Robert Kowalenko

A Curve-Fitting Approach to Ceteris Paribus-Laws

Abstract

Law-like generalisations hedged with a *ceteris paribus*-clause such as widely in use in psychology, the social and biological sciences, are best construed as incomplete strict laws. These incomplete laws can be "fleshed out" by adding a set of enabling, or completing, conditions to their antecedent. In other words, the logical form of a cp-law, *ceteris paribus* $\forall x$ (A -> B), is $\forall x$ (A & C_B -> B). The nature of C_B must be subject to non-ad hoc constraints, however, failing which all putative *ceteris paribus*-generalisations will be trivially true. Two simple and plausible constraints are that: (i) A and CB be jointly sufficient for the consequent of the law, and (ii) the relevant completer also occur in the antecedents of other laws-in other words, that there be many other law-like generalisations of the form $\forall x$ (D& C_B -> E), $\forall x$ (F & C_B -> E), etc. Apparent counterexamples to this proposal can be disarmed by interpreting the epistemology of cp-laws as a curve-fitting problem, which consists in determining the relevant nomic regularity and plotting the correct curve over a very noisy data-set that contains large numbers of outliers and anomalies. The process of specifying the content of the *ceteris paribus*-clause that is hedging a law-candidate is in fact isomorphic with the process of determining which parts of one's data are outlying and anomalous, and which are part of the regularity. I submit that statistical theorems such as the Akaike Information Criterion (AIC) are instrumental in the latter process, and therefore also in the former. AIC states that a law-hypothesis which minimises both the number of adjustable parameters and error variance (i.e. a hypothesis that achieves an optimal balance between simplicity and adequacy to the data), displays the highest estimated accuracy of prediction of future data from the same distribution. I go on to discuss how AIC in combination with conditions (i) and (ii) illustrates the fundamental difference between a *ceteris paribus*-law and a statistical law, and how it yields the distinction between spurious and genuine hedged regularities that is necessary to make cp-laws "respectable". Thus, I show how popular putative problem cases, such as "turtles live long lives" or "U.S. Supreme Court Justices are male" can be dealt with by the theory. Finally, I utilise work by Lange 2000; 2002 to deflect the criticism that cp-laws are, by their very nature incompletable, and hence indeterminate. I close by concluding that the account provides a very simple, powerful, and yet metaphysically conservative account of *ceteris paribus*-laws.

Introduction

Ceteris Paribus-generalisations are hedged, or qualified, generalisations of the form „all Fs are Gs, everything else being equal." They are a familiar feature of economics, where they were first widely used in the 19th century, as for example in the claim that "everything else being equal, the rate of wage varies with the supply of labour"; but they are of course a mainstay of most of the other sciences as well, with varying locutions serving as the hedging clause. Thus, we find generalisations such as

- 'normally, officers in the 18th-century British Navy were aristocrats,'
- 'if nothing interferes, planets travel in elliptical orbits,'
- 'ideally, the pressure (P) of a gas in a container varies with the temperature (T), volume (V), or number of gas molecules (N) according to the equation P = nRT/V.' [1]

But hedged generalisations also permeate everyday life: if I have been reliably told—say, by a good friend—that, 'usually, Paul is in the Pub on Friday nights,' then this knowledge together with the fact that it is Friday night is likely to make me believe that Paul is now in the Pub, all other things being usual:

- 'usually, Paul is in the pub on Friday nights'

Manifestly we use hedged generalisations on a daily basis to summarise, explain, and to predict events, although we do not in every context explicitly indicate that they are saddled with a *proviso* of some sort.

Their ubiquity and their epistemic significance naturally raises the question whether some hedged generalisations could be considered laws of nature. Laws are, after all, what we *also* sometimes invoke when we summarise, predict and explain events, and they are what the sciences are widely believed to discover. If we took a traditional Humean view of a law of nature as a strict but contingent regularity between property instantiations F and G, expressed in a universal conditional, $\forall x\ (Fx \rightarrow Gx)$, then *ceteris paribus*-laws would simply be non-strict laws that allow exceptional instances of an x being F and not being G, on the condition that in those instances the circumstances denoted by the expression 'all other things being equal', 'normal', 'ideal', etc. We could express this idea provisionally by prefixing our law with a cp-operator:

---

1 The second and third generalisations are stock examples of purported *ceteris paribus*-laws; the first is due to Lange 2002.

$\mathrm{cp}\ (\forall x\ (\mathrm{F}x \rightarrow \mathrm{G}x))$

One of the advantages of admitting such a type of law would be that this would yield a simple and immediate explanation of the manifest cognitive and epistemic significance of many hedged generalisations. A second, and perhaps more important advantage would be the immediate vindication of the status of the special sciences, which discover few if any strict regularities in their respective domains.

The question is of course whether the advantages outweigh the disadvantages. The problems with the notion of a cp-law are said to be numerous, but can ultimately be reduced to two main problems with the very notion of a *ceteris paribus*-law:

(1) the exact meaning of the *ceteris paribus*-clause and thus its contribution to the truth conditions of a hypothesis is unclear—we have no semantics for cp-laws.

(2) *ceteris paribusceteris paribusceteris paribus*-clauses seem to render hypotheses untestable—we have no methodology for the testing and confirmation of cp-laws.

There have been numerous proposals that attempt, in very different ways, to overcome the "no semantics" and the "no methodology" objections, and equally numerous the rebuttals and refutations of those proposals.[2] I cannot look at any of these in any detail in the space I have here. Rather I shall present my own proposal, and illustrate how I think it solves some of the problems and counterexamples that have dogged previous accounts.

A completer-account

My proposal is what Woodward 2002 calls a "completer"-account. The general argument strategy is a tried-and-tested one (and a "hopelessly misguided" one, according to Woodward...). This is to first is assume as unproblematic *strict* laws of nature for a given

[2] Pro cp-laws: Cartwright 1983, Hausman 1988, Hausman 1992, Kincaid 1990, Fodor 1991, Lange 1993 Lange 2002, Pietroski and Rey 1995, Lipton 1999, Morreau 1999, Schurz 2002, Glymour 2002, Spohn 2002; contra cp-laws: Giere 1988, Schiffer 1991, Schurz 2001, Mott 1992, Weinert 1997, Earman and Roberts 1999, Earman, Roberts et al. 2002, Woodward 2000, Woodward 2002, Mitchell 2002, Smith 2002.

domain in the form of universally quantified conditionals3, and then to claim that ceteris paribus-laws are simply incomplete strict laws of the domain—with the cp-clause denoting the missing completing condition(s) of the law. Thus, the cp-clause is nothing but a placeholder for the missing completion, and cp-laws will be legitimate to the extent that their ceteris paribus clauses can be connected in some systematic way with the underlying strict laws in the domain. Hence our condition (1):

cp $\forall x\ ((Fx \rightarrow Gx)$ if

(1) $\forall x\ ((Fx\ \&\ \mathbf{C}x) \rightarrow Gx)$

The idea here is that every exception to "All Fs are Gs", i.e. an x that is F but not G, will be due to the absence of the completer C. Now, this has of course been tried before (cf. E.g. Fodor 1991, Hausman 1988, 1992, Pietroski and Rey 1995), and it didn't work. In a nutshell,[4] the problem is that $\forall x\ (Fx\ \&\ Cx \rightarrow Gx)$ as it stands is trivially true if the completing condition C is not qualified in just the right way. For example, on the interpretation:

F*x*: *x* is spherical
G*x*: *x* is conductive,
C*x*: *x* is made of copper

the obviously spurios "*ceteris paribus*, all spherical bodies are conductors" will turn out true, because $\forall x\ (Fx\ \&\ Cx) \rightarrow Gx))$ is strictly true. (Earman and Roberts 2002). In fact, it seems as for non-probabilistic laws, for any G, *some* condition C will always be sufficient. Note however that all the work in counterexamples of this type (of which there are indefinitely many) is done by the completer C*x*, which is sufficient by itself for electrical conductivity. Hence, a good second constraint on the account seems to be Fodor's (Fodor 1995): neither F nor C can be nomically sufficient by itself for G, but must *jointly* be so.

(2) neither F nor C is nomically sufficient by itself for G.

---

[3] I am assuming here, for the purposes of this argument, that these conditionals have "nomic force", in other words they are more than accidental generalizations, and that their antecedents state nomically sufficient conditions for their consequent. In other words, I intend to focus on the specific nature of the contribution of the cp-clause to a law-statement, by helping myself to the notion of a strict law, and by presupposing that we have some general understanding lawlikeness. *Divide and conquer* is certainly the right strategy in this field, and it is rather disingenuous to charge defenders of cp-laws with the additional tasks of also providing a theory of law-likeness and/or confirmation in one fell swoop.

[4] The problems with this approach are scrutinized in detail in the literature (Schiffer, Earman and Roberts, Spohn, Mott, among others), parts of which Woodward briefly surveys.

This should be uncontentious: if either of the conjuncts in the antecedent is nomically sufficient on its own, the other conjunct will be idle, and can be left out of the law.

Yet, this has, again, been shown to be not nearly enough by a barrage of critics, who have found further counterexamples of spurious generalisations that seem to pass the test:

> "*ceteris paribus*, thirsty human beings will eat salt" (Mott)
> '*ceteris paribus*, all charged objects accelerate at 10 m/s^2'. (Woodward)

Many of the critics of cp-laws have since considered the case closed, and the completer-account of cp-laws as incomplete strict laws, to be fundamentally wrong-headed (Woodward 2002). This is, I believe, too hasty.

A curve-fitting supplement

I suggest we keep our first two conditions, and add two conditions to the account:

cp $\forall x\ (Fx \rightarrow Gx)$ if
(1) $\forall x\ ((Fx\ \&\ Cx) \rightarrow Gx)$
(2) neither F nor C is nomically sufficient by itself for G

+

(3) C occurs in the antecedents of other laws
(4) If both a hypothesis and its contrary satisfy (1)-(3), then we must chose the one that minimizes error variance, while maximizing closeness of fit to the observational data and simplicity (Akaike Information Criterion)

<u>Motivation for condition (3):</u>

The reasons for invoking (3) are quite straightforward. Candidate cp-laws and their corresponding completers clearly need to be *vouched for* in order to avoid *spurious counter-examples*, as we have seen. This is the specific problem of any theory of cp-laws: cp-laws are in danger of vacuity, as Pietroski and Rey put it, and most accounts of hedged generalisations must play a game of "avoid the counterexamples". We have seen so far that too much latitude in what we allow Cx to be results in easily reproducible nonsense. For example, on the present account, we might just as well claim that "*ceteris paribus*, turtles outrun hares" (Boghossian). After all, on an interpretation of Cx, which includes either

- divine intervention
- "ideal" conditions for turtles such that turtles have much stronger hearts, more flexible joints, longer legs, ...

we could indeed justify that claim that if Cx, then tortoises could run faster than hares. We need to restrict the range of conditions to the realm of the scientifically accessible, and a quick and efficient to do this is to require that Cx figure in the antecedents of other scientific laws. Hereby, we ensure that the purported completing condition have some scientific standing, and we encode the natural assumption that it ought always to be established science that decides what legitimate *ceteris paribus*-laws are. In other words, condition (3) provides a demarcation criterion for cp-laws from pseudo-science and plain nonsense. I am going to call completers of the „divine intervention"-type just completers, and completers that satisfy (3) and figure in existing laws, „nomological completers".

Condition (3) amounts to a boot-strapping method similar to one first proposed by Fodor 1991 for *ceteris paribus*-laws: candidate cp-laws and their relevant completers are vouched for here *via* a boot-strapping procedure whereby a cp-generalisation derives its nomic status relative to a background theory that already includes other cp-laws. (3), incidentally, also gives shape to the notion that nature's regularities 'cut across each other' and that we will always need to understand the (second-order) regularities that govern a given completer in order to understand why some of nature's regularities sometimes fail to hold in a given case.

Let's take a look at how (3) works in an actual example, where it distinguishes between a merely descriptive summary of a uniformity in events or behaviour, from a genuine *ceteris paribus* law. Take the generalisation:

> "*ceteris paribus* U.S. Supreme Court Justices are male'

I claim that this is not a proper sociological cp-law, because it is neither a strict law (there have been 2 females to date), nor does it seem to satisfy condition (3): it is not the case that for every exception (the appointment of a *female justice)* there is a nomological completer whose absence would account for the exception to the law in that particular case. There is no (lawful, nomological) Cx here, because there are presumably no stable and systematically reoccurring conditions under which male candidates will necessarily be preferred over females ones, and there are no such conditions that figure in other sociological laws. Hence it is not a legitimate law.

Now, this last claim might very well be challenged. Perhaps this is a false analysis, and that is exactly the point here: the question whether or not there are any nomological completers to a given generalisation must always be one for empirical research. So, if it turned out that the correct socio-political analysis of the appointment process implies that there are in fact systematically reoccurring, non-accidental, conditions under which men are, against the injunctions of the U.S. Constitution, systematically preferred over women, and if these conditions are theoretically promiscuous—in other words, if they also occur in other sociological laws governing, say, the election of Presidents, the appointment of Attorney Generals or Cabinet members, etc.,—then we would might have to take a second look and reconsider whether the law is not, perhaps, after all a genuine cp-law in political science.

Motivation for condition (4) :

So far so good, but (3) on its own is still not enough.[5] The problem, as it has turned out, is that by the lights of conditions (1)-(3) there will still be a rather large set of legitimate, nomological, completers, so large in fact that we can legitimize both a given hedged generalisation, and its contrary:

"It'll fly, *ceteris paribus*" <-> "it won't fly, *ceteris paribus*"

Or, to take an example closer to scientific practice:

Ceteris Paribus, nuts are healthy <-> *ceteris paribus*, nuts are lethal

Here we have a case where both a scientifically acceptable hypothesis A and its contrary seem to have a legitimate completer. Nuts are healthy: nuts have, notwithstanding their high fat content, been discovered to carry many health benefits, in particular that of being cardio-protective, So it is a roughly true generalisation of clinical nutrition that (regular) nut consumption creates positive health outcomes. Yet, evidently, there are numerous counter-instances to this "law", people with a specific allergy, N, who do experience severe and sometimes fatal anaphylactic reactions due to ingestion of nuts.

[5] Something analogous has been tried before (cf. for example Fodor 1991 and Pietroski and Rey 1995).

A sceptic can thus make the rather gratuitous claim that 'ingestion of nuts → death' is a true nomic regularity. It is just that this "law" is being "interfered" with by the more or less permanent absence in large parts of the population of a nuts allergy. Such a rogue "law" would predict that every consumption of nuts is followed by death and explain away the vast number of exceptions with the fact that in all those cases the necessary completer, absence of the allergy, happened to be not instantiated. We can see that the completer invoked here, possession of the allergy N, would be legitimate according to conditions (1) -(3), for not only is '(*x* eats nuts & *x* has N → *x* dies)' true, the allergy by itself is not sufficient for death, and we can expect there to be many other laws governing other regularities in the behaviour and properties of individuals having N.[6]

The most fruitful reply to this sort of counterexample is, I believe, to deploy a curve-fitting analysis. For the evidential reasoning of someone who on the basis of a given set of data seriously entertains the hypothesis that 'cp (x eats nuts → x dies)' must differ markedly from the reasoning of someone who on the basis of the same data comes to the contrary conclusion that 'cp (x eats nuts → x experiences health benefits).' An observer who accepts the latter is making sense of the evidence in a different way, a way we should try to make explicit and subject to objective scrutiny. Our relevant evidence is as follows: millions of people eat nuts every day, and with very few exceptions no adverse health effects are observed. In other words, we dispose of the set of data $\{Fa_1 \& Ga_1, Fa_2 \& Ga_2, \ldots Fa_n \& Ga_n\}$, where F and G express the properties 'has eaten nuts' and 'experiences health benefits', respectively, a ranges over human beings, and n is fairly large. We can display this evidence in a scatter graph, by taking instances of nut ingestion at specific times as our independent variable, and differential impact on health as our dependent one, as follows:

[6] In reality, the mere consumption of nuts together with possession of a nuts allergy is (fortunately!) not strictly sufficient for death—a number of other conditions must be fulfilled as well. But I shall ignore this complication here, and assume that N is indeed a fully specified completer of the "law" 'cp (x eats nuts-> x dies)'.

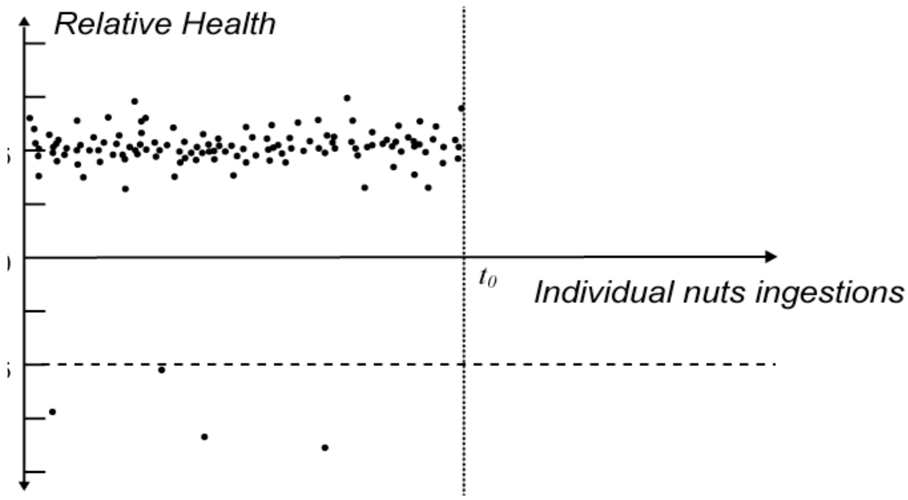

Figure 1.[7]

A very simple idea is now this: given our evidence, and in the absence of all further background knowledge, we should clearly say that "*ceteris paribus*, eating nuts is healthy". In other words, we should plot the following curve

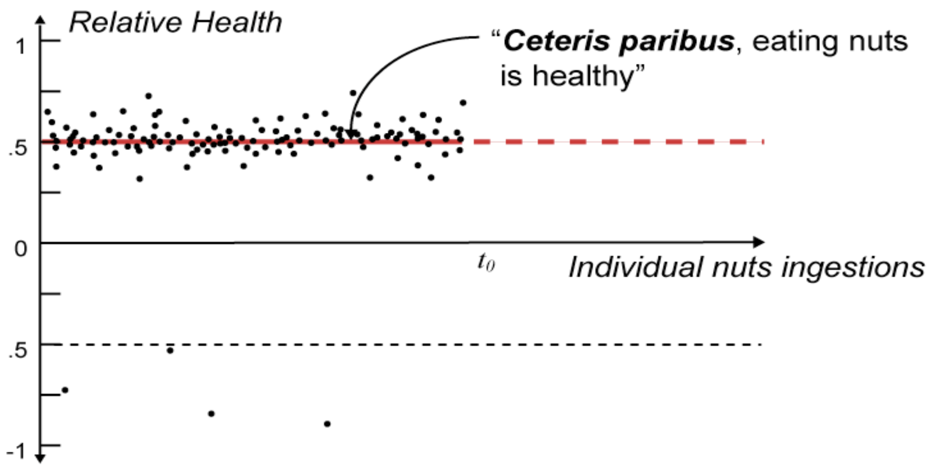

Figure 2.

---

[7] A note on the graphs. This representation is somewhat imperfect for a number of reasons: consumption is a matter of degree, and health effects are a matter of time, so I should represent the number of nuts consumed here together with health effects over a period of time. But I have assumed that the causal effect of nut consumption on health is more or less immediate and independent on quantity to simplify things, allowing me to represent each consumption via a single data point, and to spare me the need of representing the evidence as results of longitudinal studies of each individual's health over time. This still leaves me with a binary variable here of instances of consumption of nuts at different times, whereas ideally I would like to have a simple independent variable. Furthermore, although I will be plotting curves over the data, the curves should strictly speaking have no intercept, because at the intercept the value of the independent variable is of course 0, so that it cannot represent the causal influence of nuts. Finally, I'm presupposing of course that we actually have clear quantitative criteria for 'health' that I can put on my *y*-axis, and that an individual's one-time consumption of nuts would register on that scale.

On the other hand, to say that nuts are lethal, *ceteris paribus*, is somewhat perverse:

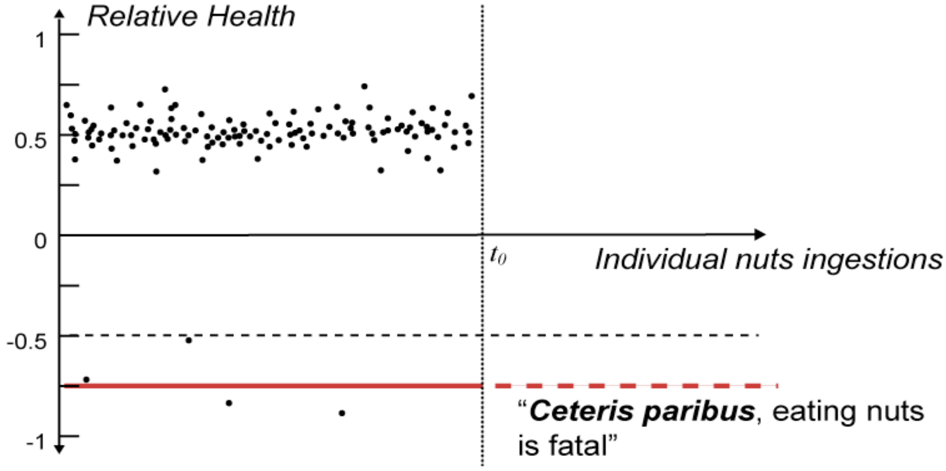

Figure 3.

By framing the task of deciding between these two alternatives as a task of choosing between two alternative curves to plot over our data, we can complete our account of cp-laws.

Let's go back to Figure 2, and the hypothesis that 'cp, nuts are healthy'. If differential impact on health and nut ingestion are the only two relevant quantities that you know of in this particular epistemic context—in other words, if you know nothing yet of the allergy—then the task of formulating a hypothesis concerning the relation between these two observed quantities becomes a trivial exercise of linear regression, familiar to all experimental scientists. In the absence of further background knowledge any experimenter would find x and y to be correlated, and the simplest mathematical model adequate to the data would be the hypothesis that the relation between nut consumption and health effects is linear and continuous, i.e. of the form $y = \beta_1 + \beta_2 x$. So we calculate the coefficients of our straight line using the least squares method by minimising the sum of the squares (SOS) of the vertical distances between each data point and our curve – and we get the straight line at y= + 0.5 (The estimate for the constant term β_1 and slope coefficient β_2 is 0.5 and 0, respectively, in other words we conclude that nut consumption is followed by a positive differential impact on health and that the size of this impact (+0.5) remains constant over time).

Now, even if we were now to learn about allergy N, then we would not change our minds on which line to plot – but we would be quite satisfied that we've found an interfering variable which can explain those very strange anomalies at the bottom of the graph.

Not so our perverse curve-fitter. Learning about anomaly N, she says the following: let's take the sparse points below − 0.5 to be the normal distribution of our data, and ignore the main bulk around + 0.5 for the calculation of the mean. In other words, let's declare all data around + 0.5 anomalous. Then, apply linear regression analysis, and plot the regression line way down in the lower part of the graph. So the assumptions of the perverse curve fitter are (1) the true regularity between the observed quantities is the one represented by the lower curve, 'nuts are lethal'; (2) this regularity is not observed more frequently because an interfering factor, the absence of N, is permanently present; (3) if the interfering factor, the absence of N, were absent - in other words, if we did have a majority of people with the allergy, then the bulk of the data would be normally distributed around $y = -0.75$. So, as things stand now, the actually observed y-values of the independent variable are highly skewed towards 0.5, but that this is merely an artefact of conditions specifically excluded by the cp-clause of 'cp (x eats nuts → x dies)'.

The moral to be drawn from this example is that cp-clauses are clearly very powerful, and they can be abused in order to arbitrarily turn on its head what is considered an exception (anomaly) and what is the norm. We hence need normative rules to constrain our interpretation of the evidence in favour of a given cp-law. My choice for such a rule is the Akaike Information Criterion.[8] AIC states that:

$$\text{Estimated (Distance from the truth of family F)} = \text{SOS (L(F))} + 2\sigma^2 k + C,$$

where a 'family' of curves F is a set given by the form of the curves as it manifests itself in their algebraic expression ('$y = ax + b$', for example, determines the family of straight lines, '$y = a + bx + cx^2$' that of parabolas, etc.); 'SOS (L(F))' stands for the sum of squares of the best-fitting curve of the family of curves F; the term 'k' denotes the number of adjustable parameters of that family; 'σ^2' represents the distribution of "er-

---

[8] The Akaike Information Critrion (AIC) was first proposed by Akaike 1973, and introduced to a philosophical audience more recently by Forster and Sober 1994. AIC, I should mention here, is contentious. Curve-fitting is a hard problem not only for scientists, but also for philosophers who are interested in the epistemic justification of the practice; not everyone believes that it is a always possible to justify our curve-fitting preferences. AIC is particularly unpopular with Bayesians, as it provides an alternative, non-Bayesian way to justify a non-deductive, ampliative, inference. Bandyopadhyay, Boik et al. 1996 claim that a Bayesian curve fitting criterion, BTC, is equivalent to AIC given the right choice of priors. I have no firm opinion on whether Bayesianist epistemology provides a better justification of curve-fitting than AIC, but this is not a debate that needs to be decided here. What I intend to do is instead to take known curve-fitting criteria and show how they can be usefully applied in a theory of cp-laws. I find that AIC yields extensionally correct results - an indirect justification of its use here, granted, but a justification nevertheless.

rors", or anything that appears as error given the hypothesis, i.e. the influence of additional factors; and 'C' is a constant.

Forster and Sober 1994 argue that AIC theorem justifies simplicity considerations in curve-fitting. When the best family of curves (or the best model) has been chosen according to the criterion provided by AIC, we can determine the best-fitting curve from that family through the method of least squares. The strength of AIC, according to Forster and Sober, is that it shows how these two steps, which are usually considered to be separate, are related. The fewer adjustable parameters (k) a model contains, the fewer parameter values need to be estimated through regression techniques; the fewer values need to be estimated, the smaller the estimation error, and the better predictive accuracy. AIC thus promises to justify our general preference for simplicity (quantified by k) without appeal to obscure metaphysical principles or mere pragmatic convenience, by reference to the data itself: AIC shows that, given the data, a hypothesis with fewer adjustable parameters will have the tangible effect of tending to lead to better predictions of future data from the same distribution (Forster and Sober 1994: 27).

How is this theorem to be applied to our problem? Well, given a certain data set, we obtain a reliable estimate of the predictive accuracy of model 'F' by adding the term '$2k\sigma^2$' to the sum-of-squares of the best-fitting curve derived from F. In other words, AIC tells us that in order to get the best predictively adequate hypothesis out of our data, we need to minimize SOS (L(F)) (distance to observed values of the dependent variable, σ^2 (error variance, amount of fluctuation in our data around the curve), and k (complexity of our curve.

In normal curve-fitting AIC compensates for potential over-fitting by penalising models that are too complex through the addition of a positive correction term proportional to their number of parameters, hence the 'k' in '$2k\sigma^2$'. For our present purposes, where we are dealing with a choice between two straight lines the degree of complexity of which is identical, the decisive factor is not k but σ. AIC penalises our perverse curve-fitter for the error variance in the data as she interprets it. σ measures the degree to which the observed data tends to fluctuate around the true curve. A greater number of outlying points will increase its value: if nuts are indeed lethal, then most of the data are outliers, and the value of σ^2 will in that case be very high, disqualifying the hypothesis.

Of course, we can imagine the defender of the "*ceteris paribus*, nuts are lethal"-hypothesis objecting as follows: "my law predicts $y = -0.5$ only for observations made under very specific conditions, and these will remain rare! Given that people will continue in their majority to survive eating nuts, I fully expect the error variance in my data to remain high in the future, and outliers (from my perspective) to remain endemic. The point is that this is provided for by the cp-clause hedging my hypothesis! Don't forget that the cp-clause in front of my law hedges not only the regularity it contains, it obviously also hedges the predictions I am liable to make on its basis. Therefore, your criterion of estimated predictive accuracy, which merely measures the difference between predicted value and observed value, does not apply to cp-laws, because it fails to capture the fact that my cp-clause will exclude most of the observed values from being taken into account in the first place."

In other words, she points out that the cp-clause hedging her "law" explicitly foresees and provides for large error variance, and that this fact is therefore not to be counted against her hypothesis.

Yet, this defence is ineffectual. The error term 'σ^2' can be interpreted in more ways than one. True, its most common interpretation is that of representing measurement inaccuracies and other random interfering factors—the noise in the data. But 'σ^2' is fundamentally just an epistemic term, which denotes *all unknown factors* that, given our working hypothesis, appear to us as 'error', no matter how that appearance actually arises. Due to our necessarily limited epistemic situation, some sort of 'error' or other in this sense is unavoidable. We can quantify the kind of 'error' arising from the action of additional variables we do not know of in just the same way as we can treat the role of measurement inaccuracies or random noise. There is no need to assume that the value of 'σ^2' is known beforehand, for it can be treated as just another adjustable parameter.

Summing up: AIC says that three factors, including error variance, must be weighed up against each other, and thus provides a quantifiable reason for avoiding hypotheses that, at no gain in simplicity, require us to interpret our evidence as packed with outliers. It is an interesting fact about many spurious cp-laws that the evidence for them is, just like in our nuts-example, indeed packed with outliers, and they hence have a large 'error' variance. (e.g. "cp, turtles outrun hares", etc.).[9] Thus, we can now conclude that cp-laws are incomplete strict laws (condition 1), whose putative completers must not be sufficient on their own for the consequent of the law (condition 2); that they must pass a scientific vetting process, which we are expressing through the requirement that they occur in other scientific laws (conclusion 3); and that while a large number of the type of rogue cp-hypotheses that threaten cp-laws with vacuity are immediately ruled out by (3), some unacceptable completers satisfy this requirement, but they do not satisfy (condition 4), namely they do not simultaneously minimize error variance, complexity, and distance from the observational data. Hence, (4) gets rid of further counterexamples.

Objections

This brings me to the last section of this paper. A theory of cp-laws is a sport of examples and counter-examples, and any proposed account must be measured in its extensional adequacy. Does it properly show how those hedged generalisations, which we accept as true, can be considered cp-laws, while explaining why those which we do not accept as true, cannot be so considered? To see whether the account tendered here does this job, I shall conclude by looking at a few more popular problem cases, as well as a very general objection to the completer-account.

Objection (1): what if the we have a case of a true cp-generalisation, where the majority of cases does not fit the law? Say, "*ceteris paribus*, sea turtles live long lives"? What is "normal" for the turtle, namely a long life, and what we would expect, ceteris paribus, is not what is statistically the case, because most turtles fall prey to predation within minutes of hatching…

Well, the first thing to note here is that the very point of cp-laws is that they are precisely not statistical laws (they are strict laws!). It is also worth bearing in mind that AIC equally is not about the statistical mean (it would make us choose Hypothesis 2 over Hypothesis 1 in Figure 4)

---

[9] AIC which constraints our curve-fitting practices in other contexts via its simplicity criterion, reduces in the case of the "curves" in Fig. 2 and 3 to the simple requirement to minimize error variance. To that extent, other theorems of descriptive statistics might take its place. My primary interest here, to reiterate, is in the application of curve-fitting norms to cp-laws, not in the question which one of the alternative curve-fitting norms is to be preferred.

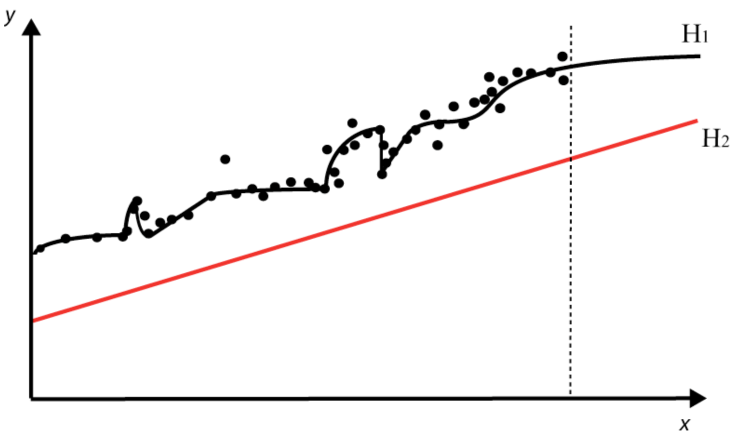

Figure 4

But the most pertinent reply to the objection is that like in the 'cp, turtles overtake hares' example, deciding whether a hedged generalisation is a valid cp-law always requires a look at the scientific details of the case, and the nature of the postulated completers. Given that most early turtle deaths are caused by the action of predators within one hour of hatching, and given that the absence of predation plays a significant role in a whole panoply of other biological cp-laws (e.g. in population biology), absence of predation can be accepted as a legitimate completer of the cp-clause in 'cp, turtles live long lives'. Thus, we can truthfully say that in the absence of predation, turtles live long lives.[10]

What about 'predation'? Can we not equally take this as our completer, and claim that the fact that turtles live short lives in the presence of predators shows the truth of 'cp, turtles live short lives'? Well, such a law with the corresponding completer 'predation' would fall out of a model according to which the turtle organism has a very short average life span even under conditions which are ideal for it. This despite the deeply paradoxical consequence that every act of reproduction would in such a model amount to an 'anomaly'... Were we to adopt such a model, our law would need to explain away every exceptional instance of a turtle older than 1h. But this doesn't make biological sense. (I am close to Gerhard Schurz's notion of 'normic laws' (cf. for example Schurz 2002). A life span of 1h is not a trait for which turtle organisms could have been selected, and therefore cases in which turtles exceed a life span of 1h do not call for an explanation-every contrary case does however call for an explanation. So we shouldn't accept predation as a completer in this case—survival until reproduction is what nature "intended", and we should expect this to be reflected in our biological laws and their

[10] Naturally, 'absence of predation' can only make up part of our completer, here. Other "ideal" conditions such as 'health', etc., would also be part of the mix. For more on the actual completability of completers, see below.

completers: it is the absence of predation, rather than its presence, which would constitute "ideal" conditions for the survival and reproduction of a given turtle organism. Thus, I would not expect to find any nomological completers for the rogue law 'cp, turtles live short lives'.

The same applies, *mutatis mutandis,* for a distinct class of objections based on a different class of counterexamples, which I have already treated above: the converse case in which the majority of cases do fit with what a generalisation predicts, and yet this is insufficient for the truth of the corresponding cp-law. (the Supreme Court example).

Objection 2: "But is, quite independently of the availability of specific counterexamples, the theory not clearly and self-evidently hopeless, because cp-laws are in principle *incompletable?"*

The completer-account, after all, might be accused of not seeing the forest for the trees, or of running down a blind alley… This is because for any interesting cp-generalisation, a fully explicit description of the conditions under which the antecedent+completer nomically necessitates its consequent, would be unstateably long or perhaps even infinitely long. None of the proposed completers, it seems, are really completers as defined, namely (jointly) sufficient conditions for the consequent. No one, certainly no cognitively limited agent, can non-vacously specify the sufficient conditions of any interesting physical event, let alone the antecedent of a law of nature.

Yet, this problem is of course old news in experimental science. The conditions a cp-clause refers to are often unfathomably complex, yes, but we should not have expected anything else. Our best representation of what it is to be 'normal', 'equal', or 'ideal' in the relevant senses will always be a heavily simplified and abstract model of the situation, which leaves out countless features. That's after all exactly how science works: it constructs simple structured representations as substitute systems for investigating and understanding the real systems they model. Without simplified representations of what we wish to understand or to do, we could not understand or do anything. Although we can never hope to complete our cp-generalisations, this circumstance alone does not stop them from fulfilling the other conditions on law-hood, and of playing the role of laws in science.

Thus, I endorse the assessment of Lange 2002, here, that the degree of certainty afforded by our incomplete knowledge of completers will not necessarily be inferior to our certainty of other, supposedly better known facts. Indeed, the very idea of a fully explicit antecedent or of a fully determinate law-statement is a myth, as Lange shows. Lange provides, in my view, a conclusive argument for the view that whether a given law is fully explicit or *ceteris paribus*, when we have a newly discovered object the question whether it is to be subsumed under that law will be equally acute in both circumstances, and will be decided in the same way. He points out, in particular, that in order to decide whether a newly discovered causal factor qualifies as interfering with normal (equal, ideal) conditions, there must be sufficient agreement within the discipline on the relevant respects of comparison, so that analogies with canonical examples supply 'compelling reasons' for deciding whether or not the factor does qualify (Lange 2002, 408-10). So we must always cash out the provisions of a given law, whether strict or *ceteris paribus*, on a case by case basis—a process 'that should be considered "business as usual" rather than symptomatic of a poisonous vagueness' (ibid.; cf. also Lange 2000).

Conclusion

A metaphysically conservative account of cp-laws in terms of completers is, contrary to the views of some, not hopeless. Simple, plausible conditions (1)-(3) in conjunction with straightforward curve-fitting criteria as provided by the Akaike Information Criterion, yield an account of cp-laws as disguised strict laws. This account provides a distinction between spurious and genuine hedged regularities, which is necessary to make cp-laws "respectable", and it also illustrates the fundamental difference between a *ceteris paribus*-law and a statistical law. The proposed theory deals successfully with popular putative problem cases, such as "turtles live long lives" or "Supreme Court Justices are male". The fact that cp-laws are, by their very nature incompletable, and hence indeterminate, is contrary to initial appearances not a serious problem, insofar as (as Lange (2002)) has shown) the type of indeterminacy exhibited by cp-laws is not fundamentally different from that exhibited by strict laws. Thus, we have a straightforward and plausibe account of cp-laws, which is essentially compatible with a simple Humean view of laws as regularities, and dispenses with the introduction of additional metaphysical entities, such as dispositions, capacities, etc. Whatever the arguments for introduction of the latter entities into the discourse of science, they can draw no additional support from the alleged problem of cp-laws.

References

Akaike, H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle. 2nd International Symposium on Information Theory. B. N. Betrov and F. Csaki. Budapest, Akademiai Kiado.

Bandyopadhyay, P. S., R. J. Boik, et al. (1996). "The Curve Fitting Problem: A Bayesian Approach." Proceedings of the Biennial Meetings of the Philosophy of Science Association **3**(Suppl): S264-S272.

Cartwright, N. (1983). How the Laws of Physics Lie. Oxford, Clarendon Press.

Earman, J. and J. Roberts (1999). "*Ceteris Paribus*, There Is No Problem of Provisos." Synthese **118**(3): 439-478.

Earman, J., J. Roberts, et al. (2002). "*Ceteris Paribus* Lost." Erkenntnis **57**: 281-301.

Fodor, J. A. (1991). "You Can Fool Some of the People All of the Time, Everything Else Being Equal; Hedged Laws and Psychological Explanation." Mind **100**(1): 19-34.

Forster, M. and E. Sober (1994). "How to Tell when Simpler, More Unified, or Less "Ad Hoc" Theories Will Provide More Accurate Predictions." British Journal for the Philosophy of Science **45**(1): 1-35.

Giere, R. (1988). Explaining Science: A Cognitive Approach. Chicago, University of Chicago Press.

Glymour, C. (2002). "A Semantics and Methodology for *Ceteris Paribus* Hypotheses." Erkenntnis **57**: 395-405.

Hausman, D. (1988). "*Ceteris Paribus* Clauses and Causality in Economics." PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association **2**: 308-316.

Hausman, D. M. (1992). The Inexact and Separate Science of Economics. Cambridge, Cambridge University Press.

Kincaid, H. (1990). "Defending Laws in the Social Sciences." Philosophy of the Social Sciences.

Lange, M. (1993). "Natural Laws and the Problem of Provisos." Erkenntnis **38**(2): 233–248.

Lange, M. (2000). Natural Laws in Scientific Practice. New York, Oxford Univ Pr.

Lange, M. (2002). "Who's Afraid of Ceteris-Paribus Laws? or: How I Learned to Stop Worrying and Love Them." Erkenntnis **57**(407-423).

Lipton, P. (1999). "All Else Being Equal." Philosophy **74**(288): 155-168.

Mitchell, S. D. (2002). "*Ceteris Paribus*--An Inadequate Representation for Biological Contingency." Erkenntnis **57**(3): 329-350.

Morreau, M. (1999). "Other Things Being Equal." Philosophical Studies **96**: 163-82.

Mott, P. (1992). "Fodor and *Ceteris Paribus* Laws." Mind **101**(402): 335-346.

Pietroski, P. and G. Rey (1995). "When Other Things Aren't Equal: Savings "*Ceteris Paribus*" Laws from Vacuity." British Journal for the Philosophy of Science.

Schiffer, S. (1991). "*Ceteris Paribus* Laws." Mind **100**: 1-17.

Schurz, G. (2001). "Pietroski and Rey on *Ceteris Paribus* Laws." British Journal for the Philosophy of Science **52**(2): 359-370.

Schurz, G. (2002). "*Ceteris Paribus* Laws: Classification and Deconstruction." Erkenntnis **57**(3): 351-372.

Smith, S. R. (2002). "Violated Laws, *Ceteris Paribus* Clauses, and Capacities." Synthese **130**(2): 235-264.

Spohn, W. (2002). "Laws, *Ceteris Paribus* Conditions, and the Dynamics of Belief." Erkenntnis **57**: 373-94.

Weinert, F. (1997). "On the Status of Social Laws." Dialectica **51**(3): 225-242.

Woodward, J. (2000). "Explanation and Invariance in the Special Sciences." British Journal for the Philosophy of Science **51**: 197-254.

Woodward, J. (2002). "There Is No Such Thing As a *Ceteris Paribus* Law." Erkenntnis **57**(3): 303-328.