

Ceteris Paribus Laws and *Minutis Rectis* Laws

LUKE FENTON-GLYNN

University College London

Special science generalizations admit of exceptions. Among the class of non-exceptionless special science generalizations, I distinguish (what I will call) *minutis rectis* (*mr*) generalizations from the more familiar category of *ceteris paribus* (*cp*) generalizations. I argue that the challenges involved in showing that *mr* generalizations can play the law role are underappreciated, and quite different from those involved in showing that *cp* generalizations can do so. I outline a strategy for meeting the challenges posed by *mr* generalizations.

1. Introduction

Many philosophers of science talk as though, if there are any non-exceptionless scientific generalizations that play at least some aspects of the law role (counterfactual support, inductive confirmation, predictive/explanatory import) tolerably well, then they belong to the class of *ceteris paribus* (*cp*) generalizations or *cp* laws. The following are representative quotations:

“A nonstrict law is a generalization that contains a *ceteris paribus* qualifier that specifies that the law holds under ‘normal or ideal conditions’, The generalizations one finds in the special sciences are mostly of this kind. In contrast, a strict law is one that contains no *ceteris paribus* qualifiers; it is exceptionless not just *de facto* but as a matter of law.” (Lepore and Loewer 1987, 632)

“*cp* lawfulness is just a species of nomological necessity, the other species of nomological necessity being strict lawfulness. What distinguishes the two species is just that *cp* laws can have . . . exceptions and strict laws can’t.” (Fodor 1991, 31–32)

“A convenient way to reconstruct the debate on [whether non-universal generalizations may constitute laws] is to understand non-universal laws as statements that are qualified by a *ceteris paribus* . . . clause.” (Reutlinger *et al.* 2011)

The copyright line for this article was changed on 8 April 2016 after original online publication

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

In each of these quotations, the notion of a non-exceptionless law is run together with that of a cp law. Several commentators on the philosophical literature on scientific generalizations indicate that this running-together is widespread:

“Philosophers of the last decades have repeatedly pointed out that most law statements, especially those in the non-physical sciences, do not express strict (i.e., universal and exceptionless) laws. Rather, they express so-called *ceteris paribus* laws.” (Schurz 2002, 351)

“Many philosophers of science think that most laws of nature ... are so called *ceteris paribus* laws, i.e., roughly speaking, laws with exceptions.” (Schrenk 2007, 221)

“Special science laws ... are usually taken to ‘have exceptions’, to be ‘non-universal’ or ‘to be *ceteris paribus* laws’.” (Reutlinger *et al.* 2011)

“[F]ew of [the generalizations of the special sciences] seem to be exceptionless ... [T]he most common strategy has been to ... search for a way of construing the generalizations of the special sciences as ‘laws’, despite appearances to the contrary. It is this strategy which motivates the *ceteris paribus* laws literature. The idea is that the generalizations of the special sciences, despite failing to state (at least explicitly) nomically sufficient conditions for outcomes, nonetheless (at least sometimes or if appropriate conditions are met) can be regarded as a special kind of law—a *ceteris paribus* law.” (Woodward 2002, 303–304)

The identification of non-exceptionless laws with cp laws can hardly be a matter of stipulative definition. The notion of a cp law is richer than the bare notion of *a law that admits of exceptions*. For one thing, the notion of a cp law is associated with a distinctive account of how exceptions arise.

Specifically, a cp law is supposed to be endowed with an implicit or explicit clause that specifies that it holds ‘other things being equal’, where this latter notion is usually construed in terms of the obtaining of ‘normal’ or even ‘ideal’ conditions (see Cartwright 1983, 46; Schurz 2002), and further explicated in terms of the absence of significant difference-making interference from outside the systems that the law in question seeks to characterize (see, e.g., Fodor 1989, 69n; Schurz 2002, 366–370). Exceptions are taken to arise due to the non-satisfaction of this cp clause.¹

After outlining the notion of a cp law (Section 2), I argue (Section 3) that it is a mistake to equate non-exceptionless laws with cp laws: there is a distinct type of non-exceptionless law—which I will call a *minutis rectis* (*mr*) law—which admits of exceptions that aren’t explained by the non-satisfaction of a cp clause. I argue (Section 4) that *mr* laws pose a distinctive set of philosophical challenges. I outline (Section 5) three potential responses to these challenges. Finally (Section 6 and

¹ I’m skating over some differences between various characterizations of cp laws that appear in the literature (for an overview, see Reutlinger *et al.* 2011). Some (e.g. Schurz 2002) distinguish various *types* of cp law. Indeed Schurz (2014) makes a distinction between cp (other things being equal) laws and (what he calls) *ceteris rectis* (other things being right) laws. My decision to call the category of generalization that I will discuss below *minutis rectis* (the details being right!) laws was partly inspired by Schurz’s terminology. But, as we will see, a *minutis rectis* law is neither a type of cp law, nor is it—as a *ceteris rectis* law is—a closely related species of non-strict law. Nevertheless, Schurz and I clearly agree that different varieties of non-exceptionless generalization ought to be carefully distinguished because they pose different challenges for those seeking to show that they are able to play (various aspects of) the law role.

Section 7), I proceed to develop the response that strikes me as the most promising: namely, that *mr* laws are approximations to strict(er) probabilistic laws.

A terminological point: talk of *cp* (and *mr*) ‘laws’ is rather awkward in the context of a discussion of whether, and to what extent, the exception-ridden generalizations of the special sciences play various aspects of the law-role. I’m somewhat sympathetic to the objections that some philosophers (e.g. Woodward 2005, Woodward and Hitchcock 2003a,b) have to such law-speak. Nevertheless, because law-speak is so common in the literature, I shall not try to forgo it in what follows, and I shall drop the jarring scare-quotes when I use it.

A (related) quasi-terminological point: Woodward (2005) and Woodward and Hitchcock (2003a,b) prefer to call those scientific generalizations that, despite being non-exceptionless, play some aspects of the law role reasonably well ‘invariant generalizations’ rather than ‘*cp* laws’ or ‘*cp* generalizations’. As shall be seen below, their characterization of invariant generalizations is somewhat different from standard characterizations of *cp* generalizations. Nevertheless, as shall also be seen below, there is at least a major overlap between those generalizations that they take to be invariant and the generalizations that other philosophers take to be *cp* generalizations or *cp* laws (compare Woodward and Hitchcock 2003a, 3; Reutlinger *et al.* 2011). We’ll see that part of the challenge posed by *mr* generalizations is precisely that they seem *not* to be invariant, in the sense of Woodward and Hitchcock.

2. Ceteris Paribus Laws

In ecology, one standard equation used for predicting population growth is the Logistic Equation (LE):

$$\frac{dn}{dt} = r_c n \left(1 - \frac{n}{K}\right) \quad (LE)$$

Here n is the number of individuals in the population, $\frac{dn}{dt}$ is the growth rate of the population (the change in n , with respect to time t), r_c is the *intrinsic per capita growth rate* of the population—that is, the per capita growth rate that obtains in the absence of intra-population competition for resources—and K is the *carrying capacity*—that is, the maximum sustainable population size given resource constraints.

LE implies that when the number n of members of a species in a particular habitat is small relative to the carrying capacity—so that there is little intra-population competition for resources—the actual population growth rate $\frac{dn}{dt}$ is close to the intrinsic per capita growth rate r_c multiplied by the number n of individuals. But, as the population grows, the actual per capita growth rate declines linearly—due to increasing competition. This decline continues until the carrying capacity K is reached, at which point population growth is 0.

Ecological generalizations—such as LE—appear to play certain aspects of the law role to a non-negligible degree. Ecologists apply LE to certain populations—especially

populations that aren't subject to significant *inter*-population competition or predation—in order to make predictions, and to give explanations.²

LE holds only *ceteris paribus* because there are possible background conditions under which it is violated (even when applied to populations concerning which, in normal circumstances, it is predictively accurate). For example, it will not hold in the event of the population being subjected to a cull, or in the event of a natural disaster that destroys (a large part of) the population. While LE may give accurate predictions about population growth *after* some such events, it won't accurately predict growth *during* such episodes. It simply doesn't include variables that represent such events.

Events like culls produce circumstances in which other things are *not* equal: interfering factors are present, so LE doesn't even approximately hold. An ecologist presumably wouldn't seek to model such factors, since they are not *ecological* factors. They interfere with the sorts of system that the ecologist seeks to model (*viz.* *ecosystems*), but come from 'outside' such systems. Perhaps this means that ecological generalizations will in principle remain *cp* generalizations (compare Davidson 1970, 94; Fodor 1989, 69n).

There have been several attempts (e.g. Lepore and Loewer 1987; Fodor 1989, 1991; Woodward and Hitchcock 2003a,b; Woodward 2005) to explain how generalizations like LE can support counterfactuals and sustain predictions and causal-explanatory relationships, despite the fact that there are possible background conditions under which they are violated. If successful, these attempts show that such generalizations play the law role to a non-negligible degree.

To illustrate, consider the account given by Woodward and Hitchcock (2003a,b). They argue that generalizations like LE support causal-explanatory relations because they are *invariant under a range of hypothetical interventions*.³ For example, considering a population that isn't subject to significant inter-population competition or predation, if we were to intervene upon the intrinsic per capita growth rate r_c of the population (e.g. by genetic engineering to increase or reduce fertility or by administering drugs that affect reproduction rate), upon the carrying capacity K (by improving or depleting the environment), or upon the population size n (by carrying out a cull or by introducing new members into the population), then the actual growth rate, $\frac{dn}{dt}$,—*after the intervention episode*—would accord with LE.

² Tsoularis and Wallace (2002) survey some successful applications of LE in ecology. Ecologists sometimes appeal to more complex equations than LE. The following discussion also applies to these more complex equations. In general, ecologists have an armory of equations (or sets of equations) for predicting population growth and other phenomena. Different (sets of) equations are more or less predictively successful with respect to different populations. The fact that such (sets of) equations apply only to some populations—and even then only approximately—may of course discline you to call them 'laws'. (Though note that there is a nuanced literature in (philosophy of) ecology about whether there are genuine ecological laws: see, e.g., Berryman 2003, Colyvan and Ginzburg 2003, Lawton 1999, Mikkelsen 2003, Murray 2000, Turchin 2001.) And since this story is repeated throughout the special sciences, you may be disinclined to admit the existence of special science laws at all (except, perhaps, in a few special cases). To reiterate: the question with which I'm concerned is not so much whether such generalizations deserve to be called 'laws', but whether and to what extent they are able to do things like predict, explain, and support counterfactuals. LE is the sort of thing that many philosophers call a '*cp* law'. And, as we'll see, it is the sort of thing that Woodward and Hitchcock (2003a,b) and Woodward (2005) take to be an '*invariant generalization*'.

³ Woodward (2005, esp. 98) gives a precise definition of the technical notion of an '*intervention*'. For present purposes, it will suffice to think of interventions as ideal experimental manipulations of variables.

The reason that LE ‘supports’ these interventionist counterfactuals is that, in evaluating them, we are considering the ‘closest worlds’ in which such interventions occur (compare Hitchcock 2001, 283; Lewis 1979; Woodward 2005, esp. 133–145). In these worlds significant interfering factors like natural disasters don’t occur.

In virtue of the fact that LE supports these interventionist counterfactuals (when it comes to the populations that it models well) it follows directly, on the account of Woodward and Hitchcock (2003a,b), that the variables on the right-hand side of LE causally explain the actual growth rate of the population $\frac{dn}{dt}$. So, on their account, the fact that there are possible background conditions under which LE fails to hold doesn’t stand in the way of its playing important aspects of the law role.⁴

3. Minutis Rectis Laws

Not all exceptions to scientific generalizations arise due to the non-fulfilment of (explicit or implicit) cp clauses. This is best illustrated with respect to a law that admits of exceptions, but that plausibly is not a cp law, viz. the *Second Law of Thermodynamics* (SLT), which states that the total entropy of an isolated system does not decrease over time. Indeed, SLT is also normally taken to imply that, if an isolated system is initially out of equilibrium (that is, if it is initially not in its maximum entropy state), then its entropy increases over time until equilibrium is reached.

SLT admits of possible exceptions. Given an initial non-equilibrium state of an isolated system, it is *possible* (though very ‘unlikely’) that the micro-state should be one that leads to a later state that is further from equilibrium. An example of SLT violation, which is nevertheless possible (i.e. consistent with the fundamental dynamical laws), is an isolated system comprising an ice cube in hot water, in which the ice cube grows larger and colder over time, while the surrounding water becomes hotter.

Such exceptions to SLT *do not* arise due to failures of an explicit or implicit cp condition to hold. SLT is not aptly construed as a cp law. Rather than a cp clause, SLT includes a precise specification of its scope of application: it applies to thermodynamically isolated systems (including the universe as a whole). Unlike LE, there’s no possibility of interference from outside the systems that SLT characterizes.

Perhaps the claim that SLT is not a cp law can be disputed. Someone might, for instance, attempt to construe its appeal to an ideal isolated system as somehow amounting to a cp clause (compare Schurz 2002, 369–370). I don’t need to insist that it’s not a cp law. What I *do* wish to insist is that there is a type of possible exception to it that has nothing to do with the violation of any cp clause. That is, there is a class of exception that is not due to the failure of its idealizations to hold. *Even assuming an ideal isolated system*, exceptions to SLT may arise just as a consequence of certain unlikely microphysical realizations of the system’s initial thermodynamic state.⁵

Laws that admit of this type of exception are what I am calling ‘*minutis rectis* (*mr*)’ laws: that is, laws that hold only when the properties that they concern are realized in

⁴ The same is true on the accounts given by Lepore and Loewer (1987) and Fodor (1989, 1991), though I focus on Woodward and Hitchcock’s account here.

⁵ It would be inapt to construe SLT as including an implicit cp condition that supposes away such microphysical realizations. That would be to construe SLT’s implicit form as something like ‘the total entropy of an isolated system does not decrease over time, except when its initial microstate is such that it does decrease’. But this comes close to rendering SLT empty when clearly it isn’t (compare Earman and Roberts 1999, 465).

the right way. SLT holds only *minutis rectis* because the macro-states that it concerns are multiply realizable by points in the underlying phase space. For a thermodynamically isolated system, the majority of points in the system's phase space (measure ≈ 1) are on non-entropy-decreasing trajectories. However, there are a very few (measure ≈ 0) that are on entropy-decreasing trajectories. SLT only holds if the initial macro-state of the system is realized 'in the right way'—viz. by one of the 'usual' points in phase space that is on a non-entropy-decreasing trajectory.

Though I have illustrated the distinction between the notion of a cp law and that of an mr law with respect to a law that's an mr law but that plausibly *isn't* a cp law—namely SLT—many special science generalizations hold *both* only cp *and* only mr. Such generalizations admit of exceptions even when their cp clauses are satisfied. Even when there's no disruptive interference from outside the systems that they characterize they may still be violated just as a consequence of the properties that they concern being realized in the 'wrong' way.

LE is an example of a cp generalization that also holds only mr. I have already argued that it holds only cp. Rather trivially it also holds only mr. LE will break down if members of a population to which it normally applies start *en masse* to exhibit SLT-violating behavior: for example, if neurotransmitters suddenly stop diffusing across their synapses, or oxygen stops diffusing in their blood streams. In such a case, the growth rate of the population will not be predicted by LE. Not for nothing does the ecologist John Lawton say that SLT is one of the "three deep universal laws that underpin all ecological systems" (Lawton 1999, 178)!

There are also more interesting reasons why LE holds only mr. For example, the *geographical distribution* of a population can make a difference to its actual growth rate by making a difference to levels of competition for resources in subregions of a habitat (Law *et al.* 2003, 252) and by making a difference to breeding possibilities (Otto and Day 2007, 591). Indeed, given that population growth can be extremely sensitive to precise initial conditions (see, e.g., Hastings 1993, May 1974), even a very small perturbation of the precise, individual-by-individual initial geographical distribution of members of a population can make a difference to whether the population grows according to LE (as normal) or sharply declines, even where the population is well below the carrying capacity (see Hastings 1993). The latter situation—in which the population is initially precisely distributed in one of those rare ways that leads to dramatically LE-violating behavior—is analogous to a thermodynamic system's being at one of those rare points in phase space that leads to SLT-violating behavior. A population's having a certain size, n , is multiply realizable by different precise individual-by-individual geographical distributions. Only if the geographical distribution is 'right' will LE approximately hold.

4. Why It Matters

The distinction between cp and mr generalizations matters because the mr nature of a generalization poses problems for its ability to support counterfactuals, predictions, and causal-explanatory relations in a way that its cp nature does not.

Consider Woodward and Hitchcock's claim that generalizations like LE are invariant (i.e., support counterfactuals about what would happen) under interventions. The argument that this is so rests upon the idea that the closest worlds in which we intervene upon (say) the population size are not worlds in which the cp condition is violated: in

such worlds there is, for example, no natural disaster that wipes out the population immediately after the intervention. Given Woodward's notion of an intervention (Woodward 2005, 98) and Lewis's suggested similarity measure over possible worlds (Lewis 1979, 472) this seems reasonably plausible. The idea is that, in these closest intervention worlds, the post-intervention growth rate is modeled by LE.

Yet the *mr* nature of LE appears to undercut its ability to support interventionist counterfactuals. Even concerning a population that is usually well-modeled by LE, it seems extremely doubtful that it is true that 'If the size of the population had been intervened upon, the post-intervention macro-state of the ecosystem (or indeed the universe) wouldn't have been realized in one of those ways that leads to the ceasing of the entropy-increasing processes necessary to the continued survival of the members of the population'. After all, it seems that the post-intervention micro-state just *might* have been one of those extremely rare ones in which the requisite entropy-increasing behavior ceases.⁶

It is also doubtful that, even where a population size is below the carrying capacity, it is true that 'If the size of the population had been intervened upon (but had nevertheless remained well above zero and well below the carrying capacity), the resulting precise individual-by-individual geographical distribution of members of the population would not have been such as to lead to a decline in the overall size of the population'. After all, it's not possible to intervene on the size of the population without impacting on the precise individual-by-individual geographical distribution (fewer or more individuals can't be distributed in the same individual-by-individual way) and, in light of the dramatic effects that slight changes in initial conditions can have on ecosystems, it seems that the post-intervention distribution just *might* have been one of those rare ones that leads to a decline in numbers.⁷

It is very doubtful that the truth of either of the counterfactuals considered in the previous two paragraphs follows from the Woodwardian notion of an intervention or the Lewisian notion of similarity among possible worlds. But if such counterfactuals aren't true, then it appears that we can't reason that, if the population size had been intervened upon, then the growth rate would have subsequently followed LE.⁸

Likewise with SLT. Consider the counterfactual 'If I had placed this ice cube into that glass of hot water, then it would have melted quickly'. SLT's *mr* nature appears to undercut its ability to support this counterfactual. It appears to be false that 'If I had placed the ice cube in the hot water, then the resulting system *would not* have been in

⁶ That this is so is credible given David Albert's point that "the subregion ... of the phase space of any thermodynamic system which is taken up by 'abnormal' microconditions, microconditions (that is) that lead to *violations* of the laws of thermodynamics, is not merely *small* ... but also *scattered*, in unimaginably tiny clusters, more or less at random, all over the place" (Albert 2000, 67; italics original). A consequence of this is that very many 'normal' microstates differ from an 'abnormal' one only by a small perturbation.

⁷ As we'll see in Section 6 below, there is evidence that 'normal' geographical distributions, that result in LE-like growth, typically differ from one of the rare ones that doesn't only by a small perturbation. This is analogous to the way—noted in Footnote 6 above—in which 'normal' points in the phase space of a system typically differ from an 'abnormal' one only by a small perturbation.

⁸ If we build into the antecedent of the counterfactual a specification of exactly *how* the intervention would have occurred—a specification that must be detailed enough to imply not only what the resulting precise geographical distribution of members of the population would be but also what the resulting point in the phase space of an isolated system of which the population is part would be—then we might get a true counterfactual. But this is not the sort of interventionist counterfactual to which Woodward and Hitchcock appeal in their account of causal explanation.

one of those rare micro-states that fails to lead to melting'. The post-intervention system *might* have been in such a micro-state, and this undercuts the assertion that the ice cube would have melted. It does not appear that the Lewisian notion of similarity combines with the Woodwardian notion of an intervention to ensure that a world in which the specified post-intervention macro-state is realized by a non-entropy-increasing micro-state is *more dissimilar* to the actual world than one in which it is realized by an entropy-increasing micro-state (compare Hájek ms).

If *mr* generalizations aren't able to sustain counterfactuals about what would happen under interventions (i.e. if they're not invariant), then this threatens to undermine their ability to underwrite causal-explanatory relations and their predictive power.⁹ This means that philosophical vindications—such as Woodward and Hitchcock's—of the causal-explanatory import and predictive power of *cp* laws¹⁰ are not *ipso facto* vindications of *mr* laws (or indeed of laws that are both *cp* and *mr*, as many special science laws appear to be).¹¹

5. Potential Solutions

There is a range of approaches that one might take in attempting to address the problems posed by *mr* laws.

First, one might consider modifying the Lewisian similarity metric so that worlds in which (e.g.) I intervene on a thermodynamic system and the post-intervention system conforms to SLT come out closer than those in which the post-intervention system does not so conform. One might similarly take conformity to special science laws like LE to make for similarity to the actual world.

Thus, for example, in response to a worry raised by Elga (2001) about whether Lewis's similarity metric delivers an asymmetry of counterfactual dependence, Dunn (2011) suggests modifying Lewis's metric so that, other things being equal, worlds obeying SLT and also the various special science laws come out closer to the actual world than those that don't. Such a proposal would seem to ensure that SLT and LE underwrite the truth of those counterfactuals—like (in SLT's case) 'If I had placed the ice cube in the hot water, then it would have melted quickly' and (in LE's case) 'If we had culled half of the population, then the per capita growth rate would now be higher'—needed to support causal/explanatory relations after all.¹²

One concern about this approach is that it appears to force upon us the truth of counterfactuals like 'If I had put the ice cube in the hot water, then the resulting system

⁹ Strictly speaking, for generalizations to be able to support predictions, they need to be able to support indicative rather than counterfactual conditionals. But the *mr* nature of SLT and LE undermines their ability to support the needed indicatives. For example, the *mr* nature of SLT undercuts its ability to support the indicative 'If I place the ice cube in the hot water, then it will melt quickly'. That's because it's seemingly false that 'If I place the ice cube in the hot water, then the resulting system will *not* be in one of those rare micro-states that fails to lead to melting'.

¹⁰ Or, rather, of those generalizations that many philosophers of science (though not Woodward and Hitchcock themselves!) call '*cp* laws'.

¹¹ Though I do not seek to show it here, the attempted vindications of *cp* laws provided by Lepore and Loewer (1987) and by Fodor (1989, 1991) fare no better than Woodward and Hitchcock's account when it comes to *mr* laws.

¹² If one adopts a possible worlds semantics for indicative as well as for counterfactual conditionals (Stalnaker 1968), then this proposal may also ensure that SLT and LE underwrite the indicatives needed to support predictions.

wouldn't have been in one of the rare entropy-decreasing microstates' and (where a particular population is well above zero and well below the carrying capacity) 'If we had culled half of the population, then the geographical distribution of the remaining population members wouldn't have been one of those rare ones that leads to population decrease'. Such counterfactuals seem much less plausible.

Perhaps there is some room for maneuver: perhaps, for instance, one could maintain that the assertion of the latter counterfactuals (i.e. the ones appearing in the previous paragraph) results in a context shift and a corresponding change in the standards of similarity (compare Lewis 1979) with the consequence that these latter counterfactual utterances assert false propositions while, in the original, ordinary context, the former counterfactuals (i.e. the ones appearing in the last-but-one paragraph) assert true propositions. I shan't explore the prospects for such a response here. Suffice to say that it appears rather *ad hoc*.

A *second* option might be to modify the Woodwardian notion of an intervention so that, for example, manipulations of a population size that result in the population being geographically distributed 'in the wrong way', don't count as 'interventions' in the relevant, technical sense. A worry about this strategy is that it is not clear that the 'wrong' sort of interventions could be specified in a systematic and non-*ad-hoc* way. Simply specifying the relevant 'interventions' in terms of precisely those counterfactual outcomes that one wants to avoid (as in 'the "intervention" on population size must not be such as to result in a precise geographical distribution that leads to a violation of LE') is *ad hoc* and unsystematic. Similarly, in the thermodynamic case, one might reasonably wonder whether there is a useful notion of 'intervention' such that manipulations of a system's macro-state that happen to result in its being in a micro-state on an entropy-decreasing trajectory fail to count as interventions. In any case, I shan't pursue this strategy further here.

A *third* option is to argue that 'deterministic' mr laws are mere approximations to probabilistic laws. For example, it is tempting to say that, while SLT is a mr law, it is an approximation to a probabilistic law that is not a mr law. Statistical Mechanics (SM), it might be said, furnishes us with an exceptionless, *probabilistic* version of SLT. And perhaps there also exists a probabilistic approximation to the deterministic LE that is stricter than LE.

One might claim that, although the probabilistic version of SLT implied by SM *does not* support counterfactuals like 'If I had placed the ice cube in the glass of hot water, then it would have melted quickly', it does support counterfactuals like 'If I had placed the ice cube in the hot water, then it would have had a *probability* of melting quickly that is *very much higher* than the probability of melting quickly that it would have had if (say) I had returned the ice cube to the freezer'. Likewise, while a probabilistic approximation to LE won't support counterfactuals like 'If we had culled half the population, then the per capita growth rate would now be higher', it might well support counterfactuals like 'If we had culled half the population, then there would have been a high probability of the per capita growth rate now being higher'. Such counterfactuals appear to be precisely the sort that are relevant to *probabilistic* causal explanation.¹³

¹³ Such probabilistic generalizations would also seem able to support the sort of indicative conditionals—such as 'If I place the ice cube in the hot water, then the probability that it will melt quickly will be very much higher than it will be if I return the ice cube to the freezer' and 'If we cull half the population, then there will be a high probability of the per capita growth rate being higher'—relevant to probabilistic prediction.

It is this third option that I will pursue in the next two sections. In Section 6, I will outline how probabilistic approximations to SLT and LE can be derived. In Section 7, I will argue that it is plausible that the probabilities that these probabilistic generalizations entail are objective chances—which is important if they are to support objective causal-explanatory relationships. And, more generally, I will argue that these probabilistic generalizations are laws, or at least play the law role to a reasonably high degree.

6. Derivation of Probabilistic Approximations to Deterministic MR Laws

As we have seen, mr laws are high-level generalizations for which there are some (unusual) realizations of the properties that they concern that (when evolved forward according to underlying dynamical laws) issue in states that realize alternatives to the behavior that they predict.

The basic trick involved in deriving a probabilistic approximation to a deterministic mr generalization is to take the state space in which the properties that the generalization concerns are realized, and impose an appropriate probability distribution over that state space. This yields probabilities for the ‘good’ realizers—namely those that issue, via the underlying dynamics, in realizers of the behavior predicted by the original mr generalization—and for the ‘bad’ realizers—that issue in alternative behaviors. The upshot is a probability for the behavior predicted by the original mr law and a probability for the alternative behavior. Bad realizers do not result in exceptions to this new probabilistic law (unlike the original mr law) because the behavior that they issue in is assigned an explicit probability by the new law (a low probability, if the original mr generalization was a good one).

In this section, I will outline how probabilistic approximations to SLT and LE can be derived in this way. It is plausible that this derivational pattern can be implemented elsewhere in the special sciences, and so we may fairly generally be able to derive probabilistic approximations to deterministic mr laws.

6.1. Derivation of a Probabilistic Approximation to SLT

The details of how SM probabilities for SLT-like behavior are derived are reasonably well known. I shall rehearse them briefly here. The state-space that is appealed to by SM, in which the possible thermodynamic states of a system are realized, is a *phase space*.¹⁴ A system’s phase space has $6n$ dimensions, where n is the number of particles that the system comprises. There are six dimensions for each particle because the values of six variables—three position and three momentum—have to be specified to give the state of each particle. A *point* in a system’s phase space thus corresponds to a precise microstate for the system (i.e. a precise state for each of the particles that it comprises).

The thermodynamic state of a system (which is specified by giving the values of macrovariables, such as the temperature, pressure, and chemical composition of various macroscopic subregions of the system) is multiply realizable by points in the system’s phase space. That is, a possible thermodynamic state of a system corresponds, not to a point, but to a *region* of the system’s phase space.

¹⁴ Where the system is quantum rather than classical, the state space is not a phase space, but rather a space of all possible quantum states of the system.

The underlying dynamics define a set of possible trajectories through this state space. In the case of Classical SM, the underlying dynamics are the deterministic Newtonian dynamics. For each point in the phase space, the Newtonian dynamics define a unique future trajectory through the space (i.e. they define how the precise microstate of the system will evolve over time given any precise starting point). In this case, a ‘good’ point in the phase space is one whose future trajectory is such that the successive points on that trajectory realize a succession of macrostates of the system (i.e. a ‘macro-trajectory’) such that each macrostate in this succession has entropy that is no lower than the previous macrostate. The ‘bad’ points in the phase space are those whose future micro-trajectories realize future macro-trajectories for the system in which entropy sometimes decreases.

The bad points of the phase space form a set of very small measure. Moreover, the bad points are scattered in the phase space, so that the measure of such points in a subregion of the phase space corresponding to a given macrostate is similar to the measure of such points in the space as a whole (Albert 2000, 67).

In SM, probabilities for SLT-like behavior are derived by imposing a probability distribution—the one that’s uniform on the Lebesgue measure—over a system’s phase space. This yields an unconditional probability of the system’s being at a good point that is equal to the (Lebesgue) measure of the good points, and likewise for the bad points. Because of the scattering of the bad points, the probability that the system is at a good point *conditional upon its being in any given macrostate* is similar to the unconditional probability of its being at a good point. Because the measure of the set of bad points is so low, conditional upon the system’s being in any particular macrostate, there is a very high probability that its entropy will not decrease over time.

One difficulty is that, because of the time-reversal invariance of the laws of Newtonian mechanics, imposing a uniform probability distribution over the phase space of a system and conditioning upon its current macrostate not only yields a high probability that the entropy of the system won’t decrease in the future, it also yields a high probability that the entropy of the system wasn’t lower in the past. And yet we know that very many thermodynamic systems had lower entropy in the past (e.g. the milk that is now mixed in my coffee was previously unmixed, a black hole which used to be more massive has now radiated away some of its mass, the universe is now cooler and less dense than it was, etc.). This is known as the ‘reversibility paradox’.

Albert (2000, Ch. 4) argues that the reversibility paradox can be solved by conditioning the uniform probability distribution over the phase space of the system, not just upon its current macrostate, but upon the fact that the system initially had low entropy. Conditional upon this additional fact, we get the desired result: a high probability that the past states of the system were lower entropy than the current state, and a high probability that future states have entropy that is no lower than the current state (higher if the system is currently out of equilibrium).

Yet this doesn’t explain where the low entropy initial states of the thermodynamic systems in question come from. Albert argues that *this* can be explained by the conjecture (supposedly supported by Big Bang cosmology) that the initial macrostate of the universe as a whole was a very low entropy one. This is the so-called ‘Past Hypothesis’ (PH). In fact, Albert (2000, see esp. 96) argues that *all of the probabilities of SM* can (in principle) be derived from an axiom system comprising PH, the fundamental dynamical laws (FD), and a probability distribution over the phase space of the universe as a whole that is uniform on

the Lebesgue measure (the so-called Statistical Postulate, or SP). The idea is that conditioning the probability distribution given by SP upon PH together with the current state of any given thermodynamic subsystem of the universe will yield the SM probabilities for that subsystem. In particular, where the current state of that subsystem is a non-equilibrium one, the resulting probability distribution will be such as to give a high probability for entropy increase in the future, but also a high probability that the system was in a lower entropy state in the past. Loewer (2012a, 16, 2012b, 124), following a suggestion of Albert's (see Loewer 2012a, 16n), dubs the conjunction of FD, PH, and SP 'the Mentaculus'.

The argument that the SM probabilities are derivable from the Mentaculus goes roughly as follows. The PH specifies a macrostate for (or at least gives some macroscopic information about) the universe as a whole at an initial time. Consider the region of the universe's phase space compatible with PH. Relative to the total volume of that region, a very large volume is taken up by micro-states that lead (by FD) to increasing entropy for the universe as a whole until thermodynamic equilibrium is reached. Consequently, conditioning the uniform probability distribution (given by SP) over the universe's phase space upon PH yields a very high probability that the entropy of the universe will increase until equilibrium is reached.

The next claim in this argument is that the fact that the universe as a whole started in a low entropy initial state means that many subsystems of the universe, if they first become (approximately) thermodynamically isolated before the universe itself reaches thermodynamic equilibrium, will themselves have non-equilibrium initial macrostates. Moreover, since a system's becoming approximately isolated is not itself correlated with its initial microstate being non-entropy-increasing, it is very likely that any such subsystem that is in initial disequilibrium will increase in entropy until it reaches equilibrium (see Loewer 2007, 302, 2012a, 124–125, 2012b, 17; Albert 2000, 81–85).

Some have expressed doubts about the derivability of SM probabilities from the Mentaculus (see, for example, Winsberg 2004; Earman 2006; Callender 2011, 99–102). While I shan't rehearse these arguments here, it is worth noting that there remains the fallback position of simply taking SM probabilities to be derivable on a system-by-system basis by imposing a uniform probability distribution over the phase space of each individual approximately isolated thermodynamic system and then conditioning upon the initial state of *that system*. This is what Callender (2011, 106–112) has dubbed a 'Localist' approach to deriving SM probabilities (see also Frigg and Hoefer 2015) as opposed to Albert's 'Globalist' approach (Callender 2011, e.g. 100). Those taking a 'Localist' approach can remain agnostic about whether the initial macrostates of the various approximately isolated subsystems of the universe (and the appropriateness of the uniform distribution over their phase spaces) are somehow to be explained (as on Globalist approaches) in terms of something like the Mentaculus (Callender 2011, 110).

Whether the SM probabilities are to be derived globally or locally, the result would appear to be a probabilistic approximation to SLT which gives a probability for entropy increase (as well as a probability for entropy non-decrease) in a thermodynamic system as a function of the macrostate of that system.

6.2. Derivation of a Probabilistic Approximation to LE

In Section 3, I noted that one of the more interesting reasons why LE holds only in that a population's having a certain size, n , is compatible with a variety of different indi-

vidual-by-individual geographical distributions of members of the population and while it might be that, according to the underlying dynamics governing the population's competition and reproduction, most initial geographical distributions lead to growth that approximately accords with the LE, there may be some that issue in different sorts of growth pattern.

In fact, we can derive a probabilistic approximation to the deterministic LE in a way that is remarkably analogous to the derivation of the probabilities of SM. Showing that this is so will involve a little bit of theoretical background, which I now set out.

Ecologists often model the spatial distribution of members of a population in a habitat using so-called *cellular automata* (see e.g. Otto and Day 2007, 591–594).¹⁵ The most basic form of cellular automaton comprises a lattice or grid structure with a finite number of cells, a specification of the state of each cell in the lattice at some initial time period t_0 , and set of deterministic update rules (i.e. recursion equations) that give the state of each cell during a time interval as a function of its own state and those of its neighboring cells in the previous time interval. In population ecology, the cells of the lattice are taken to represent sub-regions of a particular habitat and the state of the cells can be taken to represent the number of members of a population present in each sub-region. The number of members of the population in each cell in an interval may depend upon the number of members in that cell and surrounding cells in the previous interval because of the implications that this has for reproduction and local competition for resources, and because of dispersal.

Cellular automata can be generalized into more sophisticated models that treat space, time, and the state of cells as continuous, with differential equations used to model the dynamics (Hiebeler 2012, 125). They can also be generalized to allow the state of spatially remote regions to exert influence on one another, to allow for lag times in influence, and to allow for probabilistic dynamics (ibid.). Cellular automaton models and their variants can be integrated with logistic models of population growth to yield so-called 'spatial logistic equations' (Law *et al.* 2003).

Consider, for simplicity, an automaton with a finite number, m , of cells. And suppose that the state of each cell is represented by a single variable (the values of which might correspond either to the number of members of the population in that cell or alternatively to the combined *mass* of population members in that cell). We can then define an m -dimensional state space \mathcal{S} , where a point in \mathcal{S} represents a value for the variable characterizing the state of each cell. That is, a point in \mathcal{S} represents a precise geographical distribution of members of the population (or rather one that is precise up to the representational limits of a cellular automaton with only m cells). The update rules define a set of possible trajectories through \mathcal{S} (i.e. a set of possible ways in which the states of the m cells may evolve over time). In the deterministic case, the update rules are recursion or differential equations that associate each possible initial condition of the automaton (a point in \mathcal{S}) with a unique future trajectory through \mathcal{S} (i.e. a unique state of the whole automaton at each future time).

A value for the variable n in the LE—representing the number of members of the population (i.e. the sum of the number of members of the population in each of the cells) or,

¹⁵ The most famous example of a cellular automaton is John Horton Conway's 'Game of Life', described by Gardner (1970).

alternatively, their combined mass (i.e. the sum of the mass of population members in each of the cells)—corresponds to a (non-singleton) *set* of points in \mathcal{S} (i.e. to a region of \mathcal{S}), with the various points in that set corresponding to the various possible ways in which a population of that size could be distributed throughout the habitat. In this sense, a population's having a given size, n , is multiply realizable by different precise geographical distributions of the population.

For a given overall population size n —corresponding to a sub-region R of \mathcal{S} —at some initial time, t_0 , different possible geographical distributions of that population can have very different consequences for the future evolution of the overall size of the population (Hastings 1993, 1366–1367). That is, there may be pairs of points, p_1 and p_2 , within R that represent geographical distributions that, when evolved forward according to the dynamics (which for now are being assumed to be deterministic), issue in geographical distributions p'_1 and p'_2 at a later time t_1 that realize very different overall population sizes.

Hastings (1993) conducted computer simulations using a spatial logistic model, comparing what happened when different initial conditions—specifically, different precise geographical distributions of members of a population—were evolved forward according to a simple dynamics modeled by recursion equations.¹⁶ He compared the results after many iterations of the recursion equations, at a point where the standard LE would predict 0 growth for the overall population. There were essentially three types of outcome that were observed (Hastings 1993, 1364–1367). In one, the overall population size was constant (as the LE might lead one to expect). In a second there were stable oscillations. As Hastings puts it:

“[I]n the simple model examined here ... there can be very sensitive dependence of the long-term behavior on the initial conditions. ... [F]or certain parameter values there are two stable solutions: one where the total population is constant, and one where the population has a large amplitude cycle of period two.” (Hastings 1993, 1362–1363)

Hastings also found that certain (rarer) initial conditions led to unstable (possibly chaotic) behaviour (Hastings 1993, 1365).

Interestingly, the initial geographical distributions that led (under the dynamics) to different trajectories for the value of n were found to be highly scattered in the state space of the model.¹⁷ As Hastings observes:

“[T]he initial conditions that lead to ... distinct types of ... behavior ... form sets that are interspersed at any scale. ... [I]t is not possible to determine the final outcome of the solution unless the initial conditions are known precisely.” (Hastings 1993, 1370)

¹⁶ Specifically, Hastings' (1993, 1363–1365) simple discrete time model involved a two-cell automaton. Each time period comprised a 'dynamic phase'—in which the population in each cell evolved according to LE—and a 'dispersal phase'—in which a fixed proportion of the population of each cell moved to the other cell. The evolution of the overall population (i.e. the sum of the populations of the two cells) that resulted from beginning the simulation with various population numbers in each of the cells was observed.

¹⁷ Since there were only two cells in Hastings' model, the state space for the model is just the two-dimensional version of the space, \mathcal{S} , characterized in the main text above.

And, as he puts it at another stage:

“[T]he boundary between initial conditions leading to each [type of behavior] is apparently a fractal that fills up the entire set of possible initial conditions, so that effectively there is no way to specify initial conditions definitely leading to either kind of behavior.” (Hastings 1993, 1363)

This means that, if one considers the subregion R of the state space of the model corresponding to any initial overall population size, then the measure of the ‘good’ points—that lead to LE-like growth—and the measure of the ‘bad’ points—that don’t—is non-trivial (i.e. not 1 or 0). Indeed the measure of the ‘good’—and that of the ‘bad’—points remains non-trivial in almost arbitrarily small sub-regions of the state space of the model. In consequence, ‘good’ points differ from ‘bad’ points only by a small perturbation. This, of course, is reminiscent of the way in which the ‘bad’ points that lead to SLT-violating behavior are scattered throughout the *phase space* of a system.

Ecologists often model the initial geographical distribution of a population by means of a probability distribution over its *possible* initial geographical distributions (Vandermeer and Goldberg 2013, 126–142). When working with a cellular automaton model, a natural way to do this is to impose a uniform probability distribution over the state space \mathcal{S} of such a model and then to condition that distribution upon the fact that the system occupies the subregion R of \mathcal{S} corresponding to the initial overall size of the population (see Law *et al.* 2003, 254, 257; compare also Coe *et al.* 2008). This yields a probability for LE-like growth that is equal to the (non-trivial) proportion of R taken up by points that, when evolved forward by means of the (deterministic) dynamics concerning evolution through \mathcal{S} , yield LE-like growth. For the sorts of population that are normally well-modeled by LE these dynamics are presumably such that the proportion of points in any such region, R , that lead to LE-violating behavior is small,¹⁸ so that the resulting probability of LE-like growth is high.

The upshot of all this, then, is a probabilistic approximation to LE according to which the probability of LE-like growth, for the sorts of population that are normally well-modeled by LE, is high but not equal to 1. Non-LE-like behavior does not constitute an ‘exception’ to this probabilistic generalization (as it did to the original, deterministic version of LE), for such behavior is assigned an explicit (low) probability by the generalization.

So far, I have been considering the case in which our cellular automaton model of geographical distribution has deterministic update rules. As previously noted, such models can be generalized to incorporate probabilistic dynamics. But this does not affect the derivation of a probabilistic approximation to LE dramatically.

Probabilistic update rules associate each possible precise initial geographical distribution of members of the population, represented by a point in \mathcal{S} , with a (non-singleton) *set* of possible future trajectories through \mathcal{S} and a probability distribution over that set. For the sort of population that LE usually models well, an accurate model will be one in which the update rules entail, for a large proportion of points in (almost) any subregion R of \mathcal{S} corresponding to a possible initial overall size of the population, a high probability that a system at that point will take a future trajectory through \mathcal{S} that realizes an overall population growth rate that approximates the predictions of LE. The result of impos-

¹⁸ Given suitable scattering, this will be so if the proportion of ‘bad’ points in \mathcal{S} is itself small.

ing a uniform probability distribution over \mathcal{S} and then conditioning upon the fact that the system occupies such a region R will then be a high probability for LE-like behavior.

The difference from the case where the update rules are deterministic is that the resulting probability for LE-like behavior is now a weighted average (with equal weights) of the probabilities with which each of the points in R issues in LE-like behavior, with the (equal) weighting provided by the uniform distribution over R (which itself comes from conditioning the uniform distribution over \mathcal{S} upon the fact that the system is in R). Or, in more technical terms, the probability for LE-like behavior is given by a *mixture distribution* over possible trajectories for the overall size n of the population, with the mixture components being the probability distributions over possible trajectories for n that (according to the probabilistic update rules) result from the various points in R , and with the mixture weights provided by the uniform distribution over R .

What I have said so far about LE isn't quite the end of the story. For there might be other reasons—besides the possibility of rare geographical distributions of the population that lead to LE-violating behavior—why the deterministic version of LE holds only *mr.* For instance, a population's having a certain size and geographic distribution is compatible with its having a variety of *demographic structures* (e.g. age and gender profiles). The demographic structure of a population can make a difference to its growth rate. As with geographical distribution, there might be certain initial demographic structures that don't result in LE-like growth, while others do.

Ecologists seek to model the influence of demographic factors by appealing to 'structured population models' (see e.g. Otto and Day 2007, Ch. 10). In principle there's no particular difficulty in building a model that represents both population structure and geographical distribution. One way to do this would be to have the state of each cell in a cellular automaton model characterized by multiple variables: for example, variables representing the number of females aged 0–1, the number of females aged 1–2, the number of males aged 0–1, etc., in the sub-region of the habitat represented by the cell (compare Callender and Cohen 2010, 437). The update rules for such an automaton would need to be appropriately sensitive to the (now complex) state of each cell.

Where m is the number of cells, and n is the number of variables characterizing the state of each cell, such a model generates an $n \times m$ -dimensional state space, \mathcal{S}' . Probabilities for LE-like growth can be recovered by imposing a probability distribution over \mathcal{S}' . Conditioning upon the fact that the system is in some sub-region R' of \mathcal{S}' corresponding to the population's having a given size n then gives us a probability that the overall size of the population is realized by one of those points in \mathcal{S}' that, according to the update rules, yields LE-like growth.

Alternatively, if the update rules are probabilistic, the probability distribution over \mathcal{S}' , when conditioned upon R' , provides us with weights for averaging those probabilities with which, according to the probabilistic dynamics, the various points in R' yield LE-like growth. The result of this 'weighted averaging' (or 'probability mixing') procedure is a probability that the population will undergo LE-like growth given that its initial size is n .

This is still not the end of the story regarding LE since, as noted in Section 3, the possibility of violations of SLT that undermine the biochemical processes sustaining members of the population is another reason—besides the possibility of rare geographical distributions and demographic structures that lead to LE-violation—why LE might hold only *mr.* I think that the correct way to handle this is to regard the recursion or differential

equations that—in the deterministic case—entail a trajectory through \mathcal{S}' for the population (given its initial starting point) as *themselves* holding only *mr*. Unlike points in phase space, points in \mathcal{S}' do not represent fundamental physical states of a system. (Likewise—unlike the Newtonian equations that determine a classical system’s trajectory through its phase space—the equations used in an ecological model to give a trajectory through \mathcal{S}' are not fundamental laws of nature, but rather higher-level generalizations.) Points in \mathcal{S}' represent geographical distributions and demographic profiles for a population, not particle-by-particle states of the population, or of some isolated system of which it is part. A system’s being at a point in \mathcal{S}' is compatible with a thermodynamic system of which it is a part being at any one of numerous points in its phase space. That is, it corresponds to a region T of that system’s phase space.

Given the scattering of points that lead to non-SLT-like behavior in the phase space of thermodynamically systems, the measure of such points in T is similar to the measure of such points in the system’s phase space as a whole. If the system happens to be at one of these ‘bad’ points, then violations of the equations giving a trajectory of the population through \mathcal{S}' can be expected. That is to say, even a ‘good’ point in \mathcal{S}' —a point which, according to the *mr* dynamics concerning trajectories through \mathcal{S}' , results in a trajectory that realizes LE-like behavior—might be realized by a ‘bad’ point in phase space—a point which, according to the Newtonian dynamics, leads to a trajectory through phase space that does not realize LE-like behavior.

Yet the *mr* equations concerning a population’s trajectory through \mathcal{S}' approximate stricter probabilistic generalizations, which can be derived by means of a probability distribution over the phase space of a thermodynamic system of which the population is part. Conditioning that distribution upon the fact that the system is in region T yields a probability that it is at one of the bad points in phase space that lead to violations of the *mr* equations giving trajectories through \mathcal{S}' . The result is explicit probabilities for the system evolving according to these *mr* equations. In other words, the result is stricter probabilistic generalizations that can take the place of these *mr* equations.

There is a further wrinkle here. Plausibly there are other reasons, besides the possibility of non-SLT-like behavior, why deterministic recursion or differential equations concerning trajectories through \mathcal{S}' hold only *mr*. Ecologists recognize that these deterministic equations are mere approximations to more accurate probabilistic equations.

“At a microscopic scale, population growth is inherently stochastic: birth and death events occur at random . . . together with dispersal, these events lead to random variation . . . It is therefore helpful to start with a stochastic process describing behavior of individuals . . . and derive deterministic approximations from it.” (Law et al. 2003, 253)

It is plausible that the possibility of non-SLT-like behavior is not the only (or even the main) reason why a population’s trajectory through \mathcal{S}' appears stochastic. Presumably details that are more fine grained than are represented by the variables that define \mathcal{S}' but less fine grained than those represented by the variables that define a system’s phase space make a difference to whether or not the population follows a trajectory through \mathcal{S}' that is predicted by deterministic equations.

For instance, it is plausible that the *biochemical* states of members of the population (and maybe the ecosystem as a whole)—and not simply their geographical distribution and demographic profile—make a difference to population members’ fertility, mortality,

and movement around the habitat. A state space \mathcal{R} , in which the points represent a precise geographical location, a gender, age, and a biochemical state for each member of the population is one such that the points in \mathcal{S}' correspond to *regions* of \mathcal{R} . Given a set of deterministic equations concerning trajectories through \mathcal{S}' , a set of (stricter) probabilistic approximations to those equations might be recovered by imposing a probability distribution on \mathcal{R} . Given whatever equations concern trajectories through \mathcal{R} , this yields probabilities that the trajectory of the population through \mathcal{S}' will accord with the deterministic equations that are mere approximations to the true stochastic dynamics concerning trajectories through \mathcal{S}' .

If this is right, then the appropriate place to apply a uniform distribution over the phase space of a thermodynamic system of which the population is part is plausibly in seeking to derive accurate, probabilistic equations for the system's trajectory through \mathcal{R} . After all, the probabilities of SM manifestly are relevant to the biochemical processes influencing the trajectory through \mathcal{R} . For instance, diffusion and osmosis are heavily involved in all biochemical processes.

To summarize, the idea is that, to the extent that it is useful to distinguish an ordered sequence $\langle \mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_n \rangle$ of state spaces in modeling some special science phenomenon (such as population growth), where points in \mathcal{S}_i corresponding to regions in \mathcal{S}_j for each i, j such that $1 \leq i, j \leq n$ and $j - i = 1$, probabilistic equations for the dynamical evolution of a system through \mathcal{S}_i may be derived via an appropriate probability distribution over \mathcal{S}_j .¹⁹ The result is that the probabilities for the various possible trajectories through \mathcal{S}_i , given a certain starting point in \mathcal{S}_i , that are entailed by the equations governing the dynamical evolution of the system through \mathcal{S}_i are a mixture (i.e.—in the finite case—a weighted average)—with the mixture weights given by the distribution over \mathcal{S}_j —of the probabilities with which the points in the region of \mathcal{S}_j corresponding to that point in \mathcal{S}_i issue in trajectories through \mathcal{S}_j that realize those various trajectories through \mathcal{S}_i .

Fundamental probabilities enter this picture in two places. First, there are the probability distributions that, conditional upon a system being at a given point or region of \mathcal{S}_i , yield probabilities for its being located in the various subregions of the region of \mathcal{S}_j associated with that point or region of \mathcal{S}_i . Second, there are the probabilities that figure in the dynamical equations governing the system's evolution through the most basic state space, \mathcal{S}_n . Where the most basic state space is a classical phase space, the latter probabilities are the trivial (0 or 1) probabilities entailed by Newtonian mechanics.²⁰ The picture is moderately reductionist in that it supposes that the only fundamental *dynamical* probabilities (probabilities for the *evolution* of a system over time) are those that concern evolution through the most basic state space.

I have suggested that the ordered sequence of state spaces that it is useful to distinguish in modeling population growth—in the case where the fundamental dynamics are classical—is at least the sequence $\langle \mathcal{S}', \mathcal{R}, \mathcal{P} \rangle$, where \mathcal{P} is the phase space of an (approximately) isolated system of which the population is a part. The probabilities of

¹⁹ In my view, the question of what state spaces it is 'useful' to distinguish and what counts as an 'appropriate' probability distribution are ultimately questions of what state spaces and probability distributions figure in a 'best system' for the universe. The relevant notion of best systemhood will be discussed in the next section.

²⁰ Where the most basic state space is a quantum mechanical state space, they are the non-trivial probabilities of quantum mechanics. Where it is a Bohmian mechanical state space, they are the trivial probabilities entailed by the Schrödinger Equation and the Guiding Equation.

the most accurate probabilistic approximation to LE derive from (a) the (Newtonian) equations governing the trajectory of the system through \mathcal{P} together with, (b) a distribution that, when conditioned upon the fact that a population has size n , yields a probability distribution over the various subregions of \mathcal{S}' compatible with n , (c) a distribution that, conditional upon a system's being at a point or region within \mathcal{S}' , yields a distribution over the regions of \mathcal{R} compatible with that point or region of \mathcal{S}' , and (d) a distribution that, conditional upon a system's being at a point or region within \mathcal{R} , yields a distribution over the regions of \mathcal{P} compatible with that point or region of \mathcal{R} .

In this section, I have outlined how a strategy for deriving probabilistic approximations to deterministic mr laws is implemented with respect to SLT and LE. This strategy involves the imposition of probability distributions over state spaces in which the properties of concern to the original mr law are realized. My conjecture is that this strategy can be repeated elsewhere in the special sciences. To the extent that it can, the derivation of strict(er) probabilistic versions of deterministic mr special science generalizations would appear to be possible.

Yet it seems that only if the probabilities that they entail are *objective chances* can the resulting probabilistic generalizations play the law role to a reasonable degree. Specifically, only if the probabilities that they entail are objective chances can these probabilistic generalizations support objective causal-explanatory relations and underwrite predictions in the correct way.²¹ But the view that such probabilities are objective chances is plausible, not only because practicing scientists *do* take probabilities derived from distributions over underlying state spaces to play a key role in causal explanation and prediction (compare Albert 2000, 64–65; Loewer 2001, 611–612; Cohen and Callender 2009, 3), but also because the objective chancehood of these probabilities appears to follow from a family of independently plausible accounts of chance: namely the family of 'Best System'-style accounts.

Lewis (1994, 478–482, 1983, 366–368, 1986a, xi, xiv–xv, 121–131) was the first to develop such an account in detail, with related accounts proposed by Schrenk (2008), Cohen and Callender (2009), Callender and Cohen (2010), Callender (2011), Dunn (2011, 80–92), Frisch (2014), and Frigg and Hoefer (2015). The focus of the next section will be on articulating how the objective chancehood of SM probabilities and special science probabilities deriving from distributions over underlying state spaces plausibly follows from an appropriate Best System account of chance. Indeed it will also be argued in the next section that Best System accounts—which are typically advanced as accounts of laws as well as of chances—lend direct credibility to the view that probabilistic approximations to many high-level deterministic mr generalizations are full-blown laws of nature, and not merely to the view that they play the law role to a reasonably high degree.

7. Best System Analyses

According to Lewis's (1994) Best System Analysis (BSA)—which is an analysis of both laws of nature and objective chances—the laws are the theorems of that axiom system pertaining to what goes on in the universe that best balances various theoretical virtues.

²¹ Genuine laws underwrite predictions by serving as *guides* to rational credence about future outcomes. But the probabilities that they entail are not themselves credences. A generalization can serve as a guide to rational credence in the correct way by entailing objective chances.

The objective chances are probabilities that are entailed by that system. According to Lewis (1994, 478, 480), the relevant theoretical virtues are strength (or informativeness), simplicity, and fit (or likelihood).

As Lewis puts it, a system is *strong* to the extent that it says “what will happen or what the chances will be when situations of a certain kind arise” (Lewis 1994, 480). On the other hand, a system is simple to the extent that it comprises relatively few, simple axioms (e.g. linear equations are simpler than polynomials of degree greater than 1). Finally, a system *fits* the actual course of history well (has a higher likelihood) to the extent that it assigns a higher probability to the actual course of history: the higher the probability, the better the fit.²²

To an extent the theoretical virtues trade off against one another: all three can’t be simultaneously maximized. For instance, greater strength can often be achieved by adding axioms to a system, or making the axioms more complicated, and thus comes at a cost in terms of simplicity. The *Best* system is that which strikes the best overall balance between the theoretical virtues.

The principal point of difference between Lewis’s version of the BSA, and variants that have subsequently been developed is the following. Observing that the (syntactic) simplicity of an axiom system is relative to the vocabulary in which it is expressed, Lewis (1983, 367–368) takes only the simplicity of an axiom system when formulated in a language that has only perfectly natural kind terms to be a theoretical virtue for the purposes of determining the Best System. At a stroke, this essentially rules out the existence of laws or chances that are not derivable from the laws and chances of fundamental physics alone. This is because any axiom pertaining to the kinds of a higher level science is likely to be syntactically very complex when translated into a language with only perfectly natural kind terms (compare Schaffer 2007, 130; Cohen and Callender 2009, 14). Consequently, such an axiom is not likely to figure in the Best System.

Many philosophers—though apparently not Lewis (1986b, Postscript B)—find it an unpalatable consequence of Lewis’s formulation of the BSA that either there are no laws or chances of the higher level sciences, or they derive from those of fundamental physics. For instance, because of the simplicity of the Mentaculus, together with the strength and goodness-of-fit that it has on the assumption that it entails the probabilities of SM, Loewer (2001, 2007, 2008, 2012a, 2012b) has argued extensively that it is a plausible Best System for our world (see also Albert 2012).²³ If Albert and Loewer are correct that

²² This notion of fit applies only if there are only finitely many chance events and, in general, where systems don’t incorporate chance distributions over infinite sets (Elga 2004, 68; Frigg and Hoefer 2015, 554). See Elga (2004, esp. 71–72) for an extension of the notion of fit to infinite cases.

²³ If Albert and Loewer are correct about the derivability of SM from the Mentaculus, then, on Lewis’s explication of strength, the Mentaculus is stronger than a system comprising FD alone. That’s because the Mentaculus tells us what the chances will be when certain kinds of situation arise concerning which FD is silent. Consider a thermodynamically isolated system, *S*. The Mentaculus entails the chances for *S*’s future thermodynamic evolution given that a situation of the kind *S* is *in such-and-such a thermodynamic state* arises. FD (even together with suitable bridge principles) does not tell us what will happen or what the chances will be when situations of such a kind arise. Newtonian mechanics, for instance, tells us what will happen when situations of the more specific kind *S* is *located at such-and-such a point in its microphysical phase space* arise. Quantum mechanics tells us what the chances will be when situations of the kind *S* is *in so-and-so a quantum mechanical state* arise. Since thermodynamic states of systems are multiply realizable by points in phase space or by quantum states, the mere fact that *S* is in the relevant thermodynamic state is not enough information for Newtonian mechanics or quantum mechanics to provide us with any predictions about *S*’s future evolution. It is, however, enough for SM to do so.

the Mentaculus entails the probabilities of SM and that the Mentaculus is thus the Best System for the universe, then the probabilities of SM are (according to the BSA) objective chances. As such, these probabilities are able to underwrite objective causal-explanatory relationships and to play the chance role in guiding rational credence. Indeed, if Loewer and Albert are correct then, because the BSA is an analysis of laws as well as chances, any probabilistic theorems of the Mentaculus, such as a probabilistic approximation to SLT, count as full-blown laws of nature according to the BSA.

Yet because the Mentaculus, as formulated by Albert (2000, 96), includes imperfectly natural kind terms like ‘low entropy’, it seems that some modification to Lewis’s version of the BSA would be needed for the Mentaculus to even be a leading contender for Best Systemhood. After all, if translated into a language with only perfectly natural kind terms, the Mentaculus would be syntactically very complex (compare Schaffer 2007, 130).

My preferred way of modifying Lewis’s BSA so that it admits the possibility of Best System axioms invoking imperfectly natural properties is the following.²⁴ Observe that, as Lewis (1983, 347, 368) recognises, naturalness admits of degrees. If present physics is correct, then having spin 1/2 is a perfectly natural property. On the other hand a property like being green is fairly natural, but imperfectly so. Being *grue* is a still less natural property. Lewis’s predicate ‘*F*’ (defined in Lewis 1983, 367) is potentially less natural still.

As Lewis says “[n]atural properties [are] the ones whose sharing makes for resemblance” (Lewis 1983, 347). Of course, the resemblance that the sharing of a property makes for admits of degrees. If we take a set comprising all green beryl crystals that are approximately a particular size, shape, luster, clarity, and so on, and a set comprising all *grue* beryl crystals that are approximately that same size, shape, luster, clarity, and so on, then the members of the former set bear greater resemblance to one another than the members of the latter. But there is more overall resemblance even among members of the latter set than there is among members of a set comprising all *gruellow*²⁵ beryl crystals that approximate that particular size, shape, luster, clarity, etc.²⁶

²⁴ Alternative modifications designed to achieve a similar outcome are suggested by Frisch (2014, Section 5) (compare also Frigg and Hoefer 2015, Section 4) and by Cohen and Callender (2009) and Callender and Cohen (2010) (compare also Schrenk 2008, Callender 2011, and Dunn 2011). These alternatives would work just as well for present purposes. For reasons of space I won’t compare them with my own proposal here.

²⁵ Something is *gruellow* if first observed on or before New Year’s Eve 1999 and green, if first observed between New Year’s Day 2000 and New Years Eve 2019 (inclusive) and blue, or otherwise if yellow.

²⁶ Overall resemblance (in the relevant sense) among a finite set of objects should (I think) be determined by applying a loss function that scores the dissimilarity between each pair of objects in the set, with significant differences being heavily penalized—for instance, a quadratic loss function might be appropriate—and then dividing the total loss by the number of pairs in the set. The lower the resulting number, the greater the overall resemblance among members of the set.

It is of course a difficult question what exactly the correct measure of dissimilarity between a pair of objects is. I doubt that there is a single measure that is clearly superior to all alternatives. Perhaps unsurprisingly, then, the notion of resemblance and hence naturalness is liable to be imprecise. This is not the only respect in which the algorithm for determining the Best System is liable to be imprecise. After all, it is also unlikely that there are single measures of simplicity or strength (or perhaps even fit) that are clearly superior to all others. Nor is it likely that there is a single uniquely most reasonable way of balancing these virtues in the determination of an overall best system. I give my favored response to this ‘problem’ of imprecision for the BSA in Fenton-Glynn (ms.) (see also Dardashti *et al.* 2014).

My suggestion is that the naturalness of the predicates that it employs is a theoretical virtue, to be weighed alongside the simplicity, strength, and fit of a system. If an axiom system is able to achieve great simplicity and strength and fit by employing not-too-unnatural predicates then it is a candidate best system. The Mentaculus employs imperfectly natural but yet not-too-unnatural predicates such as ‘low entropy’. If Albert and Loewer are correct that the Mentaculus entails SM, then it is plausibly a fairly strong candidate for Best Systemhood.

As noted in the previous section, there are those who have expressed skepticism about the claim that the Mentaculus entails SM. In the previous section it was noted that, in contrast to ‘Globalists’, such as Albert and Loewer, who claim that SM can be derived from a probability distribution over the phase space of the universe as a whole together with an axiom concerning the initial entropy of the universe, ‘Localists’ simply claim that SM probabilities are derivable on a system-by-system basis by imposing probability distributions (that are uniform on the Lebesgue measure) over the phase spaces of the various (approximately) isolated thermodynamic systems that exist in the universe, and then conditioning upon *their* initial macrostates. Thus Callender (2011, 96) suggests that “we impose SP at (roughly) the first moment low-entropy macroscopic systems become suitably isolated”.

If the Globalists are wrong, then the Mentaculus is not as strong as we might have thought. If the Localists are right, then we can buy a lot of strength with relatively little cost in simplicity by adding to FD an axiom according to which, for each macroscopic system that is suitably isolated (including the universe as a whole), the probability distribution over its phase space is the one that’s uniform on the Lebesgue measure. The result of adding such an axiom to FD is a relatively simple, but very strong, system: it entails the SM probabilities for entropy increase in each appropriately isolated system, given the (initial) macrostate of that system. So, if the Globalist is wrong but the Localist is correct, then it remains plausible that the Best System entails the SM probabilities. As such, it remains plausible to regard the SM probabilities as objective chances and thus as able to sustain objective causal-explanatory relationships and to support predictions.²⁷ Similarly, the theorems of a Localist axiomatization of SM—which may include a probabilistic version of SLT—will plausibly count as laws of nature according to an appropriate version of the BSA.

It’s possible that an axiom system that entails SM is even stronger than I have so far suggested. That’s because Albert (2000, e.g. 94–96, 2012, esp. 28–33) and Loewer (2007, 306, 2008, 159–162, 2012a, esp. 18) have claimed that SM entails probabilistic approximations to the generalizations of the special sciences.²⁸ Loewer (2008, 160) suggests that such probabilistic generalizations can be derived in the following way (which I illustrate with respect to LE).

Take the probability distribution over possible macroscopic histories of the universe that is entailed by the Mentaculus. Condition that probability distribution upon suitable

²⁷ In this vein, Frigg and Hoefer (2015) endorse a Localist approach to SM, and combine this with an analysis of chance—which they call a ‘Theory of Humean Objective Chance’—that has a lot in common with the BSA. The upshot, they claim, is that locally-derived SM probabilities are objective chances.

²⁸ Since the Mentaculus is their preferred axiomatization of SM, they claim specifically that *the Mentaculus* entails probabilistic approximations to the generalizations of the special sciences. But a Localist could make an analogous argument for the entailment of such probabilistic special science generalizations by SM.

information about the macrostate of an ecosystem and a population (of the sort that the LE usually models well) living within it at a given time t (and perhaps additionally condition upon suitable information about the macrostate of the universe as a whole at t). Then *additionally conditioning upon n , K , and r_c taking certain values at a slightly later time t_1* yields a high probability for macro-histories of the universe in which the population's growth after t_1 is approximately equal to what the LE predicts given the values of n , K , and r_c that are conditioned upon.^{29,30}

If Albert and Loewer are right about the derivability of probabilistic approximations to the special science laws from SM alone, then perhaps it was incorrect to suppose (as I did in the previous section) that the derivation of a stricter probabilistic approximation to LE would essentially invoke state spaces other than the phase space of the universe (or perhaps some approximately isolated subsystem of it). Yet there are grounds for skepticism about Albert and Loewer's claim. More precisely, there are grounds for skepticism about whether any probabilistic generalizations concerning special science phenomena that might be derived simply from SM are the best players of the law role for the special sciences in question. Likewise, there are grounds for skepticism about whether any probabilities for special science phenomena that might be so derived are the best players of the chance role for the special sciences in question.³¹

That this is so is best seen by, firstly, noting the possibility of divergence between any probabilities for special science phenomena that may be entailed by SM alone and those that are entailed by the probabilistic generalizations that special scientists in fact formu-

²⁹ If probabilities for LE-like growth can be derived in this way then they are relativized to a macrostate that obtains at t . If we instead conditioned upon the macrostate of the universe (or at least of the ecosystem) that obtains at some different time, t^* , then we may well get different probabilities for $\frac{dn}{dt}$. As Loewer (2008, 160) admits "[o]n this account the special science laws may change over time (new ones coming into existence and old ones going out of existence)". This feature is surprising, but perhaps tolerable provided that there is typically little change in the accuracy of particular special science generalizations over short to middling periods of time (so that these generalizations can still support predictions and so that the ability of ordinary special scientific confirmatory practices to yield convergence upon (approximations to) these generalizations is not undermined).

³⁰ As Albert (2012, 30n) points out, to actually know what the values of such conditional probabilities are, one would need to know (among other things) what constraints the information about macrostates that is conditioned upon puts on the universe's microstate. That is, one must know what region of the universe's phase space is compatible with such information about macrostates. (Indeed, even if one thought that SM probabilities are only to be derived locally, one would still need to know exactly how the information about macrostates constrains the microstate of some appropriately isolated subsystem of the universe.) This is among the reasons that neither Albert nor Loewer think that the derivation of the special science laws from SM would ever be possible in practice (see Albert 2012, 32; Loewer 2008, 160). Fortunately, special scientists have other ways of discovering and confirming probabilistic special science generalizations (e.g. by fit with frequencies) than by deriving them from first principles.

It is worth noting that there are also practical obstacles to the derivation of a strict probabilistic approximation to LE (and to other special science generalizations) via the procedure that I suggested in the previous section since, according to the account outlined there, the probabilities that figure in such a generalization result from mixing probabilities that include SM probabilities. Still, as we saw there, it is possible—in practice and not simply in principle—to derive a generalization that, for instance, takes account of the various possible geographical distributions and demographic profiles of a population (by assigning probabilities to them) even without taking account of the SM probabilities. Thus the derivation of stricter—though perhaps not perfectly strict—probabilistic approximations to special science laws is perfectly possible in practice provided that we don't insist that the only fruitful procedure for doing so is to invoke a probability distribution over the phase space of a system.

³¹ Among those who express skepticism are Callender and Cohen (2010, 437–439), Callender (2011, 102–103), Dunn (2011, 84), Weslake (2014, Section 3), and Frisch (2014, 233–235).

late. The case for this possible divergence is made rather compellingly by Callender and Cohen (who themselves consider the example of ecology).

“The statistical mechanical chance is based on Lebesgue measure over phase space Ecological systems are sometimes modeled via state spaces with measures on them—sometimes even Lebesgue measure. However, . . . the state spaces are different. The (classical) physical one is parametrized with respect to position and conjugate momentum, the ecological ones are parametrized with respect to ecological variables, such as ‘number of age 1 females’, ‘number of age 2 females’, and so on. Are the generalizations that are highly probable in the one space highly probable in the other? We have no idea, and neither does anyone else. . . . Take a trajectory through one state space that is typical according to the probability measure adapted to that space. Now ‘translate’ this trajectory into another state space. Then it’s certainly possible that, with respect to a probability measure on that second state space, the resulting trajectory is atypical.” (Callender and Cohen 2010, 437–438)

If we consider our example of the LE, the point that Cohen and Callender make is that it seems perfectly possible that an appropriate probability distribution (perhaps the one that’s uniform on the Lebesgue measure) over a state space whose points represent particular gender and age profiles—and perhaps geographical distributions—for a given population should be one according to which LE-like behavior is highly likely, while an appropriate probability distribution (again, perhaps the one that’s uniform on the Lebesgue measure) over the *phase space* of some isolated system of which the population is a part is one according to which the probability of macro-histories involving LE-like growth for the population is rather different, and perhaps even *very low*. So perhaps the probabilistic approximations to LE that ecologists are likely to derive in practice (from distributions over the former sort of state space) bear little resemblance to any probabilistic principle for population growth that might be derived via the Albert-Loewer proposal, which simply involves imposing a probability distribution over phase space.

The suggestion that I made in the previous section about how probabilistic approximations to special science generalizations like LE might be derived does not require that SM itself make these generalizations probable. In the case of LE, what it requires is that the probability distribution over the state space S' whose points represent particular demographic profiles and geographical distributions for the population makes LE-like behavior probable (for the types of population to which LE normally applies). Whether this is so depends upon whether a high probability is assigned by this distribution to the region of S' comprising those points that have, according to the dynamical equations governing trajectories through S' , a high probability of resulting in trajectories that realize LE-like behavior.

According to the suggestion made in the previous section, the probabilities of SM are only taken into account when it comes to the derivation of the dynamical equations concerning trajectories through S' , or perhaps in the derivation of the dynamical equations concerning trajectories through some still more fundamental state space (such as the biochemical state space, \mathcal{R}). According to that suggestion, the ecological probabilities are not simply SM probabilities, but rather result from probability mixing, where the SM probabilities are only one element of the mix (the others being probability distributions over state spaces, such as S' and \mathcal{R} , that are more coarse grained than a phase space).

The probability for LE-like behavior entailed by SM may well diverge from that entailed by the distribution that results from this probability mixing.

Two obvious questions now arise. Firstly, are there grounds for expecting that there is significant divergence between any probabilities for LE-like growth that can be derived simply from SM and those whose derivation (additionally) involves appeal to distributions over more coarse-grained state spaces representing (e.g.) the demographic profile and geographical distribution of members of the population? Second, which of these probabilities are likely to be more accurate when used to make predictions? There is some reason for thinking that the answer to the first question is ‘yes’ and that the answer to the second question (which becomes more pressing given an affirmative answer to the first) is that it is not the probabilities that are derived simply from SM that are likely to form the basis for more accurate predictions. The foregoing answers are made plausible by the following reasoning, regarding the Mentaculus, that is advanced by Frisch.³²

“If we wanted to insist that the Mentaculus provides us with a universal physics and imagined the axioms being chosen by a Laplacian demon, who can strike the best balance among the criteria of theory choice for the entire Humean mosaic, then the domain of a higher-level science is a proper subset of that of the Mentaculus. But even the Laplacian demon could discover that [a] domain-specific theory does a better job within the very domain for which it is designed at balancing the criteria within its domain . . . than the Mentaculus, which has to strike a balance between simplicity and fit across many different domains.” (Frisch 2014, 235)

The idea here is as follows. The Mentaculus, if Albert and Loewer are correct, is an extremely simple, strong, and well-fitting system pertaining to what goes on in the universe as a whole. But the special sciences are not concerned with what goes on in the universe as a whole. Rather, they are concerned with what goes on within more limited domains. Suppose that a Laplacian special science demon wanted to formulate a system that struck the best balance between simplicity, on the one hand, and strength and fit with respect to (for example) ecological phenomena, on the other. Then, even supposing that Albert and Loewer are correct in thinking that the Mentaculus *does* entail probabilities for ecological phenomena, it is plausible that the probabilities entailed by the system that the special science demon would come up with for ecology would be even more accurate with respect to ecological phenomena. This is because achieving a pretty good overall fit across all domains (without adding more, or more complex, axioms than those that appear in the Mentaculus) will likely mean non-negligible divergence from the actual frequencies in many if not all particular domains.

Very similar reasoning applies if one prefers a Localist approach to the derivation of SM probabilities. As I suggested above, it is plausible that the best *Localist* axiomatization of SM is one that adds to FD an axiom according to which, for each (approximately) isolated system, the correct probability distribution over that system’s phase space is the one that is uniform on the Lebesgue measure. This axiom concerning probability distributions over phase spaces is designed to give accurate predictions for the frequencies of SM behavior across all thermodynamic systems in the universe (given macroscopic information about those systems). It is not designed to be as faithful as possible to the ecological frequencies.

³² See Callender and Cohen (2010, 444) and Hoefer (2007, 592–593) for similar reasoning. As I explain in the main text below, the reasoning readily generalizes to cover Localist approaches to SM too.

I suggest that the axiom system for ecology that a special science demon would devise would be one on which probabilities for ecological phenomena are derived in the way described in the previous section. That is, the axiom system would be one that involves, not just an axiom entailing probability distributions over phase spaces (and an axiom entailing the fundamental dynamics), but also axioms entailing appropriate probability distributions (i.e. probability distributions that give rise to accurate predictions) over more coarse-grained state spaces. My reason for thinking this is that practicing ecologists derive their best predictions for ecological phenomena via models that involve distributions over such coarse-grained state spaces. Moreover, they are (with due respect to them) just like cognitively limited special science demons.

Now it might be claimed that it is precisely due to their cognitive limitations that ecologists deal with coarse-grained state spaces: it would be intractable for them (unlike a special science demon) to deal with probability distributions over phase spaces in making their predictions. Yet while they don't have quite the intellect of a special science demon, practicing ecologists nevertheless have the same aim: to formulate the best possible system *for ecological phenomena*. In light of Frisch's reasoning (quoted above) it would be surprising if relieving practicing ecologists of their limitations led them simply to drop their current models (which do a pretty good job of modeling ecological phenomena) and adopt the Mentaculus-derived probabilities (which, if Frisch is correct, may well do a pretty bad job of predicting ecological phenomena) for use in their predictions.³³ It seems more likely that they would instead augment their models by incorporating SM probabilities in what appears to be their proper place: namely as influencing the evolution of an ecosystem through a fairly fundamental state space, such as the biochemical state space, \mathcal{R} , described in the previous section.

To the extent that the theorems and probabilities of the best ecological system *are* more accurate concerning ecological phenomena than those entailed by the Mentaculus (or the best Localist system for SM)³⁴, they underwrite better ecological predictions, and support more reliable counterfactual reasoning about ecological phenomena than those entailed by the Mentaculus (compare Hoefer 2007, 592–593). Thus, in cases of divergence, it appears that the theorems and probabilities of the best ecological system play the law and chance roles with respect to ecology better than do whatever generalizations and probabilities concerning ecology might be derivable from the Mentaculus.

I take it that the correct thing to say at this point is not that the Mentaculus, or perhaps some Localist axiomatization of SM, strikes the best balance between simplicity and strength and fit for what goes on in the universe as a whole, and is therefore the Best System for it, and that the more accurate generalizations and probabilities for ecology that are entailed by the best ecological system are therefore not laws and chances. This would undermine the plausibility of the BSA. After all, it would be to admit that there are generalizations and probabilities for ecology that are better players of the law and chance roles than the generalizations and probabilities for ecology entailed by the Best System. (Indeed, the generalizations and probabilities entailed by the best ecological system appear to be *very good* players of the law and chance roles, and not merely *better* players of those roles than the generalizations and probabilities for ecology entailed by SM.)

³³ Indeed, if the systems that practicing scientists devise do not at all resemble the systems that Laplacian demons would devise, then the BSA seems to imply a rather strong form of inductive skepticism.

³⁴ I will drop this qualification in what follows.

Rather, the correct thing to say is surely that the right way of balancing simplicity against strength and fit is such that a system including axioms that entail these more accurate generalizations and probabilities for ecology comes out Best. In other words the axioms (e.g. those specifying probability distributions over higher level state spaces) that are needed, in addition to those included in the best system for SM, to entail these more accurate generalizations and probabilities for ecology should be added to the best system for SM to arrive at an overall best system for the world.³⁵

Similar reasoning should of course lead us to believe that, for each of the other special sciences, the best system for the world will also include the axioms needed to entail accurate probabilistic approximations to the *mr* laws of that science, provided that, for the special science in question, those axioms aren't too many or complex. The probabilistic generalizations entailed will thus constitute probabilistic laws, and the probabilities that they entail will constitute objective chances according to the BSA.

At this point it's natural to wonder how the *cp* nature of many special science generalizations can be reconciled with their being Best System laws. After all, so far I have said nothing to suggest that a probabilistic approximation to LE, or probabilistic approximations to special science laws more generally, will not—like their deterministic, *mr*, counterparts—hold only *cp*.

Schrenk (2008) has some helpful things to say in this regard. He points out that versions of *cp* generalizations that, instead of a *cp* clause, include an explicit list of all of the situations in which they fail to apply, may well be theorems of the Best System. This is so even if those explicit lists are long and complex: after all, it is only the simplicity of the *axioms* and not of the *theorems* that is relevant to determining whether a system is *Best*. If so, Schrenk suggests, the original *cp* generalization should count as a genuine *cp law*.³⁶

³⁵ One might worry that, if SM entails probabilities for ecological phenomena and the Best System includes axioms that entail SM as well as axioms that entail more accurate probabilistic generalizations for ecology, then this leaves us with two divergent sets of chances for ecological phenomena and that this leads to contradiction via the Principal Principle (PP) (Lewis 1980).

This worry is misplaced, however. If one knew both the SM-derived probabilities for ecological phenomena and the probabilities entailed by the more accurate ecological generalizations, then it would be reasonable to set one's credences to the latter probabilities (provided that one didn't, for example, already know what the relevant ecological outcomes were). Now, according to the PP, chance guides reasonable credence in the absence of inadmissible information. So consider the claim that the SM-derived probabilities for ecological phenomena play the chance role in guiding reasonable credence. Either the more accurate probabilities for ecological phenomena are admissible or they are not. If they are admissible (which plausibly they *are* on the explication of the notion of admissibility given by Lewis 1980), then the SM-derived probabilities fail to play the PP role and so are not chances (thus only a proper subset of the lawfully-derived probabilities for ecological phenomena are chances). If, on the other hand, they are inadmissible then no contradiction is derivable (just as no contradiction is derivable via the PP from the assumption that an event that has a certain chance at one time has a different chance at another time).

For additional arguments that there is no problematic conflict between SM-derived generalizations and probabilities and those derived from axioms pertaining specifically to special sciences like ecology, see Cohen and Callender (2009, 27–30). For arguments that there is no problematic conflict between SM- and special-science-derived probabilities and any probabilities for the same phenomena that might be derived from FD *alone*, see for example Loewer (2001), Hoefer (2007), and Glynn (2010).

³⁶ Although there may well be a great many *cp* laws if Schrenk's account is correct, the ones that we actually use in making predictions and giving explanations are presumably those such that the situations in which they fail to apply are not very run-of-the-mill or at least are known not to obtain with regard to the system concerning which we are making a prediction or giving an explanation.

If what I've said so far about LE is correct, then it is plausible that the best system for our universe entails that, if there is no natural disaster or cull or nuclear war or . . . , then a probabilistic approximation to LE derived in the way suggested in the previous section holds with respect to the relevant populations. On Schrenk's rather plausible view, then, this probabilistic approximation to LE counts as a genuine cp law.

Schrenk's proposal strikes me as a satisfactory way of handling cp generalizations within the framework of a BSA-style analysis of laws and chances. Nevertheless, there is one other interesting thing to say about cp laws in the present context, which might mean that we don't even need to appeal to Schrenk's proposal.

In the previous section, I suggested that the derivation of the most accurate probabilistic approximation to LE is likely to involve, as well as probability distributions over more coarse grained state spaces, a distribution over the phase space of a thermodynamically isolated system of which a given population is a part. As indicated in Section 3, there is no possibility of interference from outside thermodynamically isolated systems.³⁷ Moreover, a probability distribution over the phase space of a thermodynamically isolated system of which a given population is a part will yield a probability that the microstate of that system is one of those that, when evolved forward under the fundamental dynamics, realizes a macro-history in which there is a natural disaster that affects the population, or there is a cull, or a nuclear war, or Thus, the most accurate probabilistic approximations to special science mr laws may no longer hold only cp, for the probabilities that they entail *incorporate* the probabilities of the events that would violate the original mr law's cp clause (in virtue of the role in their derivation played by a probability distribution over the phase space of a suitable isolated system). If this is right, then the correct response to the problem of mr laws may in fact also constitute a solution to the problem of cp clauses.

In the previous section I argued that it's plausible that SLT and LE, which are examples of deterministic laws that hold only mr are approximations to stricter probabilistic generalizations that don't hold only mr. As well as setting out the more familiar story about how a probabilistic approximation to SLT can be derived, I also described how a probabilistic approximation to LE might be derived by the imposition of probability distributions over underlying state spaces. In this section I have argued—drawing upon the BSA—that there is reason to think that the resulting probabilistic generalizations are laws and that the probabilities that they entail are chances. Of course a lot of work needs to be done to show that we can *always* replace deterministic special science generalizations that hold only mr with strict(er) probabilistic laws by imposing probability distributions over underlying state spaces. But the examples of SLT and LE establish a blueprint that is plausibly applicable elsewhere in the high-level sciences. In any case, clearly distinguishing the (non-mutually exclusive) categories of cp generalizations and mr generalizations helps to bring the challenges for science and metaphysics posed by the mr nature of many special science generalizations into sharper relief, thus enabling us to get a better sense of how these challenges might best be addressed.

³⁷ Although this is only true of genuinely thermodynamically isolated systems, as opposed to merely approximately isolated ones.

8. Conclusion

The notion of a non-exceptionless law shouldn't be equated with that of a cp law. There is another important category of non-exceptionless law that ought to be distinguished, viz. mr laws. The mr nature of high-level scientific generalizations poses distinctive challenges for those aiming to show that high-level scientific generalizations can support counterfactuals, causal explanations, and predictions. Distinguishing the two categories of non-exceptionless generalization brings these challenges into sharper relief, but also allows us to identify a possible avenue for addressing them.³⁸

References

- Albert, D. (2000). *Time and Chance*. Cambridge, MA: Harvard University Press.
- (2012). Physics and chance. In Y. Ben-Menahem and M. Hemmo (Eds.), *Probability in Physics*, pp. 17–40. Berlin: Springer.
- Berryman, A. (2003). On principles, laws and theory in population ecology. *Oikos* 103, 695–701.
- Callender, C. (2011). The past histories of molecules. In C. Beisbart and S. Hartmann (Eds.), *Probabilities in Physics*, pp. 83–113. New York: OUP.
- and J. Cohen (2010). Special sciences, conspiracy and the better best system account of lawhood. *Erkenntnis* 73, 427–447.
- Cartwright, N. (1983). *How the Laws of Physics Lie*. Oxford: OUP.
- Coe, J., S. Ahnert, and T. Fink (2008). When are cellular automata random? *Europhysics Letters* 84, 50005: 1–6.
- Cohen, J. and C. Callender (2009). A better best system account of lawhood. *Philosophical Studies* 145, 1–34.
- Colyvan, M. and L. Ginzburg (2003). Laws of nature and laws of ecology. *Oikos* 101, 649–653.
- Dardashti, R., L. Glynn, K. Thébault, and M. Frisch (2014). Unsharp Humean chances in statistical physics: A reply to Beisbart. In M. Galavotti, D. Dieks, W. Gonzalez, S. Hartmann, T. Uebel, and T. Weber (Eds.), *New Directions in the Philosophy of Science*, pp. 531–542. Dordrecht: Springer.
- Davidson, D. (1970). Mental events. In L. Foster and J. Swanson (Eds.), *Experience and Theory*, pp. 79–101. Amherst: University of Massachusetts Press.
- Dunn, J. (2011). Fried eggs, thermodynamics, and the special sciences. *The British Journal for the Philosophy of Science* 62, 71–98.

³⁸ Versions of this paper were presented at the 2013 British Society for the Philosophy of Science Annual Meeting at the University of Exeter, at the 2014 Philosophy of Science Association Biennial Meeting in Chicago, and at the 2015 Society for the Metaphysics of Science Annual Meeting at Rutgers University. Versions of this paper have also been presented at the University of Cologne, the University of Wuppertal, the Munich Center for Mathematical Philosophy, the Central European University, and to the Serious Metaphysics Group at the University of Cambridge. I would like to gratefully acknowledge the comments of the audience members present on these occasions. I would also like to thank Christopher H. Eliot, Thomas Kroedel, and Wolfgang Spohn for detailed comments. I would like to acknowledge the support of the Deutsche Forschungsgemeinschaft (Grant Number: SP279/15-1), the McDonnell Causal Learning Collaborative, and the Alexander von Humboldt Foundation. Finally, thanks to an anonymous referee for this journal.

- Earman, J. (2006). The “Past Hypothesis”: Not even false. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 37, 399–430.
- and J. Roberts (1999). Ceteris paribus, there is no problem of provisos. *Synthese* 118, 439–478.
- Elga, A. (2001). Statistical mechanics and the asymmetry of counterfactual dependence. *Philosophy of Science* 68, 313–324.
- (2004). Infinitesimal chances and the laws of nature. *Australasian Journal of Philosophy* 82, 67–76.
- Fenton-Glynn, L. (ms.). Imprecise Best System chances. Unpublished Manuscript; Available at: http://www.academia.edu/5819239/Imprecise_Best_System_Chances
- Fodor, J. (1989). Making mind matter more. *Philosophical Topics* 17, 59–79.
- (1991). You can fool some of the people all of the time, everything else being equal: Hedged laws and psychological explanations. *Mind* 100, 19–34.
- Frigg, R. and C. Hoefer (2015). The best Humean system for statistical mechanics. *Erkenntnis* 80, 551–574.
- Frisch, M. (2014). Why physics can’t explain everything. In A. Wilson (Ed.), *Chance and Temporal Asymmetry*, pp. 221–240. Oxford: OUP.
- Gardner, M. (1970). Mathematical games: The fantastic combinations of John Conway’s new solitaire game ‘Life’. *Scientific American* 223, 120–123.
- Glynn, L. (2010). Deterministic chance. *The British Journal for the Philosophy of Science* 61, 51–80.
- Hájek, A. (ms.). Most counterfactuals are false. Unpublished Manuscript; Available at: <http://tinyurl.com/ou6o63k>
- Hastings, A. (1993). Complex interactions between dispersal and dynamics: Lessons from coupled logistic equations. *Ecology* 74, 1362–1372.
- Hiebeler, D. (2012). Cellular automata. In A. Hastings and L. Gross (Eds.), *Encyclopedia of Theoretical Ecology*, pp. 123–126. Berkeley & Los Angeles, CA: University of California Press.
- Hitchcock, C. (2001). The intransitivity of causation revealed in equations and graphs. *Journal of Philosophy* 98, 273–299.
- Hoefer, C. (2007). The third way on objective probability: A sceptic’s guide to objective chance. *Mind* 116, 549–596.
- Law, R., D. Murrell, and U. Dieckmann (2003). Population growth in space and time: Spatial logistic equations. *Ecology* 84, 252–262.
- Lawton, J. (1999). Are there general laws in ecology? *Oikos* 84, 177–192.
- Lepore, E. and B. Loewer (1987). Mind matters. *Journal of Philosophy* 84, 630–642.
- Lewis, D. (1979). Counterfactual dependence and time’s arrow. *Noûs* 13, 455–476.
- (1980). A subjectivist’s guide to objective chance. In R. Jeffrey (Ed.), *Studies in Logic and Inductive Probability*, Volume 2, pp. 267–297. Berkeley: University of California Press.
- (1983). New work for a theory of universals. *Australasian Journal of Philosophy* 61, 343–377.
- (1986a). *Philosophical Papers*, Volume 2. New York: OUP.
- (1986b). Postscripts to “A subjectivist’s guide to objective chance”. In D. Lewis (Ed.), *Philosophical Papers*, Volume 2, pp. 114–132. New York: OUP.
- (1994). Humean supervenience debugged. *Mind* 103, 473–490.

- Loewer, B. (2001). Determinism and chance. *Studies in History and Philosophy of Science Part B* 32, 609–620.
- (2007). Counterfactuals and the second law. In H. Price and R. Corry (Eds.), *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*, pp. 293–326. New York: OUP.
- (2008). Why there is anything except physics. In J. Hohwy and J. Kallestrup (Eds.), *Being Reduced*, pp. 149–163. Oxford: OUP.
- (2012a). The emergence of time's arrows and special science laws from physics. *Interface focus* 2, 13–19.
- (2012b). Two accounts of laws and time. *Philosophical Studies* 160, 115–137.
- May, R. (1974). Biological populations with nonoverlapping generations: Stable points, stable cycles, and chaos. *Science* 186, 645–647.
- Mikkelsen, G. (2003). Ecological kinds and ecological laws. *Philosophy of Science* 70, 1390–1400.
- Murray, B. (2000). Universal laws and predictive theory in ecology and evolution. *Oikos* 89, 403–408.
- Otto, S. and T. Day (2007). *A Biologist's Guide to Mathematical Modeling in Ecology and Evolution*. Princeton, NJ: Princeton University Press.
- Reutlinger, A., G. Schurz, and A. Hüttemann (2011). Ceteris paribus laws. In E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/entries/ceteris-paribus/>
- Schaffer, J. (2007). Deterministic chance? *The British Journal for the Philosophy of Science* 58, 113–140.
- Schrenk, M. (2007). Can capacities rescue us from ceteris paribus laws? In M. Kistler and B. Gnassounou (Eds.), *Dispositions in Philosophy and Science*, pp. 221–247. Aldershot: Ashgate.
- (2008). A Lewisian theory for special science laws. In S. Walter and H. Bohse (Eds.), *Ausgewählte Beiträge aus den Sektionen der GAP*, Volume 6, pp. 121–131. Paderborn: Mentis.
- Schurz, G. (2002). Ceteris paribus laws: Classification and deconstruction. *Erkenntnis* 57, 351–372.
- (2014). Ceteris paribus and ceteris rectis laws: Content and causal role. *Erkenntnis* 79, 1801–1817.
- Stalnaker, R. (1968). A theory of conditionals. In N. Rescher (Ed.), *Studies in Logical Theory, American Philosophical Quarterly* (Monograph Series, Vol. 2), pp. 98–112. Oxford: Blackwell.
- Tsoularis, A. and J. Wallace (2002). Analysis of logistic growth models. *Mathematical Biosciences* 179, 21–55.
- Turchin, P. (2001). Does population ecology have general laws? *Oikos* 94, 17–26.
- Vandermeer, J. and D. Goldberg (2013). *Population Ecology* (2nd ed.) Princeton, NJ: Princeton University Press.
- Weslake, B. (2014). Statistical mechanical imperialism. In A. Wilson (Ed.), *Asymmetries of Chance and Time*. Oxford: OUP.
- Winsberg, E. (2004). Can conditioning on the “Past Hypothesis” militate against the reversibility objections? *Philosophy of Science* 71, 489–504.
- Woodward, J. (2002). There is no such thing as a ceteris paribus law. *Erkenntnis* 57, 303–328.

- (2005). *Making Things Happen*. New York: OUP.
- and C. Hitchcock (2003a). Explanatory generalizations, Part I: A counterfactual account. *Noûs* 37, 1–24.
- and — (2003b). Explanatory generalizations, Part II: Plumbing explanatory depth. *Noûs* 37, 181–199.