

# IS THERE HIGH-LEVEL CAUSATION?

LUKE FENTON-GLYNN

*University College London*

The discovery of causal relations seems a central activity of the high-level sciences, including the special sciences and certain branches of macrophysics. Those same sciences are less successful in formulating exceptionless laws. If causation must be underwritten by exceptionless laws, we are faced with a puzzle. Attempts have been made to dissolve this puzzle by showing that non-exceptionless generalizations can underwrite causal relations. The trouble is that many of these attempts fail to distinguish between two importantly different types of exception of which high-level scientific generalizations admit. Roughly speaking, one is where the values of high-level variables not represented in the generalization are abnormal: call these ‘background factor’ (*bf*) exceptions. For example, the Ideal Gas Law (IGL) may be significantly violated by a gas if a strong electric current is passed through it. Another is where the high-level states that *are* represented by variables in the generalization are *realized* in certain abnormal ways: call these ‘*mr* exceptions’ (exceptions having to do with the *multiple realizability* of high-level states). For example, the pressure of a gas may not be proportional to its temperature and volume in the way that the IGL describes if the initial macrostate of the gas is realized in a certain unusual microphysical way. While existing attempts to show that non-exceptionless generalizations can underwrite causal relations tend to work well where the generalization admits only of *bf* exceptions, they work less well when the generalizations in question admit—as most high-level scientific generalizations do—of *mr* exceptions. I argue that the best prospect for resolving the apparent problem posed by *mr* exceptions is to regard the generalizations which admit of them as approximations to *probabilistic* generalizations which don’t, and which are themselves able to support relations of probabilistic causation.

## 1. Introduction

The ‘high-level’ sciences are sciences other than fundamental physics. These include certain *macro-physical* sciences, such as thermodynamics, and also *special*

---

**Contact:** Luke Fenton-Glynn <l.glynn@ucl.ac.uk >

*sciences*, such as biology, chemistry, ecology, and economics. The search for and discovery of causes seems a central activity of these high-level sciences.<sup>1</sup> For example, we are told in school physics lessons that decreasing the volume of a container of gas will increase its pressure (if temperature and gas amount are held constant). Meanwhile ecologists tell us that a reduction in the size of a population may cause an increase in its growth rate, cellular biologists tell us that gene mutations can cause cancer, and economists tell us that the Fed's lowering of interest rates may cause a rise in inflation.

Davidson (1970) famously thought that causal relations must be underwritten by exceptionless laws. If he were right, we would have a puzzle on our hands since the generalizations formulated by practitioners of the high-level sciences are typically non-exceptionless. Yet philosophers including LePore and Loewer (1987), Fodor (1989), and Hitchcock and Woodward (2003a; 2003b) have argued that—by the lights of the most popular contemporary theories of causation—non-exceptionless generalizations *can* underwrite causation.

In this paper, I will distinguish two types of exception of which high-level generalizations typically admit. Roughly speaking, one is where the values of high-level variables not represented in the generalization in question are abnormal or changeable: call these '*background factor*' (*bf*) exceptions. For example, the Ideal Gas Law (IGL) may be significantly violated by a gas if a strong electric current is passed through it. Another type of exception is where the high-level states that *are* represented by variables in the generalization are *realized* in certain abnormal ways. These are exceptions that arise due to the *multiple realizability* of high-level states: call them '*mr* exceptions'. For example, the pressure of a gas may not be proportional to its temperature and volume in the way described by the IGL if the initial macrostate of the gas is realized in a certain unusual micro-physical way.

I will argue that, while existing arguments succeed in showing that generalizations that admit only of *bf* exceptions are able to underwrite causal relations, they do not show that generalizations that admit of *mr* exceptions are able to do so, and that *mr* exceptions pose a greater problem for the ability of generalizations to underwrite causal relations. This is important because most high-level generalizations employed by scientists admit of *mr* exceptions as well as *bf* exceptions.

The distinction between these two classes of exception is not new: similar distinctions have been made by Fodor (1991), Hofer (2004), and Fenton-Glynn (2016). The distinction is also implicit in the earlier work of Fodor (1974; 1989) and Schiffer (1991) (cf. Earman & Roberts 1999). Hofer reaches the view that the

---

1. I won't discuss whether *fundamental physics* seeks or discovers causes. For helpful discussion see Price and Corry (2007) and Frisch (2014a).

type of exception that I'm calling '*mr* exceptions' is "either in tension with, or in outright conflict with" the existence of "*genuine* causation—causation of a robust metaphysical type, with some genuine modal component, and applicable at the level of ordinary events" (2004: 113). It is this tension that I will seek to dissolve at the end of this essay.

In what follows, I will explain the distinction and relation between 'high-level' sciences and fundamental physics more precisely (Section 2). I will then more precisely explicate the notion of a *bf* exception (Section 3) and explain why—despite *prima facie* appearances—*bf* exceptions don't pose problems for a generalization's ability to underwrite causal relations (Section 4). I will then explicate the notion of an *mr* exception (Section 5) and explain why such exceptions pose more of a problem (Section 6). Finally, I will propose and defend a solution to the problem posed by the fact that high-level generalizations typically admit of *mr* exceptions (Section 7 and Section 8). Specifically, I'll argue that high-level generalizations that admit of *mr* exceptions (which is plausibly all, or almost all, of them) are approximations to probabilistic high-level generalizations that don't, and which are able to sustain relations of (probabilistic) high-level causation. Along the way (Section 7) I will respond to objections due to Hofer (2004) to the sort of solution that I pursue.

## 2. High-Level Sciences

As I indicated above, the 'high-level' sciences include certain macrophysical sciences—such as thermodynamics—as well as 'special' sciences—such as biology. These sciences are aptly described as 'high-level' because the states they concern are *multiply realizable* by the states of concern to microphysics.<sup>2</sup> Take two sets of possible states of a system: *P* and *Q*. States in *Q* are multiply realizable by states in *P* iff

(MR) For each state  $q \in Q$  there is some  $P' \subseteq P$  such that, (i) necessarily, for each state  $p_i \in P'$ , any system that is in state  $p_i$  is in state  $q$ ; (ii)

---

2. I prefer to talk about 'states' rather than 'properties' because scientific generalizations generally concern the possible states of types of system (see Friend 2016), such as ecosystems, free market economies, and thermodynamically isolated systems. (I won't attempt to give a metaphysics of systemhood—for helpful remarks, see Schurz 2001: 479—but will confine myself to giving examples of paradigm cases.) I take a state of a system just to be a complex of properties that it instantiates and/or a complex of properties instantiated by its parts. In what follows, I'll follow standard practice in taking the state of a system to be representable by a vector of variables. The value of a variable may represent whether or not the system, or one of its parts, possesses a certain property (such as viscosity), or what determinate of a determinable property (such as temperature) it possesses.

necessarily, any system that is in state  $q$  is in some state  $p_i \in P'$ ; and (iii) there is a pair of states  $p_i, p_j \in P'$  such that, possibly, some system is in state  $p_i$  but not in state  $p_j$  and, possibly, some system is in state  $p_j$  but not in state  $p_i$ .<sup>3</sup>

The notion that the possible microphysical states of a system multiply realize other of its states is drawn upon by List and Pivato (2015), who give what is perhaps the most rigorous explication to date of the notion that different sciences are concerned with different 'levels' of reality.

A system (which may be the universe as a whole, or some subsystem of it, such as a gas in a container) has a microphysical state space, which List and Pivato (2015: 121) denote  $S$ . In classical physics,<sup>4</sup> the microphysical state space is a *phase space* of  $6N$  dimensions, with three spatial and three momentum dimensions for each of the  $N$  particles that compose the system. A point in this space corresponds to a microstate of the system, that is, a precise position and momentum for each of the particles it comprises.

List and Pivato (2015: 121) take a microphysical *history* of a system to be representable by a function  $h(\cdot)$  from times into  $S$ , with  $h(t)$  being the microstate of the system at time  $t$ . On List and Pivato's account, then, a system's micro-history is a path through its phase space  $S$ . In their formalism,  $\Omega$  is that subset of all (logically) possible micro-histories that are nomologically possible (specifically, possible according to the fundamental dynamical laws). For any micro-history  $h(\cdot)$  and time  $t$ , they take  $h_t(\cdot)$  to denote the restriction of  $h(\cdot)$  to all times up to  $t$ —what List and Pivato (2015: 121) call a 'truncated history' of the system. A *continuation* of  $h_t(\cdot)$  is a micro-history  $h'(\cdot)$  such that  $h'_t(\cdot) \equiv h_t(\cdot)$ . The laws governing a system's path through  $S$  are deterministic iff for every truncation of a history in  $\Omega$ , there is only one continuation of that history in  $\Omega$ . For our purposes, it will also be useful to define the notion of a future-truncation  $h_{t+}(\cdot)$  of  $h(\cdot)$ , which denotes the restriction of  $h(\cdot)$  to all times *after*  $t$ .

---

3. This definition implies that the states in  $Q$  *supervene* upon those in  $P$ . I think this is appropriate, since it doesn't seem that a state could properly count as a realizer of some other state unless the former state necessitated the latter. I suppose that a functionalist *might* object, claiming that a state can properly be said to realize another, even though the instantiation of the former necessitates the instantiation of the latter *only when the former stands in the right causal/nomological relations to other low-level states*. In response, we could simply tweak our definition of multiple realizability to allow realizers that necessitate their realizees only given the external relations that they stand in. Nothing in what follows turns on these subtleties.

4. In what follows, I will, for the most part, assume that microphysics is classical and deterministic. This is merely for simplicity, and because nothing of substance in what follows turns on this assumption. Things could readily be re-cast to accommodate the assumption that microphysics is quantum mechanical and probabilistic. For a start, this would involve replacing the assumption that a system's basic state space is a phase space with the assumption that it is a space of all possible *quantum* states of the system.

List and Pivato (2015: 131–132) take a higher-level state to correspond to an equivalence class of microphysical states, comprising all the possible realizations of that high-level state by microphysical states of the system. They call a partition of  $S$  into a set of equivalence classes a *coarse graining* of  $S$  and denote this new coarse-grained state space  $\mathbb{S}$ , with each element  $s \in \mathbb{S}$  representing a higher-level state of the system. The elements of  $\mathbb{S}$  are thus multiply realizable by the elements of  $S$ . A history through  $\mathbb{S}$ , denoted  $\mathbb{h}(\cdot)$ , is a function from times into  $\mathbb{S}$ , with  $\mathbb{h}(t)$  being the state in  $\mathbb{S}$  that the system occupies at  $t$  (List & Pivato 2015: 132). Moreover  $\mathbb{h}_t(\cdot)$  can be used to denote the restriction of  $\mathbb{h}(\cdot)$  to times up to  $t$ : that is,  $\mathbb{h}_t(\cdot)$  is a ‘truncated history’ through  $\mathbb{S}$ . A *continuation* of  $\mathbb{h}_t(\cdot)$  is a history  $\mathbb{h}'(\cdot)$  such that  $\mathbb{h}'_t(\cdot) \equiv \mathbb{h}_t(\cdot)$ . We can also define the notion of a future-truncation  $\mathbb{h}_{t+}(\cdot)$  of  $\mathbb{h}(\cdot)$ , which denotes the restriction of  $\mathbb{h}(\cdot)$  to all times *after*  $t$ .

The special sciences, as well as certain branches of macrophysics, such as thermodynamics, are concerned with such coarse-grained states of systems (I take this to be why they are correctly characterizable as ‘high-level’). But not every coarse-grained state space,  $\mathbb{S}$ , is likely to be the concern of some or other actual high-level science.<sup>5</sup> Some may be too coarse to be of interest (e.g., if  $\mathbb{S}$  contains only one element—the state the system occupies iff some necessary truth holds); some may be too fine-grained to be of interest (e.g., if  $\mathbb{S}$  is only a slight coarse-graining of  $S$ , then the relative ease of formulating and using generalizations in terms of the states represented by  $\mathbb{S}$  may not outweigh the loss of accuracy relative to working with the states represented by  $S$ ); some may just be coarse-grained in the wrong ways (e.g., if  $\mathbb{S}$  contains elements like  $s_{grue}$  which the system occupies iff *either* it is in thermodynamic equilibrium *or* it has a sub-region which has temperature 300°K). The coarse-grained states of systems that special scientists actually concern themselves with are those which have been found to be tractable and fruitful to work with for creatures with our epistemic and technological capacities and perhaps with our predilection for certain classificatory schemes rather than others.

One nice feature of the picture presented by List and Pivato (2015: 150 Footnote) is that it doesn’t imply that the relation ‘is higher level than’ generates a total order. It may be that, for two different coarse-grainings of  $S$ ,  $\mathbb{S}_1$  and  $\mathbb{S}_2$ , it is true neither that  $\mathbb{S}_1$  is higher-level than  $\mathbb{S}_2$ , nor that  $\mathbb{S}_2$  is higher-level than  $\mathbb{S}_1$ , nor even that  $\mathbb{S}_1$  and  $\mathbb{S}_2$  are equally high-level. One way this could be the case is if there is at least one element  $s_1 \in \mathbb{S}_1$  that corresponds to an equivalence class of elements of  $\mathbb{S}_2$  *and* there is at least one element  $s_2 \in \mathbb{S}_2$  that corresponds to an equivalence class of elements of  $\mathbb{S}_1$ . Another way is if there are elements of  $\mathbb{S}_1$  and  $\mathbb{S}_2$  that ‘cross-cut’ each other (as the states of being positively charged and

---

5. Cohen and Callender (2009: 22–24) make an analogous point, but express it in terms of languages/classificatory schemes rather than coarse-grainings *per se*.

negatively charged cross-cut the gruesome states of being negatively charged and positively charged<sup>6</sup>), so that it's true of neither  $S_1$  nor  $S_2$  that all of its elements correspond to equivalence classes of elements of the other (nor are their elements identical).

Instead, the relation 'is higher-level than', as List and Pivato explicate it, merely generates a partial order. This seems to be precisely as it should be if the relation is one in which the various sciences are supposed to stand to one another. For example, the question of whether the ecological state of an ecosystem is a higher-, lower-, or same-level state as its thermodynamic state seems unanswerable (cf. Kim 2002). List and Pivato's account provides a satisfying explanation of why: it is not the case that thermodynamic states are identical to ecological states, nor is it the case the set of thermodynamic states corresponds to a set of equivalence classes of ecological states nor vice versa: one can almost certainly hold the ecological state of an ecosystem fixed while slightly varying its thermodynamic state (e.g., by slightly warming or cooling part of the ecosystem for a brief time), and one can almost certainly hold the thermodynamic state of an ecosystem fixed while slightly varying its ecological state (replace some females of a certain population with some males). Thus, List and Pivato's account overcomes some key difficulties with more traditional explications of the 'levels' of reality such as that given by Oppenheim and Putnam (1958)—which appear to suggest that these 'levels' constitute a total order.

This explication of the notion of a 'level' will be useful in what follows, particularly when we come to focus on *mr* exceptions in Section 5 and Section 6. But to begin with, we will turn our attention to what I'm calling *bf* exceptions, which will be seen to pose less of a problem for high-level causation. It will be instructive to consider why generalizations that admit only of *bf* exceptions can nevertheless support causal relations as this will allow us to see why those same reasons don't also apply to generalizations that admit of *mr* exceptions.

---

6. Something is positively (negatively) charged if it is negatively (positively) charged and its charge is first measured by a human prior to the year 2020 or otherwise if it is positively (negatively) charged. Suppose that a system comprises  $n$  particles, including both positively and negatively charged particles, with some but not all of each type first measured for charge prior to 2020. Then consider two sets  $S_1$  and  $S_2$  of states, with  $S_1$  containing all and only states corresponding to descriptions of the form 'comprises  $l$  positively charged particles and  $m$  negatively charged particles' (for all pairs of non-negative integers,  $l$  and  $m$  such that  $l + m = n$ ), and  $S_2$  containing all and only states corresponding to descriptions of the form 'comprises  $l$  positively charged particles and  $m$  negatively charged particles' (for all pairs of non-negative integers,  $l$  and  $m$  such that  $l + m = n$ ). The elements of  $S_1$  and  $S_2$  then cross-classify in the sense that the elements of neither corresponds to equivalence classes of elements of the other.



### 3. 'Background Factor' (BF) Exceptions

It is a common observation that the generalizations formulated in the high-level sciences typically admit of exceptions. A standard term for a generalization that admits of exceptions but that is at least a *prima facie* candidate for playing at least some aspects of the law role tolerably well is *ceteris paribus* (*cp*) law. On the whole, I shall avoid speaking of such generalizations as 'laws', but shall just speak of them as (scientific) 'generalizations' in order to avoid pre-judging the question of precisely *how* well these generalizations play the law role (and, in particular, whether they are capable of underwriting causal relations). I will also tend to avoid speaking of such generalizations as '*ceteris paribus*' generalizations, for I think use of that terminology can obscure the different types of exception of which high-level scientific generalizations admit.<sup>7</sup>

One category of exception of which high-level generalizations typically admit arises when certain high-level variables not represented in the generalization take values that interfere with the holding of the generalization. These are what I'm calling *background factor* (or *bf*) *exceptions*. (As will be seen in Section 5, high-level generalizations also typically admit of exceptions that arise when certain background *micro*-variables take unusual values.) In fact, there are at least two cases to distinguish, corresponding at least roughly to Schurz's (2002) distinction between *comparative ceteris paribus* laws and *exclusive ceteris paribus* laws, or what he later (Schurz 2014) calls 'literal' *ceteris paribus* laws and *ceteris rectis* laws.<sup>8</sup> A *comparative* (or 'literal') *ceteris paribus* (*cp*) law is one that holds only when the values of certain background variables remain *constant* (Schurz 2002: 351–352; 2014: 1802–1803). Boyle's Law is an example, holding as it does only when *temperature* is constant. By contrast, an *exclusive cp* law (or a *ceteris rectis* law) is one that holds only when certain background factors are *absent* or, more precisely, when *the values of certain background variables lie outside a certain range* (Schurz 2002: 352–353, 353–354; 2014: 1803).<sup>9</sup> We will see that Boyle's law is plau-

7. Indeed, although the primary distinction I will draw is between *bf* exceptions and *mr* exceptions, I will also note two other types of exception below (and comment on their implications for a generalization's ability to underwrite causal relations): exceptions that some generalizations admit of when variables represented in the generalization take values outside a certain range (cf. Hitchcock & Woodward 2003a; 2003b); and 'exceptions' that arise because the generalization only properly applies to ideal systems (as the Ideal Gas Law does) to which real world systems at best approximate.

8. Schurz (2002; 2014) is one of a number of philosophers who thus is careful *not* to run together the various types of exception of which high-level scientific generalizations admit.

9. In fact, Schurz's (2014: 1808) formal explication of *ceteris rectis* laws makes it clear that he wishes his notion of them to cover *both* cases where the generalization only holds for a certain range of values of background variables *not* represented in the generalization (2014: 1802–1803) *and* cases where the generalization only holds for a certain range of values of variables that *are* represented in the generalization (2014: 1804) *and* cases where the generalization holds only for

sibly *also* an example of what Schurz calls an *exclusive cp* law (the two categories are therefore not mutually exclusive—see Schurz 2002: 353; 2014: 1803–1804; Reutlinger, Schurz, & Hüttemann 2017: Section 3.1), since it may be violated (for example) in the presence of a strong electric field.

Boyle’s law, a simple example drawn from thermodynamics, states that the pressure and volume of an ideal gas are inversely related:

$$\text{(Boyle's Law)} \quad PV = k$$

Here  $P$  represents pressure,  $V$  volume, and  $k$  is a constant.

Of course it is well known that so stated, Boyle’s Law admits of exceptions: specifically it does not hold if temperature and/or the amount of gas is varied. It thus counts as an example of what Schurz calls a *comparative cp law*.<sup>10</sup> Of course it turns out to be possible to take explicit account of temperature and the amount of substance of the gas, and this is something that is done in the Ideal Gas Law (IGL):

$$\text{(IGL)} \quad PV = nRT$$

As before,  $P$  is pressure and  $V$  is volume, but here  $T$  is the absolute temperature of the gas,  $n$  the amount of substance of the gas (in moles), and  $R$  is the ideal gas constant. The IGL is thus a generalization of Boyle’s Law that explicitly models the temperature and the amount of gas. Correspondingly, the constant  $R$ , unlike the constant  $k$ , does not vary from system to system (the value of  $k$  in Boyle’s Law depends upon the amount of substance and temperature of the gas). The IGL thus covers cases that constituted exceptions to Boyle’s Law, namely, cases where temperature or the amount of the gas are not constant.

It is worth noting that both Boyle’s Law and the IGL properly apply only to *ideal gases*: that is gases comprising point particles that interact with container walls only via perfectly elastic collisions. Actual gases are not ideal, but many—

---

certain ranges of both sorts of variable. It is exceptions that arise due to variables not represented in the generalization not taking the ‘right’ values or—corresponding to comparative/literal cp clauses—not remaining constant that I’m calling *bf* exceptions. As I explain in Footnote 30 below, Hitchcock and Woodward (2003a; 2003b) show that a generalization’s admitting of exceptions when the values of variables that *are* represented in it take values outside a certain range does not *per se* prevent it from underwriting causal relations.

10. Earman and Roberts (1999: 461–462) claim that Boyle’s Law is not a cp generalization. However, aside from their general skepticism about cp generalizations, their specific argument for this claim merely shows that the fact that it concerns *ideal gases* does not *ipso facto* make it a cp generalization. (I shall have more to say about its appeal to ideal gases in the main text below.) Yet, as indicated in the main text, there are independent (and compelling) reasons for taking Boyle’s Law to be a cp generalization (or at least—what is more important for our purposes—to admit of *bf* exceptions).



especially lighter ones—approximate the behavior of ideal gases, especially at high temperatures and low pressures. Thus, the IGL can be used to make approximately correct predictions about them (cf. Roberts 2014: 1782–1783). More about how this relates to these generalizations' ability to support causal relations will be said in Section 4.

Although IGL is immune to some classes of exception from which Boyle's Law suffers, it will still break down given certain values of background macro-variables. For instance, if a gas is subjected to a strong electromagnetic field, a Townsend avalanche can result. This is a process whereby free electrons accelerated by the electric field collide with gas molecules, freeing further electrons in a cascade effect.<sup>11</sup> The resulting gas ions are accelerated towards the cathode (while the electrons accelerate towards the anode). In such circumstances the IGL is liable no longer to yield particularly accurate predictions. In Schurz's terminology, the IGL (and, by extension, Boyle's Law) is an *exclusive cp law*, holding as it does only when the electric current run through a gas is not too strong.<sup>12</sup>

While Townsend avalanches can be modeled by physicists, the models are complex and it's not likely that in more mundane circumstances many would be tempted, rather than using the IGL, to appeal to some more complex model (i.e., generalization or set of generalizations) that incorporates the influence of electromagnetic charge. Moreover, it's plausible that any such more complex model would itself be subject to provisos—of the *exclusive cp* and/or the *comparative cp* sort—either because there are further background factors besides electromagnetic charge that can interfere with the usual relation between gas pressure, volume, amount, and temperature, or because the influence of electromagnetic charge on gas behavior may itself admit of interference by background factors.

It's interesting to think about the relationship between Boyle's Law and the IGL in terms of List and Pivato's model. A gas in a container can be represented by a two-dimensional state space, with one dimension corresponding to the variable  $V$  and the other dimension corresponding to the variable  $P$ . Only certain temporal paths (histories) through that two-dimensional state space are compatible with Boyle's Law together with the value of  $k$  for the gas in question.

The trouble is that what's compatible with Boyle's Law isn't what's genuinely nomologically possible. As we've seen, exceptions to Boyle's Law may arise if the temperature or amount of gas is varied. This means that the gas may

11. In giving this example, we are relaxing slightly our usual assumption that the microphysics of the system is classical.

12. Alternatively, one might say that the post-ionization gas is no longer approximately ideal, so we might think of this as a case that is simply outside the scope of application of the IGL. Whether we take the behavior of the gas in the presence of a strong electric current to constitute an exception to the IGL or whether we take it to lie outside of its scope, such behavior shall be seen in Section 4 to pose a *prima facie* challenge to the IGL's ability to underwrite relations of high-level causation.

take paths through this two-dimensional state space that are incompatible with Boyle's Law.

The IGL, by contrast, lends itself to a representation of the gas by a four-dimensional state space, with dimensions corresponding, not only to pressure  $P$  and volume  $V$ , but also to temperature  $T$  and the quantity of gas  $n$ . The IGL implies constraints on the temporal paths the gas system may take through this four-dimensional state space. It also implies constraints on the temporal paths the gas may take through the original two-dimensional state space, though it has a more liberal (and accurate) view of what these nomologically possible paths are than does Boyle's Law because it allows that if temperature and gas quantity vary (corresponding to changes in the constant  $k$  in Boyle's Law), the inverse proportionality between pressure and volume described by Boyle's Law need not hold.

To the extent that the IGL itself admits of *bf* exceptions, that's because further variables not represented in the law—such as electric current—may have a bearing on the relationship between temperature, pressure, volume, and gas quantity that the generalization describes. To explicitly incorporate these variables into our model would effectively be to represent the gas system in terms of an even higher dimensional state space. Whether we can expect ever to arrive at a model that does not admit of *bf* exceptions by incorporating enough variables is a question that we'll return to below.

As a further illustration of a high-level scientific generalization that admits of *bf* exceptions (this time one drawn from a special science rather than macrophysics), consider the Logistic Equation (LE) from population ecology:<sup>13</sup>

$$(LE) \quad \frac{dn_1}{dt} = r_c n_1 \left( \frac{1 - n_1}{K} \right)$$

Here  $n_1$  is the population size,  $t$  is time (so  $\frac{dn_1}{dt}$  is the population growth rate),  $K$  is the *carrying capacity* (the maximum sustainable population size given the state of the environment), and  $r_c$  is the population's 'intrinsic per capita growth rate' (the rate at which it would grow in the absence of competition for resources). LE implies that when the number  $n_1$  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  $\frac{dn_1}{dt}$  is close to the intrinsic per capita growth rate  $r_c$  multiplied by the number  $n_1$  of individuals in the population. 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.

While no single equation for population growth models all populations at all

---

13. This is discussed as an example of a cp generalization in Fenton-Glynn (2016).

times well, there are some populations that the LE does model well (see Tsoularis & Wallace 2002). Nevertheless, even for such populations, the LE holds only given that certain background factors are absent (or constant): it will not hold in the event of the population being subject to a cull, or if a natural disaster destroys a large part of the population, or if a predatory species suddenly enters the habitat, and so on.

Of course one might devise a more complex growth model that includes variables representing some of the factors relevant to population growth that aren't modeled by the LE. For instance, Weiss (2009) describes a logistic growth model that incorporates predator-prey dynamics (call it WLM for 'Weiss's Logistic Model'). The model comprises the following two differential equations:

$$(WLM) \quad \frac{dn_1}{dt} = r_c n_1 \left( \frac{1-n_1}{K} \right) - \alpha n_1 n_2$$

$$\frac{dn_2}{dt} = -\gamma n_2 + \beta n_1 n_2$$

Here  $n_1$  is the size of the prey population,  $n_2$  is the size of the predator population,  $r_c$  is the intrinsic per capita growth rate of the prey population (the per capita growth rate that it would have in the absence of intra-species competition for resources *and predation*),  $K$  is (as before) the carrying capacity,  $\gamma$  is the predator death rate,  $\alpha$  is the capture efficiency—a measure of the effect of a predator on the per capita growth rate of the prey population—and  $\beta$  is the product of the capture efficiency and the biomass conversion efficiency—a measure of the ability of predators to convert prey into per capita growth. The model reduces to the logistic model where  $n_2 = 0$ , and thus has that model as a special case.

In terms of List and Pivato's representation, we can think of LE as modeling the evolution of a relatively simple system—a single population—through a state space of relatively few dimensions: *one-dimensional* if we think of  $n_1$  as the only true variable in LE (note that an additional dimension isn't required to represent the actual growth rate, since this is represented by the function from times to points in this state space—the *history* of the system at the level of this state space); *three-dimensional* if we wish to represent the nomological possibility of changes in  $K$  and  $r_c$ . Only certain temporal paths, or histories, through this state space are compatible with LE (given the values of  $K$  and  $r_c$ ). However, the histories, or temporal paths, through this state space that are possible by the lights of LE are only a subset of the true nomologically possible histories. That's because LE ignores factors independent of  $n_1$ ,  $K$ , and  $r_c$  that affect the population growth rate. This may be quite appropriate where there is no such influence, or where it is negligible. However, sometimes such influences do prevent approximate logistic growth.

In cases where (for example) the influence of a predator population is significant, WLM may yield more accurate predictions. WLM treats the system—or in fact a more complex system, including the predator population—as representable by means of a state space with at least one additional dimension, corresponding to  $n_2$  (more if we're prepared to think of  $\alpha$ ,  $\beta$ , and  $\gamma$  as implicitly variable). WLM also implies constraints on the evolution of the prey population through the lower-dimensional state space that lacks a dimension corresponding to the size of the predator population. However, the constraints are weaker (and more accurately reflect the possibilities) than those implied by LE because it allows that large predator numbers may lead to growth behavior for the prey population that deviates significantly from that predicted by LE.

WLM itself admits of *bf* exceptions. Any number of factors not represented in WLM—disease, war, natural disaster, continual culling, contraceptive injections, and so on—could all prevent a population growing in the manner predicted by WLM. Indeed, given the extremely large number of factors that can impact population growth, it seems that even the most sophisticated and complex population growth models that ecologists are likely to be able to devise will continue to admit of *bf* exceptions.

In general scientists rarely, if ever, formulate high-level generalizations that are entirely immune to *bf* exceptions. For one thing, attempting to do so would yield generalizations so complex as to be intractable. Indeed, the influence of some background factors may be so complex and/or unpredictable that scientists don't know, and perhaps are unlikely ever to know, how to take them into account in their models. Moreover, it is plausible that scientists typically do not even know, or can't even list, all of the relevant factors and that they certainly couldn't be listed without going outside of the usual subject matter of the discipline in question (see Lange 2002: 416–418; cf. Davidson 1970: 94, 99; 1974: 43; Fodor 1989: 69 Footnote; Schiffer 1991: 3–4; Earman & Roberts 1999: 447 Footnote, 462–463; Hüttemann & Reutlinger 2013: 183, 189–190; Reutlinger 2014: 1761, 1765–1766). In some cases, the possible interfering factors may be indefinite or even infinite (cf. Pietroski & Rey 1995: 101–102; Earman & Roberts 1999: 439, 441; Schurz 2001: 477; Lange 2002: 410–411; Hüttemann & Reutlinger 2013: 183; Reutlinger 2014: 1761).

#### 4. BF Exceptions and Causation

The lack of exceptionless laws in the high-level sciences might seem to pose a problem for accommodating high-level causation because laws that admit of exceptions might be thought unable to underwrite the relations of nomic sufficiency and counterfactual dependence appealed to by some of the leading contemporary accounts of causation.

On the face of it, it seems plausible that decreasing the volume of the container in which a gas is contained can cause the gas pressure to increase. And, indeed, Boyle's Law implies that a decrease in the volume will be associated with an increase in pressure. Yet we know that the inverse relation between volume and pressure described by Boyle's Law admits of *bf* exceptions. If the gas is cooled sufficiently as the volume of the container is decreased, the pressure won't increase. So a decrease in gas volume is not genuinely nomically sufficient for an increase in gas pressure. If a cause must be nomically sufficient for its effect then, rather counterintuitively, it appears that the decrease in gas volume does not cause an increase in gas pressure even on an occasion where the temperature of the gas is held constant and the gas pressure does increase.

A natural rejoinder is that a cause need not be nomically sufficient for its effect, but need only be a non-redundant element of a set of actually-obtaining conditions that is jointly sufficient for the effect (Mackie 1965; Wright 1985). Where the volume of a gas is decreased and its temperature is not simultaneously decreased (or at least is not decreased too much), the pressure of the gas increases, as is implied by the IGL. The decrease in volume alone is not nomically sufficient for the increase in pressure. However, it is a non-redundant element of a set of conditions—including the decrease in volume and the non-decrease (by too much) of the temperature—that is sufficient for the increase in pressure. Sophisticated (and superior) regularity accounts of causation therefore classify it as a cause.

Yet, as we have already noted, even the regularity described by the IGL admits of *bf* exceptions. Volume decrease even together with temperature non-decrease (by too much) may not be nomically sufficient for pressure increase. We might set up a fancy piece of apparatus with a cathode placed in the center of the gas and an anode placed around the container wall. Inducing a strong enough electric field through the gas would—even while the volume of the container were decreased and the gas temperature were not decreased—result a Townsend Avalanche with the gas ions accelerating toward the cathode, reducing the pressure on the container walls.

Still, one might think, we can just add the fact that—on the occasion in question—no fancy piece of apparatus like that described in the previous paragraph was present (or, more generally, that no strong electric current was passed through the gas) to the fact that the volume decreased and the temperature didn't decrease (by too much) to arrive at a set of conditions that is jointly sufficient and individually non-redundant for the increase in pressure.

Yet probably there are still further factors that might prevent volume decrease, together with temperature non-decrease and an absence of a strong electric current, being associated with pressure increase. So the question becomes, can we (or could we at least in principle) keep adding the absence of such inter-

fering background factors to our set of conditions in order to ultimately arrive at one that contains conditions that are genuinely jointly sufficient and individually non-redundant for the effect in question? This is a question that we will return to shortly, after considering our other case study: population ecology.

In population ecology, a reduction in the size of a population (perhaps by means of a cull) might be thought to cause an increase in its growth rate (due to a reduction in *intra-species* competition for resources). Such an association between population size and growth rate is described by the LE. Yet the LE admits of *bf* exceptions, meaning that a reduction in population size isn't genuinely sufficient for an increase in growth rate. For instance, if a large number of predators were introduced into the habitat of the population at the same time as the population size was reduced, then an increase in the population growth rate may not follow.

As with the previous example, the obvious response is to observe that sophisticated regularity theories do not require that a cause be nomically sufficient for its effect, but only a non-redundant element of a set of conditions present on the occasion in question which were jointly sufficient for the effect. Thus, if the size of a population is reduced and there is not an increase in a predator population (or at least not too much of an increase), the population growth rate increases, as implied by WLM. The reduction in the size of the population is thus a non-redundant element of a set of conditions (including the reduction in the size of the population and the non-increase, or non-increase-by-too-much, of any predator population) present on the occasion in question that is sufficient for an increase in the population growth rate. Sophisticated regularity theories therefore count it as a cause.

The trouble is, once again, that we in fact know that WLM itself admits of *bf* exceptions, so that population decrease even together with a non-increase in the size of any predator population is not nomically sufficient for an increase in population growth. Indeed the background factors potentially relevant to population growth might be so numerous that it will be practically—and maybe even in principle—impossible to formulate a generalization concerning population growth that doesn't admit of *bf* exceptions and thus captures a set of conditions that is jointly sufficient (and individually necessary) for an increase in population growth rate.

In spite of this, it seems that there may well be relations of nomic sufficiency that obtain between high-level states of affairs, even if scientists are unable in practice (or perhaps even in principle—for instance, if potentially interfering background factors are infinite in number) to formulate high-level generalizations that do not admit of *bf* exceptions. Providing these interfering factors are in fact absent, even infinitely many of them,<sup>14</sup> then the result is that there exists a

---

14. There is no reason to think an infinite condition cannot be satisfied. Suppose that space-time is *dense* (in the sense that, for any two space-time points, there's a space-time point between



(possibly infinite) set of conditions which are jointly sufficient and individually necessary for the effect in question.

Fodor (1989) makes an analogous point in response to Davidson's (1970) worry that a lack of exceptionless psychophysical or psychological generalizations implies a *prima facie* difficulty for accommodating mental causation. Although, in the following passage, Fodor makes the point with reference to a mental predicate *M* and a behavioral predicate *B* he takes analogous points to apply to the non-psychological special sciences:

The first – and crucial – step in getting what a robust construal of the causal responsibility of the mental requires is to square the idea that *M*s are nomologically sufficient for *B*s with the fact that psychological laws are hedged. How can you have it *both* that special laws only necessitate their consequents *ceteris paribus* *and* that we must get *B*s *whenever* we get *M*s. Answer: you can't. But what you can have is just as good: viz., that if it's a law that  $M \rightarrow B$  *ceteris paribus*, then it follows that you get *B*s whenever you get *M*s *and* the *ceteris paribus* conditions are satisfied. (Fodor 1989: 73)

Fodor's point is that, while the fact that all *M*s are *B*s only *cp* means that something's being *M* is not nomically sufficient for its being *B*, when *M* is instantiated and the *cp* conditions are satisfied then we have a set of conditions, including *M*, that are jointly nomically sufficient for *B*.

We can apply the point to our running examples. Even though the IGL admits of *bf* exceptions, a decrease in the volume of the gas may still be a non-redundant element of a set of conditions that jointly is nomologically sufficient for a rise in pressure. Some of these conditions—specifically, the failure of the temperature of the gas to be simultaneously lowered too much—are captured by variables that figure in the generalization itself. Some—such as the absence of some fancy apparatus involving a cathode placed in the middle of the gas—are not captured by variables that figure in the generalization, but at best can be thought of as covered by an implicit *cp* clause. Still if, on a particular occasion, the volume of a gas is decreased, the temperature of the gas isn't decreased too

---

them). Then consider some region *R* of space-time. And consider the claim that the world-line of the planet Earth's center of mass does not intersect *R*. This is equivalent to the infinite condition that the world-line of the Earth's center of mass does not intersect space-time point  $r \in R$  and that it does not intersect space-time point  $r' \in R$  and . . . . But even if space-time is dense clearly there can be a fact of the matter that Earth's center of mass doesn't intersect *R*. So, in general, there's no reason to think that infinite conditions can't be satisfied. Lange (2002: esp. 407–411) defends the view that there may be a fact of the matter about whether the *cp* clause associated with a high-level generalization is satisfied, even when we are unable to specify the associated conditions so as to render the law 'fully explicit'. See also Fodor (1989: 74).

much, there is no fancy apparatus involving a cathode placed in the middle of the gas, and so on (where the ‘and so on’ indicates that all those potentially interfering background factors that we haven’t even considered are absent), then the decrease in volume is a non-redundant element of a set of actually obtaining conditions that is jointly sufficient for the rise in gas pressure (or at least for the *degree* of pressure rise—if, for example, temperature is increased simultaneous with the volume decrease). By the lights of sophisticated regularity theories of causation, the volume decrease will therefore count as a cause of the pressure increase.

Similarly, although WLM admits of *bf* exceptions, a reduction in population size may still be a non-redundant element of a set of conditions that jointly is nomically sufficient for an increase in population growth rate. Some of these conditions—for instance, the failure of the carrying capacity to simultaneously decline too much—are captured by variables that figure in the model. Others—such as the absence of a continual culling of members of the population—are not, but can be thought of as covered by an implicit *cp* clause. Still, if, on a particular occasion, the size of a population decreases, the carrying capacity doesn’t decrease too much, there is no continual culling, and so on, then the population decrease is a non-redundant element of a set of actually-obtaining conditions that is sufficient for the associated increase in population growth rate. Sophisticated regularity theories of causation will therefore count the decrease in size of the population as a cause of the increase in the growth rate (or at least the *degree* of increase).

One might favor a counterfactual analysis of causation over a regularity theory.<sup>15</sup> And *prima facie* it might seem that generalizations that admit of *bf* exceptions are unable to support the counterfactual dependencies that, according to such analyses, underwrite causal relations. For instance, suppose that the volume of a gas is decreased and its pressure increases. Standard counterfactual theories of causation entail that this relation is causal if it’s true that, if the gas volume hadn’t decreased, the gas pressure wouldn’t have increased. Yet it might seem that this counterfactual is not true. After all, is it not the case that if the gas volume hadn’t decreased the gas pressure still *might* have increased—due to an increase in its temperature, or due to someone setting up a fancy anode/cathode apparatus?<sup>16</sup>

LePore and Loewer (1987: 640–642)—drawing upon a point made by Lewis

15. The classic statement of a counterfactual theory of causation is due to Lewis (1973b).

16. I take the might-counterfactual ( $\phi \diamond \rightarrow \psi$ ) and the would-not counterfactual ( $\phi \square \rightarrow \sim \psi$ ) to be *contraries*: they can’t both be true together. This is weaker than the (also plausible) view (see Lewis 1973a) that the two counterfactuals are *duals* (i.e., that  $(\phi \diamond \rightarrow \psi) \leftrightarrow \sim(\phi \square \rightarrow \sim \psi)$ ). The thesis that the two counterfactuals are at least contraries is defended (convincingly in my view) by Hájek (2017).

(1973b: 563–564)—have correctly pointed out that in fact high-level generalizations *are* able to support the sort of counterfactuals in terms of which causation is analyzed by counterfactual theories even if they admit of the sorts of *bf* exception that we have been discussing. Their point—which can be illustrated with respect to Boyle’s Law/the IGL—is that in many cases it will just be *false* that, if the gas volume hadn’t decreased there would or even might have been a rise in gas temperature that would have meant that the gas pressure increased anyway. Likewise, in many cases it will just be *false* that, if the gas volume hadn’t decreased, there would or even might have been a *bf* exception to the IGL.

The point can be put in terms of Lewis’s (1973a; 1979) possible worlds semantics for counterfactuals. Suppose that, on a certain occasion, the volume of a gas is decreased, the temperature and amount of the gas remains constant, the sorts of background factors (such as a strong electric current passed through the gas) that are apt to interfere with the association described by the IGL are absent, and the pressure of the gas rises. Then for it to be the case that, if the volume of the gas had not decreased, the temperature might have increased (or some interfering background factor might have been present), let alone that it would have increased (or that some interfering background factor would have been present), it would have to be the case that, among the closest possible worlds in which in which the gas pressure is not decreased, there are possible worlds in which the temperature of the gas increases (or some interfering background factor is present). But, on Lewis’s semantics, *closeness* is understood as *similarity*, so the closest worlds to ours in which the volume of the gas is not decreased are the most *similar* such worlds. And since the temperature of the gas doesn’t increase in the actual world, and since interfering background factors are not present in the actual world, then unless the temperature’s non-increase or the non-interference of the background factors are themselves counterfactually dependent upon the volume decrease (and there is no obvious reason to think that they typically are), this furnishes a strong presumption that the most similar worlds to ours in which the volume of the gas doesn’t decrease are also worlds in which the temperature of the gas doesn’t increase and interfering background factors are absent and in which the pressure of the gas therefore doesn’t increase. If that’s the case, then it’s true that, if the volume of the gas hadn’t decreased, then the pressure wouldn’t have increased and therefore that standard counterfactual analyses of causation class the actual decrease in the volume of the gas as a cause of its increase in pressure.

Likewise, LE and WLM seem perfectly well able to support the sorts of counterfactuals invoked by standard counterfactual theories of causation. Suppose that the size of a population is reduced (perhaps by a one-off cull) and suppose that there’s no simultaneous significant change in any predator population, in the carrying capacity or the intrinsic growth rate, or in the way in which the

predator and prey populations interact. Suppose, moreover, there's no natural or human-caused disaster, no cull after the initial reduction in the population size, and so on. The population growth rate thus increases. Again, in most circumstances, it seems plausible that none of the most similar possible worlds to our own in which the population size *isn't* reduced are worlds in which (unlike the actual world) there's a simultaneous significant change in predator population, in carrying capacity, in intrinsic growth rate or in predator/prey interaction, etc. Hence it's plausibly true that, if the size of the population hadn't been decreased, the population growth rate would have been lower and therefore that standard counterfactual analyses of causation will class the actual decrease in population size as a cause of its increased growth rate.

Hitchcock and Woodward (2003a; 2003b) and Woodward (2003) also argue that non-exceptionless high-level generalizations are able to support the counterfactuals relevant to causation. They develop the point in the context of the recent tradition of attempts to analyze causation using structural equations models (SEMs).<sup>17</sup> Structural equations express functional dependencies between variables. The LE is an example of a structural equation, as are the two equations that constitute the WLM. The equation that we have used to represent the IGL is not. That's because structural equations have a single dependent variable on the left-hand side (in the case of the LE, for example, this is the population growth rate<sup>18</sup>) and independent variables on the right-hand side (in the case of the LE these are the intrinsic growth rate, the population size, and the carrying capacity). In the case of the IGL, there is no unique dependent variable. That's because we can *either* manipulate a gas's temperature to influence just its pressure (if the gas is placed inside a rigid container) or its volume (if the gas is placed inside a perfectly non-rigid container) or both (if the gas is in an imperfectly rigid container). Given this fact, we could take one of two approaches to bringing the relationships it describes within the scope of the structural equations framework. One approach would be to define a single variable in terms of  $P$  and  $V$ —the obvious choice would be a variable whose possible values represented values of

---

17. See, for example, Glymour and Wimberly (2007), Halpern (2016), Halpern and Hitchcock (2015), Halpern and Pearl (2005), Hitchcock (2001; 2007), Pearl (2009: Chapter 10), and Woodward (2003).

18. Population size depends on population growth rate (and previous population size). But this doesn't mean that there's no unique dependent variable in LE, or in each of the equations constituting WLM. After all, firstly, it is growth rate (not population size) that ecologists interpret as the dependent variable in these models. And, secondly, while the *future derivative* of population size relative to a time  $t$  may depend upon population size at  $t$ , the population size at  $t$  does not depend on its future derivative relative to  $t$  (cf. Easwaran 2014). The correct interpretation of LE and the equations constituting WLM is as representing the dependence of the future derivative (with respect to time) of the size of  $n_1$  (and  $n_2$ ) upon the values taken by  $n_1$  and  $n_2$  (and other variables) at a given time.

$P \times V$ . Values of *this* variable could then be treated as the effect of varying a gas's temperature. An alternative would be to note that, at least on some occasions, a given system will be naturally representable as one in which pressure is the dependent variable and temperature and volume are independent variables (as where we have a gas in a rigid container), on others, a system will be naturally representable as one in which volume is the dependent variable and temperature and pressure are independent variables (as where we have a gas in a perfectly non-rigid container), and on still others, a system will be naturally representable as one where a variable defined in terms of both  $P$  and  $V$  is the dependent variable (as where we have a gas in an imperfectly rigid container). Given this fact, we can give a structural equation for each specific gas system, with the dependent variable—appearing on the LHS of that equation—determined by the nature of the system itself.

A structural equation, or a set of structural equations like WLM entails a set of counterfactuals (Hitchcock & Woodward 2003b: 183). For instance, LE tells us how the population growth rate would differ under counterfactual suppositions about the population size, intrinsic growth rate, and carrying capacity. This is determined simply by plugging into the equation the relevant values of  $n_1$ ,  $r_c$ , and  $K$ . WLM tells us how the growth rate would vary under counterfactual suppositions about these factors plus suppositions about predator population size,  $n_2$ . Structural equations accounts analyze causation in terms of the counterfactuals entailed by SEMs (a SEM simply comprises a structural equation or set of structural equations). Consequently, they are a variety of counterfactual analysis. According to typical such analyses,<sup>19</sup> relative to the SEM comprising just LE,<sup>20</sup> the actual value taken by—for example— $n_1$  on some particular occasion counts as a cause of the population growth rate,  $\frac{dn_1}{dt}$ , on that occasion because there's some alternative possible value of  $n_1$  such that LE implies that, had  $n_1$  taken that alternative value, while  $r_c$  and  $K$  had remained at their actual values, the growth rate  $\frac{dn_1}{dt}$  would have been different. Likewise, relative to the SEM comprising the two equations of WLM, typical structural equation analyses will count the actual value taken by—for example— $n_1$  on some particular occasion as a cause of the population growth rate,  $\frac{dn_1}{dt}$ , on that occasion, because there's some alternative possible value of  $n_1$  such that WLM implies that, had  $n_1$  taken that alternative value, while  $r_c$ ,  $K$ , and  $n_2$  (and  $\alpha$ ), had remained at their actual values, the growth rate  $\frac{dn_1}{dt}$  would have been different. Finally, relative to a SEM

19. For instance, the accounts of Hitchcock (2001: 286–287, 289–290), Woodward (2003: 77, 83–84), Glymour and Wimberly (2007: esp. 58), and Halpern and Pearl (2005: 853–855).

20. Structural equations analyses, in the first place, define causation relative to a SEM. Causation *simpliciter* is then typically defined in terms of the existence of causation relative to at least one 'appropriate' SEM. For discussion of what constitutes an 'appropriate' SEM, see Halpern and Hitchcock (2010). More will be said about this in Section 6.



of, say, a system comprising a gas in a rigid container (with pressure as the sole dependent variable, and gas quantity, temperature, and volume as independent variables), standard structural equation analyses will count (say) the temperature of the gas as a cause of the gas pressure because there's an alternative value of the temperature such that, had the gas volume and amount remained the same, the gas pressure would have been different.

The key point for present purposes is that, as Hitchcock and Woodward argue, the sorts of structural equations capable of supporting causal relations needn't correspond to exceptionless laws, but merely to what they call 'invariant generalizations'.<sup>21</sup> Generalizations like IGL, LE, and WLM fall short of the standards of exceptionless laws because they admit of *bf* exceptions. Nevertheless, the procedure described above for evaluating counterfactuals—simply plugging in the desired counterfactual values for the variables in the model upon which a certain variable depends and calculating the result for the dependent variable—can be thought of as a procedure for arriving at the 'closest possible world(s)' in which the independent variables take those alternative values (see Hitchcock 2001: 283). These are worlds in which significant interfering factors (such as strong electric fields in the case of IGL and natural disasters in the case of LE and WLM) are absent and so the dependent variable takes the value implied by the model. Generalizations like those encoded by IGL, LE, and WLM are thus able to support the sort of counterfactual dependencies to which structural equations analyses of causation appeal, even though they admit of *bf* exceptions.

To this point we've largely ignored the fact that IGL concerns 'ideal' gases, but that no gases are strictly ideal (though many *approximate* ideal gases, at least for moderate values of the variables represented in IGL<sup>22</sup>).<sup>23</sup> Likewise, it's

---

21. Hitchcock and Woodward prefer this term to '*ceteris paribus* laws'. This does appear *just* to be a terminological decision (see Reutlinger et al. 2017; Hitchcock & Woodward 2003a: 3; Schurz 2014: 1805) but, in any case, nothing of substance turns on the distinction here.

22. The fact that some high-level generalizations (approximately) hold only for a certain range of values of the variables represented in the generalization is discussed further in Footnote 30.

23. This sort of idealization seems to me distinct from those that 'idealize away' background influences (cf. Earman & Roberts 1999: 457, 461–462). Indeed, it seems to me potentially misleading to lump both sorts of idealization together under the heading of '*cp* condition'. The *bf* exceptions we've been discussing to this point are due to (potential) *causal influences* on the modelled behavior by background variables that are not represented in the generalization and that do not pertain to the structure of the modelled system itself. However, the present sort of 'idealization' is (a) one that pertains to the *structure* of the system being modelled—in this case the gas itself, and not merely its potential causal influences; and (b) is an ideal that few if any actual entities ever exactly conform to (in the case of IGL, *no* actual gases conform exactly to its idealization)—whereas the sorts of background factors that we have been discussing so far, and which may lead to *bf* exceptions may quite commonly be absent (or negligible) in actual fact. Moreover, as I note in the main text below, a generalization's making idealizations in the present sense seems to have different implications for its ability to underwrite causal relations than does a generalization's 'idealizing away' from the sorts of background factors discussed so far. Still, authors such as Cartwright



very plausible that LE and WLM contain at least some elements of idealization that mean that they hold only *approximately* even of those real systems that they model fairly well. For example, stochasticity in the processes of reproduction and death for members of a population are liable to mean that growth rates in real populations never perfectly conform to the smooth curve predicted by LE. One might wonder whether these sorts of idealization stand in the way of their ability to support relations of nomic sufficiency and counterfactual dependence.

Let me illustrate with respect to the IGL what I take to be the most plausible line of response to this worry.<sup>24</sup> The response I have in mind says that, while for some merely approximately ideal gas of amount  $n^*$ , having a volume  $v$  and temperature  $t$  is not sufficient (in the circumstances) for its having pressure of exactly  $\frac{n^*Rt}{v}$ , it is sufficient (in the circumstances) for its having pressure approximately equal to  $\frac{n^*Rt}{v}$ .<sup>25</sup> Depending on exactly how one wishes to frame a regularity theory of causation, one could then either say that the effect of the volume and temperature of the gas is simply *having pressure approximately equal to  $\frac{n^*Rt}{v}$* , or one could take the effect to be whatever token of the type *having pressure approximately equal to  $\frac{n^*Rt}{v}$*  the system in fact happens to possess.

If one prefers a counterfactual theory of causation (or a structural equations

(1980: e.g., 160) (cf. Pietroski & Rey 1995: 84–85, 89–90) appear to take generalizations that ‘idealize’ in the present sense as central examples of cp generalizations and I needn’t insist that they not be classified under this umbrella.

24. An alternative to the response I’m about to describe might say that generalizations that make such idealizations are approximations to, or special cases of, laws that cover non-ideal cases, and that it’s these latter generalizations that support relations of nomic sufficiency and counterfactual dependence between real high-level states. This response might work quite well in some cases. For instance, van der Waal’s Equation seeks to model the influence of particle size and attraction in real gases. The IGL is a special (and never-perfectly-instantiated) case of van der Waal’s equation where particle size and attraction is zero. The reservation I have about this approach is that it may not always be possible to formulate these more general laws (cf. Cartwright 1980: 161), and, even if it is, it may not be possible to do so without reverting to basic physics (cf. Pietroski & Rey 1995: 98). But once we do that, the danger is that the high-level states which we were taking to be potential causes and effects simply drop out of the picture.

25. Cartwright (1980: 160–161) suggests that we make assumptions like this when explaining the behavior of actual systems on the basis of idealization laws. She points out that this rests on the further assumption that physical processes are continuous (so that, for example, an *approximately* ideal gas will behave in *approximately* the way described by IGL). However, we can and do have empirical evidence for such continuities. Of course, in cases where physical processes are *dis*-continuous, idealization laws won’t underwrite causal relations about non-ideal entities/systems. Indeed, Reutlinger (2014: 1768 Footnote) suggests that there may be some idealization laws that describe behavior that isn’t approximated by any real system. If we believe that some entities/systems that don’t behave in ways that even approximate the behavior described by idealization laws are nevertheless involved in causal interactions, then we will need to look for appropriate generalizations which directly cover the non-ideal cases. In any case, many systems exhibit behavior that *does* approximate that described by high-level idealization laws. For example, many gases behave in ways that approximate the behavior described by IGL, while many populations behave in ways that approximate the behavior described by LE or WLM.

variant on such an account), then things are even more straightforward. Suppose that, for a particular system comprising an approximately ideal gas in a container, the actual values of  $P$ ,  $n$ ,  $T$ , and  $V$  are  $p$ ,  $n^*$ ,  $t$ , and  $v$ . And suppose that  $p \approx \frac{n^*Rt}{v}$ . Then clearly, because for such a gas it's generally true that  $P \approx \frac{n^*Rt}{v}$  (or at least this is true for a reasonable range of values of  $n$ ,  $T$ , and  $V$ ), there are possible values of  $n$ ,  $T$ , and  $V$  such that, had the gas system instantiated those values, then it would have been that  $P \neq p$ . That's enough for standard counterfactual theories of causation to count  $n$ ,  $T$ , and  $V$  taking values  $n^*$ ,  $t$ , and  $v$  (respectively) as causes of  $P = p$ .

So far we've seen that the fact that high-level generalizations—including generalizations of high-level physics such as IGL and special science generalizations such as LE and WLM—admit of *bf* exceptions isn't a reason for thinking that there don't exist the sort of relations of nomic sufficiency or counterfactual dependence between high-level states needed to underwrite relations of high-level causation. Yet, so far, we've not considered all classes of exception of which high-level generalizations admit. There's a certain class, to the description of which I turn in the next section, which poses more of a problem.

## 5. Exceptions Due to Multiple Realizability (*MR* Exceptions)

A class of exception that poses particular problems for the ability of high-level generalizations to underwrite causal relations arises due to the multiple realizability of the states that such generalizations concern. These are what I'm calling *mr exceptions*. To illustrate this sort of exception, consider again a system comprising a gas in a container. Note that the macrostate of such a gas—which is the type of state the IGL concerns—is specified by stating the values of macrovariables like temperature, pressure, volume, and gas amount. A high-level state space for the system can be defined in terms of these variables, with a macrostate of the system corresponding to a point in that space. As indicated in Section 2, the microstate of the system is given by specifying its location in its microphysical phase space. Macrostates of the system are *multiply realizable* by microstates. That is, a point in a high-level state space (defined in terms of macrovariables) corresponds to a *region* in its phase space. As List and Pivato put it, macrostates are *coarse-grainings* of or, in other words, correspond to *equivalence classes* of microstates.

It is well known that exceptions to thermodynamic generalizations like IGL can possibly arise due to certain unusual microphysical realizations of their macrostates. For instance, there are possible initial microstates of a system comprising a gas in a rigid container such that the pressure on the container walls spontaneously decreases without any corresponding change in the gas temper-

ature, volume, or amount, and even without the presence of any ‘interfering’ background factors like strong electric currents (cf. Roberts 2014: 1782 Footnote). Were such a spontaneous decrease to occur, it would be in violation of the IGL.

This type of exception is one of which many high-level generalizations admit. Take, for instance, a system comprising a population and its environment. As we’ve seen, a three-dimensional state space for this system can be defined in terms of the variables that appear in the LE (which we’ll focus on for simplicity):  $n_1$ ,  $r_{\sigma}$  and  $K$ . The LE has implications concerning how the system may evolve through this state space. But points in this state space are multiply realizable by points in the system’s phase space. And there are points in the system’s phase space—very unusual points—that, when evolved forward in accordance with the microphysical laws, lead to dramatic entropy decrease. It turns out that these points are highly ‘scattered’ in the phase space, so there are almost certainly such microstates that are compatible with any given macrostate (Albert 2000: 67). Yet, in the presence of dramatic entropy decrease, the LE won’t even approximately hold. After all, members of a population depend upon entropy increasing processes—such as the diffusion of oxygen in their lungs—for their very survival. Given dramatic entropy decrease, the population will therefore collapse rather than evolving in accordance with the LE.

Perhaps more interestingly, the system comprising the population and its environment can be represented by state spaces that are intermediate in grain between the three-dimensional state space defined in terms of the variables appearing in LE and the phase space of the system. For example, ecologists have noted that 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, Murrell, & Dieckmann 2003: 252) and to breeding possibilities (Otto & Day 2007: 591). To model the influence of geographical distribution upon population growth, a standard approach is to represent the population’s habitat as comprising a finite number of spatial subregions (see e.g., Otto & Day 2007: 591–594; Hastings 1993; Law et al. 2003). The growth of the population in each sub-region is then modeled as a function of the size of the population in that sub-region, the size of the population in neighboring regions (since this affects net migration levels into the region in question), and other factors such as the intrinsic growth rate of the population and the carrying capacity of the sub-region in question. A very simple such model—which we’ll consider for illustrative purposes—might simply divide the habitat into two spatial subregions, with  $n_i^t$  representing the size of the population in sub-region  $i$ , and  $K^i$  representing the carrying capacity of subregion  $i$ . A straightforward extension of the LE would then provide two equations—giving the growth rates of the subpopulations in each of the two regions—with the following form (which I label ‘SLM’ for ‘Spatial Logistic Model’):

$$\begin{aligned}
 \text{(SLM)} \quad \frac{dn_1^1}{dt} &= r_c n_1^1 \left( \frac{1-n_1^1}{K^1} \right) + \alpha n_1^2 - \beta n_1^1 \\
 \frac{dn_1^2}{dt} &= r_c n_1^2 \left( \frac{1-n_1^2}{K^2} \right) + \beta n_1^1 - \alpha n_1^2
 \end{aligned}$$

The first term on the RHS of each equation is familiar from the LE. The second and third terms represent the influence of migration between the two sub-regions. The parameter  $\alpha$  ( $0 \leq \alpha \leq 1$ ) represents the rate at which members of the sub-population in region 2 migrate to region 1, while  $\beta$  ( $0 \leq \beta \leq 1$ ) represents the rate at which members of region 1 migrate to region 2. (In this simple model, these rates are taken as fixed, though a more sophisticated model might take these rates to be functions of, e.g., the size of the sub-population relative to the carrying capacity in each of the sub-regions of the habitat.) If these parameters were both equal to zero (i.e., if there were no migration between the two sub-regions of the habitat), then the model would predict simple logistic growth of each of the sub-populations in each of the regions of the habitat.

This model allows us to capture in a simple way how population distribution influences population growth by affecting the level of competition for resources in subregions of the habitat. Of course, one could make the model more sophisticated by, for instance, partitioning the habitat into a greater number of spatial sub-regions, or by seeking to model the effects of geographical distribution on reproduction rates. But for present illustrative purposes, this simple model will do.

The total size of the population in the habitat,  $n_1$ , is the sum of the population sizes in each of the two sub-regions:  $n_1 = n_1^1 + n_1^2$ . Similarly, the maximum sustainable population size for the habitat,  $K$ , is equal to the sum of the maximum populations that can be sustained in each of its two subregions:  $K = K^1 + K^2$ . Different overall population sizes are *multiply realizable* by specific geographical distributions of the population. In this case, a single value of  $n_1$  is compatible with different combinations of values of  $n_1^1$  and  $n_1^2$ , while a combination of values for  $n_1^1$  and  $n_1^2$  determines a value for  $n_1$ . Likewise, different overall maximum sustainable population sizes are multiply realizable by different combinations of values for  $K^1$  and  $K^2$ .

As we saw, LE suggests a representation of the system comprising the population and its habitat by a state space of three dimensions—with the dimensions corresponding to the variables  $r_c$ ,  $n_1$ , and  $K$ . The SLM, on the other hand, suggests a representation by a state space comprising five dimensions: one corresponding to each of the variables  $r_c$ ,  $n_1^1$ ,  $n_1^2$ ,  $K^1$ , and  $K^2$ . Since values of the variables  $n_1$  and  $K$  are multiply realizable by combinations of values for  $n_1^1$  and  $n_1^2$ , and of  $K^1$  and  $K^2$  respectively (while the dimension corresponding to  $r_c$  exists in both state

space representations), points in the three-dimensional state space correspond to regions in the five-dimensional state space. Or, as List and Pivato would put it, points in the former state space correspond to *equivalence classes* of points in the latter. The latter state space is a *finer grained* representation of the system than the former because it represents not just total population size (and carrying capacity), but how that population (and carrying capacity) is distributed.<sup>26</sup>

Interestingly, it turns out that population growth can be extremely sensitive to precise initial conditions, including the precise geographical distribution of the population and environmental resources (see Hastings 1993; May 1974). For instance, 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 overall population grows in (approximate) accordance with LE or sharply declines, even when the population size is well below the overall carrying capacity (see Hastings 1993). When it comes to those populations that are normally well-modelled by the LE, it is presumably the case that the precise geographical distributions that lead to such sharp population declines are rare. The 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 its phase space that leads to a decrease in entropy.

For present purposes, it's a moot question whether we class *mr* exceptions as cases where the *cp* conditions associated with a generalization are violat-

---

26. The state space constructed from the variables in LE is thus a higher-level state space than that constructed from the variables in SLM. Could it also be said that a state space constructed from the variables in LE is a higher-level state space than that constructed from the variables in WLM or that a state space constructed from the variables in Boyle's Law is a higher-level state space than that constructed from the variables in IGL? I take the answer to be 'no' for the reason that the variables in WLM and IGL are simply supersets of those in LE and Boyle's Law respectively, reflecting the fact that they simply incorporate the influence of (metaphysically) independent factors that were left as background in LE and Boyle's Law rather than genuinely representing the system in some more fine-grained way. The variables in SLM, by contrast, are not a superset of those in LE. Rather, the variables  $n_1$  and  $K$  in LE are replaced in SLM by the pairs of variables  $n_1^1$  and  $n_1^2$  and  $K^1$  and  $K^2$ , respectively. Still, one might observe, we could define new variables:  $X \equiv 3 \times PV$  and  $Y \equiv 2 \times PV$ , so that  $X - Y = nRT$ . The variables appearing in the latter generalization aren't simply a superset of those in Boyle's Law. But presumably our reasons for not wanting to count IGL as representing systems at a 'lower level' than Boyle's Law apply also to this generalization. The upshot is that I think we must, either by modifying our definition of multiple realizability, or List and Pivato's characterization of what it is for one state space to be 'higher-level' than another, add the restriction that one state space does not count as 'lower-level' than another (perhaps because its states don't genuinely count as 'multiply realizing' those of that other) if it is parameterized by a set of variables that is, or is defined by logico-mathematical operations upon, a superset of the variables in terms of which that other is parameterized. (Note that one can't make logico-mathematical inferences from the values of  $n_1$  and  $K$  to the specific values of  $n_1^1$  and  $n_1^2$  and  $K^1$  and  $K^2$ , so the variables in SLM continue to count as parameterizing a more fine-grained state space than do those in LE.)



ed.<sup>27</sup> What's important is that, as we shall see in Section 6, *mr* exceptions pose quite a different set of challenges from *bf* exceptions to a generalization's ability to support causal relations—a challenge that existing accounts of how non-exceptionless generalizations can support causal relations appear unable to meet.

*MR* exceptions are quite different in nature from *bf* exceptions. *BF* exceptions—as I have characterized them—are due to the (causal) influence of factors that are representable by (macro-)variables that are metaphysically independent of those that appear in the original generalization. By contrast, *mr* exceptions are due to unusual *realizations* of the states represented by the variables in the original generalization. These lower-level states are not metaphysically independent of those represented by the variables in the original generalization. If we seek to eliminate exceptions of this type from our generalizations, we do not simply add variables, but rather *replace* the variables in our model with variables that represent more fine-grained states of affairs. Thus, in seeking to formulate a generalization that avoids the exceptions to LE that arise because of the possibility of certain ways the population may be *distributed*, we replace the variable that simply represented the overall size of the population ( $n_1$ ) with variables that represent the size of the population in different geographical sub-regions of the habitat. The overall size of the population in a habitat is not metaphysically independent from the size of the population in each of the sub-regions of the habitat: facts about the latter determine facts about the former in a stronger-than-nomological manner.

Similarly, in seeking to formulate a generalization that avoids the exceptions to IGL—or for that matter LE, WLM, or SLM—due to rare but possible *micro-physical realizations* of the macrostates of the systems these generalizations seek to model, we would need to opt for a model that, instead of representing the system's *macrostate* (using variables like temperature, pressure, and volume in the case of IGL) contains variables representing its *microstate*, that is, the positions and momenta of the particles it comprises. The macrostate of a system (such as a gas in a container) is not metaphysically independent from its microstate: the latter determines the former in a stronger-than-nomological manner.

Before discussing the problems *mr* exceptions pose for high-level causation and potential resolutions of that problem, it's worth noting that there are exceptions to high-level generalizations that are a sort of 'mixed case' between *mr* exceptions and *bf* exceptions, and in fact these may be more common than pure *mr* exceptions. This is brought out by Fodor's (1991) discussion—prompted by Schiffer

---

27. Schiffer (1991: 7) appears to view such exceptions in this way, saying "certain realizations [of the high-level state described by the generalization] may themselves be among the defeating conditions alluded to in the *ceteris paribus* clause." In Fenton-Glynn (2016), I argued that such exceptions should not be construed as cases in which the *cp* clause of the generalization is violated. However, I don't wish to take a stand on this issue here. For relevant discussion, see Earman and Roberts (1999: 463–465).



(1991)—of the distinction between what he calls ‘*mere* exceptions’ and ‘*absolute* exceptions’ to high-level generalizations. Absolute exceptions correspond closely to pure cases of what I’ve been calling *mr* exceptions; ‘*mere* exceptions’ are more of a mixed case. To make the distinction clear, a little background is needed.

Suppose that a particular high-level generalization admits of *bf* exceptions, but not *mr* exceptions. Then, following Schiffer (1991) and Fodor (1991), call the ‘background’ circumstances that must obtain for the values of the independent variables in that generalization to genuinely nomically suffice for the value of its dependent variable a ‘completer’. Ignoring, for the moment, its *mr* exceptions, a ‘completer’ for IGL—or an application of it where, say, pressure is the obvious dependent variable because the gas is heated in a rigid container—might involve the absence of strong electric currents through the gas, and so on.

One of Schiffer’s (1991: 4) key insights is that it may often be the case that there is no *high-level* state that serves as a ‘completer’ for a high-level generalization that admits of what I’m calling *bf* exceptions, or at least not one at the level of the generalization in question. Indeed, this may well be true for the IGL. For instance, suppose that a gas is in a container that is impervious to any but very high velocity molecules. Given scattering, we know that microstates of a thermodynamic system that lead to normal thermodynamic behavior typically differ from those that don’t only by a small perturbation. Consequently, enough incident high-velocity particles entering the container in just the right way could alter the gas’s microstate in such a way as to put it in a state that leads to thermodynamically abnormal behavior, with the consequence that the gas’s behavior does not (approximately) conform to IGL. The manner of incidence of high-velocity incident particles seems to be a microphysical fact that a high-level state description of the gas-in-container system, even together with a high-level state description of its environment, does not capture. Thus, a ‘completer’ for IGL would appear to involve microphysical facts.

Schiffer’s (1991: 5) second key insight is that microphysical factors (or more generally lower-level factors) may not act as ‘completers’ for the states captured by the high-level variables featuring in the original generalization. For example, the precise pattern of incidence of high-velocity particles does not combine with the facts about the gas’s *macrostate* to yield a condition that’s sufficient for behavior according to the IGL. That’s because whether the precise pattern of incidence of high-velocity particles does or does not prevent the gas from behaving in accordance with the IGL depends on the *microstate* of the gas itself—that is, upon how the macrostate of the gas is realized. So really, the facts about the precise pattern of incidence of high-velocity particles is only a completer (or at least part of a completer) for the *microstate* of the gas, that is, a condition with which the *microstate* of the gas combines to form a sufficient condition for later microstates that realize a macrostate in which (say) the pressure of the gas is higher.

Fodor (1991: 24)—drawing upon Schiffer’s insights—observes that, even if there exists a nomologically possible (micro-level) completer for every possible realizer of the gas’s macrostate, it may well be nomologically possible for a particular realizer to occur without one of its completers. In terms of our running example this would be so if the microstate of the gas were changed, because of the actual pattern of incidence of high velocity particles, into one that leads to unusual thermodynamic behavior despite the fact that alternative patterns of incidence of such particles wouldn’t have transformed it into such a state. Cases where all realizers of the state described by the independent variables in a generalization have nomologically possible completers, but where, on a particular occasion, a realizer occurs without its completer, are termed ‘*mere exceptions*’ by Fodor (1991: 24). Cases of ‘*mere exceptions*’ are really a kind of mixed case between what I have termed *bf* exceptions and *mr* exceptions. In such cases the micro-realizer of the macrostate captured by the independent variables in the generalization conspire with (*micro-level*) background factors to produce an exception to the generalization. Evidently we’ll need to be sensitive to mixed case exceptions like this, as well as to *bf* and *mr* exceptions, if we are to account for high-level causation.

Pure cases of *mr* exceptions arise where a realizer of the state described by the independent variables in a generalization is instantiated and that realizer has *no* nomologically possible completer. Fodor (1991: 24) calls exceptions that arise where such realizers are instantiated ‘*absolute exceptions*’. As Schiffer puts it, while some exceptions to high-level generalizations may result (at least in part) from the interference of background factors, “certain realizations of [a system’s macrostate] *M* may themselves be among the defeating conditions” (1991: 7).

The IGL might admit of absolute exceptions if, for example, there are possible realizers of the macrostate of a gas-in-container system such that no possible environmental interference could perturb its microstate in such a way as to lead to it evolving in (approximate) conformity with the IGL. This might be the case if, for example, the gas is in a microstate which is such that, if it is unperturbed by environmental interference, it will lead to thermodynamically unusual behavior *and* the gas is in a lead container that shields it from environmental interference thus ensuring that no such influence is possible.<sup>28</sup>

---

28. Of course, the fact that the gas is in a lead container might itself be considered a ‘background factor’. So perhaps a more clear-cut example of an absolute exception would be the exception to the Second Law of Thermodynamics (SLT) that arises if the total entropy of the universe as a whole declines for a time. There is no possibility of outside influence upon the universe as a whole. Yet its total entropy may still decline for a while if its initial microstate is of the right sort.

## 6. MR Exceptions and Causation

MR exceptions to high-level generalizations—including those of the pure, absolute, variety and those of the mixed, *mere* exception variety—pose a problem for the accommodation of high-level causal relations in a way that pure *bf* exceptions (i.e. cases where a high-level generalization admits of exceptions because of the possibility of interference from high-level background factors) do not. Fodor, Lepore and Loewer, and Hitchcock and Woodward’s arguments (reviewed in Section 4) that non-exceptionless generalizations can support relations of nomic sufficiency and counterfactual dependence apply well to generalizations that admit only of (pure) *bf* exceptions. But we have now seen that high-level laws at least often admit of *mr* exceptions too.

Suppose, for instance, that the volume of a gas is reduced while the temperature and gas amount is held constant, and suppose that this occurs in circumstances in which the gas isn’t subject to a strong electric field, or any other sort of interfering high-level background factor. And suppose that the pressure of the gas increases. We might naturally think that the reduced volume is a cause of the increased pressure. Yet it now appears that, in fact, the reduced volume is not a non-redundant element of a set of conditions present on the occasion in question that is sufficient for the increased pressure. This means that it’s not, as the arguments due to Fodor that we considered in Section 4 might have led us to believe, a cause of the increased pressure by the lights of sophisticated nomic regularity theories of causation. The reason is that, as we have seen, it’s possible for the volume of the gas to be reduced, even while the temperature and gas amount is held fixed, the gas is not subject to a strong electric field, and so on, while the gas pressure still does not rise. That’s because there are (highly unusual) micro-realizers of the macrostate of the gas system as the gas volume is reduced (perhaps together with the microstate of its environment—which may include, for example, incoming high-velocity particles) that lead, according to the microphysical laws, to precisely this result. Now, given that the gas pressure in fact increased, its microstate (together with the microstate of its environment) wasn’t in fact one of these highly unusual ones. Nevertheless, it’s only the precise microstate itself (perhaps together with the microstate of the environment), and not the macrostate of the gas system that it realizes that (in the circumstances) is sufficient for an increase in gas pressure (cf. Schiffer 1991: 7). Moreover, if the facts about the macrostate of the gas were added to the set of facts about its microstate (and that of its environment), the facts about the macrostate would be redundant, because the facts about the microstate of the gas (and its environment) are already a sufficient condition for the increase in gas pressure (cf. Hofer 2004: 106).

Drawing a similar distinction to that made here between *bf* and *mr* excep-

tions, Hoefer (2004) suggests that high-level states that are putatively sufficient in the circumstances for a given effect face

the enemy from without, and the enemy from within. The kinds of problems we have dealt with so far [viz., interfering background factors] . . . count as enemies from without. But . . . in many cases we need to exclude the enemy from within: microscopic initial- and boundary-conditions that are just perverted and ‘atypical’ enough to entail the non-production of the usual effect. . . . If the macro-level [state] proposed [as a cause] supervenes on the ‘wrong’ initial conditions . . . then all sorts of weird things may take place, including the failure of the customary effect E to ensue. (Hoefer 2004: 105)

Fodor (1991) himself doesn’t develop his account of how generalizations can support causal relations in light of the possible existence of absolute and *mere* exceptions, but only develops an account of how to reconcile the idea that a generalization admits of absolute and mere exceptions with its being both true and non-vacuous. Indeed, as should be clear from the preceding discussion, it’s hard to see how an account of how generalizations that admit of absolute and mere exceptions can nevertheless support relations of nomic sufficiency would go.

It also appears that the sort of exception under consideration makes trouble for the ability of high-level generalizations to underwrite the sort of counterfactuals in terms of which causation is analyzed by counterfactual analyses (cf. Hoefer 2004: 107–109; Hájek 2017). For instance, we might wish to say that a gas’s current volume is a cause of its pressure. But it’s possible for the gas pressure to be lower or higher than predicted by the IGL (even when there’s no strong electric field or anything like that) due to the system’s (or the system together with its environment’s) being in an unusual initial microstate. That is, there are possible micro-realizers of the gas system (plus its environment) in which a lower or a higher than usual number of the particles (or of the relatively high velocity particles) impact upon the container walls. Even if the actual microstate of the gas (plus its environment) is a relatively normal one (thus the pressure is roughly what is predicted by the IGL), it appears false that, if the gas volume had been different, the gas pressure would have been different. After all, if the gas volume had been different, the gas just might have been in one of those rare microstates that (together with the microstate of its environment) results in the pressure being exactly the same as it actually is.<sup>29</sup> The reason to think this is that

---

29. For arguments that the rarity of the microstates leading to the ‘deviant’ macroscopic behavior doesn’t mean that they’re irrelevant to the truth-values of the counterfactuals in question, see Hájek (2017) and Hoefer (2004: 109). For instance, it doesn’t seem that it can be maintained that worlds in which ‘deviant’ microstates are instantiated are less close/similar to the actual world

the gas couldn't be in a different macrostate (as it would be if the volume were different) without being in a different microstate. Moreover, it's difficult to see that there could be a fact about which specific microstate it would be in (given that many microstates are compatible with the counterfactual supposition about its macrostate). Given the 'scattered' nature, in the phase space of the system, of the microstates that lead to unusual macroscopic behavior there are, for any given macrostate, such microstates that realize it. Moreover, given scattering, it's plausible that only by specifying the full microstate of the system (plus its environment) would we specify a state upon which the gas pressure genuinely counterfactually depends. But counterfactual dependence on a microstate is not what is needed, by the lights of counterfactual analyses of causation, to give us genuine high-level causation.

The present problem also afflicts structural equations versions of the counterfactual approach to causation. Suppose we have a 'structural' equation for the gas system that expresses its pressure as a function of its temperature, volume, and amount:  $P = \frac{n^{\dagger} R t}{v}$ . We thus treat pressure as the dependent variable for this system, at least on this occasion. The trouble is that, interpreted as a structural equation, this equation entails counterfactuals of the form, 'if the amount of gas had been  $n'$ , the gas temperature had been  $t'$ , and the container volume had been  $v'$ , then the gas pressure would have been  $\frac{n' R t'}{v'}$ '. But generally such counterfactuals are false for reasons we have already seen: namely that if the amount of gas had been  $n'$ , the gas temperature had been  $t'$ , and the container volume had been  $v'$ , then the microstate of the gas *might* have been one of those unusual ones that does not give rise to a gas pressure of (approximately)  $\frac{n' R t'}{v'}$ . Encoding false (or at least not approximately true) counterfactuals is one way in which a SEM can count as 'inappropriate' (Hitchcock 2001: 287; Halpern & Hitchcock 2010) so that, even though such a model seems to imply that the gas volume is a cause of its pressure (because it implies that the former counterfactually depends on the latter), this doesn't imply that gas volume is a genuine cause of its pressure (because the counterfactuals encoded by the model are false).<sup>30</sup> To obtain a struc-

---

than those in which 'normal' microstates are instantiated, at least not on the standard, Lewisian (Lewis 1979) account of similarity among possible worlds (see Hoefer 2004: 109–110; Hájek 2017; Fenton-Glynn 2016: 279–281), especially as 'deviant' microstates are sometimes instantiated in the actual world (see Hoefer 2004: 110).

30. In fact, Hitchcock and Woodward (2003a; 2003b) in effect weaken this condition on model 'appropriateness' slightly by requiring that the structural equations in a model are accurate (or in their terminology 'invariant') over a certain *range* of possible values for the variables on the RHS of the equations (i.e., that the counterfactual values that they entail for the variable on the LHS were the variables on the RHS to take values in this range are the true counterfactual values that the variable on the LHS *would* take in such circumstances) but not that they be accurate over the whole range. (Such generalizations that are accurate over only a certain range correspond to what Reutlinger et al. [2017: Section 3.1] term 'restricted' *cp laws*. This corresponds to an additional type of exception of which a scientific generalization may admit, though not one that poses a threat to



tural equation that represents what gas pressure genuinely counterfactually depends upon, it seems that its variables (those on the RHS at least) would need to represent the microstate of the gas (plus its environment). But then structural equations analyses would only yield the result that it is the microstate of the gas (plus its environment), and not its macrostate, that is the cause of the gas pressure. They would therefore not imply that we have a case of high-level causation.

Analogous points can be made with respect to the LE and WLM. It would be very natural for an ecologist to take the small size of a population relative to its environment's carrying capacity to causally explain the high growth rate of the population (see Tsoularis & Wallace 2002). Yet, as indicated by Hastings (1993), it appears that there are some precise individual-by-individual ways a population of a given size might be geographically distributed that (even in the absence of predation, natural or manmade disasters, etc.), rather than leading to usual logistic growth, lead to quite different growth behavior (Hastings 1993: 1365). Interestingly, it appears that, in a state space that represents the precise geographical distribution of members of a population, those points that lead to abnormal growth patterns are highly dispersed or 'scattered' (Hastings 1993: 1370), just as the points that lead to unusual thermodynamic behavior are highly 'scattered' in the phase space of a system. This means that, for any given overall population size, it's likely that it is compatible with precise population distributions that lead to abnormal growth patterns. So a population's having a certain low size  $n_1$  relative to the carrying capacity  $K$  is not *sufficient* for a high growth rate (more precisely, a growth rate close to  $r_c$ ) even in circumstances where interfering background factors, like natural disasters and so on, are absent.

In order to find a state that might be sufficient (in the circumstances) for LE or WLM-like population growth, it would seem necessary to model the geographical distribution of the population (along the lines of SLM) and not simply

---

their ability to underwrite causal relations, and least not by the lights of Hitchcock and Woodward's account.) The trouble is that, given the point about the 'scattered' nature in phase space of points leading to unusual macroscopic behavior for the gas, it appears that *any* combination of values taken by the variables on the RHS will be compatible with the system's being at one of these deviant points in its phase space, so that the structural equation under consideration won't even be accurate ('invariant') over a certain range.

Another point worth noting about the structural equations approach—at least on the versions espoused by Woodward (2003) and Hitchcock and Woodward (2003a; 2003b)—is that it is associated with a specific way for evaluating counterfactuals. Namely, we're supposed to consider what would happen if their antecedents were realized by 'interventions'. The notion of an 'intervention' is a (semi-)technical term defined by Woodward (2003: 98). Roughly speaking, we can think of an 'intervention' as an ideal experimental manipulation of the variables mentioned in the antecedent of the counterfactual to set them to the required values. But appealing to this special way of evaluating counterfactuals doesn't get us off the hook. After all, no part of Woodward's (2003: 98) definition of an intervention entails, for any values of  $n$ ,  $V$ , and  $T$ , that, if the gas had been intervened on to set the values of  $n$ ,  $V$ , and  $T$  equal to those values, the gas wouldn't have been in one of those rare microstates such that its pressure was significantly different from  $\frac{n^*Rt}{v}$  (cf. Fenton-Glynn 2016).



its overall size. Yet even if the geographical distribution is not one of those rare ones that is such as would normally lead to growth quite out of keeping with LE or WLM (in circumstances where natural disasters, and so on are absent), it is doubtful whether even *it* is sufficient (in these circumstances) for LE or WLM-like growth. After all, it's possible that the *microstate* of a system comprising a population geographically distributed in that way, together with its environment, might be one of those rare ones that leads to radical entropy decrease and to population collapse. So, in order to get a condition that's truly sufficient (in the circumstances) for LE or WLM-like growth, we may again need to look to microphysics. If this is so then, by the lights of sophisticated regularity theories of causation, we do not have a case of genuine high-level causation.

Analogous points hold with respect to counterfactual theories of causation. Consider, for instance, the claim that the large size of the population relative to the carrying capacity was a cause of its low growth rate. The trouble is that, if the population size had been lower, then it just *might* have been distributed in one of those rare ways that leads to unusual growth patterns and so the population might still have grown at a low rate. To see this note that a smaller population can't be geographically distributed in precisely the same individual-by-individual way. Thus, if the population size had been different, its geographical distribution would have been different too. And, given scattering, there seems little ground for saying that the alternative geographical distribution *wouldn't* have been one of those that leads to an unusual growth pattern.

If we wished to find a state upon which the population growth rate does genuinely counterfactually depend, we might look to the more fine-grained state consisting in the population's precise geographical distribution. For it might be that there are certain alternative precise geographical distributions to the actual one (perhaps realizing lower population sizes) that definitely would have resulted in a higher population growth rate. However, even this claim is problematic. That's because one can't change the geographical distribution of a population without changing the microstate of the system of which it is part and so, given scattering, it appears that, had the precise geographical distribution of the population been different, it just *might* have been the case that the system comprising the population and its environment was realized in one of those rare microphysical ways that leads to dramatic entropy decrease and hence population collapse.

Structural equations variants of counterfactual analyses have similar difficulties accommodating the judgment that population size is a cause of population growth rate. As we have seen, interpreted as structural equations, LE and the equations that comprise WLM, entail counterfactuals. Yet, even though they imply that the population growth rate counterfactually depends on the size of the population relative to the carrying capacity (and, in the case of WLM, on the size of any predator population), it now appears that (at least many of) the coun-

terfactuals that they encode are false. This means that, even though these SEMs entail counterfactual dependence of population growth rate upon population size relative to carrying capacity, they are not apt ones for assessing whether the latter is *genuinely* a cause of the former by the lights of standard structural equations analyses of causation. It appears that, if we wished to model states upon which the population growth rate genuinely does counterfactually depend, we'd need to move to a finer-grained model representing precise geographical distributions or indeed beyond that, to one representing the microstate of the system comprising the population and its environment. But such models no longer represent *population size* as a cause of growth rate (but rather represent more fine-grained states as causes).

It thus appears that *mr* exceptions, unlike *bf* exceptions, pose a genuine difficulty for the ability of high-level laws to underwrite causal relations. In the next two sections, I'll argue that, in fact, there are high-level generalizations that do not admit of *mr* exceptions and are thus able to underwrite genuine high-level causal relations. These generalizations are *probabilistic approximations* to *mr* high-level generalizations and the causal relations that they underwrite are also probabilistic.

## 7. Probabilistic High-Level Generalizations

In Section 6, I argued that the apparent fact that high-level generalizations often admit of *mr* exceptions—both of the pure, *absolute* type, and the mixed, *mere exception*, type—poses a problem for the accommodation of high-level causal relations in a way that the fact that they typically admit of *bf* exceptions does not. I will now argue that the most promising line of response to this problem is to maintain that high-level generalizations that admit of *mr* exceptions are approximations to probabilistic generalizations that *don't* admit of *mr* exceptions, and that these probabilistic generalizations can support relations of probabilistic high-level causation.

The case is perhaps clearest when we consider the IGL. Let's start out by ignoring the possibility of microphysical influences from outside the gas-in-container system. It was observed in Section 6 that some possible micro-realizers of a gas-in-container system lead to later microstates of the system that don't realize macrostates in which  $PV=nRT$  is approximately satisfied and, likewise, changes to the values of  $P$ ,  $V$ ,  $n$ , or  $T$  will sometimes issue in microstates that don't lead to the gas approximately conforming to  $PV=nRT$ . However, we can give a precise sense to the claim that such microstates are 'rare': namely, the volume they occupy in the phase space of the system (and—given scattering—in the region of phase space associated with any initial macrostate of the sys-

tem) is very small indeed on the standard, Lebesgue, measure. Indeed, standard statistical mechanics (SM) furnishes us with a way to assign probabilities to a system being in such a microstate. It does this by invoking the Boltzmann distribution—a distribution that's uniform with respect to the Lebesgue measure. Consequently, standard SM entails not only that such microstates are 'rare', but that they are corresponding improbable. In effect, then, SM entails a probabilistic version of IGL, which entails that it's highly probable that an (approximately) ideal gas will (approximately) obey  $PV=nRT$  because it's highly probable that its initial microstate is one that leads, via the fundamental dynamical laws, to its doing so (cf. Roberts 2014: 1782 Footnote). Given scattering, this is so no matter what the initial macrostate of the system is. Cases in which the initial macrostate of a system are realized in one of those rare ways that leads to the system not approximately conforming to  $PV=nRT$  are not exceptions to this probabilistic version of IGL. Such cases are covered by the probabilistic generalization, but are simply assigned a low probability by it.

Such a probabilistic version of IGL is able to underwrite causal relations. For instance, consider the claim that the decrease in the volume of a container holding a gas held at a constant temperature was a cause of the increase in pressure. While a probabilistic IGL doesn't imply that the decrease in container volume (even in circumstances where the temperature and gas amount is held constant, and the gas isn't subject to a strong electric field, or anything like that) was *sufficient* for the gas pressure to rise (since there are possible realizers of the initial macrostate of the gas post-volume-decrease that lead to quirky macro-behavior in which the pressure on the container walls doesn't increase, at least for a time), it does imply that the decrease in container volume was sufficient (in the circumstances) for a very high (SM) probability of the gas pressure rising.

Likewise, consider situations in which the container size had not been decreased. A probabilistic IGL again implies a very high probability that the gas (approximately) obeys  $PV=nRT$ . For the most likely such situations (in which, for example, no one has significantly increased the temperature of the gas), there is thus a high probability that the pressure of the gas is relatively low. The probability of the gas pressure being relatively high is thus higher conditional upon the gas volume being decreased than it is conditional upon the gas volume not being decreased.

Similarly, while a probabilistic IGL doesn't support the (false) counterfactual 'If the gas volume hadn't been decreased, then the gas pressure would have been lower' (since the macrostate of the gas and container system just *might* have been realized in one of those rare microphysical ways that leads to unusually high pressure, at least for a time), it *does* support the (true) counterfactual 'If the gas volume hadn't been decreased, then the (SM) probability of its pressure being lower would have been high'. It supports this counterfactual because, in each of

the *closest* worlds in which the gas volume isn't decreased (which don't include worlds in which, for example, someone rapidly increases the temperature of the gas or passes a strong electric current through it<sup>31</sup>), the macrostate of the gas and container system has a much lower SM probability of being realized in a way that leads to higher pressure than does the actual macrostate of the system in which the volume of the container has been decreased.

The fact that the container volume's being decreased raises the probability of the gas pressure increasing (in both the conditional probability and counterfactual senses described in the previous two paragraphs) is just the sort of fact to which probabilistic analyses of causation appeal.<sup>32</sup> The sequence, it might thus be claimed, is a paradigm case of probabilistic causation. No wonder nomic regularity and counterfactual theories (including their structural equations variants) were unable to accommodate it: these are theories of deterministic causation being (mis)applied to a case of probabilistic causation!

So far we've been setting aside consideration of microphysical environmental influences. But handling them doesn't pose too much of a problem. Note that, where a gas-in-container system is free of the sorts of *macroscopic* interferences that have previously been described (e.g., a strong electric current being passed through the gas), there are very few initial microstates of the gas-in-container-plus-its-environment system that lead by the fundamental dynamics to a later state in which the gas does not (approximately) obey IGL. Moreover, given scattering, we know that any microstate of the gas-in-container-plus-its-environment system that leads to IGL violation for the gas differs from one that doesn't by only a small perturbation. Therefore, given whatever microstate the environment of the gas-in-container happens to be in, the SM probability that the macrostate of the gas-in-container system is realized in one of those ways that does not combine with the actual microstate of the environment in order to yield IGL violation is extremely high. The probability of conformity to the IGL conditional upon any possible macrostate of the gas-in-container is therefore high. Also, as-

---

31. The fact that a probabilistic version of IGL still admits of *bf* exceptions thus doesn't prevent it from supporting the sorts of counterfactuals about probabilities—or, indeed, as was seen in the previous paragraph, conditional probabilities—appealed to in popular analyses of probabilistic causation. Responding to a suggestion of Earman and Roberts (1999) (cf. Roberts 2014), Hüttemann and Reutlinger (2013) and Reutlinger (2014) (cf. Kowalenko 2014: 142 Footnote) provide arguments that suggest that *bf* exceptions can't in general themselves be fully modelled probabilistically. If they're right, we should not expect a probabilistic IGL that admits of *bf* exceptions to be replaceable by a probabilistic IGL that does not.

32. Probability-raising understood in the conditional probability sense figures in the probabilistic analyses of causation developed by Reichenbach (1971), Good (1961a; 1961b), Suppes (1970), Kvart (2004), and Glynn (2011). Understood in the counterfactual sense, it figures in the analyses given by Lewis (1986), Menzies (1989), and Ramachandran (2004). Twardy and Korb (2011), Halpern (2016: 46–53), and Fenton-Glynn (2017b) develop accounts of probabilistic causation that are analogues of deterministic structural equations approaches.

suming as before that the sets of nearest possible worlds in which the macrostate of the gas-in-container differs from its actual macrostate in regard of (say) the temperature of the gas don't contain worlds in which a strong electric current is passed through it (or anything like that) this means that, when we counterfactually suppose the temperature of the gas-in-container system to be varied, there remains a high SM probability that it will conform to the IGL.

It's plausible that probabilistic approximations to special science generalizations that admit of *mir* exceptions can be derived in a similar way to a probabilistic version of the IGL: namely by imposing probability distributions over more fine-grained state spaces, though these state spaces needn't be phase spaces. We've seen that the geographical distribution of members of a population can be modelled using spatial logistic models. These replace the variable in LE and WLM that represents overall population size in the habitat with variables that represent the sizes of sub-populations in spatial sub-regions of the habitat. They thus represent more fine-grained states that can only be fully captured by a more fine-grained state space than one parameterized by the variables in LE or WLM. Yet, as we've seen, in transitioning to these more fine-grained models we lose the ability to treat the coarse-grained population *size* as a cause of population growth.

Fine-grained spatial logistic models require more information as input to yield predictions as output than do more coarse-grained models: specifically, they require fine-grained information about the geographical distribution of members of the species, as opposed to just the relatively coarse-grained information about its overall size. Instead of inputting all this fine-grained information, ecologists often model the initial geographical distribution of a population by means of a probability distribution over its *possible* initial geographical distributions (Vandermeer & Goldberg 2013: 126–142). A natural way to do this is to impose a uniform probability distribution over the state space  $\mathbb{S}_{SLM}$  parameterized by the variables in a spatial logistic model (which include variables representing the size of sub-populations in spatial sub-regions of the habitat, rather than a variable simply representing overall population size) and then to condition that distribution on the fact that the system occupies the subregion  $R$  of  $\mathbb{S}_{SLM}$  corresponding to the initial overall size of the population (see Law et al. 2003: 254, 257; cf. Coe, Ahnert, & Fink 2008). For the sorts of population that are normally well-modeled by (say) LE the measure of points in any such subregion  $R$  that leads, by the dynamics, to LE-like growth is presumably (given scattering) high. The uniform distribution will therefore entail a high probability—though not equal to one—for LE-like growth. This yields a probabilistic version of LE (cf.



Law et al. 2003).<sup>33</sup> Non-LE-like behavior does not constitute an ‘exception’ to this probabilistic generalization, for such behavior is assigned an explicit (low) probability by the generalization. The suggestion is that such a generalization, like a probabilistic version of SLT, can support the sorts of probability-raising relations appealed to in probabilistic analyses of causation.

As has already been noted, the dynamics through  $\mathbb{S}_{SLM}$  are modeled by a spatial logistic model. Yet, as we have noted, a spatial logistic model is itself likely to admit of *mr* exceptions. If for no other reason, this is because points in  $\mathbb{S}_{SLM}$  correspond to regions in the system’s phase space and, given scattering, it seems that any such region will contain a set of points (very small in measure) that lead to widespread entropy decrease and—given the dependence of members of the population upon entropy increasing processes for their survival—a collapse in population numbers rather than normal growth. Yet (given scattering) the probability distribution over the system’s phase space that’s uniform on the standard measure entails a (very low) probability that the microstate is an element of this set, conditional upon whichever point in  $\mathbb{S}_{SLM}$  the system happens to be in. But then imposing such a distribution gives us a probabilistic approximation to the original deterministic spatial logistic model, one that doesn’t admit of *mr* exceptions. This generalization entails probabilistic dependencies of population growth rates on the way the population is geographically distributed, and such dependencies are precisely the sort to which standard analyses of probabilistic causation appeal. Such a probabilistic approximation to a spatial logistic model thus seems able to support relations of probabilistic causation between (for instance) the geographical distribution of a population and its growth rate.

Recall that a probabilistic version of the LE—which models a system’s dynamics through a state space parameterized by the variables  $n_v$ ,  $r_v$  and  $K$ —can be arrived at by imposing a probability distribution over the more fine-grained state space,  $\mathbb{S}_{SLM}$ , parameterized by the variables featuring in a spatial logistic model of the system. Assuming the distribution is uniform, and the dynamics through  $\mathbb{S}_{SLM}$  are deterministic, the probability of LE-like growth given that the system is at a point in the more coarse-grained state space corresponding to a region  $R$  of  $\mathbb{S}_{SLM}$  is proportional to the measure of the set of points within  $R$  that, when evolved forward according to the deterministic dynamics that govern the system’s evolution through  $\mathbb{S}_{SLM}$  (i.e., a deterministic spatial logistic model), yield LE-like growth. For a system that is normally well-modelled by the LE

---

33. Earman and Roberts (1999: 464–465) briefly suggest that being able to impose a measure over the realizers of the states described by a high-level science might be the key to giving determinate truth-conditions to high-level scientific generalizations, though they regard this as necessary for making sense of the notion that a generalization holds in ‘most of its intended applications’ rather than for deriving an explicitly probabilistic approximation to the original non-exceptionless generalization.



this measure is presumably high. But we've now noted that the true dynamics through  $\mathbb{S}_{SLM}$  are unlikely to be deterministic—but rather are given by a probabilistic approximation to a spatial logistic model. Nevertheless, the derivation of a probabilistic approximation to LE goes through almost as before: the probability of LE-like growth given that a system is at a particular point in the state space parameterized by the variables that figure in the LE is now a weighted average (or 'probability mixture') of the probabilities with which points in the corresponding region  $R$  of  $\mathbb{S}_{SLM}$  issue in LE like growth, with the weights (the 'mixture weights') given by the (uniform) distribution over  $R$ . Where a system is normally well-modeled by the LE this probability mixture presumably yields a high probability for LE-like growth. Again, such a probabilistic LE is able to support the sort of probabilistic dependence relations in terms of which probabilistic causation is standardly analyzed.

In more technical jargon, the target system (here a population and its environment) can be represented by a hierarchical model. This comprises a nested series of 'levels' or state spaces  $\langle \mathbb{S}_1, \mathbb{S}_2, \dots, \mathbb{S}_n, S \rangle$ , where  $S$  is the phase space of the system, and where  $\mathbb{S}_1, \mathbb{S}_2, \dots, \mathbb{S}_n$  are coarse-grained state spaces, with points in  $\mathbb{S}_1$  corresponding to regions in  $\mathbb{S}_2$ , points in  $\mathbb{S}_2$  correspond to regions in  $\dots \mathbb{S}_n$ , and points in  $\mathbb{S}_n$  corresponding to regions in  $S$ . The probability that the future-truncation of the system's history at the level of  $\mathbb{S}_i$  will be  $\mathbb{h}_{t+}^i$  given that the point in  $\mathbb{S}_i$  that it occupies at  $t$  is  $\mathbb{h}^i(t)$  is given as a mixture of the probabilities with which the points in the region of  $\mathbb{S}_j$  ( $j - i = 1$ ) corresponding to  $\mathbb{h}^i(t)$  lead to future-truncated histories  $\mathbb{h}_{t+}^j$  that realize the future-truncated history  $\mathbb{h}_{t+}^i$ , where the mixture weights are given by a distribution over  $\mathbb{S}_j$ . (If the dynamics of the system through  $\mathbb{S}_i$  are non-Markovian, then we might wish to ask what the probability is that the future-truncation of the system's history at the level of  $\mathbb{S}_i$  will be  $\mathbb{h}_{t+}^i$  given that its (past-)truncated history is  $\mathbb{h}_t^i$ . That would involve taking a mixture of the probabilities with which (past-)truncated histories  $\mathbb{h}_t^j$  that realize the (past-)truncated history  $\mathbb{h}_t^i$  lead to future-truncated histories  $\mathbb{h}_{t+}^j$  that realize the future-truncated history  $\mathbb{h}_{t+}^i$ , where the mixture weights are again given by a distribution over  $\mathbb{S}_j$ .) The same points apply *mutatis mutandis* to the derivation of probabilities of future-truncated histories through  $\mathbb{S}_n$  from the (Boltzmann) distribution over  $S$ .

We've been imagining our population plus environment system to be modelled by a hierarchical model comprising three levels: the fundamental level represented by the system's phase space, the level parameterized by the variables in a spatial logistic model, and the level parameterized by the variables in the LE. But it might be fruitful to conceive of it as modelable by a hierarchical model comprising even more levels (for instance, perhaps there are fecund representations of the system at levels that are higher than the phase space, but lower than the space parameterized by the spatial logistic model). In this case, derivation

of probabilities for the system's evolution through its highest-level state spaces may proceed via probability mixing with respect to these levels too.

In general, which more fine-grained levels should be invoked in deriving probabilities for the evolution of a system through a state space of a given level will depend upon how well 'fitting' the resulting probabilities are, in a sense to be discussed in the next section. As Callender and Cohen (2010: 437–438, 443–444) argue, it's rather plausible that the best-fitting probabilities for sciences like ecology aren't derived by directly imposing a probability distribution on the region of a system's phase space compatible with its instantiation of various values of ecological variables (like population size). Going 'via' state spaces of intermediate levels in deriving ecological probabilities can plausibly improve fit.

Hofer (2004: 110–113) briefly considers a response to the problem posed for high-level causation by what I have been calling *mr* exceptions that, like that developed here, takes high-level causation to be probabilistic. He objects to such a response on the grounds that the probabilities that must be invoked are not robust enough to underwrite 'genuine causation'. The nub of his central objection appears to be expressed in the following passage:<sup>34</sup>

It may be supposed that one can appeal to a 'natural' distribution [over initial conditions] on something like thermodynamic [and statistical mechanical] grounds, to shore up the idea that a certain probability  $x$  emerges naturally. . . . For simple problem set-ups like shielded coin-flippers or gases in boxes and so forth – i.e., for problems that have a uniform *micro*-description that we can handle with physical theory – this claim will sometimes be plausible. But unlike 'rigid rectangular box of volume  $V$  with a Newtonian gas of identical particles in a [Maxwell-Boltzmann] distribution at temperature  $T$ ', 'Smokes 2 packs a day, . . .' has no canonical distribution of micro-descriptions; and the boundary conditions (i.e. external influences, at the micro-level) are even less plausibly regimentable. There is just no reason to suppose that one set of micro-states of *males-smoking-2-packs-daily+environment* counts as 'normal', 'probable' or what have you, while another one does not. (Hofer 2004: 112 Footnote)

---

34. Another worry that he voices – which is that there is something philosophically problematic about assigning (objective) probabilities to initial conditions, as is done in SM (and ecology) – is one that I don't share (nor indeed does it appear to be shared by more recent temporal parts of Hofer himself – see Frigg and Hofer 2010; 2015). In a similar vein, Schiffer (1991: 8) is skeptical that generalizations admitting of what Fodor calls 'absolute exceptions' (and 'mere exceptions') can be treated as (approximations to) probabilistic laws because he thinks that no probability that isn't derived from fundamental physics could be an objective chance. In the next section I will describe a plausible metaphysical picture which vindicates probability assignments to initial conditions and the interpretation of probabilities for high-level states derived from such assignments as objective chances.

Hoefer is right that isolated thermodynamic systems admit of a natural (or at least standard) micro-characterization that takes the form of the phase space state description of those systems, and that there's a natural measure (the Lebesgue measure) and probability distribution (the one that's uniform on the Lebesgue measure) that yields probabilities for such a system's being in such-and-such a type of microstate given that it's in so-and-so a macrostate. But it's obviously not just gases in boxes that have phase spaces. Any thermodynamically isolated system—including the universe as a whole—has a well-defined phase space associated with it to which the standard measure and probability distribution can be applied. This means that, for any non-isolated system (such as a person smoking cigarettes, or an ecosystem), a probability distribution over its external micro-influences is in principle derivable given the initial macrostate of some isolated system of which it is a part (in the worst case, the universe as a whole). Although we can't derive such a probability distribution in practice, we nevertheless have evidence about its nature. For instance, the fact that gases in boxes typically conform (approximately) to the IGL is evidence that the probability of such things as high-velocity particles from outside interfering in such a way as to put the system on an entropy-decreasing trajectory is very low. The fact that certain (macroscopically) identifiable sorts of population in certain (macroscopically) identifiable sorts of circumstance typically conform to LE is evidence that the probability of micro-influence from outside the system comprising the population and its environment that is such as to lead to significant LE violation is low. And, to use Hoefer's example, the fact that smokers contract lung cancer much more frequently than non-smokers is evidence that the probability of micro-influence from outside (say) a system comprising *the world-line of a smoker together with those of the cigarettes that spatio-temporally intersect with her world-line during the period of their intersection* that is such as to prevent damage to the cells lining her lungs is low.

So I don't think there is a failure of 'regimentability' of 'external influences, at the micro-level' on such systems that poses problems for their being subject to probabilistic high-level causal relations. But Hoefer seems to have another worry too: namely, that such systems lack a 'uniform *micro-description*'. Now clearly, if we are prepared to include 'enough' environment (so that the system plus its environment) constitutes a thermodynamically isolated system, then it will have a canonical micro-description, namely the system's *phase space* state, which is well-defined when the system-plus-environment is thermodynamically isolated. Yet scientists seek to derive probabilities for the behaviors of non-isolated systems from distributions over lower-level state spaces where those state spaces are coarser-grained than phase-spaces. For example, as we've seen, recognizing that the geographical distribution of a population can make a difference to its growth rate, ecologists sometimes seek to derive probabilities for the

future growth rate of a population by imposing a probability distribution over the possible initial *geographical distributions* of members of the population given the initial population size (Vandermeer & Goldberg 2013: 126–142). Of course this is only successful because ecologists are able to recognize when the sorts of macroscopic background factors (e.g., culling programs) liable to interfere with normal growth are absent, or because these are sufficiently infrequent, and because the sorts of microscopic external influences (which presumably ecologists are unlikely to recognize unless they realize salient macro-states) liable to interfere significantly are sufficiently low probability. Yet part of Hoefer’s worry here might be that the lower-level state descriptions appealed to by special scientists, and the probability distributions over them that are invoked, have nothing like the canonical status that phase space descriptions and the Boltzmann distribution invoked in SM do.

I think the correct response, which shall be developed in the next section, is that the true chances for processes such as those under discussion are those probabilities entailed by the theorems of the ‘best’ axiom system either for the world (see Lewis 1994) or for the special sciences concerned with the processes in question (see Callender & Cohen 2009; 2011). What those chances are will depend upon the state spaces and the probability distributions over them that are invoked in such systematizations which, depending on our degree of skepticism about the current state of science, we may presume to correspond less or more closely to some of those currently invoked by scientists.

Before saying more about this, it is worth noting that Hoefer may not be entirely satisfied with this response. At a couple of points Hoefer (2007: 107, 111) suggests that Humean accounts of laws and chances—of which the Best System approaches I have alluded to are normally taken as a variety—don’t give us laws and chances that are robust enough to underwrite ‘genuine’ causal relations. To the extent that Hoefer is just expressing the common worry about Humeanism—that because Humean ‘laws’ and ‘chances’ supervene upon, but don’t ‘govern’ the mosaic they lack the modal robustness to play the law and chance role in explaining and being stable under counterfactual assumptions—I have nothing to add to the (to my mind, plausible) responses to this general worry given by Lewis (1994: 478–479), Loewer (2012: 130–132), and others. But perhaps there’s more to Hoefer’s worry. At some stages, he seems to worry that such accounts won’t yield ‘determinate’ high-level probabilities (Hoefer 2004: 111). This seems connected with his worry that there do not appear to be canonical micro-descriptions and probability distributions for modelling certain systems.

On the Best System approach, determinate chances for high-level processes will be entailed provided that (a) there is a determinate best system for our world; (b) that determinate best system entails probabilities for high-level processes; and (c) the probabilities that the determinate best system entails are

themselves determinate. In Section 8, I will argue that all serious candidates for Best Systemhood entail probabilities for high-level processes. However, I'm sympathetic to the denial of both (a) and (c). I'm sympathetic to the denial of (c) because fairly compelling arguments have been advanced to suggest that certain high-level processes are best interpreted as being subject to *imprecise*—that is, (non-singleton) *set valued*—chances, rather than precise chances (see, for example, Fine 1988 and Fierens, Rêgo, & Fine 2009 and references therein).<sup>35</sup> But, if this is correct, then it is not just *Humeans* who ought to make room for imprecise chances in their ontology. I'm also sympathetic to the denial of (a)—that is, I'm sympathetic to the view that there is not a unique best system for our world. As I've argued elsewhere (see Dardashti, Glynn, Thébault, & Frisch 2014; Fenton-Glynn 2017a; Fenton-Glynn 2017c), I don't think this is a problem for Humeans, but rather I think that this gives Humeans an *additional* reason to believe that the chances for our world are imprecise.

Still, the chances pertaining to high-level processes need not be precise for them to underwrite high-level causal relations. Standard probabilistic analyses of causation appeal to the central idea that (perhaps when the values of certain other variables are held fixed at appropriate values) causes *raise* the probability of their effects either in the sense that the probability of the effect *conditional* upon the cause (and the values of certain other variables) is greater than the probability of the effect *conditional* upon the absence of the cause (and the values of certain other variables) or in the sense that the probability of the effect *would have been* lower than it actually was if the cause had been absent (and if certain other variables had taken certain appropriate values). Note that the probabilistic facts appealed to by standard probabilistic analyses of causation are thus *qualitative*, comparative facts about certain probabilities being lower or higher than others, not precise numerical facts about their values. Even if probabilities are imprecise, these qualitative facts can still determinately hold. For instance, as arguments given in (Dardashti et al. 2014), (Fenton-Glynn 2017a), and (Fenton-Glynn 2017c) indicate, insofar as it's plausible that (high-level) chances are set valued, it's extremely plausible that the minimum chance for a gas's pressure increasing non-negligibly if its volume is decreased non-negligibly (and its amount and temperature stay constant, and no-one passes a current through the gas, etc.) exceeds the maximum probability of its pressure rising non-negligibly if its volume is *not* decreased non-negligibly (and its amount and temperature stay constant, and no-one passes a current through the gas, etc.).

I have sketched what I take to be a scientifically and metaphysically plausible account of how there can be high-level causation, even though the gen-

---

35. This may not be true only of *high-level* events/processes (see Suppes & Zanotti 1991; Hartmann & Suppes 2010).



eralizations described in the high-level sciences typically admit of *mr* (and *bf*) exceptions. Now it's possible that Hoefer may still insist that the relations that I have argued to exist, and to constitute causal relations between high-level states (roughly probability-raising relations, where the probabilities in question derive from probability distributions over lower-level state spaces, may be imprecise, and are deemed chances by Best System-like accounts) don't constitute what he refers to often as 'genuine causation' and once as "what most philosophers would call a robust, genuine form of causation" (Hoefer 2004: 111). But I don't see why we should wish to maintain this. Hoefer believes that what he calls 'genuine causation' is something that seems to be incompatible with determinism "in a complex world such as the one we inhabit" (Hoefer 2004: 99–100). Moreover, it's difficult to see that his arguments turn in any essential way upon ours being a *deterministic* world (Hoefer 2004: 110 Footnote). So I suggest that, if Hoefer is going to require that 'genuine causation' requires something more than the relations that I have described, his conception of what would constitute 'genuine causation' is too demanding. Certainly it seems to me much more revisionary of our ordinary and scientific (and indeed philosophical) ways of thinking and talking about the world to deny that there is genuine (high-level) causation than to accept the account of what that relation consists in that I have here described, which seems to me both scientifically and metaphysically plausible.

One final piece of unfinished business remains, which is to argue in more detail that the probabilities entailed by the high-level probabilistic generalizations that I've described in this section do indeed constitute genuine chances, which are therefore apt to underwrite genuine relations of high-level causation. This is the topic of the next section. As already indicated, I will draw upon Best System-style analyses of chance to argue the point.

## 8. High-Level Chances?

In the previous section I argued that there are high-level probabilistic generalizations that do not admit of *mr* exceptions and that are able to support the kind of probabilistic dependencies to which popular accounts of probabilistic causation appeal. Yet one might wonder whether the probabilities in question are genuine objective chances, given that they don't derive from the microphysical laws, but rather their derivation involves imposing probability distributions (which are not themselves derived from the fundamental dynamic laws) over underlying state spaces. It seems quite plausible that such probabilities must be objective chances if they are to underwrite genuine high-level causal relations.

In recent years, a significant number of philosophers of science have argued that there are genuine objective chances that don't derive from the fundamental



dynamic laws alone.<sup>36</sup> A particularly popular argument for this view—though by no means the only argument<sup>37</sup>—appeals to the claim that the popular Best System Analysis (BSA) of laws and chance—which received its most detailed development by Lewis (1994)—as well as variants upon it counts SM probabilities, and plausibly also probabilities associated with probabilistic special science generalizations, as genuine objective chances.<sup>38</sup>

According to the BSA, the laws are the axioms and theorems of that axiom system pertaining to what goes on in the universe that strikes the best balance between the theoretical virtues. The chances are the probabilities entailed by those axioms and theorems. The specific theoretical virtues appealed to by Lewis are simplicity, strength, and fit.<sup>39</sup> According to Lewis (1994: 480), a system is *strong* to the extent that it says “either what will happen or what the chances will be when situations of a certain kind arise”. The reason to think that adding axioms to a system that already entails the fundamental dynamic laws so that it entails probabilistic macrophysical generalizations, like a probabilistic version of IGL, and probabilistic special science generalizations, like probabilistic versions of LE and WLM, is that this increases the strength or informativeness of the system in question. The reason is that such high-level generalizations tell us what the chances will be when situations arise that are of kinds concerning which the fundamental dynamic laws are silent. The kinds of situation in question are, of course, situations of *high-level* kinds.

Take, for example, situations of the high-level kind *being an approximately ideal gas of amount  $n'$  and temperature  $t'$  in a container of volume  $v'$* . The fundamental dynamic laws don't tell us what the pressure of the gas will be, or what the probability distribution over various possible pressures will be when situations of this kind arise. They tell us only about what the pressure will be in situations of microphysical kinds like *being at such-and-such a point in phase space*. But, because of its multiple realizability, the fact that a system is of the high-level kind *an approximately ideal gas of amount  $n'$  and temperature  $t'$  in a container of volume  $v'$*  does not entail what point in its phase space the system is at. By contrast, a probabilistic IGL *does* provide us with information about what the probability distribution over possible pressures is when situations of this high-level kind arise. Consequently, a system that entails a probabilistic IGL is more informative than one that entails the fundamental dynamic laws alone.

36. A partial list includes Albert (2012), Dunn (2012), Emery (2015), Frigg and Hoefer (2010; 2015), Glynn (2010), Ismael (2009; 2011), and Loewer (2001).

37. For different arguments, see Emery (2015) and Ismael (2009; 2011).

38. Such arguments have been advanced by, *inter alia*, Loewer (2001; 2007; 2012), Callender and Cohen (2009; 2010), Dunn (2011), Frisch (2014b), Glynn (2010), and Weslake (2014).

39. A system's *fit* is the probability that it assigns to the actual course of history (Lewis 1994: 480). See Elga (2004) for a critique of, and suggested amendment to, Lewis's notion of fit. The question of the correct notion of fit needn't detain us here.

Likewise, a system entailing a probabilistic LE tells us what the chances are when a situation arises of the kind *a population with size  $n'_1$  and intrinsic growth rate  $r'_c$  in a habitat with carrying capacity  $k'$* . Again, the fundamental dynamic laws alone don't tell us what the chances are when a situation of this kind arises, since this kind of situation is multiply realizable by microphysical kinds.

In general, fundamental dynamic laws and probabilistic special science or macrophysical generalizations entail probability distributions conditional upon different sorts of proposition. High-level generalizations entail probability distributions conditional upon propositions about high-level states or kinds that a system instantiates  $P(\cdot | \mathbb{h}(t))$ , while the fundamental dynamic laws entail distributions only conditional upon propositions specifying a system's microstate  $P(\cdot | h(t))$ .<sup>40</sup> There is no conflict between divergent conditional chance distributions with different conditions. Indeed, an axiom system that entails both conditional distributions is more informative than one that entails only one (and leaves the other undefined).

It is a good question at exactly how high a price in simplicity this greater strength or informativeness is bought. The questions of exactly how much simplicity the addition of the extra axioms required to entail the probabilities of SM costs, and of whether the strength gained is worth the price, are discussed (and disputed) by Loewer (2001: 617–618), Schaffer (2007: 130–131), Hoefer (2007: 560), and Glynn (2010: 59–63).<sup>41</sup> The difficulty is that there aren't obviously most reasonable simplicity and informativeness metrics to apply (cf. Lewis 1994: 479). This, together with the fact that it's not obvious how to trade off simplicity against informativeness, makes it difficult to answer the latter question. Similar issues would obviously arise when we consider whether the best system includes axioms sufficient to entail probabilistic special science generalizations, such as probabilistic versions of LE, WLM, or SLM.

In fact, I'm inclined to think that a *a prioristic* discussion over what the right simplicity and strength metrics to apply are, and what the correct exchange rate is between these virtues, gets things backwards. A more naturalistic approach would look to science and the generalizations there that are treated as playing the law role in supporting counterfactuals, underwriting causal explanations, and entailing probabilities that are taken by scientists to play the chance role of explaining outcomes and frequencies of outcomes, constraining credences, and so forth. The idea would then be to reverse-engineer the standards of simplicity and strength, and the exchange rate between them, that scientists are implicitly committing to in their theory building. If it turns out that scientists, across disciplines, implicitly adopt similar standards, then we might regard the disci-

40. The latter distributions are trivial—that is, the probabilities are all 1s and 0s—if the fundamental dynamics are deterministic.

41. See also Callender and Cohen (2009: 10) and Dunn (2011: 91).

plines as jointly contributing—in their own separate ways—to the building (or discovery) of a best system for the universe as a whole. If instead it turns out that rather different standards are adopted in the various disciplines, then—rather than the traditional BSA—we might prefer something along the lines of the so-called Better Best System Analysis (BBSA), developed by Callender and Cohen (2009; 2010),<sup>42</sup> as our metaphysical account of laws and chances.

Briefly, Callender and Cohen’s proposal draws upon Lewis’s (1983: 367–368) observation that a system’s simplicity depends upon the vocabulary in which it is expressed. But, rather than following Lewis in restricting the systems under consideration to those whose axioms contain only perfectly natural kind predicates (more on this point in a moment), their idea is that best systemhood should be taken to be *relative* to a set of basic kinds  $K$  (or predicates  $P_K$ ). Relative to different sets of kinds, different axiom systems strike the best balance between simplicity, strength, and fit. A generalization is a law relative to  $K$  just in case it is a theorem of the Best System relative to  $K$ , and a probability is a chance relative to  $K$  if it’s entailed by a generalization that is a law relative to  $K$ .

On Callender and Cohen’s view, the generalizations of a special science (such as ecology) count as laws of that science if they are theorems of the best system relative to the science’s proprietary kinds or predicates (e.g., the ecological kinds).<sup>43</sup> Callender and Cohen (2010: 437–438) and Callender (2011: 103, 112) themselves suggest that the best axiomatizations for various special sciences will include probability distributions over underlying state-spaces (which need not be phase spaces), where those distributions closely match the frequencies with which higher-level properties are realized in the state spaces in question. As we’ve seen, including such distributions is key to deriving probabilistic approximations to generalizations like LE, WLM, and SLM. On Callender and Cohen’s view, the probabilistic theorems generated by the resulting axioms are probabilistic laws of the sciences in question, and the probabilities that they entail are chances of the sciences in question.

Although this isn’t a suggestion that Callender and Cohen explicitly make, if we wish to take a naturalistic approach to standards of simplicity, strength, and

---

42. For proposals similar to Callender and Cohen’s, see Schrenk (2008) and Dunn (2011: 88–90).

43. Weslake (2014), responding to Callender and Cohen’s proposal, suggests that axioms needed to derive SM result from a best system for the *conjunction* of the fundamental kinds and the thermodynamic kinds. It seems that Callender and Cohen (2009: 10, 28) and Callender (2011: 106–112) are sympathetic. Likewise, it seems that we would need to conjoin the ecological kinds with certain underlying kinds in order to generate a system that entails an appropriate probabilistic version of, for example, the LE. Indeed, Callender and Cohen (2009: 24) make the point that the BBSA implies that there are laws relative to *any* set of kinds relative to which a meaningful best system competition can be conducted and won (though we may simply not be interested in the laws relative to certain vocabularies—for instance, gruesome vocabularies).

balance, and we find that such standards vary from discipline to discipline, then we might take the view that the laws of a special science are determined by a best system competition relative to that special science's proprietary vocabulary *and that science's proprietary standards of simplicity, strength, and balance.*

On the other hand, if we regard the enterprise of science as more unified, with each discipline making a different contribution to the construction of an overall best system for the universe, then we may prefer the original BSA as our metaphysical picture of laws. Still, we would need to address Lewis's point that the simplicity of a system is relative to the vocabulary in which it's expressed. If we follow Lewis in requiring that a system is only simple in the pertinent respect if it's simple when expressed in perfectly natural kind terms, we're liable to rule out the existence of laws or chances that are not derivable from the laws and chances of fundamental physics alone. That's because any axiom pertaining to the kinds of a higher-level science (such as an axiom concerning temperature, pressure, and volume) is likely to be syntactically very complex when translated into a language with only perfectly natural kind terms (cf., Schaffer 2007: 130; Callender & Cohen 2009: 14). Consequently, such an axiom is not likely to figure in the Best System.

One option would be to follow Callender and Cohen in treating best systemhood as vocabulary-relative. However, a more conservative modification of the BSA to accommodate high level laws and chances is possible. To see this, observe that, as Lewis recognizes (1983: 368), naturalness admits of degrees. Naturalness of the predicates that it employs might reasonably be taken to be a theoretical virtue, to be weighed alongside the simplicity and strength of a system. If an axiom system is able to achieve significant simplicity and strength by employing not-too-unnatural predicates like 'temperature' or 'carrying capacity', then it's a plausible best system. Again, I see this as a more naturalistic approach than a *prioristic* restrictions on the relevant vocabulary (cf. Callender & Cohen 2009: 17–20). For one thing, it gives science a stronger role in determining what are the predicates that may figure in the laws of nature; for another, it accommodates scientists' actual judgments about which generalizations are lawlike (in the sense of supporting counterfactuals, predictions, and causal explanations, being confirmed by their instances, and so forth) and entail probabilities that play the chance role in guiding credence, explaining outcomes and frequencies of outcomes, and so on.

The suggestion, then, is that an appropriate version of the BSA (or, if one prefers, the BBSA) will treat high-level generalizations—including probabilistic macrophysical generalizations, such as a probabilistic IGL, and probabilistic special science generalizations, perhaps including a probabilistic version of WLM (or indeed a probabilistic LE or SLM)—as laws, or at least as lawful enough to support counterfactuals, to be such that the probabilities they entail are genuine

objective chances, and thus to support causal relations. After all, this appears to be the way that scientists themselves treat them. For example, Linquist, Gregory, Elliot, Saylor, Kremer, and Cottenie (2016: 130)—speaking about ecology—state that “current practices in the discipline . . . collectively point in the direction of causal generalizations at all levels”. Indeed, a naturalistic approach to laws of nature suggests that, even if one doesn’t think that some variant on the BSA (or BBSA) is correct, one’s account of laws should endorse certain probabilistic high-level generalizations, such as those that we have considered here, as sufficiently lawful to support causal relations (cf. Ismael 2009; 2011; Emery 2015).

## 9. Conclusion

It has been argued that the problem posed for high-level causation by the apparent absence of exceptionless high-level generalizations can be overcome. There’s one class of exception—*bf* exceptions—of which high-level generalizations admit, but that doesn’t prevent them from underwriting high-level causal relations. There’s another class of exception—*mr* exceptions—of which they appear to admit that does pose a threat to a generalization’s ability to underwrite causal relations. However, drawing upon the case studies of SM and ecology, I have argued that a strong case can be made that deterministic high-level generalizations that admit of *mr* exceptions are approximations to *probabilistic* generalizations that don’t admit of *mr* exceptions. These probabilistic generalizations are able to support the sort of objective chance dependencies (between high-level states) to which probabilistic analyses of causation appeal. To the extent that this generalizes, the apparent problem posed for high-level causation by the seeming lack of exceptionless high-level generalizations can be overcome.

## Acknowledgements

I would like to thank two anonymous referees and the Area Editor of this journal for detailed comments that led to a considerable improvement of this paper. For detailed comments on earlier drafts, I would like to thank Christopher Hitchcock, Thomas Kroedel, Wolfgang Spohn, and Joel Velasco. I would also like to thank audiences to presentations of various forerunners of this paper: namely those at the 2010 *Causation Across Levels Workshop* at the IHPST in Paris, a 2010 meeting of the *Causality and Probability Research Colloquium* at the University of Konstanz, the 2012 workshop *The Objective Reality of Causality* also at Konstanz, and a 2012 presentation at the University of St. Andrews. Important parts of the research that led to this paper were done while I held post-doctoral positions



funded by the Deutsche Forschungsgemeinschaft (SP279/15-1) and the J. S. McDonnell Causal Learning Collaborative.

## References

- Albert, David (2000). *Time and Chance*. Harvard University Press.
- Albert, David (2012). Physics and Chance. In Y. Ben-Menahem and M. Hemmo (Eds.), *Probability in Physics* (17–40). Springer. [https://doi.org/10.1007/978-3-642-21329-8\\_2](https://doi.org/10.1007/978-3-642-21329-8_2)
- Callender, Craig (2011). The Past Histories of Molecules. In Claus Beisbart and Stephan Hartmann (Eds.), *Probabilities in Physics* (83–113). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199577439.003.0004>
- Callender, Craig and Jonathan Cohen (2009). A Better Best System Account of Lawhood. *Philosophical Studies*, 145(1), 1–34. <https://doi.org/10.1007/s11098-009-9389-3>
- Callender, Craig and Jonathan Cohen (2010). Special Sciences, Conspiracy and the Better Best System Account of Lawhood. *Erkenntnis*, 73(3), 427–447. <https://doi.org/10.1007/s10670-010-9241-3>
- Cartwright, Nancy (1980). The Truth Doesn't Explain Much. *American Philosophical Quarterly*, 17(2), 159–163.
- Coe, Jonathan, Sebastian Ahnert, and Thomas Fink (2008). When are Cellular Automata Random? *Europhysics Letters*, 84(5), 1–6. <https://doi.org/10.1209/0295-5075/84/50005>
- Dardashti, Radin, Luke Glynn, Karim Thébault, and Mathias Frisch (2014). Unsharp Humean Chances in Statistical Physics: A Reply to Beisbart. In Maria Galavotti, Dennis Dieks, Wenceslao Gonzalez, Stephan Hartmann, Thomas Uebel, and Marcel Weber (Eds.), *New Directions in the Philosophy of Science* (531–542). Springer. [https://doi.org/10.1007/978-3-319-04382-1\\_37](https://doi.org/10.1007/978-3-319-04382-1_37)
- Davidson, Donald (1970). Mental Events. In Lawrence Foster and Joe Swanson (Eds.), *Experience & Theory* (79–101). University of Massachusetts Press.
- Dunn, Jeffrey (2011). Fried Eggs, Thermodynamics, and the Special Sciences. *British Journal for the Philosophy of Science*, 62(1), 71–98. <https://doi.org/10.1093/bjps/axq012>
- Earman, John, and John Roberts (1999). *Ceteris Paribus*, There Is No Problem of Provisos. *Synthese*, 118(3), 439–478. <https://doi.org/10.1023/A:1005106917477>
- Earman, John, John Roberts, and Sheldon Smith (2002). *Ceteris Paribus* Lost. *Erkenntnis*, 57(3), 281–301. <https://doi.org/10.1023/A:1021526110200>
- Easwaran, Kenny (2014). Why Physics Uses Second Derivatives. *British Journal for the Philosophy of Science*, 65(4), 845–862. <https://doi.org/10.1093/bjps/axt022>
- Elga, Adam (2004). Infinitesimal Chances and the Laws of Nature. In Frank Jackson and Graham Priest (Eds.), *Lewisian Themes: The Philosophy of David K. Lewis* (68–77). Oxford University Press. <https://doi.org/10.1080/713659804>
- Emery, Nina (2015). Chance, Possibility, and Explanation. *British Journal for the Philosophy of Science*, 66(1), 95–120. <https://doi.org/10.1093/bjps/axt041>
- Fenton-Glynn, Luke (2016). *Ceteris Paribus* Laws and *Minutis Rectis* Laws. *Philosophy and Phenomenological Research*, 93(2), 274–305. <https://doi.org/10.1111/phpr.12277>
- Fenton-Glynn, Luke (2017a). Imprecise Best System Chances. In Michela Massimi, Jan-Willem Romeijn, and Gerhard Schurz (Eds.), *EPSA15 Selected Papers* (297–308). Springer. [https://doi.org/10.1007/978-3-319-53730-6\\_24](https://doi.org/10.1007/978-3-319-53730-6_24)



- Fenton-Glynn, Luke (2017b). A Proposed Probabilistic Extension of the Halpern and Pearl Definition of “Actual Cause”. *British Journal for the Philosophy of Science*, 68(4), 1061–1124. <https://doi.org/10.1093/bjps/axv056>
- Fenton-Glynn, Luke (2017c). Imprecise Chance and the Best System Analysis. Manuscript in Preparation. Retrieved from [https://www.academia.edu/5819239/Imprecise\\_Chances\\_and\\_the\\_Best\\_System\\_Analysis](https://www.academia.edu/5819239/Imprecise_Chances_and_the_Best_System_Analysis)
- Fierens, Pablo, Leandro Rêgo, and Terrence Fine (2009). A Frequentist Understanding of Sets of Measures. *Journal of Statistical Planning and Inference*, 139(6), 1879–1892. <https://doi.org/10.1016/j.jspi.2008.08.025>
- Fine, Terrence (1988). Lower Probability Models for Uncertainty and Nondeterministic Processes. *Journal of Statistical Planning and Inference*, 20(3), 389–411. [https://doi.org/10.1016/0378-3758\(88\)90099-7](https://doi.org/10.1016/0378-3758(88)90099-7)
- Fodor, Jerry (1974). Special Sciences (or: The Disunity of Science as a Working Hypothesis). *Synthese*, 28(2), 77–115. <https://doi.org/10.1007/BF00485230>
- Fodor, Jerry (1989). Making Mind Matter More. *Philosophical Topics*, 17(1), 59–79. <https://doi.org/10.5840/philtopics198917112>
- Fodor, Jerry (1991). You Can Fool Some of The People All of The Time, Everything Else Being Equal; Hedged Laws and Psychological Explanations. *Mind*, 100(1), 19–34. <https://doi.org/10.1093/mind/C.397.19>
- Friend, Toby (2016). Laws are Conditionals. *European Journal for the Philosophy of Science*, 6(1), 123–144. <https://doi.org/10.1007/s13194-015-0131-z>
- Frigg, Roman and Hoefer, Carl (2010). Determinism and Chance from a Humean Perspective. In Friedrich Stadler, Dennis Dieks, Wenceslao González, Stephan Hartmann, Thomas Uebel, and Marcel Weber (Eds.), *The Present Situation in the Philosophy of Science* (Vol. 1, 351–371). [https://doi.org/10.1007/978-90-481-9115-4\\_25](https://doi.org/10.1007/978-90-481-9115-4_25)
- Frigg, Roman and Hoefer, Carl (2015). The Best Humean System for Statistical Mechanics. *Erkenntnis*, 80(3), 551–574. <https://doi.org/10.1007/s10670-013-9541-5>
- Frisch, Mathias (2014a). *Causal Reasoning in Physics*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139381772>
- Frisch, Mathias (2014b). Why Physics Can’t Explain Everything. In Alastair Wilson (Ed.), *Asymmetries of Chance and Time* (221–240). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199673421.003.0011>
- Glymour, Clark and Frank Wimberly (2007). Actual Causes and Thought Experiments. In Joseph Campbell, Michael O’Rourke, & Harry Silverstein (Eds.), *Causation and Explanation* (43–67). MIT Press.
- Glynn, Luke (2010). Deterministic Chance. *British Journal for the Philosophy of Science*, 61(1), 51–80. <https://doi.org/10.1093/bjps/axp020>
- Glynn, Luke (2011). A Probabilistic Analysis of Causation. *British Journal for the Philosophy of Science*, 62(2), 343–392. <https://doi.org/10.1093/bjps/axq015>
- Good, Irving (1961a). A Causal Calculus (I). *British Journal for the Philosophy of Science*, 11(44), 305–318. <https://doi.org/10.1093/bjps/XI.44.305>
- Good, Irving (1961b). A Causal Calculus (II). *British Journal for the Philosophy of Science*, 12(45), 43–51. <https://doi.org/10.1093/bjps/XII.45.43>
- Hájek, Alan (2017). Most Counterfactuals are False. Manuscript in preparation.
- Halpern, Joseph (2016). *Actual Causality*. MIT Press.
- Halpern, Joseph and Christopher Hitchcock (2010). Actual Causation and the Art of Modeling. In Rina Dechter, Hector Geffner, and Joseph Halpern (Eds.), *Heuristics, Probability and Causality: A Tribute to Judea Pearl* (383–406). College Publications.

- Halpern, Joseph and Christopher Hitchcock (2015). Graded Causation and Defaults. *British Journal for the Philosophy of Science*, 66(2), 413–457. <https://doi.org/10.1093/bjps/axt050>
- Halpern, Joseph and Judea Pearl (2005). Causes and Explanations: A Structural-Model Approach, Part I: Causes. *British Journal for the Philosophy of Science* 56(4), 843–887. <https://doi.org/10.1093/bjps/axi147>
- Hartmann, Stephan and Patrick Suppes (2010). Entanglement, Upper Probabilities and Decoherence in Quantum Mechanics. In M. Suárez, M. Dorato, and M. Rédei (Eds.), *EPSA Philosophical Issues in the Sciences: Launch of the European Philosophy of Science Association* (93–103). Springer.
- Hastings, Alan (1993). Complex Interactions between Dispersal and Dynamics: Lessons from Coupled Logistic Equations. *Ecology*, 74(5), 1362–1372. <https://doi.org/10.2307/1940066>
- Hitchcock, Christopher (2001). The Intransitivity of Causation Revealed in Equations and Graphs. *Journal of Philosophy*, 98(6), 273–299. <https://doi.org/10.2307/2678432>
- Hitchcock, Christopher (2007b). Prevention, Preemption, and the Principle of Sufficient Reason. *Philosophical Review*, 116(4), 495–532. <https://doi.org/10.1215/00318108-2007-012>
- Hitchcock, Christopher and James Woodward (2003a). Explanatory Generalizations, Part I: A Counterfactual Approach. *Noûs*, 37(1), 1–24. <https://doi.org/10.1111/1468-0068.00426>
- Hitchcock, Christopher and James Woodward (2003b). Explanatory Generalizations, Part II: Plumbing Explanatory Depth. *Noûs*, 37(2), 181–199. <https://doi.org/10.1111/1468-0068.00435>
- Hoefer, Carl (2004). Causality and Determinism: Tension, or Outright Conflict. *Revista de Filosofia*, 29(2), 99–115.
- Hoefer, Carl (2007). The Third Way on Objective Probability: A Sceptic’s Guide to Objective Chance. *Mind*, 116(463), 549–596. <https://doi.org/10.1093/mind/fzm549>
- Hüttemann, Andreas and Alexander Reutlinger (2013). Against the Statistical Account of Special Science Laws. In Vassilios Karakostas and Dennis Dieks (Eds.), *EPSA11: Perspectives and Foundational Problems* (181–192). Springer. [https://doi.org/10.1007/978-3-319-01306-0\\_15](https://doi.org/10.1007/978-3-319-01306-0_15)
- Ismael, Jenann (2009). Probability in Deterministic Physics. *Journal of Philosophy*, 106(2), 89–108. <https://doi.org/10.5840/jphil2009106214>
- Ismael, Jenann (2011). A Modest Proposal about Chance. *Journal of Philosophy*, 108(8), 416–442. <https://doi.org/10.5840/jphil2011108822>
- Kim, Jaegwon (2002). The Layered Model: Metaphysical Considerations. *Philosophical Explorations*, 5(1), 2–20. <https://doi.org/10.1080/10002002018538719>
- Kowalenko, Robert (2014). Ceteris Paribus Laws: A Naturalistic Account. *International Studies in the Philosophy of Science*, 28(2), 133–155. <https://doi.org/10.1080/02698595.2014.932527>
- Kvart, Igal (2004). Causation: Probabilistic and Counterfactual Analyses. In John Collins, Ned Hall, and Laurie Paul (Eds.), *Causation and Counterfactuals* (359–386). MIT Press.
- Lange, Marc (2002). Who’s Afraid of Ceteris-Paribus Laws? Or: How I Learned to Stop Worrying and Love Them. *Erkenntnis*, 57(3), 407–423. <https://doi.org/10.1023/A:1021546731582>
- Law, Richard, David Murrell, and Ulf Dieckmann (2003). Population Growth in

- Space and Time: Spatial Logistic Equations. *Ecology*, 84(1), 252–262. [https://doi.org/10.1890/0012-9658\(2003\)084\[0252:PGISAT\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2003)084[0252:PGISAT]2.0.CO;2)
- LePore, Ernest and Barry Loewer (1987). Mind Matters. *Journal of Philosophy*, 84(11), 630–642. <https://doi.org/10.5840/jphil198784119>
- Lewis, David (1973a). *Counterfactuals*. Harvard University Press.
- Lewis, David (1973b). Causation. *Journal of Philosophy*, 70(17), 556–567. <https://doi.org/10.2307/2025310>
- Lewis, David (1979). Counterfactual Dependence and Time’s Arrow. *Noûs*, 13(4), 455–476. <https://doi.org/10.2307/2215339>
- Lewis, David (1983). New Work for a Theory of Universals. *Australasian Journal of Philosophy*, 61(4), 343–377. <https://doi.org/10.1080/00048408312341131>
- Lewis, David (1986). Postscripts to “Causation”. In *Philosophical Papers: Volume II* (172–213). Oxford University Press.
- Lewis, David (1994). Humean Supervenience Debugged. *Mind*, 103(412), 473–490. <https://doi.org/10.1093/mind/103.412.473>
- Linquist, Stefan, Ryan Gregory, Tyler Elliot, Brent Saylor, Stefan Kremer, and Karl Cottenie (2016). Yes! There are Resilient Generalizations (Or “Laws”) in Ecology. *Quarterly Review of Biology*, 91(2), 119–131. <https://doi.org/10.1086/686809>
- List, Christian and Marcus Pivato (2015). Emergent Chance. *Philosophical Review*, 124(1), 119–152. <https://doi.org/10.1215/00318108-2812670>
- Loewer, Barry (2001). Determinism and Chance. *Studies in History and Philosophy of Science*, 32(4), 609–620. [https://doi.org/10.1016/S1355-2198\(01\)00028-4](https://doi.org/10.1016/S1355-2198(01)00028-4)
- Loewer, Barry (2007). Counterfactuals and the Second Law. In H. Price and R. Corry (Eds.), *Causation, Physics, and the Constitution of Reality: Russell’s Republic Revisited* (293–326). Clarendon Press.
- Loewer, Barry (2012). Two Accounts of Laws and Time. *Philosophical Studies*, 160(1), 115–137. <https://doi.org/10.1007/s11098-012-9911-x>
- Mackie, John (1965). Causes and Conditions. *American Philosophical Quarterly*, 2(4), 245–264.
- May, Robert (1974). Biological Populations with Nonoverlapping Generations: Stable Points, Stable Cycles, and Chaos. *Science*, 186(4164), 645–647. <https://doi.org/10.1126/science.186.4164.645>
- Menzies, Peter (1989). Probabilistic Causation and Causal Processes: A Critique of Lewis. *Philosophy of Science*, 56(4), 642–663. <https://doi.org/10.1086/289518>
- Oppenheim, Paul and Hilary Putnam (1958). Unity of Science as a Working Hypothesis. In Herbert Feigl, Michael Scriven, and Grover Maxwell (Eds.), *Concepts, Theories, and the Mind-Body Problem, Minnesota Studies in the Philosophy of Science* (Vol. 2, 3–36). University of Minnesota Press.
- Otto, Sarah and Troy Day (2007). *A Biologist’s Guide to Mathematical Modeling in Ecology and Evolution*. Princeton University Press.
- Pearl, Judea (2009). *Causality: Models, Reasoning, and Inference* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511803161>
- Pietroski, Paul and Georges Rey (1995). When Other Things Aren’t Equal: Saving Ceteris Paribus Laws from Vacuity. *British Journal for the Philosophy of Science*, 46(1), 81–110. <https://doi.org/10.1093/bjps/46.1.81>
- Price, Huw and Richard Corry (Eds.) (2007). *Causation, Physics, and the Constitution of Reality: Russell’s Republic Revisited*. Clarendon Press.

- Ramachandran, Murali (2004). A Counterfactual Analysis of Indeterministic Causation. In John Collins, Ned Hall, and Laurie Paul (Eds.), *Causation and Counterfactuals* (387–402). MIT Press.
- Reichenbach, Hans (1971). *The Direction of Time* (Maria Reichenbach, Ed.). University of California Press.
- Reutlinger, Alexander (2014). Do Statistical Laws Solve the “Problem of Provisos”? *Erkenntnis*, 79(10 Supplement), 1759–1773. <https://doi.org/10.1007/s10670-014-9640-y>
- Reutlinger, Alexander, Gerhard Schurz, and Andreas Hüttemann (2017). Ceteris Paribus Laws. In Edward Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2017 ed.). Retrieved from <https://plato.stanford.edu/archives/spr2017/entries/ceteris-paribus/>
- Roberts, John (2014). CP-Law Statements as Vague, Self-Referential, Self-Locating, Statistical, and Perfectly in Order. *Erkenntnis*, 79(10 Supplement), 1775–1786. <https://doi.org/10.1007/s10670-014-9641-x>
- Schaffer, Jonathan (2007). Deterministic Chance?. *British Journal for the Philosophy of Science*, 58(2), 113–140. <https://doi.org/10.1093/bjps/axm002>
- Schiffer, Stephen (1991). Ceteris Paribus Laws. *Mind*, 100(1), 1–17. <https://doi.org/10.1093/mind/C.397.1>
- Schrenk, Markus (2008). A Lewisian Theory for Special Science Laws. In Sven Walter and Helen Bohse (eds): *Ausgewählte Beiträge zu den Sektionen der GAP.6* (121–131). Mentis
- Schurz, Gerhard (2001). What is “Normal”? An Evolution-Theoretic Foundation for Normic Laws and Their Relation to Statistical Normality. *Philosophy of Science*, 68(4), 476–497. <https://doi.org/10.1086/392938>
- Schurz, Gerhard (2002). Ceteris Paribus Laws: Classification and Deconstruction. *Erkenntnis*, 57(4), 351–372. <https://doi.org/10.1023/A:1021582327947>
- Schurz, Gerhard (2014). Ceteris Paribus and Ceteris Rectis Laws: Content and Causal Role. *Erkenntnis*, 79(10 Supplement), 1801–1817. <https://doi.org/10.1007/s10670-014-9643-8>
- Suppes, Patrick (1970). *A Probabilistic Theory of Causality*, *Acta Philosophical Fennica*. North-Holland.
- Suppes, Patrick and Mario Zanotti (1991). Existence of Hidden Variables Having Only Upper Probabilities. *Foundations of Physics*, 21(12), 1479–1499. <https://doi.org/10.1007/BF01889653>
- Tsoularis, Anastasios and James Wallace (2002). Analysis of Logistic Growth Models. *Mathematical Biosciences*, 179(1), 21–55. [https://doi.org/10.1016/S0025-5564\(02\)00096-2](https://doi.org/10.1016/S0025-5564(02)00096-2)
- Twardy, Charles and Kevin Korb (2011). Actual Causation by Probabilistic Active Paths. *Philosophy of Science*, 78(5), 900–913. <https://doi.org/10.1086/662957>
- Vandermeer, John and Deborah Goldberg (2013). *Population Ecology* (2nd ed.). Princeton University Press. <https://doi.org/10.1515/9781400848737>
- Weiss, Howard (2009). *A Mathematical Introduction to Population Dynamics*. IMPA.
- Weslake, Brad (2014). Statistical Mechanical Imperialism. In Alistair Wilson (Ed.), *Asymmetries of Chance and Time* (241–257). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199673421.003.0012>
- Woodward, James (2003). *Making Things Happen*. Oxford University Press.
- Wright, Richard (1985). Causation in Tort Law. *California Law Review*, 73(6), 1735–1828. <https://doi.org/10.2307/3480373>