Economic theory and empirical science

1. Introduction

I argue that economics, notwithstanding its recent 'empirical turn', over-invests in orthodox theory. My reasons are not familiar philosophical complaints about idealization, social ontology, or the foundations of rational choice theory (Elster 1988, Rosenberg 1992, Lawson 1997).[1] Instead, they come from closely examining how economics achieves empirical success (when it does). To preview: in such cases, theory typically does not itself illuminate the world's causal structure and so does not directly explain. Rather, it is of benefit – when it is – only heuristically. Two further lessons follow, both of which, alas, run contrary to much contemporary practice: first, theory should be developed in close concert with empirical feedback; and second, theory has no special reason to be orthodox.[2]

A central goal of economics, like that of any science, must be empirical success. I understand the latter broadly, to include prediction, intervention, and retrospective accommodation and explanation. Other goals, such as understanding or insight, are also important but, as will become apparent, they are dependent on empirical success, so such success must come first. Economics is a heterogeneous field. Nevertheless, there is enough uniformity of method, at least in the mainstream, that general analyses can be useful. How idealized theory can connect with messy reality is a philosophical issue that requires philosophical analysis – and there has been plenty. In the space available here, I bring some of that philosophical work to bear.

Obeying the editors' request, I state some lines of criticism bluntly. But I view these criticisms as friendly amendments, concerning matters of balance and degree, and unlike some other philosophical critiques they allow for two *endorsements* of economics. First, the economist's outlook, in particular the theoretical machinery of incentives and opportunity cost and their consequences for decision-making, is useful very widely. This will be a familiar insight to any person with economic training listening to a person without such training. Second, there is of course much to economics besides theory: measurement, commercial consultancy, applied studies, teaching, and more. These can all be of great value. Especially impressive is the sophisticated use of statistical techniques, covering research design, data analysis and hypothesis testing. These techniques are perhaps the aspects of contemporary economic work most exportable to social science more widely.

2. Idealization: a red herring

Economic theory is typically highly idealized. Often, perfectly rational agents with perfect information interact in perfectly competitive markets, simple elasticities and production functions are assumed, and so on. This immediately threatens such theory's empirical prospects.

But economists have a ready reply. Many theories in other sciences too, such as physics, are highly idealized, and besides, given that any model must simplify reality, some idealization is inevitable. So, why worry?

[1] I also do not criticize economic theory here on moral or political grounds. My concerns are purely methodological.

[2] I use 'orthodox' to label a commitment to formal rational choice models, which have been the dominant theoretical form since World War Two. I will refer to such models generically as "theory".

Recent philosophy of science has studied idealization extensively.[3] There is broad agreement, but with a sting in the tail: roughly speaking, idealized theory can be endorsed, but only if it brings empirical success. For example, the most familiar form of idealization in physics is so-called Galilean idealization. Particular factors are highlighted and others ignored, as when the effect of gravity on a projectile is modelled while assuming zero air resistance. In the best case, the idealized model accurately describes the projectile's actual trajectory. If it doesn't, factors that were idealized away may be re-introduced. The point is that at some stage the theoretical framework must deliver empirical success, otherwise it will be jettisoned, often pretty quickly. Similar remarks apply to other forms of idealization in physics, such as simulations, or re-normalizations in quantum theory. To survive, they must pay their way empirically. For this reason, by and large the history of modern physics shows a close co-evolution of theory with empirical feedback. Rare exceptions, such as string theory, are controversial precisely because they are exceptions.

The conclusion: idealization is not what matters. Rather, empirical success (or lack of it) is.

3. How empirical success is achieved

The empirical record of economics is notorious. Forecasts of real GDP 18 months ahead, for example, have proved unable to beat the naïve benchmark of simply extrapolating the current GDP growth rate. One study found that of 60 recessions, i.e. instances of one year of negative growth, a recession was the consensus forecast beforehand on only three occasions. Moreover, there has been no improvement in 50 years.[4]

The fuller picture is more nuanced, though. Regarding prediction, even with GDP, forecasts for six or nine months ahead do beat the naïve benchmark. Meanwhile, businesses use economic models every day to predict supply needs. Introductory textbook examples are instantiated all the time as well, as when, say, two popular teams reach a sports final and it is correctly predicted that black market ticket prices will be higher.

That said, empirical successes are often claimed rather casually. A model may accommodate retrospectively merely part of the variance of some dataset. It can be easy to claim informally that economic theory 'explains' some phenomenon or enables us to 'understand' why something happened, but detailed and convincing case studies are rarer. It turns out to be much more instructive to look at one example in depth than to look at many superficially. Precisely this has been done in a few cases. Here, I recount a well-known and representative one: the US government spectrum auctions of 1994-6.[5] What exactly was theory's role there?

The radio spectrum is the portion of electromagnetic spectrum between 9 kilohertz and 300 gigahertz. In the USA, spectrum not needed by the government is distributed to potential users – usually telecommunications companies – by the Federal Communications Commission (FCC). In the early 1990s, the FCC acquired the right to use competitive market mechanisms such as auctions. That left it the formidable task of designing such auctions. The importance of doing this well is best illustrated by the embarrassment of doing it badly. Examples include: an Otago university student winning the license for a small-town TV station by bidding just $5 (New Zealand in 1990), an unknown outbidding everyone but then

[3] See, for instance, work by Nancy Cartwright, Daniel Hausman, Uskali Mäki, Michael Strevens, and Michael Weisberg.
[4] Betz (2006). Forecasting performance since then, such as of the 2008 crisis, has arguably been even worse.
[5] The following paragraphs draw on Guala (2005), Alexandrova (2008), and Alexandrova & Northcott (2009). See those works for references and detailed discussion.

turning out to have no money and so necessitating an expensive do-over (Australia in 1993), and collusion by four big companies to buy the four available licenses for prices only one-fifteenth of what the government had expected (Switzerland in 2000). In contrast, the FCC's series of seven auctions from 1994 to 1996 were remarkably successful. They attracted many bidders, allocated several thousand licenses, and raised an amount of money – $20 billion – that surpassed all government and industry expectations. Even the first auctions passed off without a glitch, and there was reason to believe that licenses were allocated efficiently. How was this achieved?

The government set a wide range of goals besides maximizing revenue, such as using the spectrum efficiently and intensively, promoting new technologies, and ensuring that some licenses went to favored bidders such as minority- and women-owned companies. Exactly what design would reliably achieve these goals was a formidable puzzle for teams of economic theorists, experimentalists, lawyers, and software engineers. To give a flavor of the eventual solution's complexity, geographically the country was subdivided into 492 basic trading areas, each of which had four spectrum blocks up for license. The design put all of these licenses up for sale simultaneously as opposed to sequentially, in an open rather than sealed-bid arrangement. Bidders placed bids on individual licenses as opposed to packages of licenses. When a round was over, they saw what other bids had been placed and were free to change their own combinations of bids. Bidders were also forced to maintain a certain level of activity, make upfront payments, increase the values of their bids from round to round by prescribed amounts, and obey caps on the amount of spectrum that could be owned in a single geographical area. The full rules ran to over 130 pages.

At the time, this gleaming success was hailed by the press as a triumph for game theory, which had revolutionized auction models in the 1980s. Many game theorists were hired as advisors by prospective bidders and by the FCC itself. However, the final design was not derived (or derivable) from game theory. No single model covered anywhere near all of the theoretical issues of the kind mentioned above. And in addition to the explicit and public instructions covering entry, bidding, and payment, much work also had to be put into perfecting other features such as the software, the venue and timing of the auction, and whatever aspects of the legal and economic environment the designers could control. Many experiments and *ad hoc* adjustments were crucial for the purpose of fine-tuning. These took the form of extensive testing in laboratory settings with human subjects, the results often taking designers by surprise. For example, in some circumstances – and against theoretical predictions – 'bubbles' emerged in the values of the bids. These bubbles in turn were unexpectedly sensitive to exactly what bidders knew about rival bidders' behavior. To solve the problem required investigating messy practical details.

What role did theory actually play? First, contrary to one philosophical view (Hausman 1992), the auction design was not achieved by satisfying, even approximately, theory's idealized assumptions, such as perfect rationality, no budget constraints on bidders, single units on sale (as opposed to hundreds of spectrum licenses simultaneously), and so on. And no model yielded the consequences of these assumptions not being satisfied.

Second, contrary to another philosophical view (Cartwright 1989, Mäki 1992), theory also did not identify *capacities* – causal relations that hold stably across many cases – that could be combined to design the auction.[6] Experiments demonstrated quite the opposite. The impact

[6] Something like the capacities view of theory is held by many economists (Rodrik 2015).

of any particular auction rule varied, in a way not predicted by theory, depending both on the details of how it was implemented and also on which other rules were included. In other words, a rule's effect was unstable, not stable. As a result, testing had to be holistic: because individual rules did not have stable effects across different environments, so the performance of any particular *set* of rules had to be tested as a sui generis package, and moreover tested anew with every significant change in environment. The eventual result of a complex testing process was the perfection of one auction design as a whole. This design was not a case of component capacities being stitched together, because no relevant capacities were stable enough for that.

The role of auction theory was in fact a *heuristic* one (Alexandrova 2008, Alexandrova & Northcott 2009): it suggested some useful initial ideas and categories. This was no small contribution, but of course it then left the subsequent heavy lifting to experimentalists and others, namely how to combine these and other factors into a workable design? The key to progress with that was clearly not theory. After all, much the same theoretical repertoire was available to the FCC as it had been earlier to the New Zealand and Australia authorities, yet the FCC's auction fared much better, so it wasn't new theory that made the difference. The key instead was case-specific experiments and know-how, finely sensitive to local political and economic conditions. For this reason, spectrum auction designs do not transfer easily across cases, which is why the 2000 auction in Switzerland could fail even though it was held after the successful US one. The causes of the Swiss failure were specific to Swiss circumstances.

The same heuristicist moral emerges from other case studies too, such as of attempts to apply the Prisoner's Dilemma game (Northcott & Alexandrova 2015). The Prisoner's Dilemma itself rarely predicts accurately. Its value, when it has some, is instead indirect and heuristic, by guiding our attention to the strategic incentives that encourage co-operation, and by alerting us to possible divergence between individual and social optimality.

Another moral that recurs across cases is the inadequacy of the capacities view. First, economic theory can usually derive a capacity only on the back of many idealized assumptions, so the capacity is established only in the idealized world of theory; supplementary empirical work is required to establish it in the actual world. But, second, experience shows that, by and large, this supplementary empirical work cannot successfully be done. If it could, then the capacities described by theory could be used as reliable guides for prediction and intervention in the actual world, but this is rarely the case, because causal relations established in the actual world are not stable. For example, one well-known study established a case in which raising the minimum wage increased employment (Card & Krueger 1994). But in other circumstances it no longer does – say, when the minimum wage is already high, when the increase in it is large, or when economic conditions are different. In response, crucially, rather than search for countervailing capacities that might be outweighing the original one, instead researchers just assumed that that the original capacity no longer held, in other words that raising the minimum wage no longer tended to increase employment (Reiss 2008, 173-6). This response is typical. Similar remarks apply to causal relations discovered by economic experiments (Reiss 2008, 92-6). *Economists' own practice indicates that the capacity view is not really believed*. The *form* of economic theory does suggest a world of stable causal relations, but this is misleading. In fact, causal relations are ubiquitously accepted to be fragile.

4. Methodological lessons

What do these admirable cases of empirical success teach us? First, the need for *continuous empirical refinement* when developing theory, because without such refinement theory risks drifting into empirical irrelevance (Ylikoski 2019). Empirical refinement entered the spectrum auctions story only at the stage of the final design, via the experimental testbeds. Its absence before then is why the game theoretical models contributed only heuristically.

Next, and contrary to many casual claims, economic theory *typically does not explain* (Northcott & Alexandrova 2013). The main reason is simple: economic theory is false and therefore does not identify true causal relations (or at least approximately true ones), which is what it must do in order to causally explain.[7] The warrant for having truly identified causal relations must be empirical success, but empirical success is just what economic theory lacks. In the case of the spectrum auctions, for example, warrant from empirical success accrued only to the final auction design, not to the game theoretical models.

Moreover, economic theory does not even 'partially' explain in the sense of correctly identifying *some* of the relevant causes (Northcott 2013). To see why not, consider a textbook case from physics, such as Coulomb's law: even if this Law's prediction that two positively-charged bodies repel each other is not borne out, perhaps because of interference by other forces (such as gravity), still we would have confidence that an electrostatic repulsion force was still influencing the bodies and so that Coulomb's law did explain the bodies' overall trajectory partially. Why this confidence? Because of the empirical success of Coulomb's law elsewhere, such as in laboratory experiments, combined with confidence that the electrostatic forces observed in a laboratory also operate in the relevant field environment too (Cartwright 1989). Alas, in economics the second confidence is not warranted because causal relations are typically not stable enough to be transportable. This is the instability problem for the capacities view again. The result: no warrant even for partial explanations.

The next important consequence of heuristicism is that *we lose motivation for orthodoxy*. Because economic theory typically is not causally explanatory, so we cannot say that its internal structure picks out actual causes. Without that, a traditional motivation for orthodoxy, namely that only in this way can we explain phenomena in terms of incentive structures and agent rationality, becomes moot. Meanwhile, if theory's value is heuristic, then any approach that is similarly heuristically useful will be similarly valuable, regardless of internal structure. We should therefore accept potential roles for heterodox approaches, such as Marxist or Austrian economics, behavioral economics, agent-based simulations, econophysics, or network analysis. (I do not here either endorse or reject the actual record of any of these.) For similar reasons, it is ill-motivated to insist on methodological individualism, or to insist on orthodox 'microfoundations' for macroeconomic theory.

This methodological liberalism runs counter not just to economists' expressed views but also to their widespread practice, which prizes orthodoxy above empirical accuracy. Here is one example of that (Reiss 2008, 106-22): Milton Friedman and Anna Schwartz proposed a mechanism for their famous finding that money is the main cause of changes in nominal income (Friedman & Schwartz 1963). The problem with this mechanism, in orthodox eyes, is

[7] Economic theory does not explain in non-causal ways either, such as via unification or via mathematical explanation (Northcott & Alexandrova 2013). Whether economic theory might still be explanatory in some non-standard sense has been much debated in the philosophical literature – indeed perhaps too much so. But arguably what matters more is whether, however exactly we understand explanatory success, theory development is a good way to get it (section 7). The literature has said little or nothing about this, or therefore about, say, whether we should welcome the empirical turn (section 6).

that it assumes money illusion and therefore contravenes agent rationality. Decades of effort followed to find a more acceptable alternative, culminating in (Benhabib & Farmer 2000). But although satisfying orthodox criteria, Benhabib and Farmer's ingenious new mechanism is arguably less empirically adequate than Friedman and Schwartz's old one, featuring as it does not just a standard array of idealizing assumptions but also some unusual new ones, such as a downward-sloping labor supply curve. Yet Benhabib and Farmer's paper was still seen as an important breakthrough. Why? Because orthodoxy is prized over empirical accuracy.

Heuristicism also renders illusory another often-cited advantage of theory, namely that it is *generalizable*. A great motivation for developing causal models is that they can be applied to many cases. But if theory's value is only heuristic, this advantage melts away. Heuristic value carries over to new cases much less reliably, because each time new extra-theoretical, local work is required (Alexandrova & Northcott 2009, Northcott 2015).

Heuristicism implies one further important thing: that *theoretical sophistication is no goal in itself*. Because economic theory typically has not been developed in close concert with empirical refinement, so in practice greater sophistication often leaves theory more remote from empirical application, making it *less* valuable by rendering it less heuristically useful. Many empirical successes, such as the introductory textbook ones mentioned earlier, feature very simple models indeed, sometimes nothing more than a downward-sloping demand curve. Many empirical successes in other field sciences too are also very simple theoretically. Simpler theory, without the need for extensive idealizations to enable deductive derivation, is also more likely to predict or explain successfully. If so, then simpler theory should be endorsed, regardless of methodological orthodoxy. Examples include well-evidenced historical analyses, and 'theory-free' field experiments carried out by developmental economists or by internet companies such as Google and Facebook.

All of the above runs counter to a common motivation for orthodox theory, namely that it enables us to understand phenomena in terms of the logic of incentives and rational choice. On some accounts, an economic analysis must *by definition* be couched in these terms, precisely because doing so guarantees (it is thought) such understanding. In its way, this commitment is noble and idealistic. But, alas, it is hard to defend. As noted, we have good reason to reject the claim that economic theory explains, and certainly the mere subjective feeling of understanding is no remedy for that because such a feeling is an unreliable indicator of explanation (Trout 2007, Northcott & Alexandrova 2013). We must, therefore, demonstrate some value for understanding independent of explanation. But a majority of the philosophical literature denies that this can be done (Khalifa 2012, Strevens 2013, de Regt 2017).[8]

The leading contrary view is that understanding amounts to *how-possibly* explanations, which are distinct from actual ones. Gruene-Yanoff (2009), for example, holds that Schelling's famous checker-board model of race and location serves to suggest a causal hypothesis about the actual world. It does this by establishing that such a hypothesis could *possibly* hold – in an idealized, and hence non-actual, world. Such hypotheses, as in the Schelling case, are often surprising and therefore potentially useful heuristically. But if so, then the value of understanding is just heuristic. No new third path appears, separate from explanation and heuristics, and so understanding offers no salvation for orthodoxy.

[8] Two separate errors can occur here. First, thinking that the feeling of understanding implies explanation when in fact it doesn't. And second, when faced with good arguments that it doesn't, then insisting that there must be some *other* epistemic good, distinct from explanation, indicated by the feeling of understanding.

5. Mill on field sciences

Perhaps the best available defense of economic theory's empirical shortcomings stretches back to John Stuart Mill. At its heart lies a distinction between experimental and field sciences. Roughly, a field science is one that studies uncontrolled phenomena outside the laboratory and therefore cannot run shielded experiments. Economics falls into this category. Mill argued (1843) that, unlike experimental sciences, field sciences should not adopt what he called the method of inductive generalizations, in other words they should not insist that theory predict accurately. This is because the ever-changing mix of causes in field cases makes accurate prediction a naïve goal. Instead, we should follow a *deductivist* method, according to which theory states core causal tendencies such as human agents' tendency to maximize their wealth. These causal tendencies are (roughly) what we earlier called capacities. In any particular case, we compose relevant tendencies in a deductive way and then add in as required local 'disturbing causes', in other words local factors not captured by theory. According to this view, deductivist theory delivers – even without empirical accuracy – explanation and understanding in terms of underlying causes.

Similar proposals to Mill's have recurred, with similar rationales. Examples include Weber's advocacy of ideal types, and Popper's of situational analysis. By and large, deductivism has remained the dominant method in economics, periodic rebellions notwithstanding. Something like it was the winner of the *Methodenstreit* in late-19th century Germany, and over the institutionalists in the 20th century. The recent empirical turn (section 6) is the latest chapter in the story.

According to Mill, although theory does not deliver empirical success directly, we can obtain it indirectly by adding in disturbing causes. In this way, deductivist theory is *more* empirically fruitful than inductivist alternatives because it offers (indirectly) empirical success that generalizes to many cases, by adding in different disturbing causes each time.

But does economic theory indeed play the role that Mill envisages, needing to be supplemented only by case-specific disturbing causes? We have seen, on the contrary, that this is *not* what happens. Theory does not identify Millian stable capacities but rather plays only a heuristic role. Empirical successes are built on local and empirically refined models, not by adding disturbing causes to general capacities. Mill-style theorizing therefore does not deliver explanations in terms of core tendencies. This methodological strategy does not work.

6. The empirical turn

Times are changing. In the five most prestigious journals in economics, the percentage of papers that were purely theoretical – in other words free of any empirical data – fell from 57% in 1983 to 19% in 2011 (Hamermesh 2013).[9] Not only is there more empirical work but this empirical work is also less often theory-based. Ever since the Cowles Commission at the end of the war, there has been a strong norm that econometrics should aim at testing particular theoretical models rather than more fragmented or non-theory-derived causal relations. (In my view, this norm is motivated by a mistaken capacities interpretation of theory.) But Biddle and Hamermesh (2016) report that whereas in the 1970s all microeconomic empirical papers in top-5 journals indeed exhibited a theoretical framework, in the 2000s there was some resurgence of non-theoretical studies. Citation numbers suggest

[9] It is true that there always has been much empirical work in economics, in fields such as agricultural and labor economics and in activities such as national accounting and cost-benefit analyses. Nevertheless, at a minimum, empirical work has become more prestigious (Cherrier 2016).

that the non-theoretical work is at least as influential. Angrist and Pischke (2010) also report the rise of non-theoretical practice in several subfields. There are other, more anecdotal indicators of an empirical turn too. One is that almost all recent Clark medalists have a strong empirical (albeit not non-theoretical) element to their work. Another is the rise of behavioral economics, often justified on the grounds of its greater fidelity to empirical psychology.

In my view, the empirical turn is a very positive development. I hope that it continues – for it needs to. After all, 19% of articles in the most prestigious journals are still purely theoretical, a large proportion of empirical work is still tied to testing a model derived from general theory, orthodox modelling is still considered the cornerstone of sound methodology (Rodrik 2015), and theorists still command a wage premium (Biddle & Hamermesh 2016). There is still a way to go.

7. The efficiency question
To recap: first, empirical success in economics is possible. Second, orthodox theory is not a good way to get it. Third, theory should instead be developed in close concert with empirical application and refinement, as is commonplace in other sciences. It might be that economics needs to become more of an idiographic than a nomothetic discipline. We will find out only by seeing what works, not by stipulation or wishful thinking.[10]

What is the alternative to the status quo? The answer is any mix of methods that leads more efficiently to empirical success. This is not just a matter of theory development being less abstract. In addition, once strict adherence to orthodoxy is dropped, economics may join other fields in taking advantage of a much wider range of empirical methods, generating results that in a virtuous circle then feed back into more theory development. These methods include: ethnographic observation; small-N causal inference, such as qualitative comparative analysis; other qualitative methods, such as questionnaires and interviews; causal process tracing; causal inference from observational statistics; machine learning from big data; historical studies; randomized controlled trials; laboratory experiments; and natural and quasi-experiments.[11] Each of these methods has its own strengths and weaknesses, but each is already widely practiced and has a developed and rigorous methodological literature. Turning to them is in no way a return to the fuzzy verbal analysis that is the pejorative memory of much pre-war economics. To ignore them is parochial, not to mention self-damaging.

What is the optimal balance between, on one hand, building up a library of orthodox rational choice models, and on the other hand, pursuing more contextual work and utilizing a wider range of empirical methods? Current practice is already a mixture of the two, so the question becomes: is it the right mixture? Call this the *efficiency question* (Northcott 2018). To answer it requires, so to speak, an epistemic cost-benefit analysis. The costs are the resources invested into theory, such as the training of students, and perhaps more notably the opportunity costs, such as fieldwork methods not taught and fieldwork not done. The benefits are all the cases where theory explains and predicts successfully. Similar calculations can be made for alternatives.

Of course, such calculations can only be done imperfectly. It is hard to add up explanations and predictions in an objective way, hard to weigh those versus other goals of science, and

[10] These lessons apply beyond economics: to some natural sciences, such as mathematical ecology (Sagoff 2016); to social sciences other than economics; and to heterodox economic approaches too. Thus, theory in, say, econophysics equally needs to earn its empirical keep.
[11] Of course, the latter few of these have begun to be co-opted by economics already.

hard also to evaluate the counterfactual of whether things would be better if resources were allocated differently.[12] But these calculations are *being done already* – implicitly, every time a researcher chooses, or a graduate school teaches, one method rather than another, or journals or prizes or hirers choose one paper or candidate rather than another.[13] The recent empirical turn is a large-scale example, because it is in effect a claim that resources were not being optimally apportioned before. The status quo is not inevitable. This is shown not just by the empirical turn but also by different practices in other social sciences. It is surely better to assess the matter explicitly than to leave it to inertia and sociological winds.

It might be objected that any choice here is illusory because orthodox theory and the various alternatives are too entangled to be separated. For example, a randomized trial might be testing a hypothesis derived from theory.[14] This objection is true up to a point – but not up to the point that the efficiency question can be wished away. The empirical turn itself, for example, shows that a substantive switch from theory to other methods is possible, entanglement notwithstanding.

8. Conclusion

The current emphasis on orthodox theory is inefficient. Common defenses are not persuasive: the generalizability offered by orthodox theory is illusory without empirical success, and so is the understanding in terms of agent rational choice. Empirical success comes first.

Acknowledgement

Many of the ideas in this article originated in past work with Anna Alexandrova.

[12] Two kinds of efficiency analysis are possible. The first is *global*: does the current overall allocation of resources serve economics well compared to a different allocation? This is the challenging one to assess. The second kind of efficiency analysis is *local*: what methods should be used to tackle a particular explanandum, and in what proportion? This is often much more tractable and many case studies are in part just such analyses already (e.g. Northcott & Alexandrova 2015).

[13] Of course, other factors enter these decisions too, such as what best serves one's own career. Nevertheless, an implicit efficiency analysis is certainly one important component.

[14] It is a truism that all empirical work assumes *some* 'theory' in the form of background assumptions. The issue here is whether these background assumptions must include those of economic orthodoxy.

References

Alexandrova, A. (2008). 'Making models count'. *Philosophy of Science* 75, 383-404.

Alexandrova A., and R. Northcott (2009). 'Progress in economics: Lessons from the spectrum auctions', in H. Kincaid & D. Ross (eds.), *The Oxford Handbook of Philosophy of Economics*, Oxford University Press, 306-337.

Angrist, J., and S. Pischke (2010). 'The credibility revolution in empirical economics: How better research design is taking the con out of econometrics', *Journal of Economic Perspectives* 24, 3-30.

Benhabib, J., and R. Farmer (2000). 'The monetary transmission mechanism', *Review of Economic Dynamics* 3, 523–550.

Betz, G. (2006). *Prediction or Prophecy?* (Wiesbaden: Deutscher Universitaets Verlag)

Biddle, J., and D. Hamermesh (2016). 'Theory and measurement: emergence, consolidation and erosion of a consensus', NBER Working Paper No. 22253.

Card, D., and A. Krueger (1994). 'Minimum wages and employment: A case study of the fast food industry in New Jersey and Pennsylvania', *American Economic Review* 84, 772–93.

Cartwright, N. (1989). *Nature's Capacities and Their Measurement*. Oxford.

Cherrier, B. (2016). 'Is there really an empirical turn in economics?', Institute for New Economic Thinking blog 29th September 2016 https://www.ineteconomics.org/perspectives/blog/is-there-really-an-empirical-turn-in-economics

Elster, J. (1988). 'The nature and scope of rational-choice explanation', in E. Ullmann-Margalit (ed.), *Science in Reflection*, Springer, 51-65.

Friedman, M., and A. Schwartz (1963). 'Money and business cycles', *Review of Economics and Statistics* 45, 32–64.

Gruene-Yanoff, T. (2009). 'Learning from minimal economic models', *Erkenntnis* 70, 81-99.

Guala, F. (2005). *Methodology of Experimental Economics.* Cambridge, England: Cambridge University Press.

Hamermesh, D. (2013). 'Six decades of top economics publishing: who and how?' *Journal of Economic Literature* 51, 162-172.

Hausman, D. (1992). *The Inexact and Separate Science of Economics.* Cambridge, England: Cambridge University Press.

Khalifa, K. (2012). 'Inaugurating understanding or repackaging explanation?' *Philosophy of Science* 79, 15-37.

Lawson, T. (1997). *Economics and Reality*. Routledge

Mäki, U. (1992). 'On the method of isolation in economics', *Poznan Studies in the Philosophy of the Sciences and the Humanities* 26, 319-354.

Mill, J. S. (1843). *A System of Logic.* London: Parker.

Northcott, R. (2013). 'Degree of explanation', *Synthese* 190, 3087-3105.

Northcott, R. (2015). 'Opinion polling and election predictions', *Philosophy of Science* 82, 1260-1271.

Northcott, R. (2018). 'The efficiency question in economics', *Philosophy of Science* 85, 1140-1151.

Northcott, R., and A. Alexandrova (2013). 'It's just a feeling: why economic models do not explain', *Journal of Economic Methodology* 20, 262-267.

Northcott, R., and A. Alexandrova (2015). 'Prisoner's Dilemma doesn't explain much', in M. Peterson (ed.) *The Prisoner's Dilemma* (Cambridge), 64-84.

de Regt, H. (2017). *Understanding Scientific Understanding*. New York: Oxford University Press.

Reiss, J. (2008). *Error in Economics: Towards a More Evidence-Based Methodology*. Routledge.

Rodrik, D. (2015). *Economics Rules: the rights and wrongs of the dismal science*. Norton.

Rosenberg, A. (1992). *Economics: Mathematical Politics or Science of Diminishing Returns?* Chicago.

Sagoff, M. (2016). 'Are there general causal forces in ecology?' *Synthese* 193, 3003-3024.

Strevens, M. (2013). 'No understanding without explanation', *Studies in History and Philosophy of Science* 44, 510–515.

Trout, J. (2007). 'The psychology of scientific explanation', *Philosophy Compass* 2, 564–591.

Ylikoski, P. (2019). 'Mechanism-based theorizing and generalization from case studies', *Studies in the History and Philosophy of Science* 78, 14-22.