desirable" (2004, 417).

In my view, it should be uncontroversial that – everything else being equal – a version of the argument *without* technical notions would be clearly preferable. The intuitive force of van Inwagen's argument undeniably stems from the intuitive plausibility of its premises. It is far from clear, however, whether any version of the argument that has yet been proposed offers a natural explication of its key notions without lacking some other desirable feature (e. g. logical validity, or intuitive plausibility). In fact, I'm inclined to think that no such version has ever been developed. Schnieder (2004, 418–421, 2008, 105–107), for instance, suggests to interpret van Inwagen's notion of being able to falsify a proposition in terms of explanatory notions.[5] And I concede that his suggestion leads to a rather natural and plausible interpretation of the notion. But I have argued elsewhere that, given this (and similar) explications of the notion, rule beta turns out to be invalid and the argument fails (see Hausmann 2018).

As I see it, we should, as Lewis suggests, simply try to figure out whether our explication of being able to falsify a proposition – technical, or not –, is suitable for the purposes of van Inwagen's argument. We should ask, therefore, whether our explication is such that:[6]

(i) rule beta is valid
(ii) nobody is able to falsify 'p_0'
(iii) nobody is able to falsify 'l'

5 In other papers, van Inwagen's choice of words suggests an interpretation of his operators in terms of explanatory notions (2008, 452, 2015, 19). See Hausmann (2018) for critical discussion. In this paper, I focus on van Inwagen's papers in which his wording does not suggest an interpretation in terms of explanatory notions.

6 For a similar approach see Huemer (2000, pp. 529–530) and Pruss (2013, p. 432).

In what follows, I am going to argue that van Inwagen's revised interpretation of the operator (and so van Inwagen's technical explication of being able to falsify a proposition) satisfies the first condition. As I shall show in a moment, van Inwagen's technical explication enables me to present a strong argument in favor of the validity of rule beta. Unfortunately, I will have to leave it open whether van Inwagen's revised interpretation of the operator satisfies the second and the third condition as well (given that a discussion of this question would be beyond the scope of this paper).[7]

4.1. An argument for Widerker's version of rule beta

That one might develop a strong argument in favor of van Inwagen's new version of rule beta is not as obvious as it might look like. For given that the original version of rule beta turned out to be invalid, van Inwagen himself now confesses that "although he *accepts* the revised rule, he assigns its validity a rather lower subjective probability than the near certainty he once so confidently assigned to the validity of the original β" (2017a, 13). At least with respect to his original interpretation of beta, van Inwagen went even so far to say that: "I can think of no generally accepted inference rules that seem relevant to (β), much less of any that it can be deduced from. The prospect of *showing* (β) to be valid, therefore, appears to be bleak [...]" (1983, 97).[8]

I think that van Inwagen should be far more confident with respect to the validity of his revised rule. For I shall argue that his revised rule follows from the standard logic of counterfactual conditionals (provided that we add some rules for the quantifiers).[9] However, in order to argue for van Inwagen's version of rule beta, I shall first recapitulate how Michael Huemer and Alexander Pruss modified the original version of the Consequence Argument. It will become clear that their proposal lays the ground for a strong argument in favor of van Inwagen's version of rule beta.

It is in fact interesting to note, in this respect, that van Inwagen's revised interpretation of the operator turns out to be equivalent to an interpretation of the operator that has been suggested by Huemer. Huemer (2000, 538–539) suggests (this is a translation of what he suggests into the terms of van Inwagen's definition) that a truth is untouchable if and only if p and no matter

7 For a lengthy discussion of this question, see Hausmann (2024).

8 Fischer draws a similar conclusion with respect to his version of rule beta: "although I have argued that the Transfer Principle [= Fischer's version of rule beta] cannot be disproved in certain ways, it is not evident to me that it can be proved" (Fischer 1994, 45)

9 My argument for van Inwagen's new version of rule beta may be seen as an attempt to improve upon Carlson's (2000) and Gustafsson's (2017) informal arguments for versions of rule beta (in which they occasionally refer to principles of propositional counterfactual logic). One of the main differences between their arguments and my argument is that their arguments only appeal to principles of *propositional* counterfactual logic (and not to principles of *quantified* counterfactual logic). Unlike Carlson and Gustafsson, I am therefore able to develop a rigorous formal proof (and not only an informal sketch) in favor of van Inwagen's new version of rule beta (or in favor of other versions of rule beta).

what anyone does, it would still be the case that p.[10] For if we use a formal language to represent Huemer's suggestion, we get the following proposal:

(*Huemer*) $\mathrm{U}p \leftrightarrow (p \;\&\; \forall x\, \forall y\, (x \text{ is able to do } y \rightarrow (x \text{ does } y \;\Box\!\!\rightarrow p)))$

I leave it to the reader to verify that van Inwagen's and Huemer's proposal actually turn out to be equivalent. What is important is that, as Pruss has shown, given Huemer's interpretation of the operator, we may use Widerker's (1987, 41) instead of van Inwagen's version of rule beta in order to show that the argument is logically valid. For if we take Widerker's version of rule beta[11] (and if we interpret it according to van Inwagen's or Huemer's proposal), we get the following rule of inference:

Rule β': $\mathrm{U}\, p, \Box\, (p \rightarrow q) \vdash \mathrm{U}\, q$

As Pruss (2013) has shown, given Lewis' semantics for counterfactual conditionals, it seems to follow that this is a valid rule of inference. His argument rests on the idea that if, no matter what anyone does, it would still be the case that p, and if it is necessary that, if p, then q, then, no matter what anyone does, it would still be the case that q. In fact, according to Lewis' semantics, the rule of weakening for counterfactual conditionals holds:[12]

Weakening: $p \;\Box\!\!\rightarrow q, \Box\, (q \rightarrow r) \vdash p \;\Box\!\!\rightarrow r$

And given that weakening holds, we might simply argue as follows: Assume that it is an untouchable truth that p. And assume, further, that it is a necessary truth that if p then q:

(1)	$\mathrm{U}\, p$	assumption
(2)	$\Box\, (p \rightarrow q)$	assumption

It follows, given Huemer's (or van Inwagen's) interpretation, that p and no matter what anyone does, it would still be the case that p:

10 In Huemer's own words: "(No matter what S does, p) = (p, and for each action, A, that S can perform, if S were to perform A, it would still be the case that p)" (2000, 538).

11 Widerker's (1987) original idea was (unlike van Inwagen's or Huemer's idea) *not* to replace the original interpretation of van Inwagen's operator with a new interpretation, but rather to replace van Inwagen's rule beta (according to which, roughly, if nobody has a choice about whether p and nobody has a choice about whether if p then q, then nobody has a choice about whether q) with a new version of rule beta (according to which, roughly, if nobody has a choice about whether p and *it is necessary that* if p then q, then nobody has a choice about whether q). Pruss' strategy is to combine van Inwagen's (or Huemer's) idea with Widerker's idea.

12 Pruss explains: „In Lewis' semantics, weakening is easy to show. If p is impossible, then $p \;\Box\!\!\rightarrow r$ holds trivially. If p is possible, then $p \;\Box\!\!\rightarrow q$ holds if and only if there is a world w_1 in which p and q hold and which is closer to the actual world than any world at which p and ~ q hold. Let w_1 be such a world. Since q entails r, we must also have r holding at w_1. Let w_2 now be any world at which p and ~ r hold. Because q entails r, this is a world at which p and ~ q hold and hence this world is further from the actual world than w_1 is, by choice of w_1. Thus $p \;\Box\!\!\rightarrow r$ holds" (2013, 433). For more on Lewis's semantics see Lewis (1973, 1–43).

(3) p & $\forall x\ \forall y$ (x is able to do y → (x does y □→ p)) follows from (1)

But if we omit the universal quantifiers, we get:

(4) a is able to do b → (a does b □→ p) follows from (3)

And given that the rule of weakening holds, it clearly follows that:

(5) a is able to do b → (a does b □→ q) follows from (2) and (4)

Thus, if we reintroduce the quantifiers, it seems safe to infer that it is an untouchable truth that q:

(6) $\forall x\ \forall y$ (x is able to do y → (x does y □→ q)) follows from (5)

(7) q & $\forall x\ \forall y$ (x is able to do y → (x does y □→ q)) follows from (2), (3) and (6)

(8) U q follows from (7)

Hence, at least at first glance, Widerker's version of rule beta seems to follow from the standard semantics for counterfactual conditionals. Pruss goes even so far to say that Widerker's "version of the beta principle is a theorem given one plausible axiom for counterfactuals (weakening)" (2013, 430). Therefore, we might simply modify the original version of the Consequence Argument using Widerker's instead of van Inwagen's version of rule beta. For if we let 'p_0' be a proposition that expresses the state of the world in the remote past, if we let 'p' be an arbitrary truth and if we let 'l' be the conjunction of all laws of nature then we might argue as follows:

(1) □ ((p_0 & l) → p) follows from determinism

(2) U (p_0 & l) premise

(3) U p follows from (1), (2) and rule β'

As has been shown above, given Huemer's (and van Inwagen's) interpretation of the operator, the crucial step from (1) and (2) to (3) seems perfectly valid now. According to Pruss, this result suggests that "[i]nstead of being about transfer principles, the debate should be over whether the distant past and laws are up to us" (2013, 430). It remains to show, given this improved version of the argument, that the argument is not only logically valid, but sound.

4.2. An argument for van Inwagen's version of rule beta

In what follows I am going to expand on Pruss' proposal and I am going to develop an argument in favor of van Inwagen's – and not only in favor of Widerker's – version of rule beta. In doing so, I anticipate and remove a worry that one might otherwise have with respect to Widerker's version of the argument. For Widerker's version, unlike van Inwagen's version, rests on the assumption that it is an untouchable truth that p_0 & l (as opposed to van Inwagen's two assumptions that it is an untouchable truth that p_0 and an untouchable truth that l). The worry is that Widerker's assumption might, as Finch and Warfield put it, well turn out to be "formally stronger" (1998, 523) than the conjunction of van Inwagen's two assumptions (if the principle of agglomeration turns out to be invalid)[13] and appears, therefore, to be in need of additional justification – unless, of course, van Inwagen's version of rule beta (and, thereby, the principle of agglomeration) can be shown to be valid. In arguing for van Inwagen's version of rule beta, I might, therefore, not only strengthen the case for van Inwagen's but also for Widerker's version of the argument.

The main idea of my argument is as simple as it is intuitive: If, no matter what anyone does, it would still be that case that p, and if, no matter what anyone does, it would still be the case that that, if p, then q, then, no matter what anyone does, it would still be the case that q. Therefore, if it is an untouchable truth that p and if it is an untouchable truth that, if p, then q, then it is also an untouchable truth that q.

My argument for van Inwagen's version of rule beta will, unlike Pruss' argument, not depend on Lewis's whole *semantics* for counterfactual conditionals but rather on a small part of Lewis's *logic* for counterfactual conditionals. This will enable me to proceed from widely accepted and fairly undisputed assumptions about counterfactual conditionals (and their logic) – assumptions that, unlike the assumptions that Pruss has made, are perfectly compatible with a whole range of different semantics for counterfactual conditionals.[14]

However, the main reason why my argument will draw on definitions, rules and axioms of counterfactual *logic* (rather than on principles of counterfactual *semantics*), is that it will not become evident how to argue for van Inwagen's version of rule beta as long as we do not appeal to counterfactual logic (rather than to counterfactual semantics). To this end, let me first introduce a couple of definitions, rules and axioms that are familiar from Lewis' logic for counterfactual conditionals (see Lewis 1973, 22, 132). First of all, as is well known, Lewis defines the necessity operator in terms of the counterfactual connective:

(*Definition*) $\Box A \leftrightarrow (\sim A \mathrel{\Box\!\!\rightarrow} A)$

Second, Lewis assumes that the following rule of deduction within conditionals holds:

13 In fact, all I have done so far is to *defend* the principle of agglomeration (in view of a counterexample). I have not *shown* that, given van Inwagen's revised interpretation of the operator, the principle of agglomeration is valid. For all that has been shown so far, agglomeration might still fail (because of a different counterexample).

14 For instance, the assumptions of propositional counterfactual logic in my argument for van Inwagen's version of rule beta are perfectly compatible with a probabilistic semantics for counterfactual conditionals as developed by Leitgeb (2012a, 2012b).

(*Deduction within Conditionals*)
$\vdash (B_1 \,\&\, \ldots \,\&\, B_n) \rightarrow C$
$\vdash ((A \,\Box\!\!\rightarrow B_1) \,\&\, \ldots \,\&\, (A \,\Box\!\!\rightarrow B_n)) \rightarrow (A \,\Box\!\!\rightarrow C)$

And finally, Lewis assumes interchange of logical equivalents and he takes any truth-functional tautology to be an axiom.

Of course, even though Pruss does not put too much emphasis on this, all these definitions, axioms and rules won't suffice to establish the desired conclusion: we obviously won't be able to derive van Inwagen's or Widerker's rule from Lewis' logic for counterfactual conditionals, if we are not prepared to add some introduction and elimination rules for the universal quantifier.[15] And, as is well known, this might be achieved in a variety of ways.[16] In what follows, I will mainly focus on two alternatives.

On the one hand, we might simply add the rules of classical quantification theory. We then get the following pair of rules:

(∀-Elimination) $\forall x$ A ⊢ A [c/x]
(∀-Introduction) A [c/x] ⊢ $\forall x$ A   where c is not in $\forall x$ A nor in any of its hypotheses

On the other hand, we might prefer to add the quantification rules of free logic. We then arrive at the following pair of rules for the universal quantifier (given that we introduce 'E' as an abbreviation for the existence predicate):

(∀-Elimination) $\forall x$ A ⊢ E(c) → A [c/x]
(∀-Introduction) E(c) → A [c/x] ⊢ $\forall x$ A   where c is not in $\forall x$ A nor in any of its hypotheses

In what follows, I am going to rely on the latter (and not on the former) pair of rules for the universal quantifier. That is, for the sake of argument, I am going to assume that the quantification rules of free logic are valid. But it is evident that, once I adopt the rules of classical quantification theory, I will be able to obtain *the same results*.

With these logical assumptions in hand, I am in a position to develop a strong argument in favor of van Inwagen's (and not only in favor of Widerker's) version of rule beta:

Rule β: U p, U (p → q) ⊢ U q

For, to begin with, suppose that it is an untouchable truth that p. And suppose, further, that it is an untouchable truth that if p then q:

15 Therefore, it is at best misleading to say, as Pruss does, that Widerker's "version of the beta principle is a theorem given one plausible axiom for counterfactuals (weakening)" (2013, 430).

16 For a detailed discussion of quantified counterfactual logic see Priest (2008, 399–420).

(1) U p — assumption

(2) U (p → q) — assumption

It follows that it is an untouchable truth that q. For, given Huemer's (or van Inwagen's) interpretation of the operator, it certainly follows that p and, no matter what anyone does, it would still be the case that p:

(3) p & $\forall x\,\forall y$ (x is able to do y → (x does y □→ p)) — follows from (1)

And likewise, we may certainly infer that if p then q, and that, no matter what anyone does, it would still be the case that if p then q:

(4) (p → q) & $\forall x\,\forall y$ (x is able to do y → (x does y □→ (p → q))) — follows from (2)

It remains to show that q and that, no matter what anyone does, it would still be the case that q. To this end, we should first drop the quantifiers (according to our elimination rule for universally quantified statements in free logic):

(5) E(a) → E(b) → (a is able to do b → (a does b □→ p)) — follows from (3)

(6) E(a) → E(b) → (a is able to do b → (a does b □→ (p → q))) — follows from (4)

We should then pursue a conditional proof strategy and introduce a couple of assumptions:

(7) E(a) — assumption conditional proof

(8) E(b) — assumption conditional proof

(9) a is able to do b — assumption conditional proof

This obviously yields the following results:

(10) a does b □→ p — follows from (5), (7), (8) and (9)

(11) a does b □→ (p → q) — follows from (6), (7), (8) and (9)

But now consider the following truth-functional tautology:

(12) (p & (p → q)) → q — tautology

Now given that we may take any truth-functional tautology to be an axiom, we may simply apply the rule of deduction within conditionals in order to get:

(13) (a does b □→ p & a does b □→ (p → q)) → a does b □→ q

Thus, it certainly follows that:

(14) a does b □→ q — follows from (10), (11) and (13)

We should now close our conditional proof, in order to reintroduce the quantifiers (according to our introduction rule for the universal quantifier in free logic):

(15) $E(b) \rightarrow (\text{a is able to do b} \rightarrow (\text{a does b} \mathrel{\Box\!\!\rightarrow} q))$	follows from (8) to (14)
(16) $\forall y\ (\text{a is able to do y} \rightarrow (\text{a does y} \mathrel{\Box\!\!\rightarrow} q))$	follows from (15)
(17) $E(a) \rightarrow \forall y\ (\text{a is able to do y} \rightarrow (\text{a does y} \mathrel{\Box\!\!\rightarrow} q))$	follows from (7) to (16)
(18) $\forall x\ \forall y\ (\text{x is able to do y} \rightarrow (\text{x does y} \mathrel{\Box\!\!\rightarrow} q))$	follows from (17)

We may now easily infer that it is an untouchable truth that q:

(19) $q\ \&\ \forall x\ \forall y\ (\text{x is able to do y} \rightarrow (\text{x does y} \mathrel{\Box\!\!\rightarrow} q))$	follows from (3), (4), (18)
(20) $U\ q$	follows from (19)

Thus, given (a small part of) the axioms and rules of Lewis's counterfactual logic, not only Widerker's but also van Inwagen's version of rule beta can be clearly seen to be valid (once introduction and elimination rules for the universal quantifier are added).

I conclude that, unlike in the case of the original version of the argument, the logical validity of the modified version of the Consequence Argument should not bother us anymore – at least not as long as we are ready to accept the fairly innocuous logical assumptions that have been made in order to derive van Inwagen's version of rule beta.[17]

5. Conclusion: Beta has Left the Beta Phase

Let me take a step back. Without doubt, Peter van Inwagen's Consequence Argument remains one of the most important and fascinating arguments in the debate. However, the original version of van Inwagen's argument employed an inference rule that was shown to be invalid given his interpretation of the operators. In response, van Inwagen suggested a new version of his argument based on a revised interpretation of his operators. Arguably, van Inwagen's new version depends on the notion of being able to falsify a proposition (at least implicitly, given that it depends on the notion of an untouchable truth).

17 I think it is fair to say that Lewis's logic for counterfactual conditionals is still the standard logic for counterfactual conditionals. There is, however, a debate about whether Lewis's rule of deduction within conditionals is valid. For more on this debate see, for example, Nolan (1997), Williamson (2018) and Berto et al. (2018). Berto et al. (2018) argue against Lewis's rule on the grounds that, in their view, there are false counterfactual conditionals with antecedents that are impossibly true. Their suggestion is, accordingly, to restrict Lewis's rule to counterfactual conditionals with antecedents that are possibly true.

Note that, in the argument for van Inwagen's rule beta, Lewis's rule has only been applied to counterfactuals whose antecedent states that a does b and note that, according to the assumptions of the argument, a is *able* to do b (see p. 148). Thus, given that ability implies possibility, in the argument for van Inwagen's rule beta, Lewis's rule has only been applied to counterfactuals with antecedents that are possibly true. Thus, the argument for van Inwagen's rule beta would still go through even if it turned out that Lewis's rule has to be restricted to antecedents that are possibly true.

Hence, if van Inwagen's argument is supposed to be successful, it should rest on an explication of this notion that is at least such that:

(i) rule beta is valid
(ii) nobody is able to falsify 'p_0'
(iii) nobody is able to falsify 'l'

What we have seen in this paper, is that there is a convincing argument to the conclusion that van Inwagen's new version of the argument satisfies the first condition. For, given a few principles of quantified counterfactual logic, van Inwagen's argument turns out to be logically valid. We had to leave it open, however, whether van Inwagen's new version of the argument satisfies the second and the third condition as well. For all we know, there might still be a serious flaw in the premises of van Inwagen's argument. All that has been shown is that, once compatibilists and incompatibilists reach an agreement about a few principles of quantified counterfactual logic, they may safely ignore questions about the validity of van Inwagen's argument, and they may safely focus on the task of finding a flaw in the premises of van Inwagen's argument.

References

Berto, F., French, R., Priest, G. & Ripley, D. 2018. "Williamson on Counterpossibles." *Journal of Philosophical Logic* 47: 693–713.

Blum, A. 2000. "'N'." *Analysis* 60: 284–286.

Blum, A. 2003. "The Core of the Consequence Argument." *Dialectica* 57: 423–429.

Carlson, E. 2000. "Incompatibilism and the Transfer of Power Necessity." *Noûs* 34: 277–290.

Crisp, T. and Warfield, T. 2000. "The Irrelevance of Indeterministic Counterexamples to Principle Beta." *Philosophy and Phenomenological Research* 61: 173–184.

Finch, A. and Warfield, T. 1998. "The *Mind* Argument and Libertarianism." *Mind* 107: 515–528.

Fischer, J. M. 1994. *The Metaphysics of Free Will. An Essay on Control*. Oxford: Blackwell.

Gustafsson, J. 2017. "A Strengthening of the Consequence Argument for Incompatibilism." *Analysis* 77: 705–715.

Hausmann, M. 2018. "The Consequence Argument Ungrounded." *Synthese* 195: 4931–4950.

Hausmann, M. 2024. *Anderskönnen trotz Determinismus*. Tübingen: Mohr Siebeck.

Huemer, M. 2000. "Van Inwagen's Consequence Argument." *The Philosophical Review* 109: 525–544.

Kapitan, T. 1996. "Incompatibilism and Ambiguity in the Practical Modalities." *Analysis* 56: 102–110.

Kapitan, T. 2002. "A Master Argument for Incompatibilism." In *The Oxford Handbook of Free Will*, edited by R. Kane. 127–157. Oxford: Oxford University Press.

Leitgeb, H. 2012a. "A Probabilistic Semantics for Counterfactuals. Part A." *The Review of Symbolic Logic* 5: 26–84.

Leitgeb, H. 2012b. "A Probabilistic Semantics for Counterfactuals. Part B." *The Review of Symbolic Logic* 5: 85–121.

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

Lewis, D. 1981. Are we Free to Break the Laws? *Theoria* 47: 113–121.

McKay, T. and Johnson, D. 1996. "A Reconsideration of an Argument against Compatibilism." *Philosophical Topics* 24: 113–122.

Nolan, D. 1997. "Impossible Worlds: A Modest Approach." *Notre Dame Journal of Formal Logic* 38: 535–572.

O'Connor, T. 1993. "On the Transfer of Necessity." *Noûs* 27: 204–218.

Priest, G. 2008. *An Introduction to Non-Classical Logic*. Cambridge: Cambridge University Press.

Pruss, A. 2013. "Incompatibilism Proved." *Canadian Journal of Philosophy* 43: 430–437.

Schnieder, B. 2004. "Compatibilism and the Notion of Rendering Something False." *Philosophical Studies* 117: 409–428.

Schnieder, B. 2008. "On What We Can Ensure." *Synthese* 162: 101–115.

Van Inwagen, P. 1983. *An Essay on Free Will*. Oxford: Clarendon Press.

Van Inwagen, P. 2008. "The Consequence Argument." In *Metaphysics. The Big Questions 2nd ed.*, edited by P. van Inwagen and D. Zimmerman. 450–456. Oxford: Blackwell.

Van Inwagen, P. 2013. "A Dialogue on Free Will." *Methode* 3: 212–221.

Van Inwagen, P. 2015. "Some Thoughts on An Essay on Free Will." *The Harvard Review of Philosophy* 22: 16–30.

Van Inwagen, P. 2017a. "Introduction: Van Inwagen on Free Will." In *Thinking about Free Will*. 1–19. Cambridge: Cambridge University Press.

Van Inwagen, P. 2017b. "Free Will Remains a Mystery." In *Thinking about Free Will*. 90–110. Cambridge: Cambridge University Press.

Van Inwagen, P. 2017c. "Ability." In *Thinking about Free Will*. 210–229. Cambridge: Cambridge University Press.

Vihvelin, K. 2013. *Causes, Laws, and Free Will. Why Determinism Doesn't Matter*. Oxford: Oxford University Press.

Widerker, D. 1987. "On an Argument for Incompatibilism." *Analysis* 47: 37–41.

Williamson, T. 2018. "Counterpossibles." *Topoi* 37: 357–368.