



***The Costs and Benefits of Generality: A Trade-off Based
Argument Against the Unconditional Value of Generality
in Economic Modelling***

Stefan Mendritzki

Working paper 15.02

Eindhoven Centre for Innovation Studies (ECIS),
School of Innovation Sciences,
Eindhoven University of Technology, The Netherlands

The Costs and Benefits of Generality: A Trade-off Based Argument Against the Unconditional Value of Generality in Economic Modelling

Stefan Mendritzki¹

Philosophy and Ethics
School of Innovation Sciences
Eindhoven University of Technology (TU/e)

Abstract. Generality is a defining modelling virtue of the economic tradition. It is often assumed that generality is not only a normative standard but also a costless commitment. This paper takes up a criticism of this assumption: degree of generality of a model is a choice which comes with costs under certain conditions. It is argued that if economic systems are heterogeneous, then generality trades-off with empirical fidelity. Trade-offs have the net effect of supporting virtue pluralism. The degree to which trade-offs bite, and thus the degree of virtue pluralism, depends on the degree of ontological heterogeneity of economic systems. Generality as a normative standard is undermined because it is a gamble on ontology.

¹ e: smendritzki@gmail.com

1 Introduction

Recent philosophy of science, and particularly philosophy of social science, has increasingly focused on the differences between fields. One outcome of this project has been to highlight that there are several different epistemic virtues in play (e.g. generality, simplicity, empirical fidelity) and that different fields seem to value these virtues differently.

In economics, several virtues have been identified as being important. In a comparative analysis of network models from physics, sociology, and economics, Marchionni (2013) associates economic models with virtues of simplicity, generality, and microfoundations.² There are surely other virtues which could be added to this list, such as formal rigour (cf. Blaug 2002). Among these virtues, one seems to stand out as particularly important: generality. While generality plays a part in methodological discussions in most social sciences, it is given a much more prominent role in discussions in economics. There is traditionally *a* market, *a* representative agent, and *a* price setting mechanism. One of the main complaints of critics of economics is that it fails to provide convincing empirical accounts, especially when applied to non-standard markets and goods (cf. Røgeberg & Melberg 2011 on

² Microfoundations is less clear than the other two virtues, but can be thought of as a particular sort of multilevel explanation which has bottom-out entities at the (rational) agent level of the causal hierarchy.

Rational Addiction Theory). The complaint goes that the economist only has one tool in her tool belt and assumes that it should apply generally.

The complaint against giving priority to generality – that it somehow undermines empirical aspects of explanation – cannot be evaluated within the project of describing disciplinary conventions. Instead, it requires a normative project of asking if the convention is a *justified* and *non-arbitrary* standard. Disciplinary conventions may be legitimately enforced if they are normative in the sense of justified and non-arbitrary standards but not if they are merely descriptive of current practices. The important question is whether the normative interpretation can be supported: is the convention of giving priority to generality in economics a justified and non-arbitrary standard?

To answer this question, this paper will explore different possible systematic relationships between generality and empirical fidelity. It will ask which conditions would impede the simultaneous achievement of these virtues. The most important condition will be drawn from discussions in the philosophy of biology on the idea of trade-offs between virtues. It has been argued that ontological heterogeneity in biological systems induces trade-offs. Trade-offs, I argue, have a decisive influence on the debate about normative standards because they imply *virtue pluralism* with respect to models. In a situation of virtue pluralism, giving priority to generality is only one, arbitrarily chosen solution to the trade-off. The implication is that the status of the normative standard of generality hinges on empirical beliefs about the ontological heterogeneity of economic systems. I outline the main dimensions of the heterogeneity

question and argue that heterogeneity of some sort is supported by many economists and that the key open question is the type and degree of heterogeneity in economic systems.

This paper proceeds as follows. Section 2 examines the possible relationships between generality and the virtue of empirical fidelity, with a focus on which relationships would support a normative standard of generality. In Section 3, the literature on heterogeneity-induced trade-offs is introduced and generalized to domains beyond biology. Section 4 presents the implications of trade-offs, as they relate to virtue pluralism. Section 5 argues that, unlike the case of biology, fundamentally different views exist on the degree of heterogeneity in economics. Section 6 discusses the gamble that ensues in supporting the priority of generality in economic modelling when heterogeneity is a matter of dispute. Section 7 concludes.

2 The (in)dependent value of generality

One way of going about justification of a normative standard of generality would be to focus on the positive features that flow therefrom. This would be in line with much current analysis which treats a given virtue as an independent good and then specifies methodological implications based on this virtue in isolation. In this tradition, the strongest arguments for generality are those which link it to explanatory power. General models may explain the behaviour of many different targets, implying an economy of inference. More pragmatically,

generality may be very useful for the process of developing and distributing knowledge (e.g. understanding, impact of findings, pedagogy) and in the social organization of science (e.g. in developing common disciplinary assumptions or research traditions).

This approach can undoubtedly yield advances in understanding why generality is a modelling virtue. However, because the analysis proceeds by analyzing generality in isolation, these arguments only speak to the *ceteris paribus* good of generality. The claim being supported is that, all other things being equal, it is better for a model to be more general. This is a fairly uncontroversial claim but it is also shaky ground on which to build a normative standard. The problem is seen in the conflicting advice which can arise from analyzing different *ceteris paribus* virtues in isolation. When generality is the only virtue taken into consideration, the resultant methodological advice downplays consideration of detailed differences between targets. However, one could equally start from the *ceteris paribus* virtue of empirical fidelity. All other things being equal, it is better for a model to track particular targets in a detailed way. Cashing this commitment out in a similar way leads to the advice of downplaying the importance of tracking multiple targets (generality). Conflicting advice arises from analysing virtues in isolation because of neglect of possible interactions between virtues. The only way to determine if different *ceteris paribus* virtues are in fact compatible is to consider virtues jointly.

The important question for any virtue which is claimed as a normative standard is then whether this virtue induces costs in terms of other virtues. In other words, is there a

systematic relationship between the virtues in question? Ylikoski and Kuorikoski (2008) speak of the difference between explanatory and evidential virtues. Explanatory virtues are about desirable properties of an explanation assuming that it is true (or in comparative terms, what makes one true explanation better than another). Evidential virtues are only about the empirical support for an explanation independently of whether the explanation is satisfying in terms of explanatory virtues. The approach I will take is to start by accepting the *ceteris paribus* normative good of generality as an explanatory virtue. Generality is the number of systems to which a model applies. Empirical fidelity refers to the detailed agreement between a model and a target.³ It is treated here as an evidential virtue. The question is then what happens when the virtues are treated jointly (i.e. when the *ceteris paribus* conditions are mutually relaxed). There are three possible relationships between empirical fidelity, and generality: ampliative, neutral, and attenuative.

For the promoter of generality, the ampliative view would be the most satisfying. This would imply that achieving generality has some positive effect on the pursuit of empirical fidelity. Consider the ampliative claim from the perspective of some particular target T_i . This would be the claim that: increasing the intended generality of model M by claiming that it also

³ Empirical fidelity is a covering term for different specifications of the correspondence relations between models and targets. One could substitute more specific virtues such as precision or empirical adequacy without substantially changing the structure of the argument.

applies to another target T_j would imply – simply because of the increase in generality - an increase in empirical fidelity between M and T_i .⁴ Consider an example of the claim in comparative terms: model X is empirically superior to model Y with regard to intermediate consumption simply because X , unlike Y , is also meant to apply to final consumption. More generally, the empirical fidelity of T_i is directly related to the number of T 's which the model addresses. This option is difficult to accept. There are no grounds provided for supposing that the quality of an account relative to one target increases the quality of the explanation relative to another target. Whatever the explanatory good of generality, pursuing it does not directly aid in achieving evidential virtues.

This leaves two views of the relationship between generality and evidential virtues. There is the neutral view in which the empirical quality of the account of T_i is independent of whether a model also applies to T_j . Or, there is the attenuative view, in which the empirical quality of the account of T_i is diminished by the fact that the model also needs to apply to T_j . The question is whether a model can maintain the same level of empirical fidelity while applying to many targets.

⁴ This should not be confused with the claim that there is some well explained T_j and a similarity relation between T_i and T_j that warrants the transfer of inferences from T_i to T_j .

Suppose that the criticism of a lack of incorporation of evidential virtues in economic models is taken seriously. Under the neutral account, this has simply led to a neglect of other virtues. There are then relatively clear potential remedies: taking steps to incorporate evidential virtues in existing general models or building new models that simultaneously incorporate both sets of virtues. Under the attenuative account, the situation is much less clear. The choice to give priority to generality is then an indirect choice to accept a cost in terms of empirical fidelity. To continue to support generality to the exclusion of other virtues, the paramountcy of generality needs to be asserted. A paramountcy account would claim that despite interference with other virtues, a given virtue ought to be given priority. In the current case, a paramountcy claim would need to argue that even if the pursuit of generality impedes the pursuit of empirical fidelity, it is better for a model to be more general. This is a strong claim which would need to be supported with strong arguments.

The important question then revolves around the neutral and the attenuative views: Are generality and empirical fidelity independent virtues or does the degree of one set a limit on the achievement of the other? Is it necessarily the case that requiring generality in economics will exact a price in terms of empirical fidelity? In order to answer these questions, an account is required of the conditions under which attenuation occurs. This is the topic of the recent debate in philosophy of biology about modelling trade-offs.

3 Trade-offs and Heterogeneity

To answer the question of the independence of generality and empirical fidelity, an account of relationships between virtues is needed. To be useful, the account must provide hinge conditions which distinguish between cases where attenuation is expected to hold and where it is not. I will propose that the literature around trade-offs in the philosophy of biology is fit for this purpose. The focus will be on accounts which propose that trade-offs occur when the targets of modelling are of a certain kind, namely heterogeneous systems.

The philosophical discussion of trade-offs derives from an attempt to substantiate a claim from the model practitioner literature. This began with Levins's (1966) seminal paper, in which he introduces trade-offs to motivate his abstract modeling approach against more descriptive approaches. He argues that a trade-off existed between the virtues of generality, precision, and realism. Focusing on any two virtues would preclude the third. Though the philosophical literature on trade-offs has moved away from arguing about Levins's particular formulation (e.g. the particular virtues and a three-way trade-off are no longer widely discussed), the essential motivation and logic remains the same. Modelling trade-offs refer to the systemic relationship of attenuation between virtues; getting more of virtue X implies sacrificing performance on virtue Y. If trade-offs can be asserted then the question can be answered in favour of attenuation; conversely, an absence of trade-offs implies neutrality. The issue, then, is the conditions under which trade-offs occur.

There has been significant interest recently on the topic of trade-offs (e.g. Orzack & Sober 1993; Orzack 2005; Weisberg 2006; Matthewson & Weisberg 2009), but much of the literature argues for or against an unconditional application of trade-offs.⁵ An exception to this is Matthewson (2011), who presents the idea that trade-offs are induced by the nature of the target being modelled. The key idea is that the trade-offs may be caused by the nature of the systems being represented and may thus be variable across different disciplines. This approach is motivated by the observation of “the difference in how hard the trade-offs bite in different disciplines” (Matthewson 2011, 330). He argues that trade-offs are more prevalent in biology relative to other physical sciences and that this is due to the heterogeneity of relevant entities and causal forces. In support of the different nature of the targets, he adduces well known problems with the attempt to analyze biology in the same way as other physical sciences (e.g. the lack of universal generalizations, historical dependence, sensitivity to initial conditions). In this view, the severity of trade-offs is variable, with more homogenous systems being much less susceptible to trade-offs than more heterogeneous ones.

⁵ An alternative view, argued for in Odenbaugh (2003), is that trade-offs are a pragmatic matter. The interest here is in principled trade-offs, though this is fully compatible with the additional constraint of pragmatic trade-offs.

Though Matthewson's focus is on biology, it seems that the target-induced account has desirable properties for expansion to apply to other domains, such as the social sciences. First, it is inherently a theory about differences between the systems studied by different disciplines. This should make it open to extension to other disciplines. Second, the relevant differences between biological systems and the other physical sciences are not biology-specific. The argumentative structure turns on general difficulties with modelling heterogeneous systems.

The intuitive argument for heterogeneity-induced trade-offs is quite appealing. When modelling electrons, the numerical identity of different electrons lends itself to increasingly precise models of the entire class. When modelling ecosystems, there cannot be the same confidence that precisely capturing the dynamics of a desert ecosystem can be generalised to a tidal ecosystem. This can be read as an argument that the 'classical physics' relationship to modelling is the exception, not the rule. It was made possible due to the fortuitously homogenous nature of particle-based systems.

A more formal argument could also be constructed by combining the appeal to heterogeneity with Matthewson's previous work with Michael Weisberg. In a model-focused approach, trade-offs are caused by the very nature of model-based representation (Matthewson &

Weisberg 2009).⁶ Here the discussion focuses on two particular virtues: generality (the number of targets represented) and precision (specifically, parameter precision). The question Matthewson and Weisberg pose is: what effect does increasing parameter precision have on generality? The effect is considered in terms of two senses of generality: the set of all logically possible targets and the subset of actual targets. The case of all logically possible targets is considered first to anchor the argument. Increasing parameter precision - by reducing parameter uncertainty - is argued to directly reduce the number of systems to which a model may possibly apply. The more precise parameter has the effect of picking out some systems in greater detail but also of excluding other targets which the new parameter uncertainty range excludes. The direct attenuative relationship between precision and possible generality (p-generality) is referred to as a strict trade-off. With the argument anchored in p-generality, the more complicated case of actual generality (a-generality) trade-offs can be considered. Here the argument from logical possibility is modified by the fact that the subset of actual systems may be such that increasing parameter precision does not reduce the set of actual systems to which a model applies. This is the case just where the additional detail in question applies to all actual targets. Strict-tradeoffs do not apply in this case. The

⁶ The following account somewhat simplifies and reconstructs the account of Matthewson and Weisberg. For example, the distinction between individual models and model families is neglected.

most that can be said is that increasing performance with regard to one virtue either decreases performance with regard to another virtue or leaves it the same (i.e. both cannot be simultaneously increased); this is called an increase trade-off. This is similar to the claim that there may be either attenuative dependence or independence (but not ampliative dependence) between virtues. In summary: models face increase trade-offs in the best case and strict trade-offs in the worst case.

I will not follow the arguments of this account in detail. Instead I propose reconstructing the argument starting from considerations of target heterogeneity. The interest here is in a-generality (the standard notion of generality) and whether particular models exhibit strict trade-offs (henceforth, simply trade-offs) in the context of some particular set of targets.⁷

Consider a model which applies to a set of targets. Then consider the possible status of the model with respect to the same targets after adding some additional property. The status will depend on whether the model continues to apply to all targets. This will be achieved exactly if the property chosen is a common property of all the systems in the set. More formally: consider some antecedent model M and a modified model M_{+P} which is just M plus some new property P ; if P is a common property among the set of targets T then there is not a trade-off in

⁷ Increase trade-offs are not of interest here because they do not differentiate between independence and attenuative dependence.

going from M to M_{+P} . The existence of trade-offs is then conditional on the homogeneity of the set of target systems.⁸

Analysis in terms of common properties can be extended to consider the full set of properties of a given model. The degree to which trade-offs bite is determined by the properties of the relevant targets. Consider the two extreme cases of perfect homogeneity and perfect heterogeneity. In the former, all properties are common while in the latter there are no common properties. In the perfectly homogeneous case there are no trade-offs at all. A more and more detailed representation can be constructed without losing empirical fidelity with regards to any of the targets. In the perfect heterogeneity case, trade-offs are at their worst because every property added to the model will have the effect of reducing empirical fidelity with regard to some targets.

To be useful in practice, the common-properties approach also needs to cover intermediate cases between the ideals of perfect heterogeneity and perfect homogeneity. To analyze intermediate cases, a connection needs to be made between the degree of heterogeneity and

⁸ Trade-offs are a limiting relationship between virtues. Individual models will often do worse than the theoretical maximum enforced by trade-offs; the theoretical maximum is to actual model performance as thermodynamic efficiency is to actual engines.

strength of trade-offs. This can be formalized in a numerical framework, the intent of which is not to provide a directly practicable method but to briefly and precisely show the extensibility of the common properties-approach.

Suppose we want to calculate the degree of heterogeneity (H) of a set of targets. In the perfect homogeneity case $H=0$, for perfect heterogeneity case $H = 1$.⁹ The simplest approach to analyze intermediate cases is as a binary normalized sum over a set of properties with the truth condition being whether a property is common to all targets:

$$(1) H = \sum_i \begin{cases} 1/i, & P_i \text{ is common} \\ 0, & \text{else} \end{cases}$$

Then $H=0.5$ means that half the properties of the targets in question are common. A more sophisticated approach is to consider the commonness of a property itself to be a normalized variable at the level of individual properties. Take C_i to be the proportion of targets (T_j) to which a property P_i applies. Then $C_i=0.5$ means that a property applies to half of a set of targets. This gives:

⁹ There are obvious problems of in directly measuring such a ratio in terms of the infinite number of possible systems and properties involved. The numerical examples are only intended to show the conceptual coherence of the notion of degree of heterogeneity.

$$(2) H = \sum_i C_i = \sum_i \sum_j \begin{cases} 1/(i * j), & P_i \text{ applies to } T_j \\ 0, & \text{else} \end{cases}$$

Further extension could even address the problem of the binary nature of ‘applies to’. One could consider the degree to which a particular property fits a particular target (F_{ij} = normalized degree of fit between P_i and T_j).¹⁰ Then $F_{ij}=0.5$ means that P_i has an intermediate degree of fit with respect to T_j :

$$(3) H = \sum_i \sum_j F_{ij} / (i * j)$$

While the exact specification of the metrics can certainly be debated, this seems a clear proof of concept. These formulations capture, at increasing levels of specificity, the degree of heterogeneity of a set of target systems. This would allow the trade-off argument to be gradated, with more heterogeneity leading to stronger trade-offs. Though this is not pursued in detail here, the important result is that the common-properties approach does not break down when going beyond the ideals of perfect heterogeneity and homogeneity. Analyzing heterogeneity as a matter of degree allows for a nuanced answer to the question of the

¹⁰ F is notably more vague than the other variables involved and needs further consideration to be even conceptually operational.

attenuative vs. neutral relationship between virtues. Degree of heterogeneity determines the degree to which the relationship between generality and empirical fidelity is attenuative.

The common-property approach to target-induced trade-offs provides a general account of how heterogeneity leads to trade-offs. It provides more precise reasons for the heterogeneity-induced trade-offs suggested by Matthewson. The specific arguments from the biology-physical sciences case are just one instance of a general argument about domains of science. The important factor in determining the extent of trade-offs is the degree of heterogeneity, which could easily be applied to any domain of scientific modelling.

4 From Trade-offs to Virtue Pluralism

The existence of trade-offs suggests that modellers of significantly heterogeneous systems need to choose how much empirical fidelity and generality to pursue. This constrains the normative stance which should be taken when trade-offs hold. However, the fuller implications have not been thoroughly explored in the trade-off literature, which focuses more on the existence and causes of trade-offs. The view defended here is that the main implication of trade-offs is that a certain sort of pluralism emerges, namely virtue pluralism. Using only an indifference principle between virtues, several of the counter-arguments to virtue pluralism can be diffused.

To understand the implications of trade-offs it is helpful to start with their intended argumentative function. The most common use of trade-off arguments has been as a localized attempt to defend a minority position against an entrenched virtue. In the case of biology, this was to licence more abstract ecological modelling against a prevalent virtue of empirical fidelity. While the legitimization of alternative modelling virtues is a valid implication of trade-offs, it leaves an unclear normative situation. Given the structure of the trade-offs, the same argument which licenses the minority position against the entrenched majority equally licenses the majority against the minority. Trade-offs cut both ways. In order to go beyond local legitimization, a general account of relationships between different virtues is needed.

The direct implication of the trade-offs is that different models do well against different virtues. A more specific reading is that different models do well against evidential vs. explanatory virtues. Since there is no clear way for a model to serve all virtues simultaneously, it is natural to turn to a pluralist interpretation. This is termed virtue pluralism to mark it out from the many other pluralist approaches (cf. Kellert et al. 2006). This is the view that multiple, incompatible virtues may each be legitimate targets of modelling but that no virtue or combination thereof constitutes a justified and non-arbitrary normative standard. All that is required to move from trade-offs to virtue pluralism is a weak negative principle, which I call the indifference principle. This is the principle that, in the absence of justified

and non-arbitrary arguments for a general normative difference between virtues, these virtues ought to be treated as having equal normative standing. In a situation of ideal trade-offs, different models are required to satisfy evidential and explanatory virtues. If we additionally assume that the indifference principle is in force, then it is not possible to rank virtues and thus not possible to rank models that take different virtues as their objective.¹¹ Virtue pluralism is in force.

To demonstrate the argumentative power of the indifference principle and thus virtue pluralism, it is useful to consider views which attempt to argue for one of the constituent virtues over the other. There are many attempts to compare virtues relative to some *ceteris paribus* good. Here some examples are provided from economics. In favour of empirical fidelity, consider the arguments of Shapiro (2005). He compares approaches based on generality and empirical fidelity in terms of a ‘what knowledge is gained if it fails’ criterion. He argues that a focus on generality is prone to catastrophic failure. In the case that such a model fails, one is left with a conceptual and mathematical framework which has no target, rendering it practically useless. On the other hand, a focus on empirical fidelity is much more robust to the heterogeneity threat. If there is a failure to create a systematic explanation, then

¹¹ This leaves open the possibility that different models may well be ranked relative to a given virtue.

one is still left with a series of specific explanations which do have targets. However, the opposite conclusion is reached if one accepts Mäki's arguments about the explanatory ambition of economics (e.g. Mäki 2009a). General explanations help economists seek out deep causes of economic events instead of merely describing surface causes. This view of the benefits of successful general analysis is the mirror image of the 'what-if-it-fails' line of reasoning.

These examples clearly show the inability of most arguments about virtues to diffuse virtue pluralism. They proceed by only addressing the benefits of a given virtue and/or the drawbacks of competing virtues. However, they do not consider the virtues jointly. If both Shapiro and Mäki are correct relative to their chosen virtue, then both only provide a part of the overall judgement. In effect, each rightly points out one side of the trade-off without considering the other side; this is like doing only half of a cost-benefit analysis. Given that the trade-off argument is the conjunction of the claims that evidential *and* explanatory virtues are important, simply affirming one premise has no effect on the argument. To overcome the indifference principle one would need to argue that one virtue or set of virtues is systemically more valuable than another.

The requirements for a viable counterargument against the indifference principle have been clarified: one must provide justified and non-arbitrary reasons for the differential valuation of

either particular virtues or sets of virtues. There are at least two counterarguments which meet this standard. The first of these would attempt to resolve the trade-off by focussing on the purposes of different models. A key claim of Uskali Mäki's pragmatic account of models (e.g. Mäki 2005) is that purposes resolve many apparent problems in modelling. Different virtues might be amenable to certain purposes and thus there is no need to be a pluralist about virtues if one is a pluralist about purposes. The second objection would be that the implications of trade-offs are not so pluralist after all but rather point towards the superiority of mid-level modelling. Mid-level modelling aims to balance virtues and might be argued to be the best of both worlds. This argument accepts existence of trade-offs but substitutes a moderation principle for the indifference principle. These will be addressed in turn.

Pluralism of purposes is often evident in modelling practice. A certain purpose may act as a way to define appropriate virtues for a given modelling exercises. So, to use an economic example, if a model is meant to give policy advice at the European level, then this purpose has the effect of setting the required level of generality. The argument is then that, according to the given purpose, a certain level of generality is a more appropriate solution than those of a higher or lower level of generality. This seems to argue against the indifference principle.

In fact, this argument misses the point by conflating the pragmatic reasons for which purposes are imposed with justified and non-arbitrary reasons. Suppose that one accepts the prevalence

and pragmatic utility of purposes in modelling practice. The argument from purposes is that *given* a certain purpose, different valuations can be assigned to sets of virtues. This certainly offers reasons against indifference in the context of a particular purpose; what is less clear is if these reasons are justified and non-arbitrary epistemic reasons. Purposes, in the standard usage, are not the sort of thing which can be independently epistemically justified, but are rather specified on a pragmatic basis. To revisit the European-level example, one may well have pragmatic reasons to want such a model. However, there is no justified epistemic reason to privilege this choice of purpose. Then purposes do nothing to solve the problem of trade-offs in general; at best, they side-step it. This side-step can be a pragmatically justified focussing tool but it carries the danger of obscuring the fact that a trade-off has been made at all. The situation can be clarified as follows: Trade-offs constrain the set of possible solutions about which one should be pluralist. Purposes may be imposed which make a pragmatic second selection from this set based on fitness to pragmatic purpose. None of this argues against talking in terms of pluralism of purposes or implies that pluralism about purposes and virtues is incompatible. It does, however, highlight that those imposing a certain purpose ought to be aware that their choice of purpose indirectly imposes a pragmatic selection on trade-off constrained alternatives. Conflating pragmatic and trade-off induced selections is what lends the appeal to purpose its intuitive appeal.

The second objection to the indifference principle is that one ought to prefer a moderation

principle. This asserts that a balance of virtues is superior to highly specialized approaches (in the direction of either virtue). One way of spelling this intuition out is as an argument towards mid-level modelling. An argument for the superiority of mid-level modelling is that there is a ‘sweet spot’ in modelling which achieves intermediate values of both evidential and explanatory virtues rather than specializing in one or the other. This can be justified by the sort of logic behind judicial balancing principles: in the absence of the ability to simultaneously achieve right A and right B, they ought to be balanced against each other.

It is certainly the case that trade-offs legitimize a mid-level strategy. Balancing of virtues is one way to solve the trade-off. Furthermore, there are likely many different mid-level strategies, if mid-level is taken to be the broad class of non-specialist strategies. It is also the case, as discussed in the previous objection, that a certain balance of virtues may be optimal relative to some framing purpose. This is a natural place to perform mid-level modelling. However, legitimization aside, this is not an argument to the privileged status of the mid-level.¹² The most that can be said then is that, when considering the many possible weightings

¹² A case for privileged status must rely on some way of assigning differential normative evaluations to balanced and specialist methods. For example, one would need to argue for the existence of decreasing returns to pursuing a particular virtue. This is what drives the intuition that there should be a ‘sweet spot’.

of virtues, the extremes of exclusively preferring one virtue are a small portion of the set of possible solutions. Most solutions will be ‘mid-level’ in the sense of not being at either extreme but not ‘mid-level’ in the sense of assigning equal weighting. Unlike the analogy from the judicial balancing principle, there is room for multiple approaches which do not directly interfere with each other. In the absence of reasons for supporting an equal weighting of virtues, the indifference principle remains in force.

The relatively weak indifference principle has been shown to be sufficient to deal with counter-arguments to the proposition that trade-offs imply virtue pluralism. It implies that where evidence for trade-offs is clear, there should be an expectation of different legitimate weightings of virtues. In the context of the larger argument, it is the final step in an argument which connects heterogeneity to trade-offs to virtue pluralism. This line of argument implies that much is at stake in making and assessing claims about the heterogeneity of a given domain. In the case of economics, it highlights the importance of beliefs about the heterogeneity of economic systems.

5 Are economic systems heterogeneous?

The tactic of this paper has been to argue that introducing considerations of ontological heterogeneity affects the ability of generality and empirical fidelity to sit well together. To

recap the argument thus far, virtue pluralism is the implication of trade-offs which occur under conditions of heterogeneity. This section returns to the specific implications of these general arguments for the question of whether generality is a justified and non-arbitrary standard in economics. Even if one accepts heterogeneity as the hinge condition between neutrality and attenuation in general terms, it remains to be argued into which side of the hinge economic systems fall. This sets the focus on the nature of economic systems and the degree of heterogeneity that they exhibit.

The central problem with arguing about ontological heterogeneity is the difficulty of assessing a priori how serious the problem is. In order to have explicit epistemic access to such information, one would already need clear knowledge of the system in question, rendering the methodological question moot. This leaves two avenues available for continuing this debate: via general principles and via analysis of scientific practice. A general approach would specify general conditions under which assertion of significant heterogeneity is plausible and argue whether economic systems fulfil these conditions. A practice-based approach would leverage the views of researchers closest to the systems in question to come to conclusions as to whether heterogeneity is a significant feature of economic systems.

In his treatment of biological systems, Matthewson finds the practice-based option more convincing. The specific strategy employed is to engage biologists' background commitments

by connecting heterogeneity to the principle of variation in evolutionary theory. To summarize his argument: if you accept Darwinism, then you are already committed to heterogeneity (via the principle of variation). This approach has the advantage of connecting heterogeneity to the biological vernacular and avoiding the difficulties of direct philosophical arguments. The power of the argument in the case of biology is the near unanimous commitment to something like the principle of variation. The force of the argument would be diminished in the case of conflicting background commitments. A further complication which is not considered by Matthewson is the need to distinguish between different degrees of heterogeneity. These can range widely from slight divergences from natural types to the total lack of commonality between any two token system behaviours. Since degree of heterogeneity determines the degree to which trade-offs bite, distinguishing between degrees of heterogeneity is crucial.

The approach in this section will be to follow the practice-based strategy with a focus on identifying differing views on heterogeneity. To assist in identifying diversity of opinion, I propose a set of different positions on heterogeneity and attempt to populate these with proponents (both economists and philosophers of economics).¹³ These positions are (in

¹³ Of course, not all views will be neatly accommodated (e.g. those instrumentalists who profess to make no meaningful assumptions about the nature of the system under study). The

ascending order of commitment to heterogeneity): lack of heterogeneity, surface heterogeneity (acceptance of common underlying mechanisms), variable heterogeneity (acceptance of local generalizations), and pervasive heterogeneity. If, as in the case of biology, a near unanimous commitment to one of the positions arises, then the question can be resolved in a similar manner.

One extreme view is to discount heterogeneity entirely (View 1 hereafter). This would imply viewing economic systems as akin to those of classical physics. The limited number of adherents to this view would be historical: namely, those early pioneers of the attempt to boost the scientific credentials of economics by emulation of physics. This view might be ascribed to some historical figures reviewed by Mirowski (1989). However, this view can be seen as not even covering the early modern tradition despite its characteristically heavy reliance upon homogeneity assumptions. For example, Hodgson (2013) has argued that Marshall was already very engaged by the heterogeneity of actual economic systems, despite the strong homogeneity assumptions in his models. If any major figure ever accepted this view, it is no longer seriously maintained. The main lines of argument have moved on from the existence of heterogeneity to the nature and import of heterogeneity.

goal is not to be exhaustive, merely to provide some insight into those who do take a position.

A more relaxed version of the homogeneity thesis accepts some degree of heterogeneity while maintaining the existence of deep homogeneity (call this View 2). Such an argument would claim that there is some variety in the details but an underlying homogeneity in the most causally important elements. One way to parse this out in terms of entities is as diversity centred on natural kinds. From a causal perspective, one might say that there are common mechanisms, instances of which differ in their details. In such a ‘centred diversity’ view, one might argue that the heterogeneity impacts on exactness but that does not affect empirical fidelity in the broad strokes. Such a position has many more supporters, both among philosophers of economics and economists. An obvious philosophical proponent of this view is Uskali Mäki (2009b) with his mechanistic realism. Similar commitments seem to be taken seriously by modern economic thinkers. Though economists are not as explicit in their commitments as philosophers, they present arguments that are at least consistent with this view.¹⁴ Take, for example, Akerlof’s (1970) famous model of information asymmetry which was used to describe of dynamics of the market for used vehicles. The focus was on describing the core mechanism of information asymmetry. A common interpretation of the relationship to economic targets is that the mechanism developed explained the main

¹⁴ A fuller parsing of what the ‘real’ commitments of particular economists is a task for history of economic thought. What is important for present purposes is that there is at least a substantial minority of economists who would subscribe to the view in question.

dynamics of many markets. Though more precise accounts would require certain details to be added, these would not change the core mechanism. If this view were in fact true of economic systems, then trade-offs would only have a modest effect of precluding highly general models from making detailed empirical claims. One could then argue that a generality-focussed approach can do most of the work with some filling in of details for application to specific cases.

The next strongest version of heterogeneity views makes a substantial concession to the existence of heterogeneity (View 3 hereafter). This view holds that while there are no universal regularities across economic systems, there may be some more limited domains in which relative homogeneity can be found. A kind of second order heterogeneity holds: the degree of heterogeneity is variable across and/or within given domains. That is, while there are some elements which exhibit deep heterogeneity, others exhibit more shallow heterogeneity. Another way of parsing this idea out is in terms of local regularities. This is the claim that within a particular frame (spatial, temporal, institutional), there is a relatively low level of heterogeneity. To see the contrast with the common mechanisms view (View 2), consider the differences between so-called New and Old Institutional economists. In contrast to New Institutionalists who hold a centred diversity view (that institutional structures are explained by a central mechanism of transaction costs), Old Institutionalists would fit more comfortably into this variability view. They would argue that it is exactly where institutions have formed

that pockets of homogeneity are to be observed. This view is also likely to find some philosophical support. One could read Nancy Cartwright's Dappled World view (1999) as being consistent with the view in question. In this interpretation, using the term 'dappled' implies both first and second order heterogeneity. The implications of accepting this view are already significant in terms of trade-offs. In at least some domains, one can expect trade-offs to bite rather hard. In any case, it is always open to debate to what degree a particular domain is subject to trade-offs. There may be some better-than-zero sum combinations of virtues but trade-offs are usually being made to some extent. Even where local regularities exist, there is still some concession of generality at work; it is exactly in accepting reduced generality that the local regularity is isolated.¹⁵

A final view is the most extreme with regards to heterogeneity (View 4). This is the view that heterogeneity is the rule. This view does not take the co-existence of generality and empirical fidelity very seriously. There are not even pockets of regularity beyond the level of individual contexts. This view has not been expressed very often in the context of economics. The closest available philosophical support would be an adaptation of the methodological

¹⁵ Where local generalizations are present, empirical fidelity does not need to be sacrificed to obtain additional generality. Approaches which are less general than the level of the regularity miss out because more general approaches do not give anything up for added generality.

localism of Little (2009). In a less explicit sense, some economic anthropologists seem to adopt something like this view when they approach each economic context without the goal of generating any general theory of economic behaviour. This view on heterogeneity tends to be expressed through single virtue approaches. It seems that this view is too strong for most economists and philosophers of economics.

Though there is surely additional descriptive work to be done, some preliminary conclusions can already be drawn. In terms of implications of trade-offs, View 4 would imply taking trade-offs at face value; View 3, that trade-offs are a serious problem; View 2, that trade-offs only affect making detailed claims; and View 1, that trade-offs are irrelevant. The major break point for the implications of trade-offs is between views 2 and 3, which have been argued to have at least a substantial minority of support. In light of these initial results, there seems little prospect of the unanimity seen among biologists.

6 The Gamble of Priority

The different positions on the nature of heterogeneity in economics allow evaluation of the original question of this paper: the justifiability of the normative standard of generality in economics. Recall that this judgement was argued to be conditional on the relationship between generality and empirical fidelity: the standard is arbitrary if generality stands in an

attenuative relationship with other virtues but might be justified in the case of a neutral relationship. Judgement of attenuation is in turn conditional on beliefs about ontological heterogeneity, which is a matter of dispute in economic practice. Given that there is no clear consensus on the hinge condition, the question is what general conclusions can be drawn about the normative standard of generality in economics.

One subset of views on the heterogeneity of economic systems, namely Views 1 and 2, largely supports the neutrality of generality with respect to other virtues. These views are consistent with giving priority to generality. Since there is no significant principled interaction between virtues in this case, there is no need to make a principled argument to justify giving priority to generality. Instead, it would be sufficient to claim that there will be due attention paid to all virtues in the full course of future research. One could make a pragmatic argument about fruitfulness of giving priority to generality. If one could successfully argue that a generality-first approach yielded better progress than a generality-later approach, then priority could justifiably be given to generality.

The situation is much more difficult in the case of Views 3 and 4, which accept significant heterogeneity of economic targets. The virtue pluralism that is the consequence of accepting these views implies that giving priority to generality is just one of many solutions to the trade-off. Systematically enforcing this single solution goes beyond mere neglect of evidential

virtues to an indirect choice to exclude such concerns. Recall from Section 2, that in this case one would need to argue for a paramountcy claim. This is a significant challenge indeed, all the more so given that a paramountcy clause would need to overcome the indifference principle described in Section 4.

Given that the descriptive approach of the last section was unable to decide between these two cases, the best that can be given is a conditional answer to the justifiability question. In the absence of independently justified paramountcy claims, the justifiability of the norm of giving priority to generality depends on beliefs about ontology. However, given the strong nature of the concept of justification, even this conditional answer has some force. If current norms embed assumptions about the low degree of ontological heterogeneity of economic systems, then these norms are based on a gamble regarding the nature of the world. If giving priority to generality is a gamble, is a poor basis for justified norms.

To clearly see the risks involved, one only needs to consider the consequences of assuming one view when the other in fact holds. Suppose that View 2 is assumed but View 3 is in fact the case. Here, there would be priority for generality-focused research to the exclusion of research focussing on empirical fidelity. This would be based on the mistaken premise that empirical fidelity could be added *post hoc* to generality-focused research. The opposite situation also has risks (i.e. where View 3 is assumed but View 2 in fact holds). Then time and

resources would be spent analyzing many tokens more than necessary, with the opportunity cost in investments that could have been made in the discovery of underlying types. If both views entail risks, then neither is a suitable basis for justified norms.

If one is committed to the norm of giving priority to generality and one accepts the lack of practice-based consensus, the only way to avoid the gamble is to argue that something like View 2 is in fact the case. One needs to argue that despite the variety of beliefs about ontological heterogeneity, a low degree of heterogeneity is in fact the case. An obvious place to look for arguments about which systems should be expected to exhibit significant heterogeneity would be Matthewson's arguments for heterogeneity in biology. After all, it is his argument about the differences between specific fields (biology vs. the physical sciences) which was used to motivate this general framework. The first set of arguments that Matthewson gives relate to fruitful markers of heterogeneity. Candidates are: the lack of universal generalizations, historical dependence, and sensitivity to initial conditions. These conditions do not seem to offer a clearer case for View 2. Consider universal generalizations. The evidence for universal generalizations in economics is not promising. The strongest arguments for a law-like interpretation posit that economic theories are 'inexact' (Hausman 1992), which is some distance from universal generalizations. Despite the current paucity of universal generalizations, one might hope that the problem might be overcome in the future (e.g. there are universal generalizations but we just have not found them yet). These sorts of

disciplinary immaturity arguments only replace one gamble with another and one that only gets weaker with time.¹⁶ In sum, while normative considerations are not definitive they do point in the direction of the plausibility of heterogeneity assumptions. If anything, normative considerations only make the gamble on View 2 seem all the riskier.

If one takes a dim view of the gamble on giving priority to generality, it raises the question of the standing of the fruits of the research done under its auspices. What is the status of previous economic work in the generality tradition? That some models will strive for generality is defensible as a legitimate position within the trade-off. However, this does not license the argument that all models should be expected to make the same trade-off. Thus the main target of the argument is not carrying out research within the generality tradition but the imposition of the generality tradition on all research. Hindriks (2008) claims that, in economics, models should be seen as engines for developing explanations rather than cameras for capturing targets in detail. Claiming models are “an engine, not a camera” is explicitly imposing a focus on explanatory virtues over evidential virtues; it is exactly the sort of claim which the gamble of priority calls into question. Of course, the opposite claim is equally

¹⁶ In probability terms, the immaturity argument is that the ratio of explored to unexplored candidate generalizations is low. As more possibilities are considered and rejected the ratio increases, making the immaturity argument less probable.

unjustified. A principle of tolerance holds: all solutions to the trade-off warrant equal hearing.

The pluralist interpretation of trade-offs is then a partial victory for the critical view that sees generality as being too dominant an impulse in economics. It is a victory, in that it forces the issue on the priority of generality. The burden of proof is then on those who view generality to have priority in economics to provide convincing arguments for either homogeneity or paramountcy. It is also a victory in that it legitimizes the kind of empirically-oriented modelling that the critical view is fond of. However, these victories are only partial in that the principle of tolerance also applies to approaches which give priority to generality. The trade-off view follows the critical view on the problems in the way that economics has dealt with differing modelling virtues but does not follow it in its rejection of generality-focussed research. The result, as is so often the case with arguments to pluralism, may leave neither side particularly satisfied.

7 Conclusion

In traditional philosophy of science, different fields were largely treated as instances of the same type. At the least, a strong argument can be made that the historical trend was to take classical physics as a regulative ideal. The observation that, in at least some fields, generality

and empirical fidelity sit well together would then make the neutrality view plausible. In light of the arguments presented here, this default assumption has ignored the nature of the targets being studied. In assuming neutrality, one takes on an implicit commitment to the homogenous nature of the system in question. Under conditions of ontological heterogeneity, generality will have costs in terms of empirical fidelity. The conditionalized value of generality undermines the claim that the disciplinary convention of giving priority to generality is a justified and non-arbitrary standard.

The clearest implications of this analysis are for those who self-identify as developing economic models (or theories) that embrace heterogeneity, such as the sub-field of Evolutionary Economics. Here the message is that the sort of heterogeneity assumed matters a great deal. The common refrain is that homogeneity only comes in one form but heterogeneity comes in many. New approaches in economics which attempt to engage more fundamental sorts of heterogeneity (e.g. those described in View 3 or 4), should expect to encounter strong trade-offs. These trade-offs will imply a need for significant methodological reflection to negotiate the complexities of virtue pluralism.

Thinking in terms of trade-offs casts fields which study heterogeneous systems as much messier than the classical view. A plurality of models will co-exist. An open question is what one can say in a normative sense about the interaction between different approaches in the

same field. Are there still reasons to think that collaboration will be fruitful and/or necessary?

These considerations are also likely to have some impact in considering inter-field relationships.

In the wider context of the philosophy of the social sciences, the idea that economics has traditionally solved the trade-off in favour of generality helps clarify its position both in a descriptive and a normative sense. Descriptively, the *ceteris paribus* good of generality and the neutrality view provide a reconstruction of how economics became much more focussed on generality than neighbouring social sciences. It could also be an interesting tool to re-analyze debates over the imperialism of economics in relation to neighbouring social sciences. Perhaps at least part of the conflict is not only at the level of applying a specific theory (i.e. Rational Choice Theory), but also at the level of the 'correct' level of generality.

At a normative level, the trade-off view pushes back against attempts to describe any ideal level of generality. While there do not seem to be strong reasons to favour mid-level models exclusively, there is a strong case for expanding the range of legitimate solutions to trade-offs. This opens up the opportunity and the challenge of modelling the economic world at many different levels of generality.

References

- Akerlof, George. 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *The Quarterly Journal of Economics* 84 (3): 488–500.
- Blaug, Mark. 2002. "Ugly Currents in Modern Economics." In *Facts and Fiction in Economics*, ed. Uskali Mäki, 35–56. Cambridge, UK: Cambridge University Press.
- Cartwright, Nancy. 1999. *The Dappled World: A Study of the Boundaries of Science*. Cambridge, UK: Cambridge University Press.
- Hausman, Daniel. 1992. *The Inexact and Separate Science of Economics*. Cambridge, UK: Cambridge University Press.
- Hindriks, Frank. 2008. "False Models as Explanatory Engines." *Philosophy of the Social Sciences* 38 (3): 334–60.
- Hodgson, Geoffrey. 2013. "Come Back Marshall, All Is Forgiven? Complexity, Evolution, Mathematics and Marshallian Exceptionalism." *The European Journal of the History of Economic Thought* 20 (6): 957–81.
- Kellert, Stephen, Helen Longino, and Kenneth Waters, eds. 2006. *Scientific Pluralism*. Minneapolis, MN: University of Minnesota Press.
- Levins, R. 1966. "The Strategy of Model Building in Population Biology." *American Scientist* 54 (4): 421–31.
- Little, Daniel. 2009. "The Heterogeneous Social: New Thinking about the Foundations of the Social Sciences." In *Philosophy of the Social Sciences. Philosophical Theory and*

Scientific Practice, ed. Chrysostomos Mantzavinos, 154–78. Cambridge, UK: Cambridge University Press.

Mäki, Uskali. 2005. “Models Are Experiments, Experiments Are Models.” *Journal of Economic Methodology* 12 (2): 303–15.

———. 2009a. “Realistic Realism about Unrealistic Models.” In *Oxford Handbook of the Philosophy of Economics*, ed. Harold Kincaid and Don Ross, 68–98. Oxford, UK: Oxford University Press.

———. 2009b. “Unrealistic Assumptions and Unnecessary Confusions: Rereading and Rewriting F53 as a Realist Statement.” In *The Methodology of Positive Economics: Reflections on the Milton Friedman Legacy*, ed. Uskali Mäki, 90–116. Cambridge, UK: Cambridge University Press.

Marchionni, Caterina. 2013. “Playing with Networks: How Economists Explain.” *European Journal for Philosophy of Science* 3 (3): 331–52.

Matthewson, J., and M. Weisberg. 2009. “The Structure of Tradeoffs in Model Building.” *Synthese* 170 (1): 169–90.

Matthewson, John. 2011. “Trade-Offs in Model-Building: A More Target-Oriented Approach.” *Studies In History and Philosophy of Science Part A* 42 (2): 324–33.

Mirowski, P. 1989. *More Heat than Light: Economics as Social Physics, Physics as Nature's Economics*. Historical Perspectives on Modern Economics. Cambridge, UK: Cambridge University Press.

- Odenbaugh, Jay. 2003. "Complex Systems, Trade-Offs, and Theoretical Population Biology: Richard Levin's 'Strategy of Model Building in Population Biology' Revisited." *Philosophy of Science* 70 (5): 1496–1507.
- Orzack, S.H. 2005. "Discussion: What, If Anything, Is The Strategy of Model Building in Population Biology? A Comment on Levins (1966) and Odenbaugh (2003)." *Philosophy of Science* 72 (3): 479–85.
- Orzack, Steven Hecht, and Elliott Sober. 1993. "A Critical Assessment of Levins's The Strategy of Model Building in Population Biology (1966)." *The Quarterly Review of Biology* 68 (4): 533–46.
- Rogeberg, Ole, and Hans Olav Melberg. 2011. "Acceptance of Unsupported Claims about Reality: A Blind Spot in Economics." *Journal of Economic Methodology* 18 (01): 29–52.
- Shapiro, I. 2005. *The Flight from Reality in the Human Sciences*. Princeton, NJ: Princeton University Press.
- Weisberg, M. 2006. "Forty Years of 'The Strategy': Levins on Model Building and Idealization." *Biology and Philosophy* 21 (5): 623–45.
- Ylikoski, Petri, and Jaakko Kuorikoski. 2010. "Dissecting Explanatory Power." *Philosophical Studies* 148 (2): 201–19.