Undergraduate Philosophy Journal of Australasia www.upja.online
VOL. 5, FEBRUARY 2024

Robustness Analysis as a Procedure for Determining Difference-Makers

JAN ZEBROWSKI[1]

UNIVERSITY OF CAMBRIDGE

Abstract

Model-based science—the style of theoretical work dominant in many social sciences, including economics—studies complex real-world systems indirectly through highly idealized model systems. The viability of model-based science pivots on the possibility of determining 'difference-makers' for every system it studies. However, economic systems are neither clearly circumscribed nor 'closed' in the sense that any outcome studied by economists is, to a greater or lesser extent, causally influenced by an infinitely complex network of factors. This makes salient the question: How do economists determine which factors are explanatorily relevant to any given outcome and should be included in its explanation? This is the problem of explanatory relevance. In this paper, I try to make headway towards solving this problem using robustness analysis (RA)—a well-known procedure in theoretical economics by which modellers gauge the sensitivity of their models' results to assumptions that fuel the derivation of these results.

[1] Jan Zebrowski was awarded a BSc (with first class honours) in Philosophy, Logic and Scientific Method by the London School of Economics and Political Science (LSE) and is currently reading for the MPhil in History and Philosophy of Science and Medicine at the University of Cambridge. His research interests fall into three research areas: general philosophy of science, especially epistemology of science, philosophy of biology and philosophy of economics.

1. Introduction

Model-based science—the style of theoretical work dominant in many social sciences, including economics—studies complex real-world systems indirectly through highly idealized model systems.[2] Much of model-based science is about mapping the components, activities, properties, and organizational features of these model systems onto the corresponding components, activities, properties, and organizational features of their target, usually real-world, systems. If the former resemble the latter in relevant respects, these model systems are said to *represent* their target systems and may be used to explain and understand them.[3] Model-based explanations are successful, on the causal-mechanical model of explanation, if they accurately represent all factors that make a difference to their target systems and abstract away all factors that do not. The viability of model-based science, then, pivots on the possibility of determining 'difference-makers' for every system it studies.

However, economic systems are neither clearly circumscribed nor 'closed,' in the sense that any outcome studied by economists is, to a greater or lesser extent, causally influenced by an infinitely complex network of factors.[4] This makes salient the question: How do economists determine which factors are explanatorily relevant to any given outcome and should be included in its explanation? This is the problem of explanatory relevance.

In this paper, I try to make headway towards solving this problem using robustness analysis (RA)—a well-known procedure in theoretical economics by

---

[2] Godfrey-Smith, Peter (2006) "The Strategy of Model-based Science", *Biology and Philosophy*, **21**(5): 725–740; Weisberg, Michael (2006a) "Forty Years of 'The Strategy': Levins on Model Building and Idealization", *Biology and Philosophy*, **21**(5): 623–645; Weisberg, Michael (2007a) "Who Is a Modeler?", *The British Journal for the Philosophy of Science*, **58**(2): 207–233.

[3] Giere, Ronald (1988) *Explaining Science: A Cognitive Approach*. Chicago, IL: Chicago University Press; Godfrey-Smith, Peter (2009) "Models and Fictions in Science", *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, **143**(1): 101–116.

[4] Davidson, Donald (1970) "Mental Events." In: L. Foster and J. W. Swanson (eds), *Experience and Theory*. Oxford: Clarendon Press, 207–224.

which modellers gauge the sensitivity of their model's results to assumptions that fuel the derivation of these results.[5]

The paper is divided into two sections. Section 1 is expository: I explain why the problem of explanatory relevance besets the causal-mechanical model of explanation and why it persists in the face of the pragmatic dimension of explanation, such as the context in which an explanation occurs and interests of those providing and receiving it. Section 2 is polemical: I introduce the eliminative procedure for determining difference-makers and bring up two problems that, although not damning, show that it does not give us much of a handle on the problem of explanatory relevance. Then, I make a case for RA as a procedure for determining difference-makers, compare RA with the eliminative procedure and conclude that it is better geared to the style of theoretical work dominant in economics.

2. Explanation and Explanatory Relevance

2.1 The Causal-Mechanical Account of Explanation

The causal-mechanical model of explanation is based on the intuitive idea that factors cited in the explanation must fit into a causal nexus with the outcome to be explained (the explanandum-outcome).[6] It measures the success of an explanation (i) by how well it represents causal mechanisms that bring about its explanandum-outcome and (ii) by how well it discriminates the factors that make a difference to its explanandum-outcome from the rest of the factors that

[5] Woodward, James (2006) "Some Varieties of Robustness", *Journal of Economic Methodology*, **13**(2): 219–240.

[6] Lewis, David (1986) "Causal Explanation" in: *Philosophical Papers*, Vol. II. Oxford: Oxford University Press; Railton, Peter (1978) "A Deductive-Nomological Model of Probabilistic Explanation", *Philosophy of Science*, **45**(2): 206–226; Salmon, Wesley C. (1984) *Scientific Explanation and the Causal Structure of the World*. Princeton, NJ: Princeton University Press.

played a role in producing it.[7] At least some explanations that score a success in this way are model-based, in the sense that they rely on highly idealized models to accurately represent the causal mechanisms that bring about their explanandum-outcomes. Such models explain if their components, activities, properties and organizational features correspond to, or map onto, the components, activities, properties and organizational features of their target systems.[8] They provide understanding of real-world phenomena not by showing "*that* [these real-world phenomena] fit into a *nomic* nexus" but by showing "*how* [these real-world phenomena] fit into a *causal* nexus."[9] But what is it to show how real-world phenomena fit into a causal nexus?

2.1.1 Woodward's Manipulationist Framework

Woodward[10] introduces a 'manipulationist' framework for analyzing causal relationships that has the rare merit of showing causation's place in the circle of interrelated concepts that includes "cause," "counterfactual dependence," "explanation," and "explanatory relevance." The basic idea is that to show how X and Y fit into a causal nexus is to show how they fit into a pattern of counterfactual dependance. Is manipulation of X a way of manipulating Y? If changes in X produced by interventions are systematically associated with changes in Y, X and Y fit into a pattern of counterfactual dependance. If this pattern of counterfactual dependance is sufficiently invariant and continues to hold under a range of interventions in X, X and Y fit into a causal nexus. A

[7] Bechtel, William and Richardson, Robert C. (2010) *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Cambridge, MA: MIT Press; Machamer, Peter, Darden, Lindley and Craver, Carl F. (2000) "Thinking About Mechanisms", *Philosophy of Science*, **67**(1): 1–25; Strevens, Michael (2009) *Depth: An Account of Scientific Explanation*. Cambridge, MA: Harvard University Press.

[8] Kaplan, David M. and Craver, Carl F. (2011) "The Explanatory Force of Dynamical and Mathematical Models in Neuroscience: A Mechanistic Perspective*", *Philosophy of Science*, **78**(4): 601–627.

[9] Salmon, Wesley C. (1984) *Scientific Explanation and the Causal Structure of the World*. Princeton, NJ: Princeton University Press.

[10] Woodward, James (2003) *Making Things Happen*. New York, NY: Oxford University Press.

necessary and sufficient condition for a relationship between X and Y to be causal is therefore that it be invariant under a range of interventions.

What about explanation and explanatory relevance? It is a point made by Hitchcock[11] that our intuitive judgements concerning explanatory relevance correspond to our judgements concerning the truth of counterfactuals. Woodward[12] takes this point further and explicitly analyzes explanatory relevance in terms of counterfactual dependence. He argues that explanations must provide true 'what-if-things-had-been-different' information: information about how changes produced by interventions in the factors cited in the explanation are systematically associated with changes in its explanandum-outcome. For example, the explanation of the length of a shadow cast by a flagpole in terms of the elevation of the sun and the height of the flagpole shows how this explanandum-outcome fits into a causal nexus with the elevation of the sun and the height of the flagpole in that it does provide true 'what-if-things-had-been-different 'information: it correctly suggests that, other things being equal, had the elevation of the sun and/or the height of the flagpole been different, the length of a shadow cast by a flagpole would have been different. It bears emphasizing that true 'what-if-things-had-been-different' information is causal information in that it is, by definition, invariant under a range of interventions: if things had been different, it would continue to hold. So, Woodward's manipulationist framework collapses the distinction between providing explanation and providing information about causal mechanisms and identifies explanatory relevance with causal relevance.

2.2 The Problem of Explanatory Relevance

[11] Hitchcock, Christopher R. (1995) "Salmon on Explanatory Relevance", *Philosophy of Science*, **62**(2): 304–320.
[12] Woodward, James (2003) *Making Things Happen*. New York, NY: Oxford University Press.

The snag with Woodward's concept of explanatory relevance is that it is too weak. It is a familiar point that social systems in general and economic systems in particular are neither clearly circumscribed nor 'closed,' in the sense that any outcome studied by economists is, to a greater or lesser extent, causally influenced by an infinitely complex network of factors.[13] These include environmental stresses acting on individuals, gravitational forces exerted by distant stars acting on households and firms etc. And yet, when explaining economic outcomes such as, say, economic growth in the UK in the fourth quarter of 2022 (E), economists neglect environmental stresses (C_1) and gravitational forces exerted by distant stars (C_2), instead attuning their models to changes in technology (A), physical capital stock (K) and human capital stock (L). How do economists account for this neglect? Surely, changes in C_1 and C_2 produced by interventions would be systematically associated with changes in E. For example, it is a safe bet that changes in C_1 produced by interventions that increase viral mutation rates would be systematically associated with slowdowns in E. It follows that C_1 and E fit into a pattern of counterfactual dependance. If this pattern of counterfactual dependance was sufficiently invariant and continued to hold under a range of interventions in C_1, C_1 and E would fit into a causal nexus. Supposing that it was, would it be reason enough to conclude that C_1 is causally *and* explanatorily relevant to E? But then, an analogous argument could be run with C_2, C_3, …, C_n, leading to the untoward conclusion that economic models targeting E would have to be infinitely complex to accurately represent causal mechanisms underlying it. So, there are grounds for diffidence about identifying explanatory relevance with causal relevance.

2.2.1 The Pragmatics of Explanation

[13] Davidson, Donald (1970) "Mental Events." In: L. Foster and J. W. Swanson (eds), *Experience and Theory*. Oxford: Clarendon Press, 207–224.

It might be objected that explanation is not a two-term relation between an explanandum-outcome and a model that explains it, but a three-term relation between an explanandum-outcome, a model that explains it and a range of pragmatic factors, such as the context in which it occurs and interests of those providing and receiving it. For van Fraassen[14] and Lewis,[15] explanations are answers to why-questions and therefore require irreducible reference to such pragmatic factors. Why-questions are often explicitly contrastive: a request for the explanation of P often takes the form "Why P rather than Q?" But even when why-questions are not explicitly contrastive, they may be construed as "Why P rather than other members of its contrast class X?", where the contrast class X is a class of alternatives to P specified by the context in which they occur. Garfinkel[16] and Lipton[17] argue that what counts as an explanatorily relevant factor depends not only on P but also on its contrast class X. So, returning to our example, whether or not C_1 and C_2 are explanatorily relevant to E depends on the context in which the explanation of E occurs.

Explanation may very well be context- and interest-relative but specifying the context in which it occurs and interests of those providing and receiving it will not give us much of a handle on the problem of explanatory relevance. In most contexts, when explaining E, the influence of such spurious causal factors as C_1 and C_2 on E will be straightforwardly ruled out courtesy of pragmatic considerations. Granted. However, pragmatic considerations alone will not do much to discriminate the factors that make a difference to E from the rest of the factors that played a role in producing it. They will rule out C_1 and C_2 as explanatorily irrelevant, but they will neither rule out nor help us gauge the

[14] Van Fraassen, Bas (1980) *The Scientific Image*. Oxford: Clarendon Press.
[15] Lewis, David (1986) "Causal Explanation" in: *Philosophical Papers*, Vol. II. Oxford: Oxford University Press.
[16] Garfinkel, Alan (1981) *Forms of Explanation: Rethinking the Questions in Social Theory*. New Haven, CT: Yale University Press.
[17] Lipton, Peter (1990) "Contrastive Explanation", in D. Knowles (ed.), *Explanation and Its Limits*. Cambridge: Cambridge University Press, 247–266; Lipton, Peter (1991). *Inference to the Best Explanation*. London and New York: Routledge.

relative magnitude of a motley bunch of other factors that, to a greater or lesser extent, causally influence E. These other factors include "knowledge externalities" featured in Romer's[18] growth model, "learning externalities" featured in Tamura's[19] growth model, "international integration" featured in Rivera-Batiz and Romer's[20] growth model etc. All these factors are causally relevant to E, in the sense that, if we could but manipulate them, it is a safe bet that changes in them produced by interventions would be systematically associated with changes in E. And, in most contexts, all these factors will be explanatorily relevant to E. My worry is that since a successful explanation must accurately represent causal mechanisms that bring about its explanandum-outcome, in most contexts explaining E will involve citing a network of causal factors that, though not infinite, will be very complex and intractable indeed. And since it is unlikely that any one model will represent all of them, explaining E will likely involve a battery of models, each requiring its own set of assumptions to fuel the derivation of E. I submit that our understanding of economic outcomes would be aided greatly if we could determine precisely which factors are necessary for their causal production.

3. In Search of Difference-makers

Strevens[21] goes a long way towards addressing the problem of explanatory relevance. Like Hitchcock[22] and Woodward,[23] he argues that difference-making is a necessary condition for explanatory relevance: only factors that make a difference to an explanandum-outcome are explanatorily relevant to this

[18] Romer, Paul M. (1986) "Increasing Returns and Long-Run Growth", *Journal of Political Economy*, **94**(5): 1002–1037.
[19] Tamura, Robert (1991) "Income Convergence in an Endogeneous Growth Model", *Journal of Political Economy*, **99**(3): 522–540.
[20] Rivera-Batiz, Luis A., and Romer, Paul M. (1991) "Economic Integration and Endogenous Growth", *The Quarterly Journal of Economics*, **106**(2): 531–555.
[21] Strevens, Michael (2004) "The Causal and Unification Approaches to Explanation Unified: Causally", *Noûs*, **38**(1): 154–176.
[22] Hitchcock, Christopher R. (1995) "Salmon on Explanatory Relevance", *Philosophy of Science*, **62**(2): 304–320.
[23] Woodward, James (2003) *Making Things Happen*. New York, NY: Oxford University Press.

outcome. However, unlike Hitchcock[24] and Woodward,[25] he offers a procedure for determining precisely which factors are necessary for the causal production of particular outcomes. He calls this procedure "the eliminative procedure for determining difference-makers."[26]

3.1 The Eliminative Procedure for Determining Difference-makers

According to Strevens,[27] difference-makers for E may be determined courtesy of the following procedure:

1. Take the infinitely complex network of factors causally relevant to E, and "find a part [of the network] that [is] in itself sufficient to causally produce E."[28]
2. Remove from it all factors that are not necessary to causally produce E. This means removing all factors that do not "play a role in the entailment of E."[29]

The eliminative procedure is disarmingly simple. A moment's consideration shows, however, that it will not go a long way towards determining difference-markers for overdetermined economic phenomena and that it runs afoul of Duhem's non-separability thesis.

Firstly, the eliminative procedure consists of two steps but I do not see exactly how the first step gets us any closer to determining difference-makers for E. Strevens[30] glosses causation in terms of the semantic relation of entailment and stipulates that a set of conditions is sufficient to causally produce E "just in case the conditions jointly entail the causal production of E." So if $\mathbf{C} = \{C_1, C_2, \ldots, C_n\}$

[24] Hitchcock, Christopher R. (1995) "Salmon on Explanatory Relevance", *Philosophy of Science*, **62**(2): 304–320.
[25] Woodward, James (2003) *Making Things Happen*. New York, NY: Oxford University Press.
[26] Strevens, Michael (2004) "The Causal and Unification Approaches to Explanation Unified: Causally", *Noûs*, **38**(1): 154–176.
[27] Ibid.
[28] Ibid.
[29] Ibid.
[30] Ibid.

was a set of conditions the description of which entailed the description of E, **C** would be sufficient to causally produce E. But then, I do not see exactly how causal sufficiency is different from causal relevance. My first worry is that the set of factors sufficient, in Strevens'[31] sense, to causally produce E and the set of factors causally relevant, in Woodward's[32] sense, to E will be very similar in terms of their respective cardinalities—numbers of elements—and just as intractable.

My first worry is exacerbated by the fact that social phenomena in general and economic phenomena in particular are multiply realizable.[33] Think of all realizing conditions of economic outcomes, such as economic growth in the UK in the fourth quarter of 2022 (E). We are talking millions of economic agents making billions of transactions on a daily basis, we are talking knowledge and learning externalities, international integration etc. If that was not enough, each of these realizing conditions is itself multiply realizable. Think of all realizing conditions of these transactions, think of all realizing conditions of knowledge and learning externalities, international integration etc. Because they are multiply realizable, economic phenomena are overdetermined, in the sense that more than one set of conditions is sufficient to causally produce them. For example, E might be realized by $\mathbf{C} = \{C_1, C_2, \ldots, C_n\}$, where C_1 is the Bank of England raising interest rates to 3.50%, C_2 is the Bank of England reversing quantitive easing etc., but it might also be realized by $\mathbf{C'} = \{C'_1, C'_2, \ldots, C'_n\}$, where C'_1 is the Bank of England raising interest rates to 3.75%, C'_2 is the Bank of England not reversing quantitive easing etc. But then, if the description of both **C** and **C'** entailed the description of E, both **C** and **C'** would be sufficient, in Strevens'[34] sense, to causally produce E. And I would hazard a guess that

[31] Strevens, Michael (2004) "The Causal and Unification Approaches to Explanation Unified: Causally", *Noûs*, **38**(1): 154–176.
[32] Woodward, James (2003) *Making Things Happen*. New York, NY: Oxford University Press.
[33] Fodor, Jerry (1974) "Special Sciences (Or: The Disunity of Science as a Working Hypothesis)", *Synthese*, **28**(2): 97–115; Searle, John R. (1995) *The Construction of Social Reality*. London: Allen Lane; Searle, John R. (2005) "What Is an Institution?" *Journal of Institutional Economics*, **1:** 1–22.
[34] Strevens, Michael (2004) "The Causal and Unification Approaches to Explanation Unified: Causally", *Noûs*, **38**(1): 154–176.

there are more than two sets of conditions the description of which entails the description of E; such sets are a dime a dozen. If I am right, the first step of the eliminative procedure, that is, discriminating the set of factors sufficient to causally produce E from the set of factors causally relevant to E, would leave us where we started, namely with the set of factors causally relevant to E. It would not go a long way towards determining difference-markers for E.

Secondly, and more importantly, supposing that one put a finger on a set of conditions sufficient to causally produce E, it is not clear exactly how one is to remove from it all factors that do not "play a role in the entailment of E." It is a familiar point, known as Duhem's non-separability thesis, that the empirical content of a hypothesis or a theory or any other empirically significant unit cannot be 'parcelled out' among its parts.[35] This is because a conjunction of two hypotheses or theories or any other empirically significant units may very well entail a result that is not entailed by either of them taken in isolation.[36] For example, a sufficiently detailed list of facts about the behaviour of nerve cells in my brain entail propositions about my consciousness, i.e., about me being conscious of my body, my self, the world at large etc. However, no single fact about the behaviour of nerve cells in my brain entails such propositions. So if **C** = $\{C_1, C_2, \ldots, C_n\}$ entails E, there is, and can be, no guarantee that C_1 alone entails E or that C_2 alone entails E or that C_n alone entails E.

Strevens might reply that all factors that do not "play a role in the entailment of E" might be removed 'by hand.' If **C** = $\{C_1, C_2, \ldots, C_n\} \vDash$ E, we might remove C_1 from **C** 'by hand' and see if the resulting set entailed E. If it did, we might

[35] See Ariew, Roger (1984) "The Duhem Thesis", *The British Journal for the Philosophy of Science*, **35**(4): 313–325; Duhem, Pierre (1917) "Liste des Publications de P. Duhem" and "Notice sur les Travaux Scientifiques de Duhem," *Mémoires de la Société des Sciences Physiques et naturelles de Bordeaux*, 7, 41–169. English translation of Parts 2 and 3 of "Notice" in Duhem (1996); Duhem, Pierre (1954) *The Aim and Structure of Physical Theory*. Princeton, NJ: Princeton University Press; Quinn Philip L. (1974) "What Duhem Really Meant", in: *Methodological and Historical Essays in the Natural and Social Sciences*. Dordrecht: Springer.

[36] Quine, Willard V. O. (1981) *Theories and Things*. Cambridge, MA: Harvard University Press.

conclude that C_1 is not necessary to causally produce E. The same might be done with C_2, C_3, all the way to C_n.

But then, suppose that after we had removed C_1 from **C** 'by hand,' we found that the resulting set did not entail E. Does this finding license the conclusion that C_1 is necessary to causally produce E? The answer is an emphatic "no." For all we know, C_1 might be an 'enabler' of C_2, in the sense that it might be a factor, one of many, without which C_2 is causally inert. If C_2 had other enablers, C_1 would not be necessary to causally produce E. In general, the members of **C** might 'interlock' to such an extent that removing all those that do not "play a role in the entailment of E" might be little more than guesswork.[37] This is why I am skeptical about the reply that all factors that do not "play a role in the entailment of E" might be removed 'by hand.'

I argued that the eliminative procedure will not go a long way towards determining difference-markers for overdetermined economic phenomena and that it runs afoul of Duhem's non-separability thesis. I grant that these problems are not damning. For one thing, not all phenomena are as overdetermined as economic phenomena. For another, Duhem's non-separability thesis does not always hold. It was floated with theoretical physics in mind and applies only to those sciences that do not "observe facts directly, but substitute for them measurements […] of magnitudes that only a mathematical theory has defined."[38] So the eliminative procedure may apply to some cases. My point is that, although it is advertised as "a procedure for determining difference-makers," it does not apply 'across the board' to all cases and therefore does not give us much of a handle on the problem of explanatory relevance.

---

[37] Quine, Willard V. O. (1981) *Theories and Things*. Cambridge, MA: Harvard University Press.
[38] Duhem, Pierre (1917) "Liste des Publications de P. Duhem" and "Notice sur les Travaux Scientifiques de Duhem," *Mémoires de la Société des Sciences Physiques et naturelles de Bordeaux*, 7, 41–169. English translation of Parts 2 and 3 of "Notice" in Duhem (1996).

3.2 Robustness Analysis as a Procedure for Determining Difference-makers

In this Subsection, I try to make headway towards solving the problem of explanatory relevance using robustness analysis (RA)—a well-known procedure in theoretical economics by which modellers gauge the sensitivity of their models' results to assumptions that fuel the derivation of these results. I compare RA with the eliminative procedure and conclude that it is better geared to the style of theoretical work dominant in economics. Then, in Subsections 2.2.1 and 2.2.2, I consider two objections to RA as a procedure for determining difference-makers and try to meet them.

Recall that economics is a model-based science in that it studies complex real-world systems indirectly through highly idealized model systems. These model systems are "highly idealized" in that descriptions specifying them contain (descriptively false) idealizing assumptions introduced to isolate causal mechanisms underlying their target systems and (descriptively false) tractability assumptions introduced for reasons of mathematical tractability.[39] For example, the description specifying Howitt's[40] growth model contains the assumption that labour force growth rates are equal across countries. This assumption is descriptively false: labour force growth rates are not equal across countries. For example, according to CEIC data, Bangladesh's labour force participation rate increased to 58.8% in Dec 2022, compared with 58.2% in the previous year; by contrast, Bahrain's labour force participation rate increased to 71.7% in Dec 2022, compared with 71.0% in the previous year. But then, the assumption that labour force growth rates are equal across countries is introduced "not because [it is thought] accurate for describing what is happening now, but because [it is thought] a convenient fiction for a steady state model to explore international spillovers."[41]

---

[39] Wimsatt, William C. (1987) "False Models as a Means to Truer Theories", in: M. Nitecki and A. Hoffmann (eds), *Neutral Models in Biology*. Oxford: Oxford University Press, 23–55.

[40] Howitt, Peter (2000) "Endogenous Growth and Cross-Country Income Differences", *The American Economic Review*, **90**(4): 829–846.

[41] Klenow, Peter J. and Rodriguez-Clare, Andrés (2005) "Externalities and Growth", In: P. Aghion and S. Durlauf (eds), *Handbook of Economic Growth*. North-Holland: Elsevier, 817–861.

Depending on how they are specified, different models maximize different theoretical desiderata of model building. Weisberg[42] argues that there are three such desiderata, namely generality, realism and precision, and shows that they trade off against each other. For example, some models are more general than others, in the sense that they can be applied to more target systems than others. However, generality trades off against precision: general models are those that leave vague the magnitude of the causal forces they describe. The three-way trade-off between generality, realism and precision explains why economic models are in no short supply and why model-based explanations enjoy such a wide currency in economics. And I submit that it is this feature of economic models—abundance—that makes it particularly worthwhile to analyze their results for robustness.

What does it mean to analyze the results of economic models for robustness? Let **M** = {M_1, M_2, …, M_n} be a set of diverse models of some economic phenomenon, say, economic growth. Suppose that each member of **M** can be broken down into a causal core (C_n) and a belt of auxiliary assumptions (A_1, A_2, …, A_n). And suppose that two members of **M** are diverse if, and only if, they can be broken down into logically non-equivalent auxiliary assumptions. Given that **M** is a set of *diverse* models, we know that each of its members contains logically non-equivalent assumptions and maximizes different theoretical desiderata of model building so our background knowledge does not favour any member of **M** over its competitors. In a classic paper, Levins[43] argues that if these models, despite their logically non-equivalent assumptions, converge on similar results, this would justifiably increase the modeller's confidence that their converging on similar results depends not on the details of their assumptions but on the 'essentials' shared across them:

---

[42] Weisberg, Michael (2003) "When Less is More: Tradeoffs and Idealization in Model Building." Dissertation, Stanford University; Weisberg, Michael (2006a) "Forty Years of 'The Strategy': Levins on Model Building and Idealization", *Biology and Philosophy*, **21**(5): 623–645.

[43] Levins, Richard (1966) "The Strategy of Model Building in Population Biology", *American Scientist*, **54**(4): 421–431.

> [I]f these models, despite their different assumptions, lead to similar results, we have […] a robust theorem that is relatively free of the details of the model. Hence, our truth is the intersection of independent lies.[44]

A "robust theorem" is Levins'[45] term of art for a conditional statement, sometimes prefaced by a qualifying *ceteris paribus* clause, that links the 'essentials' shared across a set of diverse models with the results they converge on. So analyzing the results of economic models for robustness means searching for such theorems.[46]

Examples of robust theorems are few and far between, but consider the following one. The Classical Growth Model (M_1), the Malthusian Growth Model (M_2) and the Solow Growth Model (M_3) are members of **M**. They are diverse, in the sense that they contain logically non-equivalent assumptions, some of which are descriptively false: M_1 assumes saving-investment equality, M_2 assumes that the standard of living and the total fertility rate are directly proportional and M_3 assumes full employment of capital and labour. If M_1, M_2 and M_3, despite their logically non-equivalent assumptions, converged on similar results, say, a particular value of economic growth in the UK in the fourth quarter of 2022 (E), this convergence would be either an artefact of some causal structure shared across them or "a remarkable coincidence."[47] So after determining that M_1, M_2 and M_3 converge on a particular value of E, they might be analyzed for a common causal structure. What M_1, M_2 and M_3 have in common is that they all construe economic growth (Y) as a function of three variables: technology (A), physical capital stock (K) and human capital stock (L). So it looks like we might have latched onto a robust theorem: Y is a function

[44] Ibid.
[45] Ibid.
[46] Weisberg, Michael (2006b) "Robustness Analysis", *Philosophy of Science*, **73**(5): 730–742.
[47] Kuorikoski, Jaakko. et al. (2010) "Economic Modelling as Robustness Analysis", *The British Journal for the Philosophy of Science*, **61**(3): 541–567.

of A, K and H, and, *ceteris paribus*, changes in A, K and L are associated with changes in Y. Notice robust theorems are (qualified) patterns of counterfactual dependance: to say that changes in A, K and L are associated with changes in Y is to say that A, K and L fit into a pattern of counterfactual dependance with Y.

I therefore submit that RA lends itself admirably to determining the factors that make a difference to a given explanandum-outcome and distinguishing these difference-makers from the rest of the factors that played a role in producing it. Although it would be wishful thinking to assume that RA is a sure-fire procedure for determining difference-makers, I think that the odds are good that the 'essentials' shared across a set of diverse models of some economic phenomenon (E) are factors that matter, counterfactually, to E: if we could but manipulate them, this would be a way of manipulating E. And I think that adding new members to this set would increase these odds: the greater the cardinality—number of elements—of a set of diverse models, the greater the odds that the 'essentials' shared across them fit into a pattern of counterfactual dependance with E. How does RA compare with the eliminative procedure? I argued that the eliminative procedure will not go a long way towards determining difference-markers for economic phenomena because they are overdetermined, in the sense that more than one set of conditions is sufficient to causally produce them. What I argued is a crippling handicap for the eliminative procedure is a valuable asset to RA. Overdetermination of economic phenomena ensures the steady supply of diverse models targeting them. If these models, despite their logically non-equivalent assumptions, converged on similar results, they might be analyzed for a common causal structure.

3.2.1 Objection 1: Models Do Not Decompose That Way

I also argued that the eliminative procedure runs afoul of Duhem's non-separability thesis. One might balk at RA and object that it does so too. I

acknowledge this objection and think it is well-taken. There is no denying that analyzing diverse models for a common causal structure assumes that models can be broken down into parts.[48] However, I think there is a difference between RA and the eliminative procedure. Showing that a set of diverse models converges on a particular result (E), analyzing these models for a common causal structure (C) and then attributing E to C is a way of 'parcelling out' the empirical content of a model among its parts. Granted. However, I argued that it would be wishful thinking to assume that RA is a sure-fire procedure for determining difference-makers. It is not. C may be a difference-maker for E or not, but it is possible that it is. My contention, then, is that RA allows us to make educated guesses about difference-makers in the face of Duhem's non-separability thesis. By contrast, Strevens[49] argues that "in order for an event C to qualify as a difference-maker for an event E, it is necessary and sufficient that C appear in a kernel Strevens'[50] term of art for the end product of the eliminative procedure] for E." Appearing in a "kernel" for E may be sufficient for qualifying as a difference-maker for E, but it is not necessary. Duhem's non-separability thesis, as long as it holds, gives the lie to this claim. One cannot strip a model of all parts that do not "play a role in the entailment of E." This is because these parts might 'interlock' to such an extent that, removing any one of them would make the model causally inert and its empirical consequence class empty. So, if appearing in a "kernel" for E was necessary for qualifying as a difference-maker for E, I have a hunch that no event would qualify as a difference-maker for E.

3.2.2 Objection 2: Models Do Not Represent Causal Mechanisms

[48] Rice, Collin (2019) "Models Don't Decompose That Way: A Holistic View of Idealized Models", *British Journal for the Philosophy of Science*, **70**(1): 179–208.
[49] Strevens, Michael (2004) "The Causal and Unification Approaches to Explanation Unified: Causally", *Noûs*, **38**(1): 154–176.
[50] Ibid.

One might also object that since economic models are shot through with false assumptions, they misrepresent their target systems and therefore are not explanatory. For example, Odenbaugh and Alexandrova[51] argue that if a model contains false assumptions, the causal mechanism it represents "cannot be what the model says it is." Similarly, Alexandrova and Northcott[52] argue that models "do not qualify as causal explanations because they are false and therefore do not identify any actual causes." But then, it is a familiar point that *all* models contain false assumptions and are false in this sense. If only true models could identify actual causes, this would lead to the untoward conclusion that none of them identifies actual causes and therefore none of them qualifies as a causal explanation. Indeed, Kuorikoski et al.[53] argue by *reductio ad absurdum* (RED) that if every false assumption entering a given model-based explanation had to be discharged or de-idealized with some true assumption for it to accurately represent causal mechanisms, no model-based explanation, but only "reality itself," would be "capable of such a feat." And yet, the argument goes, at least some model-based explanations are explanatory. Kuorikoski et al.[54] conclude, by RED, that it is not the case that every false assumption entering a given model-based explanation has to be discharged or de-idealized with some true assumption for it to accurately represent causal mechanisms.

I grant that if no model represented causal mechanisms, RA as a procedure for determining difference-makers would be a 'no go.' But the claim that models containing falsehoods do not represent causal mechanisms strikes me as plain

[51] Odenbaugh, Jay and Alexandrova, Anna (2011) "Buyer Beware: Robustness Analyses in Economics and Biology", *Biology and Philosophy*, **26**(5): 757–771.
[52] Alexandrova, Anna and Northcott, Robert (2013) "It's Just A Feeling: Why Economic Models Do Not Explain", *Journal of Economic Methodology*, **20**(3): 262–267.
[53] Kuorikoski, Jaakko. et al. (2012) "Robustness Analysis Disclaimer: Please Read the Manual Before Use!", *Biology and Philosophy*, **27**(6): 891–902.
[54] Ibid.

wrong. Hausman[55] and Mäki[56] argue that *all* models misrepresent their targets, but this does not automatically bar them from representing causal mechanisms and being explanatory. More specifically, the fact that a model contains falsehoods "does not preclude employing the model in giving explanations if the explanations do not rely on the falsehoods."[57] Since it is not always the case that explanations in economics rely on the falsehoods contained in models, the premise that only true models can identify actual causes and be explanatory is not a good one. The explaining may very well be done by accurate claims about causal mechanisms, should models contain any. In other words, models containing falsehoods may accurately represent *some* causal mechanisms underlying their targets, and therefore they may be used to explain their targets, insofar as these false assumptions are not "driving the results."[58]

4. Conclusion

RA is a well-known procedure in theoretical economics by which modellers gauge the sensitivity of their models' results to assumptions that fuel the derivation of these results. I argued that the applicability of RA is wider than that and made a case for RA as a procedure for determining difference-makers for overdetermined economic phenomena. Although RA is not a sure-fire procedure for determining difference-makers, it allows us to make educated guesses about what factors matter, counterfactually, to a given explanandum-outcome. It also compares favorably with the eliminative procedure, though, to be sure, both procedures are beset with similar problems, especially Duhem's non-separability thesis, and neither of them applies 'across the board' to all cases. In general, I think RA may have the edge over the

[55] Hausman, Daniel M. (2013) "Paradox Postponed", *Journal of Economic Methodology*, **20**(3): 250–254.
[56] Mäki, Uskali (2013) "On a Paradox of Truth, or How Not to Obscure the Issue of Whether Explanatory Models Can Be True", *Journal of Economic Methodology*, **20**(3): 268–279.
[57] Hausman, Daniel M. (2013) "Paradox Postponed", *Journal of Economic Methodology*, **20**(3): 250–254.
[58] Kuorikoski, Jaakko. et al. (2010) "Economic Modelling as Robustness Analysis", *The British Journal for the Philosophy of Science*, **61**(3): 541–567.

eliminative procedure when the explanandum-outcome is a *type*-phenomenon. It may give us a handle on the problem of explanatory relevance when we want to determine difference-makers for economic growth *in general* or inflation *in general*. By contrast, the eliminative procedure may have the upper hand when the explanandum-outcome is a *token*-phenomenon. It may give us a handle on the problem of explanatory relevance when we want to determine difference-makers for a *particular* economic growth rate or a *particular* inflation rate.

References

Alexandrova, Anna and Northcott, Robert (2013) "It's Just A Feeling: Why Economic Models Do Not Explain", *Journal of Economic Methodology*, **20**(3): 262–267.

Ariew, Roger (1984) "The Duhem Thesis", *The British Journal for the Philosophy of Science*, **35**(4): 313–325.

Bechtel, William and Richardson, Robert C. (2010) *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Cambridge, MA: MIT Press.

Davidson, Donald (1970) "Mental Events." In: L. Foster and J. W. Swanson (eds), *Experience and Theory*. Oxford: Clarendon Press, 207–224.

Duhem, Pierre (1917) "Liste des Publications de P. Duhem" and "Notice sur les Travaux Scientifiques de Duhem," *Mémoires de la Société des Sciences Physiques et naturelles de Bordeaux*, 7, 41–169. English translation of Parts 2 and 3 of "Notice" in Duhem (1996).

Duhem, Pierre (1954) *The Aim and Structure of Physical Theory*. Princeton, NJ: Princeton University Press.

Duhem, Pierre (1996) *Essays in History and Philosophy of Science*. Indianapolis, IN: Hackett.

Garfinkel, Alan (1981) *Forms of Explanation: Rethinking the Questions in Social Theory*. New Haven, CT: Yale University Press.

Giere, Ronald (1988) *Explaining Science: A Cognitive Approach*. Chicago, IL: Chicago University Press.

Godfrey-Smith, Peter (2006) "The Strategy of Model-based Science", *Biology and Philosophy*, **21**(5): 725–740.

Godfrey-Smith, Peter (2009) "Models and Fictions in Science", *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, **143**(1): 101–116.

Fodor, Jerry (1974) "Special Sciences (Or: The Disunity of Science as a Working Hypothesis)", *Synthese*, **28**(2): 97–115.

Hausman, Daniel M. (2013) "Paradox Postponed", *Journal of Economic Methodology*, **20**(3): 250–254.

Hesslow, Germund (1983) "Explaining Differences and Weighting Causes", *Theoria*, **49**(2): 87–111.

Hesslow, Germund (1988) "The Problem of Causal Selection", In: D. J. Hilton (ed.), *Contemporary Science and Natural Explanation: Commonsense Conceptions of Causality*. New York, NY: New York University Press.

Hitchcock, Christopher R. (1995) "Salmon on Explanatory Relevance", *Philosophy of Science*, **62**(2): 304–320.

Howitt, Peter (2000) "Endogenous Growth and Cross-Country Income Differences", *The American Economic Review*, **90**(4): 829–846.

Kaplan, David M. and Craver, Carl F. (2011) "The Explanatory Force of Dynamical and Mathematical Models in Neuroscience: A Mechanistic Perspective*", *Philosophy of Science*, **78**(4): 601–627.

Klenow, Peter J. and Rodriguez-Clare, Andrés (2005) "Externalities and Growth", In: P. Aghion and S. Durlauf (eds), *Handbook of Economic Growth*. North-Holland: Elsevier, 817–861.

Kuorikoski, Jaakko. et al. (2010) "Economic Modelling as Robustness Analysis", *The British Journal for the Philosophy of Science*, **61**(3): 541–567.

Kuorikoski, Jaakko. et al. (2012) "Robustness Analysis Disclaimer: Please Read the Manual Before Use!", *Biology and Philosophy*, **27**(6): 891–902.

Levins, Richard (1966) "The Strategy of Model Building in Population Biology", *American Scientist*, **54**(4): 421–431.

Lewis, David (1986) "Causal Explanation" in: *Philosophical Papers*, Vol. II. Oxford: Oxford University Press.

Lipton, Peter (1990) "Contrastive Explanation" in: D. Knowles (ed.), *Explanation and Its Limits*. Cambridge: Cambridge University Press, 247–266.

Lipton, Peter (1991) *Inference to the Best Explanation*. London and New York: Routledge.

Machamer, Peter, Darden, Lindley and Craver, Carl F. (2000) "Thinking About Mechanisms", *Philosophy of Science*, **67**(1): 1–25.

Mäki, Uskali (2013) "On a Paradox of Truth, or How Not to Obscure the Issue of Whether Explanatory Models Can Be True", *Journal of Economic Methodology*, **20**(3): 268–279.

Odenbaugh, Jay and Alexandrova, Anna (2011) "Buyer Beware: Robustness Analyses in Economics and Biology", *Biology and Philosophy*, **26**(5): 757–771.

Quine, Willard V. O. (1981) *Theories and Things*. Cambridge, MA: Harvard University Press.

Quinn Philip L. (1974) "What Duhem Really Meant", in: *Methodological and Historical Essays in the Natural and Social Sciences*. Dordrecht: Springer.

Railton, Peter (1978) "A Deductive-Nomological Model of Probabilistic Explanation", *Philosophy of Science*, **45**(2): 206–226.

Rice, Collin (2019) "Models Don't Decompose That Way: A Holistic View of Idealized Models", *British Journal for the Philosophy of Science*, **70**(1): 179–208.

Rivera-Batiz, Luis A., and Romer, Paul M. (1991) "Economic Integration and Endogenous Growth", *The Quarterly Journal of Economics*, **106**(2): 531–555.

Romer, Paul M. (1986) "Increasing Returns and Long-Run Growth", *Journal of Political Economy*, **94**(5): 1002–1037.

Salmon, Wesley C. (1984) *Scientific Explanation and the Causal Structure of the World*. Princeton, NJ: Princeton University Press.

Searle, John R. (1995) *The Construction of Social Reality*. London: Allen Lane.

Searle, John R. (2005) "What Is an Institution?" *Journal of Institutional Economics*, **1:** 1–22.

Strevens, Michael (2004) "The Causal and Unification Approaches to Explanation Unified: Causally", *Noûs*, **38**(1), 154–176.

Strevens, Michael (2009) *Depth: An Account of Scientific Explanation*. Cambridge, MA: Harvard University Press.

Tamura, Robert (1991) "Income Convergence in an Endogeneous Growth Model", *Journal of Political Economy*, **99**(3): 522–540.

Van Fraassen, Bas (1980) *The Scientific Image*. Oxford: Clarendon Press.

Weisberg, Michael (2003) "When Less is More: Tradeoffs and Idealization in Model Building". Dissertation, Stanford University.

Weisberg, Michael (2006a) "Forty Years of 'The Strategy': Levins on Model Building and Idealization", *Biology and Philosophy*, **21**(5): 623–645.

Weisberg, Michael (2006b) "Robustness Analysis", *Philosophy of Science*, **73**(5): 730–742.

Weisberg, Michael (2007a) "Who Is a Modeler?", *The British Journal for the Philosophy of Science*, **58**(2): 207–233.

Weisberg, Michael (2007b) "Three Kinds of Idealization", *The Journal of Philosophy*, **104**(12): 639–659.

Wimsatt, William C. (1987) "False Models as a Means to Truer Theories", in: M. Nitecki and A. Hoffmann (eds), *Neutral Models in Biology*. Oxford: Oxford University Press, 23–55.

Woodward, James (2003) *Making Things Happen*. New York, NY: Oxford University Press.

Woodward, James (2006) "Some Varieties of Robustness", *Journal of Economic Methodology*, **13**(2): 219–240.