Venetian Sea Levels, British Bread Prices and the Principle of the Common Cause: A Reassessment

Iñaki San Pedro*

Contents

1 Introduction

It is still a controversial issue whether Reichenbach's Principle of the Common Cause (RPCC) is a sound method for causal inference. In fact, the status of the principle has been a subject of intense philosophical debate. An extensive literature has been thus generated both with arguments in favor and against the adequacy of the principle.

A remarkable argument against the principle, first proposed by Elliott Sober (Sober, 1987, 2001), consists on a counterexample which involves correlations between bread prices in Britain and sea levels in Venice. The following quote somehow summarises the spirit of the example:

*Universidad Complutense de Madrid. (`inaki.sanpedro@filos.ucm.es`).

> Because both quantities have increased steadily with time, it is true that higher than average sea levels tend to be associated with higher than average bread prices. [...] we do not feel driven to explain this correlation by postulating a common cause. Rather, we regard VSL and BBP as both increasing for endogenous reasons. [...] Here, postulating a common cause is simply not very plausible, given the rest of what we believe.(Sober, 1987, 2001)

There have been different attempts to deal with examples of the kind of Sober's 'Venetian sea levels and British bread prices' —two of which I shall review in what follow. It seems striking though that none of these make use of recent development regarding the formal structure of the probability spaces where the relevant events are defined, the corresponding correlations among these and the common causes that might be involved in a potentially adequate causal explanation of such correlations. In particular, Hofer-Szabó et al. have shown formally that screening-off events exist for any correlation:

> [...] *every* classical probability space $(\mathcal{S}, \mu)$ is common cause completable with respect to *any finite* set of correlations [...] given *any finte* set of correlations in a classical event structure, one can *always say* that the correlations are due to some common causes, possibly 'hidden' ones, i.e. ones that are not part of the initial set $\mathcal{S}$ of events.(Hofer-Szabó et al., 1999)

Both quotations above seem to contain opposite claims and appear then to convey two quite incompatible views. The aim of this paper is to put into perspective criticisms to RPCC of the kind of Sober's 'Venetian sea levels and British bread prices' in the light of such recent formal results. In other words, I shall make use of the so-called *extendability* and *completability* theorems to reassess whether examples such as Sober's still constitute a thread to RPCC.

The structure of the paper is as follows: Section 2 provides a very brief review of the main ideas of Reichenbach's Principle Common Cause (RPCC) along with the main issues regarding its significance and philosophical status. The ideas of correlation and Reichenbachian common cause will be central here. Section 3 introduces the ideas of *extendability* and common cause *completability*. In Section 4 I review the mains of Sober's argument against RPCC and discuss on what grounds it may be taken as a potential thread to the principle. Two recent reactions to Sober's example are reviewed here as well in Section 5. In Sections 6 and 7 I suggest two alternative solutions to save RPCC from criticism of the kind of Sober's, along with some concluding remarks.

2 Reichenbachian Common Causes

The idea of common cause goes back to Reichenbach and has its origins at the observation of apparently unrelated events that nonetheless take place simultaneously with a certain regularity:[1]

> *If an improbable coincidence has occurred, there must exist a common cause.*

Reicehnbach's original proposal may be formalised so that the concept of common cause can be analysed in detail within the framework of classical probability theory.[2] The notion of correlation is central, for it is assumed to capture the Reichenbach's intuitions about *improbable coincidences*. Let us then define correlation as follows:[3]

Definition 1. *Let $(\mathcal{S}, p)$ be a classical probability measure space with Boolean algebra $\mathcal{S}$ representing the set of random events and with the probability measure p defined on $\mathcal{S}$. If $A,\ B \in \mathcal{S}$ are such that*

$$p(A \wedge B) - p(A) \cdot p(B) > 0, \tag{1}$$

then the events A and B are said to be (positively) correlated, and we write $Corr_p(A, B)$.

Within this formal framework a *Reichenbachian* common cause[4] is then defined as:

Definition 2 (Reichenbachian Common Cause)**.** *An event C is said to be a Reichenbachian common cause if the following independent conditions hold:*

$$p(A \wedge B|C) = p(A|C) \cdot p(B|C) \tag{2}$$

$$p(A \wedge B|\neg C) = p(A|\neg C) \cdot p(B|\neg C) \tag{3}$$

$$p(A|C) > p(A|\neg C) \tag{4}$$

$$p(B|C) > p(B|\neg C), \tag{5}$$

where $p(A|B) = p(A \wedge B)/p(B)$ denotes the probability of A conditional on B and it is assumed that none of the probabilities $p(X)$ $(X = A, B, C, \neg C)$ is equal to zero.

[1] Cf. (Reichenbach, 1956, p. 157). The simultaneity requirement is crucial for Reichenbach to rule out direct causal explanations and thus suggest the existence of a common cause.

[2] I follow here the work by Hofer-Szabó et al. in late 1990's and early 2000. See (Hofer-Szabó et al., 1999, 2000a,b) and (Rédei, 2002) for the main results of the program.

[3] This definition is of positive correlation. A completely symmetrical definition may be given for negative correlations. Distinguishing between positive and negative correlations will not be important in what follows and positive correlations will thus be assumed.

[4] The term *Reichenbachian* is useful if we are to distinguish this particular type of common causes from other events which could in principle be considered to be common cause as well but do not fulfil this definition.

With both the notions of correlation and Reichenbachian common cause above at hand Reichenbach's Principle of the Common Cause (RPCC) may be restated as follows:

Definition 3 (RPCC)**.** *For any two (positively) correlated event types A and B ($Corr_p(A,B) > 0$), if A is not a cause of B and neither B is a cause of A, there exists a Reichenbachian common cause C of A and B, i.e. there exist an event C such that conditions* (2)-(5) *hold.*

The definition above consists of two distinct and independent claims. The first is a claim at the ontological level, regarding the existence of common causes, and the other at the methodological level which provides a concrete characterisation of the postulated common causes (through equations (2)-(5)).

A proper distinction of these two claims is crucial for the assessment of the status of RPCC. In particular, since each of the two claims in RPCC is logically independent of the other, arguments aimed at criticising the characterisation of common causes through expressions (2)-(5) may very well leave untouched the existential claim about common causes. In fact, while it is part of the received view that equations (2)-(5) do not constitute neither a necessary nor a sufficient condition for common causes, there have been prominent defenders of the existence of common causes.[5] On the other hand, arguments devised to deny the very existence of common causes may be completely compatible with the claim that common causes, whenever they exist, are to satisfy equations (2)-(5).

Despite the controversies, endorsing RPCC may be motivated by at least two reasons. First, note that for Reichenbach the role of the principle as a whole, and of the screening-off condition in particular, is mainly explanatory. Reichenbach explicitly points out that the four statistical relations *explain* the correlations between A and B in two senses. First, he notes that the four relations entail that A an B are (positively) correlated, i.e. $Corr(A,B) > 0$. On the other hand, a common cause C satisfying these four relations explains the correlation by rendering A and B statistically independent.[6] The explanatory power of screening-off common causes may thus be taken as a good reason to support the adequacy of RPCC for the inference of causal relations from probabilistic facts, even if it can be patently shown no to hold as a necessary nor as a sufficient condition for causation.

Second, recent results show that, at least formally, it is always possible to provide a Reichenbachian common cause for any given correlation.[7] These

[5]Salmon (1984) and Cartwright (1987) are perhaps the most influential proposals for the generalisations of Reichenbach's original criterion for common causes.

[6]Cf. (Reichenbach, 1956, p. 159).

[7]While this is not true in general for any sort of Reichenbachian common cause, it holds for the particular type of common cause we are interested here. See (Hofer-Szabó et al., 1999, 2000b,a) for details.

results build on the intuition that any probability space $\mathcal{S}$ containing a set of correlations and which does not include (Reichenbachian) common causes of these, may be extended in such a way that the new probability space $\mathcal{S}'$ does include (Reichenbachian) common causes for each of the original correlations. These intuitions are formalised in so-called *extendability* and *common cause completability* theorems.

3 Common Cause Completability

It is not difficult to find examples of probability spaces containing correlated events which do not feature however any event which satisfies the definition of Reichenbachian common cause. We shall call such probability spaces *Reichenbachian common cause incomplete* spaces.[8]

One seems to have two alternatives when dealing with Reichenbachian common cause incomplete probability spaces. Either we go for a weakening of the common cause criterion —this is for instance the case in both Salmon's 'interactive forks' and Cartwright's generalisation of the fork criterion—, or we may simply embark on the search for screening-off common causes, hoping that such events exist but have remained somehow 'hidden' to us all along. Here I shall only pay attention to the second alternative.

So how should we face this search? Note first that we need to be searching new events (i.e. the Reichenbachian common causes) which are not contained in the original (incomplete) probability space $(\mathcal{S}, p)$. Intuitively, we need some notion of extension than could be applied to our probability space. This is formally achieved as follows:[9]

Definition 4 (Extension)**.** *The probability space $(\mathcal{S}', p')$ is called an extension of $(\mathcal{S}, p)$ if there exist a Boolean algebra embedding h of $\mathcal{S}$ into $\mathcal{S}'$ such that $p(X) = p'(h(X))$, for all $X \in \mathcal{S}$.*

Extendability allows then for the enlargement of the original probability space so that new events are included. In a second step, we should be able to set up a procedure to enlarge our common cause incomplete probability space such that the new extended probability space contains common causes for the original correlations.[10] This intuition is formalised through the idea of Reichenbachian common cause completability:

[8] Reichenbachian common cause incomplete probability spaces are quite usually found when describing usual experimental data. Indeed, most examples aimed to rule out screening-off as a necessary condition for common causes exploit such incompleteness. Further more formal examples of such common cause incomplete spaces are provided in (Hofer-Szabó et al., 2000a).

[9] Cf. (Hofer-Szabó et al., 2000a).

[10] Indeed, Definition 4 ensures that the extension operation be consistent with the old event structure $(\mathcal{S}, p)$. In particular, correlations stay invariant under the extension operation, that is $Corr(A, B) \equiv Corr_p(A, B) \equiv Corr_{p'}(A, B)$. Without going into details, such consistency is achieved by defining the embedding h such that the initial probabilities

Definition 5 (RCC Completability). *Let $Corr(A_i, B_i) > 0$ ($i = 1, 2, \ldots, n$) be a set of correlations in $(\mathcal{S}, p)$ such that none of them possess a common cause in $(\mathcal{S}, p)$. The probability space $(\mathcal{S}, p)$ is called Reichenbachian common cause completable with respect to the set $Corr(A_i, B_i)$ if there exists an extension $(\mathcal{S}', p')$ of $(\mathcal{S}, p)$ such that $Corr(A_i, B_i)$ has a Reichenbachian common cause C_i in $(\mathcal{S}', p')$ for every $i = 1, 2, \ldots, n$.*

Completability is hence the key for successfully searching Reichenbachian common causes. The question is now whether any incomplete probability space $(\mathcal{S}, p)$ can be extended such that it is (Reichenbachian) common cause completable. We may ask, in other words, when is a probability space $(\mathcal{S}, p)$ Reichenbachian common cause completable?

Hofer-Szabó et al. answer this question with the following proposition:[11]

Proposition 1. *Every classical probability space $(\mathcal{S}, p)$ is common cause completable with respect to any* finite *set of correlated events.*

The proposition shows, in other words, that given a Reichenbachian common cause incomplete probability space an extension $(\mathcal{S}', p')$ may *always* be performed such that it contains (Reichenbachian) common causes for all the original correlations.

Common cause completability hence constitutes a very powerful tool if we are to provide common cause explanations of generic correlations. It nevertheless faces its own problems, especially when it comes to the physical interpretation of either the enlarged probability space $\mathcal{S}'$ or the new common causes contained in it. In particular, it seems a fair criticism to the whole program to claim that common cause completability is a merely a formal device, which is likely to lack physical meaning in many (perhaps too many) cases. I shall retake the issue later on and just point out for now that such criticisms may be avoided.[12] At any rate, the notion of completability as it stands seems to provide further methodological grounds for RPCC to be applicable.

4 Venetian Sea Levels and British Bread Prices

Examples such as Sober's 'Venetian Sea Levels and British Bread Prices' (VSL & BBP) are devised to refute Reichenbach's Common Cause Principle (RCCP) by criticising the metaphysical content of it. Sober however draws

---

and correlations are maintained under the new probability measure p'. See (Hofer-Szabó et al., 2000a) for details.

[11]Cf. (Hofer-Szabó et al., 1999, p. 384).

[12]I point the reader to (San Pedro and Suárez, 2009) for a recent assessment of the significance of common cause completability, the possible criticisms to it and possible strategies to avoid them.

some methodological consequences from the existence of such correlations — e.g. between sea levels in Venice and bread prices in Britain— which cannot be accounted for in terms of common causes. This, he argues, favors the use of the Likelihood Principle instead of RPCC. Although the intuitions behind Sober's argument are in my view quite strong, the conclusions seem to clash with the results I just reviewed concerning common cause completability. I shall thus try to clarify where the incompatibility lies and what are the possible alternative ways to take. Let us first start with the details Sober's example itself.

It is the case that the sea level in Venice (VSL) and the cost of bread in Britain (BBP) have been (monotonically) increasing during a given period of time. Table 1 displays such trends in the values of Venetian sea levels and British bread prices in accordance with Sober's actual example. From the data displayed, we are told that 'higher than average values' of Venetian sea levels and those of British bread prices are correlated:[13]

> As I claimed initially, higher than average bread prices *are* correlated with higher than average sea levels.

Let us denote the event 'the Venetian sea level in year i is higher than average' by the expression '$\mathrm{VSL}_i > \langle \mathrm{VSL} \rangle$'. (Similarly for bread prices in Britain.)

What Sober seems to have in mind when claiming that 'absolute values' of VSL and BBP are correlated is the following. On the one hand, the probability of observing a 'higher than average' Venetian sea level in year i can be seen (directly from the data displayed in Table 1) to be

$$p\,(\mathrm{VSL}_i > \langle \mathrm{VLS} \rangle) = 1/2.$$

Similarly, for British bread prices one has that

$$p\,(\mathrm{BBP}_i > \langle \mathrm{BBP} \rangle) = 1/2.$$

On the other hand, one may also calculate the joint probability of both:

$$p\,[(\mathrm{VSL}_i > \langle \mathrm{VSL} \rangle) \wedge (\mathrm{BBP}_i > \langle \mathrm{BBP} \rangle)] = 1/2.$$

These three probabilities entail that:

$$p\,[(\mathrm{VSL}_i > \langle \mathrm{VSL} \rangle) \wedge (\mathrm{BBP}_i > \langle \mathrm{BBP} \rangle)] \\ -p\,(\mathrm{VSL}_i > \langle \mathrm{VLS} \rangle) \cdot p\,(\mathrm{BBP}_i > \langle \mathrm{BBP} \rangle) > 0. \qquad (6)$$

[13] Cf. (Sober, 2001, p. 334). The appeal to 'higher than average values' rather than just 'values' is mainly motivated by some critiques to an earlier version of the counterexample (Sober, 1987). We do not need to review such arguments here since they will not play any important role in the foregoing discussion. The important point is that Sober's later formulation of the counterexample —involving 'higher than average values'— stands. Sober also refers to 'higher than average values' as 'absolute values' and I shall use these two expressions indistinctly.

Venetian Sea Levels and British Bread Prices									
Year (i)	1	2	3	4	5	6	7	8	$\langle \text{Year} \rangle = 4.5$
VSL	22	23	24	25	28	29	30	31	$\langle \text{VSL} \rangle = 26.5$
BBP	4	5	6	10	14	15	19	20	$\langle \text{BBP} \rangle = 11.625$

Table 1: Sober's Venetian sea levels and British bread prices data (Sober, 2001, p. 334).

Thus, the argument goes on, 'higher than average values' of Venetian sea levels and British bread prices are in fact (positively) correlated.

The question is then how this correlation is to be explained away. Sober points out that there are three possible ways to go:[14]

(i) To postulate the existence of an unobserved common cause.

(ii) To claim that the data sample is unrepresentative.

(iii) To claim that the data arises from a mixing of populations with different causal structures and correspondingly different probability distributions.

Considering these three options in turn shows, according to Sober, that RPCC fails. The argument is as follows. First, Sober dismisses option (ii) by pointing out that the correlations in his example do not come out of an unrepresentative sample since data could be spread over a larger period of time and the correlations would still be there —I completely agree with this and I will also dismiss option (ii) altogether. Second, option (i) is false in the example *ex hypothesis*. Consequently, Sober takes option (iii) to provide the right (causal) explanation of the correlation. This therefore constitutes a failure of RPCC.

5 Is VSL & BBP a Genuine Counterexample?

But does Sober's VSL & BBP really constitute a genuine counterexample to RPCC? In order to answer this particular question we need to address two further questions, I think. First, we need to know whether the VSL & BBP correlation is indeed genuine (as defined formally in Section 2). Second, in case it is so, we need to ask whether the VSL & BBP correlation really

[14] These three possible explanations had been already suggested by Meek and Glymour after Yule (1926). See also (Sober, 2001, p. 332) and references therein.

cannot be explained away in terms of a common cause. In other words, if the counterexample is to stand, we need to make sure that no common cause at all may be provided for the correlation. I will tackle these two questions in turn

Kevin Hoover and Daniel Steel have both recently tried to diffuse Sober's example at some level, each with quite different arguments and thus reaching different conclusions. The main issue these two reactions differ in is in fact the answer to our first question, i.e. whether the VSL & BBP correlations are indeed genuine correlations or mere associations of the sample.

As we have seen, the probabilities of 'higher than average values' of Venetian sea levels and British bread prices display what seems to be a probabilistic dependence —by means of expression (6) on page 7. However, if we are to conclude that 'higher than average values' of sea levels and bread prices are correlated, we need first make sure that the probabilities involved refer to the same and only probability space.[15] In other words we need to check that 'absolute values' of sea levels and bread prices are events of the very same probability space.

But nothing in the data set tells us the probability measure should be the same. In fact, the probabilities for each quantity are derived quite independently (from the relative frequencies of the corresponding measured sea levels and bread prices over a time span). Strictly speaking we should perhaps have initially written them as $p^1(\mathrm{VSL}_i > \langle\mathrm{VSL}\rangle)$ and $p^2(\mathrm{BBP}_i > \langle\mathrm{BBP}\rangle)$, i.e. as referring to two different probability measures p^1 and p^2. Similarly for the joint probability we should perhaps have written it relative to yet another probability measure $p^3\left[(\mathrm{VSL}_i > \langle\mathrm{VSL}\rangle) \wedge (\mathrm{BBP}_i > \langle\mathrm{BBP}\rangle)\right]$. With these assumptions, relation (6) becomes then:

$$p^3\left[(\mathrm{VSL}_i > \langle\mathrm{VSL}\rangle) \wedge (\mathrm{BBP}_i > \langle\mathrm{BBP}\rangle)\right] \\ -p^1(\mathrm{VSL}_i > \langle\mathrm{VSL}\rangle) \cdot p^2(\mathrm{BBP}_i > \langle\mathrm{BBP}\rangle) > 0. \qquad (7)$$

Thus, the question whether the expression above reflects a correlation between sea levels and bread prices may be now restated in terms of these three different probability measures. That is, are p^1, p^2 and p^3 in fact one and the same probability measure?

These remarks are somehow related to Hoover's[16] arguments in reaction to the VSL & BBP case. Hoover distinguishes correlations from mere associations of the sample. Very succinctly, while associations are a property of the sample, correlations are a property of the probabilistic space used to model it. Hoover assumes that it is only correlations that can reveal 'real' properties of the system. In our case then, if the probability measures p^1, p^2 and p^3 would be different, expression (7) could only be said to reflect some

[15]Note that this is explicitly required in the formal definition of correlation (Definition 1).

[16]Cf. (Hoover, 2003).

degree of association between sea levels and bread prices, but not a correlation. In order for it to represent a correlation a consistent probabilistic model —with a single probabilistic measure, that is— must be constructed such that the example's data may be embedded in it. This is not the case in the case at hand, in Hoover's view. He thus concludes that the VSL & BBP scenario does not constitute a counterexample to the RPCC.

Hoover's case might find support in Sober himself. For, as it is claimed in the original argument, each data series belongs to different causal structures —this was option (iii) in Sober's argument. Which can then be seen to justify the claim that the two probability spaces need to be different. But is it right to claim that the 'Venice-Britain' scenario cannot really be described in a whole single probability space such that the corresponding values or sea levels and bread prices are correlated? In other words, why could it not be the case that data in Table 1 give rise to correlations? I think that this is indeed an option. In particular, while I share Hoover's view regarding the difference between mere associations and genuine correlations, I do not see why the data in Table 1 may not be embedded, or modelled if we like, in a single probability space.

In fact, an argument along these lines is provided by Steel[17]. Steel's model takes advantage of a well known mathematical result so-called the 'mixing theorem'. In brief, the 'mixing theorem' provides us with information about the behaviour of the probability distribution resulting from the mixing of the distributions from two populations, each of which with probabilistically independent variables. The theorem tells us, in particular, under what conditions the variables of such 'mixed' probability distribution are probabilistically independent. The theorem then shows that a probability distribution may display dependencies just because it is the result of the mixing of two other probability distributions. Steel claims this is the case in Sober's 'Venice-Britain' example, and constructs a model from two initial sets of data (of both VSL and BBP), each corresponding to different (distant) time spans. If the probability distributions from these two populations are mixed, the resulting distribution displays probabilistic dependencies in just the manner suggested by Sober.[18]

Summing up, I see no convincing reason why correlations such as those in the 'Venice-Britain' example would not be *genuine*. Sober's kind of examples may then very well be counterexamples to RPCC. Once the question as to whether VSL & BBP are (genuinely) correlated has been positively answered, we shall turn to our second question. Is it really the case that no common cause explanation can be given of the correlation between sea levels in Venice and bread prices in Britain?

When answering this question, we seem to have three possible outcome

[17]Cf. (Steel, 2003).

[18]See (Steel, 2003) and references therein for further details.

scenarios. First, we may find out that no common cause whatsoever may be provided that explains the correlation. In that case VSL & BBP kind of example would, as claimed by Sober, constitute a genuine counterexample to RPCC. Alternatively, we might find out, by making use of the common cause completability results, that it is indeed possible to provide a common cause explanation of the correlation. Finally, Sober's criticism could also be avoided if we could show the question to be non-applicable. For instance, even if we take the VSL & BBP correlation to be genuine we might want to argue that it needs no causal explanation after all, perhaps due to the fact that it does not reflect any feature of the system itself. Clearly, the first option only seems to make sense once the other two have been ruled out. Let us then consider the two last options in turn.

6 Screening-off Events Exist

Recall that, by common cause completability, an appropriate extension of a common cause incomplete probability space guarantees that there exist a screening-off event. For what it has been said up to now, there is no reason why common cause completability should not also work in our case at hand. A quite obvious such extension would include 'time events'.

Suppose for instance our new model, i.e. our extension, contains events of the type '$\mathrm{Y}_i > \langle\mathrm{Year}\rangle$', which we may call, following Sober's terminology, 'higher than average time values', 'higher than average values of years', or 'absolute values of years', never mind how strange this may sound. We may then assign probabilities to such events in exactly the same way as we did for 'higher than average' values of sea levels and bread prices, that is by referring to their relative frequencies. (Only, we need to make sure that the probability measure is the same for all three values.) Thus, we might write, again from the data on Table 1,

$$p\,(\mathrm{Y}_i > \langle\mathrm{Year}\rangle) = 1/2.$$

If we now take conditional probabilities we obtain, also looking at the data in Table 1,

$$p\,(\mathrm{VSL}_i > \langle\mathrm{VSL}\rangle \mid \mathrm{Y}_i > \langle\mathrm{Year}\rangle) = 1, \tag{8}$$

$$p\,(\mathrm{BBP}_i > \langle\mathrm{BBP}\rangle \mid \mathrm{Y}_i > \langle\mathrm{Year}\rangle) = 1. \tag{9}$$

It is also easy to check that

$$p\,[(\mathrm{VSL}_i > \langle\mathrm{VSL}\rangle) \wedge (\mathrm{BBP}_i > \langle\mathrm{BBP}\rangle) \mid \mathrm{Y}_i > \langle\mathrm{Year}\rangle] = 1. \tag{10}$$

It becomes now clear that as soon as we consider the event '$\mathrm{Y}_i > \langle\mathrm{Year}\rangle$' the correlation will vanish. This is because the dependence of the original series washes out conditional on '$\mathrm{Y}_i > \langle\mathrm{Year}\rangle$'. In particular if we define a new probability measure $p^Y = p(\cdot \mid \mathrm{Y}_i > \langle\mathrm{Year}\rangle)$, the above equations yield

$$p^Y\left[(\mathrm{VSL}_i > \langle\mathrm{VSL}\rangle) \wedge (\mathrm{BBP}_i > \langle\mathrm{BBP}\rangle)\right] \\ - p^Y(\mathrm{VSL}_i > \langle\mathrm{VSL}\rangle) \cdot p^Y(\mathrm{BBP}_i > \langle\mathrm{BBP}\rangle) = 0. \quad (11)$$

This example is of course specific for the case at hand, and the 'trick' has been quite the obvious one, since I transformed both the original non-stationary data series into stationary ones by finding the probability space in which all probabilities are one. In this case the probability space with such 'nice' properties is particularly easy (and obvious) to find since the time dependence of both sea levels and bread prices' higher than average values is exactly the same. However obvious and specific this example might be, I hope it illustrates sufficiently how a screening-off event may be provided.

This is not quite a causal explanation yet, since the event '$\mathrm{Y}_i > \langle\mathrm{Year}\rangle$' does not seem capable of a causal interpretation in an obvious or straightforward way. The question is more specifically whether we can make sense of events '$\mathrm{Y}_i > \langle\mathrm{Year}\rangle$' as (common) causes. I must admit that I do not have an answer to this question. For, in what relevant sense is time a causal factor in the VSL & BBP example? Indeed, I find it hard to understand '$\mathrm{Y}_i > \langle\mathrm{Year}\rangle$' as a cause event.

Of course, such problematic issues will be present whenever a time dependent event is to be interpreted as a cause but the point of the example above it to show that it there is actually good methodological ground to suggest that screening-off common causes may be provided for any given correlation, included those in the 'Venice-Britain' example. Again, it is true that we might face problems of interpretation once a screening-off event has been identified after a particular extension. But perhaps further conceptual innovations may at some stage provide an adequate framework so as to be able to interpret such time dependent events as (common) causes. On the other hand, it should be noted that the extension of a common cause incomplete probability space is not unique. Thus, even in the case the problems of interpretation proved too severe, common cause completability, it seems to me, still leaves the door open for other, perhaps more easily interpretable, screening-off events to be provided by extending (and completing) in other different ways the original probability space.

7 Purely Formal Correlations

The problems with the interpretation of some screening-off events may suggest that the physical significance of both such screening-off events and the corresponding correlations we are trying to explain is at least dubious.

Indeed, going back to our case at hand, a more thorough analysis of the 'absolute year values' events suggests that the VSL & BBP correlation arises solely due to the shared time dependence of the systems evolutions. Put it the other way around, we may ask, more specifically, what does the correlation between sea levels and bread prices really say, if anything at all, about the level of the sea in Venice and the price of the bread in Britain?

A closer look to the kind of events we are dealing with, i.e. 'higher than average values', reveals that these are defined relative to the average of the corresponding quantity over a certain period of time. Also, note that the values carry a label that identifies a specific instance of time. It is thus clear that the correlated events in the 'Venice-Britain' example have some sort of time dependence.

Time dependent data are also commonly known as non-stationary data. It is also well known that non-stationary data display dependencies that do not always reflect the system's inner structure. For instance, Steel[19] points out that it is a consequence of the so-called 'mixing theorem' of probability theory that two sets of non-stationary data display probabilistic dependencies even if each of them refers to a completely different historical period.[20] Also as a consequence of the 'mixing theorem', if the probabilistic dependence of two data series is due to them being non-stationary the correlation will vanish as soon as we describe the data in a probabilistic model in which one of them is no longer non-stationary. This is in fact what happens with Sober's Venetian sea levels and British bread prices, as we have seen in the model-example above.

What the above suggests thus is that sea levels and bread prices in the VSL & BBL case are only correlated in virtue of telling us something about time. This in turn suggests that correlations that arise due to the non-stationary properties of the data do not provide any information whatsoever about the underlying (physical) structure of the system, if there is any system we can speak of. To the contrary, they seem to be a case of what we can call *purely formal* correlations, i.e. correlations that arise solely as a product of formally modeling experimental data. These considerations are very much along the lines of the argument in Steel (2003). Steel concludes that, although RPCC cannot be applied to non-stationary data series it constitutes a genuine counterexample is genuine. For there is a genuine correlation which is not to be explained in terms of a common cause.

Steel's conclusion however, does not seem to me it fits with the spirit of the Reichenbach's original ideas on common causes. For, although, as I said Reichenbach's notion of 'improbable coincidence' is to be captured by the idea of correlation, I take it that not all correlations can be said to reflect 'improbable coincidences'. In particular, *purely formal* correlations —an example of which are those in the 'Venice-Britain' scenario— do not seem to be 'improbable' in any sense. Quite the contrary, the structure

[19]Cf. (Steel, 2003).

[20]See (Steel, 2003) for further details and references.

of the model formally *entails* such coincidences at any rate. In this view thus, correlations that arise purely formally from the model structure are not required an explanation of any kind, be it causal or not, and Sober's example therefore would not constitute a genuine counterexample to RPCC either.

Bibliography

Cartwright, N. (1987). How to tell a common cause: Generalizations of the conjunctive fork criterion. In J. H. Fetzer (Ed.), *Probability and Causality*, pp. 181–188. Reidel Pub. Co.

Hofer-Szabó, G., M. Rédei, and L. E. Szabó (1999). On reichenbach's common cause principle and reichenbach's notion of common cause. *The British Journal for the Philsophy of Science 50*, 377–399.

Hofer-Szabó, G., M. Rédei, and L. E. Szabó (2000a). Common cause completability of classical and quantum probability spaces. *International Journal of Theoretical Physics 39*, 913–919.

Hofer-Szabó, G., M. Rédei, and L. E. Szabó (2000b). Reichenbach's common cause principle: Recent results and open questions. *Reports on Philosophy 20*, 85–107.

Hoover, K. D. (2003). Nonstataionary time series, cointegration, and the principle of the common cause. *The British Journal for the Philosophy of Science 54*, 527–551.

Rédei, M. (2002). Reichenbach's common cause principle and quantum correlations. In T. Placek and J. Butterfield (Eds.), *Non-locality, Modality and Bell's Theorem*, pp. 259–270. Kluwer.

Reichenbach, H. (1956). *The Direction of Time*; Edited by Maria Reichenbach. Unabridged Dover (1999) republication of the original University of California Press (1956) publication.

Salmon, W. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton University Press.

San Pedro, I. and M. Suárez (2009). Reichenbach's principle of the common cause and indeterminism: A review. In J. G. Recio (Ed.), *Philosophical essays in physics and biology*. Georg Olms.

Sober, E. (1987). The principle of the common cause. In J. H. Fetzer (Ed.), *Probability and Causality: Essays in Honor of Wesley Salmon*, pp. 211–228. Reidel.

Sober, E. (2001). Venetian sea levels, british bread prices, and the principle of the common cause. *The British Journal for the Philosophy of Science 52*, 331–346.

Steel, D. (2003). Making time stand still: A response to sober's counter-example to the principle of the common cause. *The British Journal for the Philosophy of Science 54*, 309–317.

Yule, G. U. (1926). Why do we sometimes get nonsensical relations between time series? a study of sampling and the nature of time series. *Journal of the Royal Statistical Society 89*, 1–64.

end