# ERASMUS JOURNAL FOR PHILOSOPHY AND ECONOMICS
## VOLUME 8, ISSUE 1, SPRING 2015


The Erasmus Journal for Philosophy and Economics (EJPE) is a peer-reviewed bi-annual academic journal supported by the Erasmus Institute for Philosophy and Economics, Faculty of Philosophy, Erasmus University Rotterdam. EJPE publishes research on methodology of economics, history of economic thought, ethics and economics, and the conceptual analysis of inter-disciplinary work relating economics to other fields. EJPE is an open-access journal. For additional information, see our website: <http://ejpe.org>. All submissions should be sent via e-mail to: <editors@ejpe.org>




http://ejpe.org/pdf/8-1-2015.pdf ISSN: 1876-9098

ERASMUS JOURNAL FOR PHILOSOPHY AND ECONOMICS
VOLUME 8, ISSUE 1, SPRING 2015

TABLE OF CONTENTS

**PHD THESIS SUMMARIES**

*Erasmus Journal for Philosophy and Economics,*
*Volume 8, Issue 1,*
*Spring 2015, pp. 1-23.*
*http://ejpe.org/pdf/8-1-art-1.pdf*

# Equality of resources, risk, and the ideal market

LARS LINDBLOM
*Umeå University*

**Abstract:** Ronald Dworkin's theory of equality of resources makes extensive use of markets. I show that all these markets rely on one specific neoclassical conception of the ideal market in full equilibrium, as analyzed by Debreu. This market must be understood as operating under circumstances of certainty, and this is incompatible with several components of Dworkin's account. In particular, it does not allow one to hold people responsible for their option luck, and it implies a high social safety net rather than insurance schemes for addressing brute luck. I conclude by outlining an interpretation of equality of resources that takes the ideal market seriously.

**Keywords:** The market, risk, Dworkin, distributive justice, hypothetical insurance, equality of resources, luck egalitarianism

**JEL Classification:** A12, B21, D41, D80, D49

When Ronald Dworkin introduced his theory of *equality of resources* back in 1981, his first claim was that "an equal division of resources *presupposes* an economic market of some form, mainly as an analytical device but also, to a certain extent, as an actual political institution" (Dworkin 2000, 66, my emphasis).[1] He then went on to distinguish his use of the market from the two standard ways in which the market has been justified in the debate on economic justice: as an engine of efficiency or as a guarantor of liberty. Dworkin took a different approach: "[T]he idea of an economic market, as a device for setting prices for a vast variety of goods and services, must be at the center of any attractive theoretical development of equality of resources"

[1] Dworkin's classic 1981 papers were reprinted as part of his 2000 book, *Sovereign virtue*, which elaborated his account into a full-fledged theory. I shall follow convention by referring to this book throughout.

**AUTHOR'S NOTE:** I wish to thank Sara Belfrage, Kalle Grill, Sven Ove Hansson, Niklas Möller, and Philippe van Parijs for penetrating and helpful comments.

(Dworkin 2000, 66). Over the following thirty years Dworkin continued to refine his theory, but the market remained a fundamental aspect of his thought on equality. In *Justice for hedgehogs*, published in 2011, he stated that “[a] free market is not equality’s enemy, as is often supposed, but indispensable to genuine equality” (Dworkin 2011, 357). For Dworkin, then, it is necessary to make use of the notion of the market in order to explicate the ideal of equality. The subject of this paper is Dworkin’s idea that the market is essential to the normative ideal of equality.

What I shall do is to investigate what the assumptions that Dworkin makes about the market imply for other parts of his theory of equality. I will argue, in section 2, that the market at the core of equality of resources is a neo-classical market in full equilibrium, as analysed by Gérard Debreu. Moreover, I will show that this market must be understood as operating under circumstances of certainty. In section 3, I will begin to spell out the implications of this interpretation of the market for Dworkin’s theory of equality. Starting from Dworkin’s notion that a just distribution of resources mimics the distribution produced by an ideal market, I will show that the idea of option luck, the normative axiom of individual responsibility for choices under risk usually considered a corner stone of Dworkin’s theory, is incompatible with the goal of mimicking such a market. There are no choices under risk when there is complete certainty. The last sections of this article develop the implications of the argument that the goal of justice should be to mimic an ideal market *under certainty*. In section 4, I argue that this implies a rather high social safety net. Section 5 discusses Dworkin’s theoretical solution to the question of how much to redistribute to persons who have been unlucky in the natural lottery. I argue that the thought experiment of insurance purchases behind the veil of ignorance becomes problematic, since there can be no insurance in a market under certainty.

What emerges from this analysis of Dworkin’s theory is a new reading of the idea of equality of resources. This interpretation, which may be called the certainty interpretation of equality of resources, is more faithful to the basic motivations of the theory than Dworkin’s own version, and retains its most attractive features while alleviating the arguably callous aspects that have been criticized by, for instance, Elizabeth Anderson (1999).

The aim of this paper is constructive rather than critical. I accept the assumptions that Dworkin makes and do not present any changes to his idea of equality of resources other than those that follow from taking the ideal market seriously. In order to show that these implications follow, I will have to engage in some exegetical work. This paper, then, follows a different strategy from well-known critical contributions by Joseph Heath (2004), John Bennett (1985), and Colin McLeod (1998), which all try to poke holes in Dworkin's ideal market foundations. I do not dispute that there may be problems that have yet to be solved within ideal market accounts. Moreover, I do not in any way dispute that there has been tremendous progress made in economics since Debreu wrote in the late fifties (see Camerer 2003; Bowles 2004). My goal is to investigate the account of the market that Dworkin has placed at the core of equality of resources. My focus is the question of what the ideal market actually commits us to if we accept the Dworkinian framework, and I will try to show that the implications of equality of resources are both different and in many ways more attractive than they appear at first sight.

## II. DWORKIN'S MARKETS

Dworkin's idea of equality of resources has many attractive features. It provides an interpretation and synthesis of two compelling principles: that it is equally important that each life goes well, and that persons are responsible for their ambitions. It presents a plausible conception of equality in the sense of equal treatment, while, as G. A. Cohen famously pointed out, incorporating the enemy's most dangerous weapon by making a place for responsibility within egalitarian theory (Cohen 1989, 933). By modelling justice on the workings of the ideal market under fair starting conditions, it also connects equality with efficiency. When Dworkin's market has done its work the end result is a distribution of resources that is just, that follows from the operation of egalitarian principles, and an economy in socially optimal equilibrium. What is there not to like?

In this section, I will outline the core of Dworkin's theory and try to show how the different markets it uses relate and then proceed to argue that the central market is the neo-classical market as systematised in the groundbreaking works of Gérard Debreu (1959), and of Kenneth Arrow and Frank Hahn (1971). This is the kind of market that is modelled in what Dworkin calls the 'pre-auction', and which I argue should be understood as taking place under conditions that imply that the parties face choices under certainty and not risk. I shall try to show that this is a view to which Dworkin commits himself, and demonstrate how his emphasis in later writings on the importance of the *ex ante* perspective, and hence choice under risk, is inconsistent with some core ideals of equality of resources.

The central device that Dworkin uses to explain equality of resources is a specific thought experiment, which starts in the aftermath of a shipwreck. The thought experiment is first developed in the 1981 articles which become chapters 2 and 3 of *Sovereign virtue* (Dworkin 2000; see also Dworkin 2002; 2004; 2006; 2011). The survivors come ashore on a deserted island and find it, luckily, full of resources. They all agree that no one has a prior claim to any of these resources and decide to hold an auction in order to distribute them fairly to each. An auctioneer is appointed to be responsible for finding the fairest way to price the resources on the island. This job is done in two stages. At the first stage, which is called the pre-auction, the auctioneer is assumed to have perfect foresight and is thus able to predict perfectly the resulting distribution of resources on an ideal market based on the preferences of the immigrants (Dworkin 2000, 155-158). This pre-auction is the fundamental market of Dworkin's theory. It provides the standard against which all other markets are to be evaluated and corrected. I will refer to this market as the ideal market or the pre-auction depending on context.

At the second stage, we come upon slightly more realistic circumstances. The auctioneer creates property rights, with the aim of mimicking under more realistic circumstances the distribution of resources produced by the ideal market of the pre-auction. This means that the outcome of the ideal market is the criterion of just distribution in equality of resources. To achieve her goal the auctioneer must take into account issues, such as externalities, coordination failures, and transaction costs that impede the working of less than ideal markets. This is why the auctioneer "imagines a purer, pre-auction auction in which the participants have perfect knowledge and predictive power, and in which there are no organizational costs" (Dworkin 2000, 158).

The auctioneer uses two principles to come up with the correct bundles to auction off. The principle of abstraction, which insists that

> an ideal distribution is possible only when people are legally free to act as they wish except so far as constraints on their freedom are

> necessary to protect security of person or property, or to correct certain imperfections in markets (Dworkin 2000, 148).

It is meant to establish "a strong presumption in favor of freedom of choice" (Dworkin 2000, 148). And the principle of correction, which says that "[c]onstraints on freedom of choice are required and justified [...] if they improve the degree to which equality of resources secures its goal, which is to achieve a genuinely equal distribution measured by true opportunity costs" (Dworkin 2000, 157). These true opportunity costs are identified by the ideal market—the value of what each person has is its value to others. This principle corrects the imperfections of real world markets alluded to in the principle of abstraction by putting restrictions on how resources may be used.[2] We identify these imperfections by comparing the ideal market (the pre-auction) to more realistic markets. Note that this means that how this ideal market is defined becomes paramount for understanding the implications of Dworkin's theory.

The ideal market is thus fundamental to equality of resources. Dworkin aims to show that the market is indispensable to the theory of justice, and this is the account of the market he places at the centre of his theory of justice. All other markets of the theory are to be corrected if they result in outcomes that diverge from the outcomes generated by the ideal market. Moreover, as we have seen, Dworkin starts out from two motivating principles, one of which is that each person is responsible for his or her ambitions. Now, if the outcomes of a market are the result of externalities or uncertainty, they would not be fully ambition-sensitive, but rather endowment-sensitive, in the sense that they result from the person's situation rather than from his or her ambitions.[3] This is a further reason for why the ideal market is fundamental to equality of resources.

---

[2] The example that Dworkin uses to explain how the principle of correction works involves zoning regulations concerning the architectural style in which houses can be built in a given town or area (2000, 156), but it seems that the implications are rather wide since the auctioneer is supposed to solve all kinds of coordination problems. For example, the principle of correction would seem to mandate taking some goods out of the market altogether, such as those services provided by the military, the judiciary, and the police.

[3] The prices on the ideal market will also be decided by other things than the individual's ambitions, namely other people's ambitions. This might make price levels seem like a kind of externality. However, Dworkin argues that we should not think of

Returning to the shipwreck thought experiment, each so-called immigrant survivor gets an equal amount of clamshells to use as markers in the auction. The immigrants can then bid for the resources they individually prefer, with their whole future lives in mind. The immigrants use the 'envy test' to check that equality is preserved. This notion was developed by economists searching for an alternative to the standard Pareto criterion for use in welfare economics (see Foley 1967; Varian 1974) and says that a justified division is achieved when no one prefers anybody else's bundle of goods to her own. The test is satisfied by the clamshell procedure because if a person envies (i.e., prefers) what someone else has he or she is free to bid for it. When the market has cleared, no one will want to exchange their bundle for anyone else's. In this way every person bears the true opportunity cost of his or her choice of lifestyle, while equality is preserved. Thus, we get a theory that incorporates equality, efficiency, and the value of people taking responsibility for their ends. The resulting allocation satisfies the envy test and is consistent with the optimality thesis of general equilibrium analysis.

The ideal market of the pre-auction, then, is crucial to Dworkin's undertaking. How does it work? We are told to think of the ideal market in the manner outlined by Debreu:

> I mean to describe a Walrasian auction in which all productive resources are sold. I do not assume that the immigrants enter into complete forward contingent claims contracts, but only that markets will remain open and will clear in a Walrasian fashion once the auction of productive resources is completed. I make all the assumptions about production and preferences made in G. Debreu, *Theory of Value* (New Haven: Yale University Press, 1959). In fact the auction I describe here will become more complex in virtue of a tax scheme discussed later (Dworkin 2000, 478, fn2).

Before turning to the assumptions of Debreu's model, let me just point out that the tax scheme mentioned above is based upon the hypothetical insurance markets that will be discussed in later sections. Debreu analyzes a market where resources are defined by their physical attributes, their location, and the time of delivery. This is not quite the everyday concept of resources, but it is close enough for the purposes of

---

them in terms on luck. Prices indicate true opportunity costs and should be thought of as information (see Dworkin 2000, 68-69).

equality of resources. Agents, either in their role as producers or as consumers, are rational in the sense that they conform to what comes down to two conditions: *complete preordering* (the relation of non-strict preference holds over all prospects) and *continuity* (there is a continuous relationship between physical characteristics and desirability) (Harsanyi 1977). Markets are conceived of as taking place at a point in time where all goods, future and present, are sold (the market clears). As the definition of resources includes the time of delivery, there is no use for any further markets. All possible transactions will have been made. From his assumptions, Debreu proves that, so specified, the market will reach an optimal equilibrium.

This market has several features that would seem attractive from an ethical perspective. It satisfies the Pareto principle and the envy test.[4] Moreover, since there are full property rights for all goods, there are no externalities. The assumptions also make strategic interaction impossible. And, of course, it fits into Dworkin's more general theory of political morality.

There are no probabilities in this model; agents do not face decisions under risk. Instead, the ideal market should be understood as taking place under circumstances of certainty. In fact, the reason that Arrow and Hahn—in that other classic statement of neoclassical general equilibrium analysis *General competitive analysis* (1971)—re-introduce the Walrasian auctioneer is to make this point about full information. This neo-classical auctioneer will be recognized from Dworkin's island. On the subject of risk, Debreu says, in the introduction to *Theory of value*, that he assumes that it is not necessary to point out the limitations of the model:

> One may stress here the certainty assumption made, at the level of interpretations, throughout the analysis of Chapters 2 to 6, according to which every producer knows his future production possibilities and every consumer knows his future consumption possibilities (and his future resources if resources are privately owned—otherwise only the future *total* resources need be known). This strong assumption is weakened, albeit insufficiently, in the last chapter (Debreu 1959, xi).

[4] Outcomes that satisfy the envy test must also satisfy the Pareto principle. The reason is that if there is no allocation of goods that anyone would prefer, it is also the case that there is no way to make anyone better off without making anyone else worse off.

This quotation raises an interpretative question about Dworkin. Is he following the Debreu of chapters 2 through 6, in which the analysis proceeds under the assumption of choice under certainty, or is he working out his thought experiment on the basis of the slightly more realistic chapter 7, where risk is introduced? I suggest that we should interpret Dworkin according to the first alternative. The textual basis for this interpretation is that Dworkin points out that he does not assume that there are contingent claims contracts, and chapter 7 of Debreu's book deals with such contracts. There are also theoretical reasons. The importance of leaving contingent contracts out of the core of equality of resources can be brought out by considering how Arrow and Hahn present their way of incorporating risk into general equilibrium theory: "Each commodity now must be interpreted as a contingent claim […] a promise to supply one unit of commodity *i* if state of the world *s* occurs and nothing otherwise" (Arrow and Hahn 1971, 124). Debreu puts the same point in the following way:

> A contract for the transfer of a commodity now specifies, in addition to its physical properties, its location and its date, an event on the occurrence of which the transfer is conditional. This new definition of a commodity allows one to obtain a theory of uncertainty free from any probability concept and formally identical with the theory of certainty developed in the preceding chapters (Debreu 1959, 98).

This is an elegant move that brings risk into the analysis by redefining goods. Debreu's initial definition of resources seemed suitable for a conception of equality that takes the perspective of equality as measured over whole lives. But now we have a new definition of commodities, where what is traded are contingent claims. On this new interpretation of the theory, goods are no longer concrete resources, but rather lotteries. The metric of justice becomes inherently risky. It seems clear that Dworkin should prefer the certainty interpretation. The reason for this is that if Dworkin accepted the move from concrete resources to this kind of commodity, the equalisandum of equality of resources would change from resources, understood in the commonsensical way, to lotteries. This would be a whole different theory.

There is, however, some reason to suspect that Dworkin would not agree with this. In a more recent article he says: "True equal concern requires *ex ante*, not *ex post*, equality" (Dworkin 2002, 124). This implies that people should be held responsible for how they approach risky choices as a fundamental feature of equality of resources, and it suggests that Dworkin would prefer to include risk in the model of the ideal market. In other words, this would amount to a risk interpretation of the ideal market.

To see why such an inclusion of risk is problematic for strictly Dworkinian reasons, we must ask why we ought to take resources, and not welfare for instance, as the equalisandum of justice. Dworkin argues that the notion of fair shares of resources is needed to make sense of the ideal of equality, which is shown by the failures that he finds in welfare conceptions of egalitarianism (Dworkin 2000, chapter 1). For instance, we do not want to compensate people for having frustrated preferences for unfairness. Furthermore, with resources as our metric we can hold people responsible for their choices in an appropriate way, for instance in expensive taste cases. A third reason is brought out in this quotation:

> People make their choices, about what sort of a life to lead, against a background of assumptions about the rough type and quantity of resources they will have available with which to lead different sorts of lives. They take that background into account in deciding how much of what kind of experience or personal relationship or achievement of one sort must be sacrificed for experiences or relationships or achievements of another (Dworkin 2000, 28-29).

This implies that having at least a rough idea of what resources will be under one's command is a precondition for the responsible agency demanded by his theory of justice. Dworkin argues that welfarist theories cannot take this point into account. The distribution of resources in society will be a function of the distribution of preferences in society. If preferences change, then so must the distribution. People need some sense of security with regards to what resources will be available to them in order to make decisions on how to live their lives and consequently what choices to make and what preferences to form. But if this is true then the risk interpretation of the ideal market turns out to distribute the wrong things. If the auctioneer is going around distributing lottery tickets then it will be as difficult under equality of resources as under equality of welfare to plan one's life. If such uncertainty counts against equality of welfare it must also count against the *ex ante* approach. On the basis of Dworkin's own argument, then, we should prefer the *ex post* approach, since this gives us a more appropriate equalisandum. In other words, for Dworkinian reasons we should prefer the certainty interpretation of the ideal market.

To be clear, the question we are investigating now is *what* a theory of justice should distribute, not *how* it ought to distribute it.[5] But including risk in our description of the market, through the definition of resources, means that there can be some disturbing inequalities in resources, *independently of any choice*. Such inequalities will not depend on differences in people's ambitions, but on the luck of the draw. Including risk in the equalisandum severs the connection between responsibility and resources.

These two different approaches to risk and certainty suggest the following distinction between theories of equality: *equality of resources* is concerned with the distribution of concrete things (such as oranges). A theory of *equality of commodities* would concern the distribution of contingent claims (such as the probability of having an orange on Tuesday in Brussels). It is fairly evident that these theories will evaluate distributions rather differently. Equality of commodities has some unappealing consequences. In terms of a decision tree, equality of commodities is in principle consistent with all of your actual resources being situated on branches that are never instantiated. Or if we put this in terms of an Edgeworth box, achieving an envy-free allocation of commodities from equal starting points is consistent with the actual distribution of, say, apples and oranges being at the origin of one of the players.[6] Or, finally, in everyday terms, perfect equality of commodities is consistent with perfect inequality of resources. Equality of commodities is both inegalitarian, since these differences in allocation do not depend on choice but on what we have chosen as the equalisandum of the theory of justice, and in violation of the demand that people should have a secure sense of what resources they will have available when planning their lives. It would undermine Dworkin's theory of equality to incorporate risk into the ideal of the market in this way.

[5] And note that what we are discussing here is the fundamental criterion of justice and not the issue of how to implement justice. It would be quite unrealistic to think that we could avoid uncertainty when we are trying to achieve justice in the world, but this does not mean that we should valorise uncertainty at the level of fundamental principle. I return to the issue of implementation in the later part of section III.

[6] For a quick explanation of Edgeworth boxes, see Bowles 2004, 209-210; and for decision trees, see Peterson 2009, 17-19.

To sum up, the way Dworkin introduces and explains the ideal market suggests that it ought to be understood as taking place under certainty. Furthermore, if risk is included in the model of the market this has two unfortunate consequences for equality of resources. First of all, it changes the equalisandum of the theory from actual resources to Debreu's commodities. Second, if one takes this step to commodities, one allows inequalities in terms of resources that are not the result of people's choices or ambitions, since, as we have seen, this kind of commodity is in fact a lottery ticket. A move to lottery tickets would mean that Dworkin must give up on ambition-sensitivity. For these reasons, I submit that if Dworkin wants to use the ideal market as characterized in Debreu's analysis, then he is, or ought to be, committed to model the market on choices under certainty, and not under risk.[7] In the following sections, I will draw out the implications of this certainty interpretation of Dworkin's theory.

## III. OPTION LUCK IN THE IDEAL MARKET

We turn now from the question of what should be distributed, to how these things should be distributed. One of the most path-breaking aspects of equality of resources is that it introduced a way for egalitarian theories to take risk into account, through the concept of option luck. This is the part of Dworkin's theory where risk is actually introduced as a consideration. To see how, let us return to the island thought experiment.

After the auction, the immigrants get on with their lives, and further trade ensues. Here we come upon real—or at least realistic—markets. This is then the third type of market in equality of resources, which is, of course, constrained by the two previous markets. As time goes on, some people will have worked more than others, some will have fallen sick, some will have been lucky in business, or unlucky, and some will have had accidents. As a result, the envy test will no longer be satisfied. Since the goal of equality of resources is that the distribution of resources is ambition-sensitive, but not endowment-sensitive, the

[7] Another conclusion one could draw is that neoclassical economics is not the best way to model what Dworkin wants to capture; perhaps an account of the market that takes transaction costs into consideration would be more suitable. It would certainly be more realistic, but it would almost as certainly be less compelling. Why would justice demand that we take the lack of knowledge or the existence of strategic action into consideration at the level of principle? These phenomena make the world, if anything, worse, rather than better.

inequalities resulting from such contingencies must be rectified before the market can achieve a fully justified envy-free distribution of resources. This is why the hypothetical insurance scheme is needed. In fact there are several such schemes; there are schemes for general health, unemployment, and handicap insurances as well as a scheme that deals with the taxation of inheritance. I will not discuss the inheritance tax insurance scheme, as it seems to me that the inequalities that result from differences in bequest are best understood as externalities, and as such they fall directly under the domain of the principle of correction. Letting inherited wealth affect the price mechanism means that the prices will be distorted in comparison with the baseline, where each person has equivalent purchasing power, in the form of an equal amount of clamshells. If justice demands that we mimic the distribution of an ideal market, and that this distribution should be endowment-insensitive, we cannot let these kinds of endowments play a role. I will deal with the special issues that handicap insurance raises in a later section, and health and unemployment insurance in this section.

It is here that Dworkin invokes the famous distinction between brute luck and option luck. Option luck has to do with how deliberate gambles turn out: for instance if a person decides to play the stock market and loses his or her money—or, for that matter, gets very rich. Brute luck concerns unforeseeable events, such as being hit by a meteorite, or in the case of brute good luck, stumbling upon a lost treasure. The idea is that the effects of brute luck ought to be rectified, but that it is fair that people are held responsible for their choices, and hence for their option luck. Insurance provides a bridge between brute and option luck. If it is available, then, Dworkin says, brute luck is converted into option luck, since the choice of whether to insure or not is a deliberate gamble in the right way.[8] Thus, the envy test can still be satisfied *ex ante*. Someone who prefers to play it safe can buy insurance, while risk-takers can choose to go without a safety net. After a person has received his or her equal endowment, he or she is fully responsible for the outcomes that ensue. Ambition-sensitivity and responsibility for option luck seem to overlap.

The immigrants, in possession of equal shares of resources, are asked to consider a situation where the risk distribution of something

[8] But see Michael Otsuka 2004 for important qualifications.

bad, for instance serious disease, remains as it is, but where they are placed behind a veil of ignorance, so that they do not know if they will be unlucky or not. Next, they are asked what level of insurance they would buy against illness when the veil is lifted. In line with this choice, their resources are then taxed and redistributed to those who qualify for payments. This scheme has several virtues, Dworkin claims, such as that it gives reasonably precise answers to the question of how much should be redistributed and that it permits rational trade-offs between health and other goals in life. The immigrants will, if rational, not spend all their resources on insurance, because they will have to be concerned with having a decent life after making their insurance payments. The scheme also takes into account the costs of running the system, in that the model incorporates the demand for profits under which insurance firms would operate. The implication of the last two points is that the envy test can only be approximately satisfied, because the envy at the start of trading will not be eradicated by the transfers of the insurance scheme, but it will, Dworkin argues, come close enough for the purposes of equality of resources.[9]

There is much to admire about this attempt to handle risk within an egalitarian theory. However, I shall argue that the notion of option luck cannot play a fundamental role in equality of resources since it is incompatible with the characteristics of an ideal market. If justice demands that we mimic the ideal market, then holding people responsible for their choices in the manner indicated by the argument from option luck lacks a fundamental motivation in the theory of

[9] It is not always clear if the redistribution effected by the insurance market should be understood as taking place before the pre-auction, which would then ensure that each immigrant had equal resources at the start of the auction, or if we are to think of insurance as patching up the market after the fact. As can be seen from the way I formulate this query, I prefer the first alternative. I have the feeling that Dworkin has the second way of thinking in mind; the recurrent use of the term 'compensation' would indicate this. In recent work, the insurance market seems to have become even more central to Dworkin's position, whereas other markets have dropped from view. In Dworkin's "Sovereign virtue revisited" (2002) and *Is democracy possible here?* (2006), one gets the impression that the only demand that the theory presents is that there should be a social insurance system that mimics the ideal insurance market, and that one need not ensure that people start out with roughly equal shares. This does not seem to me to be a fully egalitarian position and perhaps a name like 'the compensation view of resources', rather than equality of resources, would be more appropriate. I shall therefore continue to interpret Dworkin as requiring a role for equality in all markets, not just the insurance market. At any rate, this seems to me a much more attractive view.

equality of resources. Ambition-sensitivity and option luck do not overlap. Consider how Dworkin introduces the concept of option luck:

> Option luck is a matter of how deliberate and calculated gambles turn out—whether someone gains or loses through accepting an isolated risk he or she should have anticipated and might have declined. Brute luck is a matter of how risks fall out that are not in that sense deliberate gambles (Dworkin 2000, 73).

The idea of option luck, then, is characterized by analogy with gambling. There are several possible actions that you choose between based on the odds and the values you attach to the different states of the world that might result. Gambling is decision-making under risk, but, as we have seen, in the ideal market people face decisions under certainty. There can be no such deliberate gambles where there is no risk. And if there are no such gambles, then there is no room for the application of the concept of option luck. Obviously, there are options—one can choose between apples and oranges, say—but these have nothing to do with luck. If the ideal of the market is to be retained, then the idea of option luck cannot be fundamental to the theory of equality of resources. Full ambition-sensitivity can only be achieved by dropping option luck. The point of Dworkin's theory is to hold people accountable for their ambitions, not to hold them responsible for having to choose under conditions of risk. He wants to capture choices concerning what we want to do with our lives: if we prefer to work hard or enjoy our leisure, or if we want to, in Kenneth Arrow's well-known terms, drink pre-phylloxera claret and eat plovers' eggs, rather than fulfil more mundane tastes (Arrow 1973, 254). Such ambition sensitive distinctions remain valid, even if we find that the idea of option luck should not play a role in equality of resources. If the ideal market is indispensable, then we cannot dispense with it here.

This is where the distinction between the pre-auction, the auction and actual markets makes itself useful. The pre-auction is the core market of equality of resources, and it is here that we should not let risk play any role. There are no gambles in the pre-auction for later markets to mimic. On the other hand, it is obvious that we encounter risk and uncertainty in our daily lives, and that at least sometimes we feel that people should be held accountable for how they approach risky situations. The auctioneer, using the principle of correction, must try to devise a market that mimics the ideal market as closely as possible.

Never holding people in real markets responsible for how they approach risk would be an inefficient way of achieving this. Moreover, in real markets there will be, and should be, insurance. The argument that there is no uncertainty in the ideal market does not mean that there will not be risky purchases in actual markets. But from the idea that when designing *institutions* it is reasonable, for efficiency reasons, to incorporate a concern for responsibility for risky choices, it does not follow that such a responsibility should be thought of as a fundamental aspect of the normative theory. The pre-auction takes place under certainty, but when the auctioneer has developed the set of property rights to auction off in the actual market, as it were, then we can take something resembling the *ex ante* approach.

To return to the lottery metaphor, in realistic settings there will be lotteries, but their prizes should be such that the outcome mimics the distribution on the ideal market. The pre-auction decides the prize structure. Furthermore, note that saying that risk is not fundamental to equality of resources is not the same as saying that responsibility for one's preferences cannot be fundamental to equality of resources. Responsibility for ambitions is fundamental for equality of resources, but responsibility for risk is not. If one finds oneself facing a decision under risk, one finds oneself with a deficit of that highly valuable resource: knowledge concerning how the world will turn out.[10]

How about brute luck? This concept is also defined in a way that seems to include risk. Does this mean that it must also be dropped from equality of resources? It does not seem so because it is, I believe, better to interpret the term risk in the definition of brute luck as unwelcome outcomes, rather than in the way that term is understood in the theory of decision under risk. The concept of brute luck does not concern the choices you face, but rather the fortune that faces you. Even under certainty there will be outcomes that are tragic; uncertainty and unfortunate outcomes are distinct phenomena. That it is certain that you will become ill does not make it any less of a misfortune. Perhaps brute (mis)fortune would be a better term than brute luck, but the concept can be applied regardless of whether certainty or risk is at issue.

[10] Not only does having full information help with preference satisfaction, it also makes strategic interaction impossible and internalizes externalities.

To see the implications of this certainty interpretation of equality of resources and of dropping option luck, it is instructive to turn to the charge of callousness made by Elizabeth Anderson (1999). The core of her criticism can be brought forth by considering three of her best known counterarguments to luck-egalitarian theories like equality of resources. *The abandonment of negligent victims* points out that it is an implication of equality of resources that ambulances should pass by the uninsured victims of accidents. Since only bad brute luck should be compensated, it follows that the results of bad option luck are of no concern to justice. A related problem is *the abandonment of the prudent*. This is where a person makes every reasonable choice, but has bad luck repeatedly and ends up unable to pay for insurance. These problems present theoretical difficulties for equality of resources, even if it is hard to think of realistic scenarios where they would appear.

If, however, I am correct in saying that equality of resources is best understood under the certainty interpretation, then these counterarguments are blocked. In both abandonment problems the trouble starts from taking or having to take risks—and then insisting on holding people responsible for their option luck. But the idea of option luck does not make sense as a fundamental aspect of equality of resources, since there is no room for the application of that concept on the ideal market. If we are to mimic the ideal market, then we cannot let option luck influence the distribution of resources. This means that these two charges cannot be made to stick against the certainty interpretation—a further reason to prefer this interpretation of equality of resources.

There is, however, a third aspect of Anderson's charge of callousness that cannot be answered in this way. This is the problem of *the lack of a safety net*. Even if option luck is dropped from the fundamental level of the theory, there is still no guaranteed minimum outcome. It might still be the case that a person makes choices that lead to destitution, and then it would seem that justice according to equality of resources would demand that this person is left destitute. If we find a theory of equality that allows for some people having absolutely nothing unattractive, then we will feel that little ground has been gained by excluding option luck. In the next section we will turn to the question of a social safety net.

## IV. RATIONALITY AND A SOCIAL SAFETY NET IN THE IDEAL MARKET

I believe that under the certainty interpretation the notion of mimicking the market gives us an answer to Anderson's safety net argument. To see how the fact that the ideal market operates under certainty does this, we have to look into the different implications of the assumptions of choice under certainty and under risk.

Imagine a situation where Adrian can choose between two alternatives. In A he is guaranteed €5, and in B he can either end up with €100 or nothing, with a probability of 0.1 and 0.9 respectively. Assuming a linear relationship between utility and money, and risk neutrality, the expected utility of option A is 5 and for B it is 10. A rational person would then choose B, as this affords the highest level of expected utility. This seems to imply that if we are trying to mimic the market we should see to it that option B is implemented. Most of the time, then, justice would demand that Adrian receives nothing. This, however, is not the choice that Adrian would face in the neo-classical world described in Debreu's model of a market with perfect foresight; there are no probabilities in that world. The theory of rational choice under certainty says that if one knows with certainty what will happen, then one should chose the option with the outcome that one prefers the most. Nine times out of ten (not the best way of putting it, but the charitable reader will understand what I mean) he would stand before these options: A = 5 and B = 0. In these cases, mimicking the market would mean implementing A and not B.

A rational individual such as Adrian would not choose B unless A is no longer a live option. This example presents two considerations in favour of a social safety net. First, equality of resources says that a justified distribution is the result of trade from equal starting points in an ideal market; and second, it is clearly a violation of Pareto efficiency to choose to make oneself worse off than the baseline. Any transaction that left the traders worse off than the baseline of equality would not take place and therefore should not be permitted by real world institutions attempting to mimic the operations of the ideal market.

Adrian's decision at the auction would obviously be much more important than the choice we just discussed since he is making a one-off choice that will affect the rest of his life. But his choice will be based on the same principle. He has been given an equal amount of clamshells, and, since he does not face any choice under risk, he is in fact guaranteed the corresponding level of resources in terms of the prices set at the auction. Any trade that he makes will have to be an improvement. When the auctioneer is applying the principle of correction she will have to take this baseline as a floor that limits the level of inequality.

An assumption of Dworkin's is that the parties to the auction trade with the goal of making the whole of their lives as valuable for themselves as possible. Under certainty, this seems to require that they will see to it that they have a sufficient level of resources for food, shelter, health care, and so on throughout their lives. "When I am 36 I will not need food" would clearly be irrational if one planned to live to 37 or longer. A rational individual choosing under certainty, and endowed with an equal share of purchasing power, would plan his or her life so as to have a sufficient level of resources at any time during his or her life.

The argument makes some mild assumptions about the content of peoples' preferences: they care about their futures. But these are the same assumptions that are needed to get the 'right' answers in Dworkin's original argument. If a person is prudent enough to buy hypothetical insurance then he or she will choose to get the same level of protection in a choice under certainty. However, note that it is not irrational to not care about the future. It is difficult to think that a person would spend all his or her resources in youth and then spend fifty years in poverty if this choice were made with full information, including information about the effects of poverty. But, of course, rationality in itself does not rule out such preferences. If Dworkin's assumptions about how people care about their own futures are wrong, then equality of resources, even under the certainty interpretation, would not require a high safety net. However, similar assumptions are shared by all theories that give preferences this kind of role when it comes to risk; it is not a problem specific to the certainty interpretation. The aim of this paper is to draw out the implications of Dworkin's theory rather than to develop an external critique. Therefore, the argument will continue under the assumption that Dworkin is right about how people would make these kinds of choices.

Dworkin's principle of correction, then, implies that the auctioneer should set up property rights so that the immigrants never fall below the level of sufficient resources. Furthermore, it would take unwelcome outcomes out of the market if they are of a type that no one would rationally choose. For example, the auctioneer will have to design property rights that leave each immigrant with guaranteed health care and sufficient resources if unemployed. If the result of trade on the ideal market is that people would have access to resources of this kind over their whole lives, then this is something that the auctioneer must replicate when designing property rights.

The social minimum will therefore be rather high.[11] The distribution will be ambition-sensitive; there is room for the trade-off between work and leisure, or between beer and pre-phylloxera wine. The parties will have to take into consideration the true opportunity costs of their choices. The one thing that the distribution cannot vary with is ignorance.

This argument does the same work as the insurance schemes for health care and employment do and therefore these two insurance schemes are not needed in the theory anymore. The attractive features of equality, responsibility, and efficiency are retained in a version of equality of resources that drops the ideas of option luck and these types of insurance. We have yet, however, to discuss the hypothetical insurance market that is devoted to the issue of handicaps.

## V. INSURANCE, HANDICAPS, AND THE IDEAL MARKET

In the previous two sections, I have made two points. First, that option luck is inconsistent with the idea of mimicking the ideal market, and second, that several of Dworkin's insurance thought experiments are *unnecessary* under the certainty interpretation of equality of resources. In this section, I will make a related but distinct claim, namely that the notion of insurance itself is *incompatible* with the ideal market.

The story of the immigrants has up until now assumed equality of personal resources, i.e., equal talents and no handicaps. However, in Dworkin's scheme insurance does more than bring personal responsibility for risk into equality of resources; it also redistributes. It provides a way of fixing inequality of *personal* resources, such as physical and mental capacities. In order to achieve full equality of resources, some process of justice must equalise the combined value of personal and *impersonal* resources. Dworkin argues that a hypothetical insurance market could be used to derive answers to the question of what redistribution is required for true equality. Note that this is a

[11] The exact amount will, however, vary according to the size of a society's total social product. It will be higher in rich countries and lower in less well off societies.

further use of the ideal market in Dworkin's theory, which means that this insurance market must be compatible with the characteristics of the ideal market. This hypothetical insurance market is designed to redistribute impersonal resources so that we make up for any inequality in personal resources. Even if insurance is superfluous with regards to the problems we discussed above, it could be necessary for dealing with the problem of inequality in personal resources.

However, there is a problem. Insurance and the ideal market do not go very well together. As Ronald Coase pointed out: "when there are no costs of making transactions, it costs nothing to speed them up, so that eternity can be experienced in a split second" (Coase 1988, 15). The importance of this point, for our purposes, becomes more evident when we consider a quotation from Stigler on the previous page in Coase's book: "The world of zero transaction costs turns out to be as strange as the physical world would be without friction. Monopolies would be compensated to act like competitors, and insurance companies would not exist" (Stigler 1972, 12; see also Stigler 1966). There cannot be any insurance if the future is fully known. You will never be able to insure your car if it is certain that you will crash. The idea of insurance lacks any foundation if there is no risk. We cannot consistently use the ideal market to argue for the hypothetical insurance scheme as a fundamental feature of equality of resources. If justice demands that we mimic the ideal market of the pre-auction, then we cannot, at the same time, mimic a market where there is insurance, and hence uncertainty. This thought experiment is inconsistent with how an ideal market works.

We should not confuse the uncertainty that is introduced by the veil of ignorance with the kind of uncertainty needed to get the notion of insurance off the ground. The veil introduces risk concerning your identity in order to define an impartial—fair—standpoint between ourselves and others; insurance concerns individual vulnerability to future contingencies. Just introducing the veil of ignorance does not lead us to a coherent idea of insurance.[12] To see why, consider two

[12] It should probably be noted that it is even controversial whether the veil of ignorance works as Dworkin supposes. John Roemer argues that the parties behind the veil would allocate more resources to those who are most able to use them efficiently in the pursuit of utility, and so the healthy would get all the money (Roemer 1996, chapter 7). Marc Fleurbaey (2002) makes a similar point. Dworkin, however, has replied that this kind of argument relies on an account of the motivation of the parties that is far from self-evidently true. People care not only about the expected aggregate level of utility, but also about their own futures.

different choices behind a veil of ignorance. In the first case you know the following: Peter and Paul will both get $100, but you do not know if you are Peter or Paul. In the other case you are either Penny or Paula, but you do not know which. Penny will either get $100 or nothing, and Paula will get either $20 or $80. In both cases, there are veils of ignorance, but only in the second case are there contingencies. Therefore, contingencies and the veil of ignorance are distinct.

Clearly, the auctioneer could design a system of handicap insurance for the auction, but this would not make the idea of insurance any more fundamental to equality of resources than zoning restrictions are. The pre-auction would still be the standard that we should use to evaluate the workings of this insurance market. Dworkin's argument is that an insurance market could, in principle, solve the problem of how to redistribute in order to compensate for brute luck, but it seems to me that it is the market, rather than insurance, that is essential to his solution. More precisely, what is essential is that the auctioneer finds a way of measuring the size of the personal inequality deficit, so that she can redistribute impersonal resources in a way that achieves net equality.

The ideal market is essential for Dworkin's theory, whereas the insurance model is not. The hypothetical insurance scheme is best understood, I believe, as *a tool* for measuring the equality deficit brought about by the initial inequality of personal resources. It puts a price on personal resources, and redistributes impersonal resources until rough net equality ensues. It sets this price by asking the immigrants how much of their impersonal resources they would be willing to forgo in order to avoid ending up in an unwelcome situation. This understanding of the insurance mechanism—as a tool that helps us set a price on inequalities of personal resources—is still open to us under the certainty interpretation.[13] But, if Dworkin's hypothetical insurance markets are best understood as tools, then we should ask if there are better tools available. Again we have come to the conclusion that the aspects of equality of resources that have to do with risk are best understood as means to achieve the equal division of goods that would be the outcome of an ideal market.

[13] Understood this way, the debate between Roemer and Dworkin is not about whether the mechanism is utilitarian or not, but about the correct value of the inequality deficit.

## CONCLUDING REMARKS

My purpose in this paper has been to show that if we take the ideal market seriously, the implications of equality of resources are different and perhaps more attractive than is usually thought. The argument presented has the following form: according to Dworkin, mimicking the ideal market from equal starting points is fair. In this market, as described by neoclassical economics, trade takes place under full information, which rules out choices under risk. Therefore, there can be no such thing as option luck in the ideal market. Consequently, one cannot and should not hold people responsible for option luck in equality of resources. Moreover, mimicking this market implies that the principle of correction will direct us to set up a social safety net. Given Dworkin's assumptions about the motivations of rational agents, it follows that rational agents choosing under certainty would make sure to have access to enough resources at each point in time to carry out their life plans. This makes several of the hypothetical insurance schemes unnecessary to equality of resources. In relation to inequality in personal resources, I have also argued that the idea of insurance is incompatible with the conditions of certainty that define the ideal market.

This certainty interpretation of equality of resources would not be vulnerable to various well-known criticisms from callousness. In addition, the revised theory is more true to the core tenets of Dworkin's equality of resources than his original formulation—especially the central notion that the market should play a fundamental role in the theory of equality. It is based on a more thorough and consistent understanding of the ideal market, and still makes room for a difference between endowments and ambitions. The market continues to be crucial to the theory of equality of resources, but it no longer allows risk and option luck to play roles in deciding what is just.

## REFERENCES

Anderson, Elizabeth S. 1999. What is the point of equality? *Ethics*, 109 (2): 287-337.

Arrow, Kenneth. 1973. Some ordinalist-utilitarian notes on Rawls's *Theory of justice*. *Journal of Philosophy*, 70 (9): 245-263.

Arrow, Kenneth J., and Frank H. Hahn. 1971. *General competitive analysis*. San Francisco: Holden-Day Publisher.

Bennett, John. 1985. Ethics and markets. *Philosophy and Public Affairs*, 14 (2): 195-204.

Bowles, Samuel. 2004. *Microeconomics: behavior, institutions and evolution*. Princeton: Princeton University Press.

Camerer, Colin. 2003. *Behavioral game theory: experiments in strategic interaction*. Princeton: Princeton University Press.

Coase, Ronald H. 1988. *The firm, the market, and the law*. Chicago: The University of Chicago Press.

Cohen, G. A. 1989. On the currency of egalitarian justice. *Ethics*, 99 (4): 906-944.

Debreu, Gérard. 1959. *Theory of value*. New Haven: Yale University Press.

Dworkin, Ronald. 2000. *Sovereign virtue: the theory and practice of equality*. Cambridge (MA): Harvard University Press.

Dworkin, Ronald. 2002. Sovereign virtue revisited. *Ethics*, 113 (1): 106-43.

Dworkin, Ronald. 2004. Reply to critics. In *Dworkin and his critics*, ed. Justine Burley. Malden (MA): Blackwell Publishing, 339-395.

Dworkin Ronald. 2006. *Is democracy possible here?* New Jersey: Princeton University Press.

Dworkin, Ronald. 2011. *Justice for hedgehogs*. Cambridge (MA): The Belknap Press.

Fleurbaey, Marc. 2002. Equality of resources revisited. *Ethics*, 113 (1): 82-105.

Foley, Duncan K. 1967. Resource allocation and the public sector. *Yale Economic Essays*, 7 (1): 45-98

Harsanyi, John. 1977. *Rational behavior and bargaining equilibrium in games and social situations*. New York: Cambridge University Press.

Heath, Joseph. 2004. Dworkin's auction. *Politics, Philosophy and Economics*, 3 (3): 313-335.

MacLeod, Colin M. 1998. *Liberalism, justice, and markets: a critique of liberal equality*. New York: Oxford University Press.

Otsuka, Michael. 2004. Equality, ambition, and insurance. *Proceedings of the Aristotelian Society (Supplement)*, 78 (1): 151-166.

Peterson, Martin. 2009. *An introduction to decision theory*. Cambridge: Cambridge University Press.

Roemer, John E. 1996. *Theories of distributive justice*. Cambridge: Harvard University Press.

Stigler, George J. 1966. *The theory of price*, 3rd ed. New York: Macmillan.

Stigler, George J. 1972. The law and economics of public policy: a plea to the scholars. *Journal of Legal Studies*, 1 (1): 1-12.

Varian, Hal R. 1974. Equity, envy, and efficiency. *Journal of Economic Theory*, 9 (1): 63-91.

**Lars Lindblom** is a senior lecturer at Umeå University. He works on issues in the intersection between political philosophy and economics. He has written on topics such as invisible hand explanations in Nozick's political philosophy, incomplete contract theory and consent in business ethics, and responsibility-catering prioritarianism and risk management.

Contact e-mail: <lars.lindblom@umu.se>

*Erasmus Journal for Philosophy and Economics,*
*Volume 8, Issue 1,*
*Spring 2015, pp. 24-41.*
*http://ejpe.org/pdf/8-1-art-2.pdf*

# Resolving the small improvement argument: a defense of the axiom of completeness

JACK ANDERSON
*University of Utah*

**Abstract:** This paper defends the axiom of completeness against a particular incomparabilist objection, the small improvement argument (or SIA). In my view, a theory of choice must admit of a number of folk psychological assumptions, most importantly, that agents conceive of choice options as simplified possible worlds and have preferences between such worlds. In addition, this paper argues that an additional folk psychological assumption allows a trimodal theory of choice to satisfactorily address the concerns about preference-indifference intransitivity raised by the SIA. This additional claim is that agents resolve their consideration of choice options to varying degrees. In my view, the SIA can be answered without abandoning or modifying the axiom of completeness.

**Keywords:** axiom of completeness, comparability, folk psychology, preference, utility theory

**JEL Classification:** B40, B41, D80

This paper defends an assumption in utility theory, specifically the assumption that agents either hold a strict preference over two options, or that they are indifferent between them. I will call this sort of comparison 'trimodal'. This assumption of comparability will be defended against a particular sort of objection, the small improvement argument (SIA), perhaps most famously presented by Ronald De Sousa as the problem of the 'fairly virtuous wife' (1974, 544).[1] The fairly virtuous wife appears to be indifferent "between keeping her virtue for nothing and losing it in Cayucos for $1,000" (1974, 545). The fairly virtuous wife, however, also appears to be indifferent between keeping

[1] The term "small improvement argument" is from Ruth Chang (2002).

AUTHOR'S NOTE: I am grateful to Chrisoula Andreou, Elijah Millgram, and Mariam Thalos for their comments on earlier drafts. I would also like to thank the anonymous referees and the *EJPE* editors for their helpful critiques.

her virtue and losing it for $1,500, which presents a problem for utility theorists. For them “indifference, like preference, in terms of which it is defined, is a transitive relation” (De Sousa 1974, 545), and the wife’s rankings are a case of preference-indifference intransitivity.[2] While it is a failure of the assumption of transitivity that brings the problem into focus, De Sousa holds that the options presented to the fairly virtuous wife are actually incomparable. In general, philosophers have shared this interpretation of the choice problem. Joseph Raz, for example, refers to failures of transitivity as “the mark of incommensurability” (1985, 120).

The structure of the SIA exposes an inconsistency between the assumption of completeness and the assumption of transitivity. If the wife’s deliberative stances—by which I mean her attitudes about the two options (absent any particular theoretical account of choice being applied to those attitudes)—are understood as preference rankings, then those rankings are intransitive and leave the wife vulnerable to a money pump. The response to that apparent inconsistency between completeness and transitivity advanced in this paper involves the claim that agents may consider an option in more or less detail depending on what that option is being compared to. I will argue that the objection to comparability illustrated by the SIA can be answered without abandoning a trimodal approach to explaining choice, provided that the approach also assumes that agents are able to resolve[3] choice options at finer or coarser grains—which is to say, that the number of details considered when assessing a choice option and, importantly, the precision with which agents consider those details may vary.[4] I will also

[2] See Gustafsson and Espinoza’s “Conflicting reasons in the small-improvement argument” (2010) for a detailed account of how preference-indifference intransitivity allows for a money pump type problems to arise.

[3] This use of “resolution” is similar to the manner it is employed by Nien-He Hsieh in the paper, “Equality, clumpiness, and incomparability” (2005). Both Hsieh and I argue that the resolution at which options are compared will vary. However, there are significant differences between Hsieh’s conception and the one I will be suggesting. For Hsieh, the variation in resolution occurs because the “covering considerations” with respect to which the options are assessed are themselves clumpy (Hsieh 2005, 181). For example, one grading scale might clump student papers into As, Bs, Cs, etc., while another, more fine grained grading scale, might clump papers in to As, A-minuses, Bs, etc. And, Hsieh understands “comparison to be distinct from choice” (2005, 199). In contrast, I examine the role resolution might play in a utility theoretic explanation of choice, an explanation which does not necessarily involve the notion of covering considerations at all. In my account, resolution is a fundamental feature of how agents mentally represent choice options as opposed to a feature of certain types of covering considerations.

[4] This claim depends on the notion that there is a large degree of variability in terms of what an agent might believe about choice outcomes, i.e., it is a response that depends

argue that the costs to agents of making comparisons will vary depending on the resolution at which the comparisons are made. For example, the representation of an outcome as "I receive a bag of oranges" is less finely resolved than the representation of an outcome as "I receive a bag containing 11 oranges"; a fortiori, generating that less finely resolved representation is less costly as I do not have to count the oranges.

I will explain the process of resolving in the context of an axiomatic, subjective, folk psychological theory of rational choice, and will provide an account of that utility theory below. However, this paper is not meant to provide a *tout court* defense of comparability but, rather, a response to a very specific sort of objection particular to the SIA. And, as many examples of the SIA, like De Sousa's, conflate that specific sort of objection with various other objections to comparability, I first want to isolate the particular problem I mean to solve.

## THE PARTICULAR PROBLEM POSED BY THE SIA

In examples such as De Sousa's, at least part of the reason for focusing the objection on the assumption of comparability—rather than on the assumption of transitivity—is the idea that the two options are "qualitatively different" (De Sousa 1974, 545). Sinnot-Armstrong illustrates preference-indifference intransitivity using choices between death and amounts of pain, and the problem is often illustrated via choice situations between various sorts of careers, such as the choice between becoming a lawyer or a clarinetist (Raz 1985, 126). However, examples such as these, which involve such qualitatively different options, actually conflate two separate sorts of objections to the notion of comparability. The first sort of objection is simply that such

on making adjustments to the choice options which then account for apparent cases of preference-indifference intransitivity. John Broome is dubious of "refining the individuation of outcomes" in this fashion. He states that, "if this sort of individuation is always allowed, transitivity will truly be an empty condition" (1991, 101). However, while I do claim that the notion of resolution does eliminate the apparent inconsistency between the assumptions of comparability and transitivity illustrated by the SIA, I do not claim that all instances of intransitive preferences can be eliminated in this way. I do not, for example, dispute that perceptual thresholds can result in the sorts of intransitive preferences described by W. S. Quinn in "The puzzle of the self-torturer" (1990), and, the notion of resolution as presented here does not leave transitivity as "an empty condition".

qualitative differences necessarily render certain options incomparable.[5] In De Sousa's presentation of the SIA, for example, the force of this objection stems from the intuition that virtue simply cannot be priced in dollars. The second sort of objection, the sort particular to the SIA, stems from the intuition that the wife's deliberative stances are plausible and reasonable. In what follows, I will be concerned with answering the second sort of objection rather than the first—this is for two reasons: (1) because the second objection applies to a much wider range of choice situations (among them are the sorts of choice situations routinely addressed by economists); (2) because it is this second sort objection that actually arises from the structure of the SIA (whereas in the first sort of objection the structure of the SIA is just a consequence of the prior intuition—that some options are evaluatively different and that such differences make trimodal comparisons impossible). Ruth Chang (2002) presents the SIA as a choice between tea and coffee, where the agent is supposed to be indifferent both between a cup of tea and a cup of coffee, and between a slightly improved cup of tea and the same cup of coffee. This example, which I will examine in some detail below, shows that the problem of preference-indifference intransitivity arises not just in choice situations that involve hard choices between things like virtue and money (or death and pain), but even in the simplest situations involving choices between what Chang calls "mere market goods" (2002, 96).

Again, one might object that it is impossible to compare things when the options are qualitatively different. One can quite sensibly take the position that, in certain hard cases, the assumption of comparability is prima facie (or for any number of reasons[6]) false, and that things like virtue cannot be compared to things like money. But, one need not *begin* with the intuition that virtue and money are somehow inherently incomparable to note that the wife's three deliberative stances are, taken together, intuitively sensible. Even the trimodal comparabilist that is absolutely convinced that there is no such thing as qualitatively

[5] Or, at least that such options cannot be compared trimodally. For an account of how the existence of such evaluatively different options might be compared using a tetramodal comparative approach, see Chang 2002.

[6] For example, one might be convinced by an argument from the diversity of values—that "some items are 'so different' that there is no 'common basis' on which a comparison can proceed" (Chang, 2002, 72). Or, one might be convinced by the claim that certain options are constitutively incomparable, where the constitutive features of certain options prevent those options from being compared in certain cases (Raz, 1986). Again, however, replies to these objections are available (see Chang, 2002).

different options, or that such qualitative differences simply have no effect on an agent's ability to compare, still faces the problem illustrated by the SIA if that same comparabilist nonetheless intuits that deliberative stances like the virtuous wife's are plausible and reasonable.

## A TRIMODAL THEORY OF CHOICE

That the force of the SIA is intuitive is significant. The account of rational choice advanced here should be understood as what Alexander Rosenberg describes as "folk psychology formalized" (2008, 80).[7] It assumes that agents not only choose, but that choices are motivated by an internal preference set which is both complete and transitive. Such an account is vulnerable to objections which appeal to intuitions given that the process of formalization needs to account for intuitive judgments about the nature of agents' mental states. If it seems at least plausible that the fairly virtuous wife has the deliberative stances that she does and, at the same time, that she is rational, then the SIA presents a problem for a trimodal theory of choice which also assumes that a rational agent's preference rankings must be transitive.

Understood in the context of a folk psychological account of choice, the assumption of comparability involves claims about agents' capacities. And, per the utility theoretic account of choice forwarded here, agents' choices are entirely motivated by their preferences, where 'preference' is understood as an agent's all-inclusive, subjective judgment about which of two options she wishes to consume.[8] So, I will defend a trichotomy of choice where the agent either prefers A to B, or prefers B to A, or is indifferent between them (where indifference can be understood as the agent being willing to say "you choose", i.e., the agent is willing to accept either option on offer). These three modes are derived from utility theory's axiom of completeness: for *all* $X_1$ and $X_2$ in

[7] Revealed preference theorists will object to this approach, but I regard the arguments forwarded by Daniel Hausman (2012) and Alexander Rosenberg (2008) as to why economics must understood as having to do with the mental states of agents as convincing.

[8] My use of "preference" is similar to Hausman's, who defines "preferences" as the agent's "total comparative subjective evaluations" (2012, 34). That the judgments are entirely subjective is significant and means that there is a distinction in the choice theory between objects of preference and the outcomes which will actually obtain. The theory holds that a preference for a mental representation of some actual outcome motivates the agent to choose the actual outcome represented. The theory is silent both in regard to *how* the internal mental state of preferring motivates actual choice and in regard to how mental representations come to be *about* actual outcomes.

**X**, either $X_1 \succcurlyeq X_2$ or $X_2 \succcurlyeq X_1$, where **X** is the consumption set, and $X_1$ and $X_2$ consumption options within that set. Expressions of strong preference, for example "the agent prefers A to B", formally represented as $(A \succ B)$, are derived from pairs of weak preference relationships, $((A \succcurlyeq B) \& \sim(B \succcurlyeq A))$. Again, in this choice theory 'preference' is understood as the mental state which motivates an agent to choose, and whatever motivates preference is, as per the usual economic approach, exogenous to the theory. So, the virtuous wife's preferences can be given as follows:

virtue ≈ \$1000

virtue ≈ \$1500

\$1500 ≻ \$1000

As noted above, these preferences are problematic because they violate utility theory's axiom of transitivity: for any three elements in the consumption set **X**: $X_1$, $X_2$, $X_3$, if $X_1 \succcurlyeq X_2$, and $X_2 \succcurlyeq X_3$, then $X_1 \succcurlyeq X_3$.

In addition to the axioms of completeness and transitivity, this folk psychological conception of rational choice involves another assumption about agents' capacities, one rarely formally recognized. It is usually omitted that the options compared by agents are not the actual outcomes that obtain.[9] These options are, rather, mental representations, which I will refer to as "simplified possible worlds".[10] Call the assumption that agents mentally represent choice outcomes as simplified possible worlds the "axiom of subjectivity": **X** is the consumption set of simplified possible worlds as conceived of by the choosing agent.

I further assume that the simplified possible world that an agent conceives of as representative of some particular actual outcome can vary in resolution. This last assumption suggests that the fairly virtuous wife's preferences given above are composed of comparisons made at varying resolutions, and that the failure of transitivity appears to be a consequence of comparing differently resolved options—a comparison the virtuous wife herself never actually makes.

[9] Of course, expected utility theory explicitly involves agent beliefs, but I mean to point out that choice options must be considered in this manner even when agents are unconcerned with assigning probabilities to various outcomes.

[10] These options are sometimes called states of affairs. Matthew Adler uses the term "simplified possible worlds" (2012, 514) to refer to the choice options faced by choosing agents, and that term seems apt.

To reiterate, these assumptions, that agents prefer, that agents mentally represent options, are to be understood as a folk psychological in nature. Theoretical terms such as 'prefer' are, therefore, "definable functionally, by reference to their causal roles" (Lewis 1972, 207). Though, I do not hold that the functional roles such terms play in a folk psychological interpretation of utility theory actually reflect the ordinary folk understanding of those terms. Rather, the theory being deployed here supposes that the functional roles described by theoretical terms ("to prefer" and "mentally represent") are roles agents are actually capable of performing. Agents believe (mentally represent, somehow) things about alternative outcomes. Given beliefs about alternatives, agents are able to weakly prefer (or not) a mental representation (a collection of beliefs about some alternative) of some outcome to another. The epistemological justification for assuming completeness is that it is possible for agents to actually think in the manner described by the axiom. Agents can weakly prefer one option to another, and weakly preferring is a thing agents do in their heads by comparing "alternatives they *believe* to be available" (Hausman 2012, 15; emphasis added).

## RESPONDING TO THE SIA

There are two distinct argumentative lines of the SIA: a practical line and an abstract one. The abstract line is meant to present the objection without allowing for replies which simply posit agent error, as such replies are, arguably, sufficient responses to practical examples of the SIA. However, by abstracting completely away from any actual choice situation, the abstract line of the argument also loses quite a bit of intuitive force. I will show that the abstract line depends, not on the intuition that the deliberative stances presented in the SIA are plausible and reasonable, but rather, on the prior intuition that certain options are, for whatever reason, incomparable. Without this prior intuition, some actual choice options must be posited in order to get any sort of objection off the ground. I will proceed by explaining how a trimodal comparabilist might respond to the abstract line of the SIA. I then show how the capacity to resolve any given choice outcome with varying degrees of precision and detail allows agents to rationally navigate the difficulties presented by the practical line without abandoning a trimodal approach to choice.

***The abstract line of the SIA***

The abstract line of the SIA rests upon the claim that certain sorts of small improvements simply cannot make a difference to an agent's preferences. Ruth Chang presents the argument quite clearly. Though the trichotomy that she considers ('better', 'worse', 'equally good') departs from the trichotomy of preference that I am interested in defending, the distinction makes no difference in terms of the comparabilist response. It will be helpful to consider her presentation of the abstract line of the SIA in some detail.

Chang describes the abstract intuition as "in general, for evaluatively very different sorts of items, certain small improvements—given by a dollar, a pleasurable tingle, and so on—cannot effect a switch from an item's being worse than another to its being better" (2002, 128). She accurately notes that,

> […] if this intuition is correct, then it follows that the trichotomy of relations sometimes fails to hold. For take an arbitrary pair (r, s) of evaluatively diverse items. We can create a spectrum of r-items by successively adding or subtracting dollars (or pleasurable tingles, etc.) from r. If we add enough dollars, we get an r-item, r+, that is better than s, and if we subtract enough dollars, we get an r-item, r-, that is worse than s. Now, according to our abstract intuition, adding a dollar, pleasurable tingle, etc., cannot make a difference to whether one item is better or worse than another item evaluatively different from it. Therefore, there must be some r-item, r*, in the spectrum that is neither better nor worse than s. But what relation holds between r* and s? Suppose one of the trichotomy [for our purposes that the agent either prefers r to s, or vice versa, or is indifferent] always holds. Then since r* is neither better nor worse than s, it and s must be equally good [the agent must be indifferent between them]. According to our intuition that a dollar cannot make a difference, however, this is impossible. For if we add fifty cents to r*, we get an item that is better than s; if we take away fifty cents from r*, we get an item that is worse than s. And the difference between r*-plus fifty cents, which is better than s, and r*-minus fifty cents, which is worse than s, is a dollar. Thus r* and s cannot be equally good (Chang 2002, 128).

The main issue that I want to address here is the notion of "our abstract intuition" and that intuition's role in the subsequent development of the abstract line of the SIA. Chang's presentation is quite precise. If one begins with the assumption that qualitatively different options exist, and that "certain small improvements cannot [in choice situations involving such qualitatively different options] effect a switch from" one preference to another, then, as Chang shows, intransitive preference rankings such as the sort exhibited by the virtuous wife can be shown to follow as a consequence of that initial assumption. Formally, in the abstract line as given by Chang, an agent can be shown to have the following preferences that conform to the usual, problematic, SIA pattern:

r* ≈ s

(r*- plus fifty cents) ≈ s

(r*- plus fifty cents) > r*

Again, these preferences are intransitive and the agent having such preferences can be subject to a money pump. Chang's solution to this problem is to question the propriety of classifying the relationships between s and r* and between s and r*- plus fifty cents as 'indifference' in the usual utility theoretic fashion.[11] But, as her presentation of the abstract version of the SIA suggests, there is an alternative, straightforward, response available to the comparabilist presented with the abstract line of the SIA—namely, to reject the foundational abstract intuition. Without the abstract intuition, that "certain small improvements" cannot effect a comparative difference between options which are qualitatively different, there is no particular reason to regard the above abstract preference rankings of r*, s, and r*-plus fifty cents as plausible; therefore the abstract line can simply be put aside.

Of course, the abstract intuition is abstracted from somewhere, and in Chang's presentation of the abstract line of the SIA, it is developed through examples of the usual sorts of hard choices which are often assumed to be the clearest examples of qualitatively different options: "a career in hang-gliding and one in accounting, a Sunday afternoon in the amusement park and one at home with a book, a zero-tolerance policy towards crime and one that aims only to reduce homicides, and so on" (2002, 128). Choice theorists differ on how convincing such examples are in terms of establishing the notion of 'qualitatively different'; incomparabilists may assert that it is impossible to price virtue in dollars, while comparabilists may assert that it is quite possible and that so-called qualitatively different options can be compared in the

[11] Chang suggests a fourth comparative relationship—"on par" (2002).

same fashion as comparisons between mere market goods, such as tea and coffee. However, what is really significant about the SIA is that even if the comparabilist dismissal of the possibility of qualitatively different options which make for hard choices stands, and all choices are ultimately as simple as the choice between tea and coffee, the argument *still* presents a trenchant objection to trimodal accounts of choice. To show exactly how that objection runs, and how I think the notion of resolution answers it, I will now consider one final instantiation of the SIA, the practical example proposed by Chang of a choice between a cup of tea and a cup of coffee.

***The practical line of the SIA***

Hopefully, the structure of the practical line of the SIA is at this point familiar. It consists of three plausible deliberative stances, all held by a single agent. Those deliberative stances are often presented and meant to be understood as outside the context of any particular theoretical description, as the SIA is meant to present evaluative judgments to which the standard trimodal descriptions do not apply. However, as noted above, the force of the SIA can be demonstrated quite simply by applying a trimodal theoretical description to the plausible evaluative judgments and then proceeding to illustrate exactly how such a description entails a failure of transitivity. For example, one might propose that Abby the agent has the following preferences:

(i) tea $\approx$ coffee

(ii) tea+ $\approx$ coffee, where tea+ is a slightly improved version of tea

(iii) tea+ $>$ tea

Again, per utility theory, each of the following axioms applies:

**Axiom of subjectivity**: **X** is the consumption set consisting of $X_1$, $X_2$ …, $X_n$, where any $X_i$ is some simplified possible world as mentally represented by the agent.

**Axiom of completeness:** for *all* $X_1$ and $X_2$ in **X**, either $X_1 \geqslant X_2$ or $X_2 \geqslant X_1$

**Axiom of transitivity:** for any three elements in the consumption set **X**: $X_1$, $X_2$, $X_3$, if $X_1 \geqslant X_2$, and $X_2 \geqslant X_3$, then $X_1 \geqslant X_3$

The preference relationships given in (i), (ii), and (iii) are problematic for this axiomatic theory because, if those relationships are as described, then the axiom of transitivity fails to hold. And, the "tea or coffee" example constructed by Chang illustrates two important features of the SIA. First, as noted above, the objection clearly applies to choice situations involving mere market goods, and the problem clearly applies to a very wide array of choice situations. Second, the alternatives on offer are immediately and fully comprehensible, unlike De Sousa's (or any other example which involves a large amount of uncertainty, such as a choice between a career as a lawyer and a career as a clarinetist). Whereas the fairly virtuous wife might reasonably be thought to be facing a choice situation best explained with an expected utility model, the "coffee or tea" problem does not seem to involve anything other than a straightforward trade-off between two choice options that can be known with as much certainty as anything can.

Interestingly, it also seems quite reasonable that Abby *is* actually indifferent (willing to say, "you choose") between the two options in the cases where she does not express a strict preference for one over the other. Abby not caring whether she gets tea or coffee seems plausible. However, Abby not caring whether she keeps her virtue or gets $1000 seems somewhat less plausible. This points to a problem with examples from De Sousa, Raz, and Sinnott-Armstrong that attempt to present practical situations which are meant to strongly invoke incomparable intuitions prior to any consideration of an agent's preferences. Such examples involve momentous choices. From a practical perspective, an agent being genuinely indifferent between such significant options seems suspect unless one intuits that, for example, the fairly virtuous wife when presented with the choice between either $1000 or $1500 and her virtue is content to say to her suitor, "you choose". While such a conclusion is certainly possible, it seems so unlikely that it invites practical dismissals of the problem, most obviously that the fairly virtuous wife's lack of preference for either the money or her virtue does not indicate any sort of fixed deliberative stance at all, but rather that she is still thinking about it. The trimodal comparabilist can simply admit that a trimodal description of the wife's deliberative stances does not apply because the wife has not actually reached any evaluative judgments. Abby's preferences, in contrast, do not invite this sort of dismissal, and, nonetheless, they exhibit preference-indifference intransitivity. The practical line of the SIA is, I think, best illustrated with mundane choices. This is not to say that examples of the SIA that involve hard choices cannot be understood as manifesting the particular objection that I am concerned with answering here (that the agent's preference rankings appear plausible yet intransitive); rather, such examples may conflate various incomparabilist objections, and such hard case examples of the SIA permit the objection to be evaded rather than addressed. That said, the response to the SIA suggested here will work just as well in such hard case examples, provided that the objection is understood as the objection arising from the structure of the SIA. Again, if one comes to such hard case examples of the SIA already intuitively convinced that certain options simply cannot be compared, no answer to *that* intuition is offered here.[12]

The apparent inconsistency between Abby's preferences and the axiom of transitivity can be clearly seen if Abby's preferences are described slightly more formally:

i) (tea $\geqslant$ coffee) & (coffee $\geqslant$ tea)

ii) (tea+ $\geqslant$ coffee) & (coffee $\geqslant$ tea+)

iii) (tea+ $\geqslant$ tea) & ~(tea $\geqslant$ tea+)

And now, in violation of the axiom of transitivity, it is plain that, while (tea $\geqslant$ coffee, from (i)), and (coffee $\geqslant$ tea+, from (ii)), it is also the case that (~(tea $\geqslant$ tea+), from (iii)). Nonetheless, it seems very reasonable that, if Abby is indifferent between tea and coffee, then she would be indifferent, as well, between tea+ and coffee. Given that failures of transitivity are more difficult to explain than failures of completeness,[13] even when prior incomparabilist intuitions are put aside, the problem exposed by the SIA still suggests that either Abby's preferences are not, in fact, complete, or that the meaning of 'completeness' is not, exactly, as described by the axiom of completeness. Given the nature of Abby's preferences, and given that they seem perfectly sensible, the problem is often regarded, as it is by Chang, as a problem with the notion of

[12] Those interested in such replies can find a multitude of them in Ruth Chang's *Making comparisons count* (2002), where she argues that the SIA is, in effect, the last objection to trimodal comparability left standing.

[13] There are a number of alternatives on offer that might allow a rational agent to choose without referencing a complete preference set (or, indeed, without preferring at all, see, for example, Chan 2010), or that propose that the notion of completeness be adjusted (see, for example, Chang 2002).

indifference. In such accounts, Abby's perfectly sensible preference relationships which (i) and (ii) attempt to describe, are not instances of indifference between the options therein considered, but rather, some other type of comparative relationship or the absence of any comparative relationship whatsoever.

By contrast, the account proposed here suggests that Abby's preference relationships can be explained by a trimodal theory of choice. That theory must assume that agents have the capacity to resolve choice options in various ways. Given the capacity to represent outcomes as simplified possible worlds, resolution is, I think, a capacity agents quite obviously possess. The SIA simply shows that it is a capacity that must also be theoretically acknowledged. Once it is acknowledged, the objection raised by the SIA can be answered in a straightforward manner.

To illustrate how the process of resolution works, I will include resolutions with Abby's preferences.

i) tea ≈ coffee (at resolution alpha)

ii) tea+ ≈ coffee (at resolution alpha)

iii) tea+ > tea (at resolution beta)

In the first choice problem (i) Abby must decide between tea and coffee. Abby considers her options at resolution alpha, and she is indifferent between the two options. In the second choice problem (ii) Abby must decide between tea+ and coffee. Again, Abby considers her options at resolution alpha, and she is indifferent. In the third choice problem (iii) Abby must decide between tea and tea+. In this case, Abby considers her options at a different resolution, beta, at which she notes the superior aroma of tea+ as compared to tea. Abby prefers tea+. But, she is considering the simplified possible world that will result if she picks tea+ differently in case (iii) than she does in case (ii). Abby's preference rankings will, to some extent, vary depending on the resolution Abby uses to consider her choices.

The question of why Abby considers case (iii) at a different resolution than cases (i) and (ii) admits of a straightforward and intuitive answer. It is less costly to compare two types of tea than it is to compare tea with coffee, so smaller differences can be taken into account in pursuit of smaller benefits. As incomparabilists tend to raise objections to the axiom of completeness precisely because of this intuition—that some comparisons are more difficult than others—I do not think it needs much defending here. However, in this instantiation of the SIA, the explanation might be that the options in (i) and (ii) are considered relatively crudely by Abby as "a cup of tea" or "a cup of coffee", with no attention being paid to finely grained details (such as the very slightly superior aroma of tea+), because the costs of resolving the choice problem in a manner that takes such small details into account exceed the benefits Abby might reasonably expect to get by noticing them. In (iii), the items under consideration are fundamentally the same, tea. This circumstance lowers the costs of considering such small details. This low cost makes it more likely that Abby will use a fine resolution as she can expect to receive benefits that exceed her comparison costs. Comparing tea+ to tea is less costly than it is to compare tea+ to coffee because Abby can take advantage of the fundamental sameness to avoid the costs associated with generating a simplified possible world populated with details about tea entirely. There is no need for Abby to consider how tea tastes compared to tea+, for example, as they taste the same. The only comparison Abby actually makes in (iii) is to note that tea+ offers a '+' and tea does not. In effect, Abby simply disregards everything about the two options that is the same, and chooses between what is left. Her choice in (iii) amounts to the choice between the 'aroma improvement' (the '+') or 'nothing'. Even though Abby is using a more fine-grained resolution in (iii) in order to be able to consider the improvement, the costs of comparing in (iii) are still, I think, likely to be far lower than in (i) and (ii), as there are far fewer details that Abby needs to include in the simplified possible worlds she compares in (iii).[14] In general, any change to an agent's mental representation of an outcome can be considered a matter of resolution. A simplified possible world which includes the sort of office chair that a career as a lawyer would have me sitting in is more finely resolved than the simplified possible world that just has me sitting in some chair, and the simplified possible world which includes details

[14] The reader will have noticed that, throughout this paragraph, I have been discussing the choice problem as a choice between actual things in the world, tea and coffee, rather than between simplified possible worlds. This is purely a matter of grammatical convenience. As always, the choice options are more accurately described by the, admittedly cumbersome, "the simplified possible world that the agent believes will result if…" construction.

about how sitting in that particular chair might actually feel is more finely resolved still.

By my account, *at any particular resolution* Abby's preferences are complete and transitive. If she considered all three comparisons, (i), (ii), and (iii), at resolution alpha, then, in (iii), Abby would be indifferent between tea+ and tea and no violation of the axiom of transitivity would occur. If she considers all three options at resolution beta, the only thing certain is that she will prefer tea+ to tea. Taking small details like particular aspects of aroma into account, Abby may prefer tea to coffee, prefer coffee to tea, or remain indifferent. If she does remain indifferent between tea and coffee at resolution beta, she will, at resolution beta, prefer tea+ to coffee.

It might be thought that Abby's indifference between tea+ and coffee at resolution alpha must be an error in judgment on her part, if, as argued here, she has the capacity to discern qualities that could cause her to prefer tea+. Especially if we allow that Abby is permitted a sip of each beverage before choosing, it seems reasonable to wonder, given the simplicity of the choice situation, how Abby might fail to notice some feature of tea+ at resolution alpha that she does notice at resolution beta. But, even simple experiences like sipping tea can be attended to more or less closely. I might, for example, appreciate the same sip as "warm tea", or as "warm tea with a soft, sweet flavor, and ginseng accents". This variation in how objectively identical experiences may be perceived translates quite naturally to variation in how simplified possible worlds are resolved.

Of course, the same sorts of concerns that apply to agents making adjustments to the partitions of states in an expected utility model of choice apply here as well. The same choice situation considered at different grains of resolution may result in the agent making different choices. As described above, indifference may resolve into strict preference, and there is no particular reason to disallow outright preference reversals. Abby, might, for example, prefer a cup of coffee to a cup of tea, but, were she to examine the options at some finer grain of resolution, she might find the aroma of tea (a detail she had not considered at all at the coarser resolution) so delightful that once this aroma is considered at some finer level of resolution she finds the tea preferable to the coffee.

Such preference changes might appear problematic. If more finely resolved choice options provide Abby with "a fuller and more realistic picture" of her choice situation (Joyce 1999, 70), then it seems as if Abby, knowing she has the capacity to resolve choice options more finely, rationally should pursue that fuller, more realistic picture in order to establish as accurate preference rankings as possible. The notion of costs, however, goes some distance towards answering such concerns. Abby may be well aware that if she took the time and effort to consider her sample sips of tea and coffee at a finer degree of resolution her preference would change and she would cease to be indifferent between the two options. But, there is the matter of cost, the extra time and effort. While Abby might prefer tea to coffee when she considers the choice situation at resolution beta, unless the benefits of choosing tea over coffee exceed the extra costs of considering the choice situation at a finer resolution, Abby should use the coarser resolution. Therefore, Abby should only use resolution beta to compare coffee and tea when she suspects that, for example, she will not just prefer one option to the other at that resolution, but that she will prefer the simplified possible world where she gets the now preferred option and pays some extra costs (the cost of comparing at resolution beta minus the cost of comparing at resolution alpha) to the world where she gets the lower ranked option and does not pay the extra cost.

## CONCLUSION

The SIA shows that intuitively plausible deliberative stances may constitute an objection to the assumption that agents can compare by establishing one of three comparative relationships between any two options. Directed at a utility theoretic account of choice, the objection shows that if the axiom of completeness is an accurate account of such preferences, then the axiom of transitivity cannot be an accurate account of them, as the intuitively plausible preferences display preference-indifference intransitivity.

However, a more complete account of choice also assumes that choice options are simplified possible worlds, mentally represented by agents; I call this assumption the axiom of subjectivity. An agent's ability to represent alternative outcomes as choice options includes the capacity to vary the amount and precision of details included in the simplified possible worlds. The capacity to resolve choice options to a finer or coarser degree answers the SIA by showing that as long the agent's preferences are all described at the same degree of resolution, the inconsistency between the claims made in the axiom of completeness and the axiom of transitivity is eliminated. So, the objection is illustrated by a failure of the axiom of transitivity, directed at the axiom of completeness, and answered by referring to the axiom of subjectivity.

I maintain that the force of the objection presented by the SIA depends on comparing choice options in a manner that does not correspond to a reasonable folk psychological account of how agents actually go about comparing. Agents resolve different choice problems at varying grains of resolution, and have good reasons (namely costs) for doing so. If one compares a simplified possible world that has been appraised by an agent at a certain grain of resolution with a simplified possible world that has been appraised at some other grain of resolution, one is making a mistake. Absent this sort of mixing and matching of differently resolved simplified possible worlds, the SIA does not illustrate any inconsistency between the axioms of completeness and transitivity

## REFERENCES


Adler, Matthew D. 2012. *Well-being and fair distribution.* Oxford: Oxford University Press.

Broome, John. 1991. *Weighing goods.* Cambridge: Blackwell Publishers.

Chan, David K. 2010. Reasoning without comparing. *American Philosophical Quarterly*, 47 (2): 149-160.

Chang, Ruth. 2002. *Making comparisons count.* New York: Routledge.

De Sousa, Ronald B. 1974. The good and the true. *Mind*, 83 (332): 534-551.

Gustafsson, Johan E., and Nicolas Espinoza. 2010. Conflicting reasons in the small-improvement argument. *Philosophical Quarterly*, 60 (241): 754-763.

Hausman, Daniel. 2012. *Preference, value, choice, and welfare.* Cambridge: Cambridge University Press.

Hsieh, Nien-hê. 2005. Equality, clumpiness and incomparability. *Utilitas*, 17 (2): 180-204.

Joyce, James M. 1999. *The foundations of causal decision theory.* New York: Cambridge University Press.

Lewis, David. 1972. Psychophysical and theoretical identifications. In *Readings in philosophy of psychology, vol. 1*, ed. Ned Block. Cambridge (MA): Harvard University Press, 207-222.

Quinn, Warren S. 1990. The puzzle of the self-torturer. *Philosophical Studies*, 59 (1): 79-90.

Raz, Joseph. 1985. Value incommensurability: some preliminaries. *Proceedings of the Aristotelian Society*, 86: 117-134.

Raz, Joseph. 1986. *The morality of freedom.* Oxford: Clarendon Press.

Rosenberg, Alex. 2008. *Philosophy of social science* [3rd ed.]. Boulder: Westview Press.

Sinnott-Armstrong, W. 1985. Moral dilemmas and incomparability. *American Philosophical Quarterly*, 22 (4): 321-329.

**Jack Anderson** is a doctoral candidate in the department of philosophy at the University of Utah. His research focuses on the areas of practical reasoning and the philosophy of economics.
Contact e-mail: <jack.anderson@utah.edu>

*Erasmus Journal for Philosophy and Economics,*
*Volume 8, Issue 1,*
*Spring 2015, pp. 42-60.*
*http://ejpe.org/pdf/8-1-art-3.pdf*

# Lesser degrees of explanation: further implications of F. A. Hayek's methodology of sciences of complex phenomena

SCOTT SCHEALL
*Arizona State University*

**Abstract:** F. A. Hayek argued that the sciences of complex phenomena, including (perhaps especially) economics, are limited to incomplete "explanations of the principle" and "pattern predictions". According to Hayek, these disciplines suffer from—what I call—a *data problem*, i.e., the hopelessness of populating theoretical models with data adequate to full explanations and precise predictions. In Hayek's terms, explanations in these fields are always a matter of "degree". However, Hayek's methodology implies a distinct *theory problem*: theoretical models of complex phenomena may be underspecified so that, even when all data is available, a full explanation could not be inferred from the model. Where the sciences of complex phenomena are subject to both the *data* and *theory* problems, explanations and predictions will be of even lesser "predictive degree". The paper also considers how to interpret Hayek's claim that pattern predictions are falsifiable.

**Keywords:** Hayek, explanation, prediction, economics, methodology, complexity

**JEL Classification:** A12, B31, B41, B53

From the early-1950s on, F. A. Hayek was concerned, among his several other interests, with the development of a methodology of sciences that study systems of *complex* phenomena.[1] According to Hayek, complex

[1] The various disciplines that Hayek counted under the heading of sciences of complex phenomena include theoretical psychology (1952), "cybernetics, the theory of automata or machines, general systems theory, and perhaps also communications

AUTHOR'S NOTE: This project was completed while I was a research fellow with Duke University's Center for the History of Political Economy. Many thanks to Bruce Caldwell, Gene Callahan, Catherine Herfeld, Joseph Ranweiler, Solomon Stein, two anonymous reviewers, and Luis Mireles-Flores (*EJPE* co-editor) for their several constructive comments and criticisms. All remaining errors of omission and commission are, unfortunately, my own.

phenomena consist of a large number of elements interconnected (both to each other and to the external environment) in such a way as to give rise to an emergent order that possesses "certain general or abstract features which will recur independently of the particular values of the individual data, so long as the general structure [...] is preserved" (Hayek 1967 [1964], 26). The scientist of complex phenomena investigates these emergent orders and their properties, which cannot be fully reduced to the properties of the particular elements involved (Hayek 1967 [1964], 39). However, the knowledge that can be acquired about such orders is limited—in virtue of their complexity (and the comparatively narrow boundaries of human cognitive faculties)—as compared to the knowledge that scientists of simpler phenomena can acquire about the objects of their analyses. In particular, Hayek (1967 [1964], 27) argued that the number of elements of such complex systems is so large as to constrain the capacity of the scientist of complex phenomena to populate theoretical models with data sufficient to generate any but circumscribed explanations ("explanations of the principle") and predictions ("pattern predictions").[2]

The present paper aims to draw out and clarify a number of further implications of Hayek's methodology of sciences of complex phenomena that have heretofore been unspecified in the primary and secondary literature on Hayek. In particular, the paper seeks to elucidate the implications of Hayek's methodology with respect to the specific dimensions along which the scientist's knowledge of some complex phenomena may be limited. Hayek's fallibilism—i.e., the epistemological position according to which knowledge is never complete and, in any case, always revisable in the light of new evidence—was an essential (if not always explicit) aspect of his arguments against the defenders of both socialism (1948 [1935]; 1948 [1940]) and countercyclical monetary policy (1978 [1975]). Yet, despite the fact that his conceptions of both complex phenomena and the methodology appropriate to their investigation imply that ignorance might beset the scientist in respects beyond the aforementioned difficulties of data collection, he never explicated these latter implications of his methodology.

theory" (1967 [1955], 20); as well as economics, linguistics (1967, 72), geology, evolutionary biology, and the branches of astrophysics that investigate the formation of stars and galaxies (1967, 76).

[2] Hayek (1967 [1955], 9n) understood explanation and prediction to be two sides of the same coin. The two terms are used interchangeably here.

Predictive capacities are limited wherever such ignorance is rife. More to the point, the specificity of a scientific prediction depends on the extent of the relevant scientist's (or scientific community's) knowledge concerning the phenomena under investigation. The paper offers an account of the considerations which, according to Hayek's methodology, determine the extent to which a theory's implications prohibit the occurrence of particular events in the relevant domain. This theory of "predictive degree" both expresses and—as the phenomena of scientific prediction are themselves complex in Hayek's sense—*exemplifies* the intuition that the specificity of a scientific prediction depends on the relevant knowledge available.

## I

According to Hayek's epistemology, knowledge comes in two varieties: there is "scientific" (or "theoretical") knowledge ("knowledge of general rules") and there is empirical knowledge ("knowledge of the particular circumstances of time and place") (Hayek 1948 [1945], 80).[3] The possibility of a "full explanation" or a "precise prediction of particular events"[4] requires that the scientist possess both kinds of knowledge to a sufficient extent: "[s]uch prediction will be possible if we can ascertain […] all the circumstances which influence those events. We need for this both a theory which tells us on what circumstances the events in question will depend, and information on the particular circumstances which may influence the event in which we are interested" (2014 [1961]).[5]

[3] Hayek's epistemology includes knowledge of which we may not be "explicitly aware", but which we "merely manifest […] in the discriminations which we perform" (Hayek 1952, 19). This is "tacit" knowledge (Polanyi 1966) or "knowledge how" as opposed to "knowledge that" (Ryle 1946). Tacit knowledge is not, for Hayek, a third class of knowledge. Rather, we can both have tacit knowledge of general rules of conduct and only tacitly recognize the particular circumstances in which these general rules are relevant.

[4] A "full" explanation need not be complete in the sense of encompassing *every detail* of the phenomena under investigation: an explanation "can never explain everything to be observed on a particular set of events" (Hayek 1952, 182). The concept of explanatory "fullness" should be thought of as sensitive to scientific context.

[5] In addition to those sciences of complex phenomena mentioned in footnote 1 above, Hayek (1967 [1964]) offers two specific examples of theories of complex phenomena, namely, Darwin's theory of evolution by natural selection—"[p]robably the best illustration of a theory of complex phenomena which is of great value, although it describes merely a general pattern whose detail we can never fill in" (1967 [1964], 31)—and Walrasian general equilibrium theory, with respect to which Hayek writes:

> […] economic theory is confined to describing kinds of patterns which will appear if certain general conditions are satisfied, but can rarely if ever derive from this knowledge any predictions of specific phenomena. This is seen most clearly if we

Hayek defines pattern predictions as the implications of theories that would suffice to generate detailed predictions of particular events if only the parameters[6] of the theory could be filled in with sufficient empirical data: "[e]very algebraic equation or set of equations defines in this sense a class of patterns, with the individual manifestation of this kind of pattern being particularized as we substitute definite values for the variables" (1967 [1964], 24). Hayek writes that

> [s]uch a theory[,] destined to remain 'algebraic', because we are in fact unable to substitute particular values for the variables, ceases then to be a mere tool and becomes the final result of our theoretical efforts. Such a theory…enables us to predict or explain only certain general features of a situation which may be compatible with a great many particular circumstances […] [I]n many fields this will be for the present, or perhaps forever, all the theoretical knowledge we can achieve (Hayek 1967 [1964], 28-29).[7]

Thus, Hayek's definition of a pattern prediction implies that, other things equal, the *degree* of a prediction—i.e., the specificity "of the events mentioned […] in the prognosis" (Hayek 1967 [1955], 8)—is positively related to the extent of the available data.[8] We can also say that, a maximally-specific prediction, i.e., a prediction that rules out the occurrence of every possible event in the relevant domain but one (if you like, a prediction of degree 1 or, in Hayek's verbiage, "a precise

> consider those systems of simultaneous equations which since Léon Walras have been widely used to represent the general relations between the prices and the quantities of all commodities bought and sold. They are so framed that if we were able to fill in all the blanks…we could calculate the prices and quantities of all the commodities. But, as at least the founders of this theory, clearly understood, its purpose is not [quoting Pareto 1927, 223-224] 'to arrive at a numerical calculation of prices', because it would be 'absurd' to assume that we can ascertain all the data (Hayek 1967 [1964], 35).

[6] Hayek uses three different terms seemingly interchangeably to speak of the theoretical elements of explanations, namely, "variable", "element" (1967 [1955]), and "parameter" (1964 [1967]). For the most part, I have adopted the latter locution, though I may occasionally use one of the other two terms merely to avoid excessive repetition.

[7] Also see Hayek (1952, 183): "[t]he distinction between the explanation of the principle on which a wide class of phenomena operate and the more detailed explanation of particular phenomena is reflected in the familiar distinction between the 'theoretical' and the more 'applied' parts of the different sciences".

[8] Relatedly, predictive degree increases (other things equal) as the possible range within which the value of some variable might lie is narrowed. This is particularly relevant given the well-known vagaries of empirical measurement: *ceteris paribus*, the degree of a prediction increases as observational error is minimized (see Hayek 1967 [1955], 9-10).

prediction of particular events"), is an implication of a fully-specified theory, and a complete (relative to the specified theoretical parameters) and precisely-measured data set. Stated more plainly, according to Hayek's epistemology, a scientist will be able to precisely predict particular events only if she knows everything of a theoretical and empirical nature relevant to the phenomena in the given context.

Moreover, Hayek's methodology of sciences of complex phenomena implies that *for every theory (T) adequate to imply precise predictions of particular events*, there is some number greater than 0 (the degree of a statement devoid of empirical content) and less than 1 (the degree of a precise prediction of particular events), which is the degree of what we might call a *mere* pattern prediction, i.e., an implication of the conjunction of an *empty* data set and a theory that is completely specified relative to the requirements of a full explanation in the relevant scientific context. Naturally, other things equal, the predictive degree of such a conjunction approaches 1 as the data set is increasingly populated.

Summarizing, a bit more formally, a precise prediction of particular events specified to the extent desired in the given scientific context requires both knowledge of a (positive and "large") number $p$ of theoretical parameters and knowledge of the particular value that each of these parameters assumes at the time relevant to the prediction. If (and only if) the scientist's knowledge satisfies both conditions, then her prediction will be of degree 1. However, if the first condition is satisfied, i.e., if the scientist possesses knowledge of $p$ theoretical parameters, but she does not know the value of any of the parameters at the time relevant to the prediction, then she will only be able to make a mere pattern prediction, the degree of which will be greater than 0 and less than 1, but which will approach 1 as she acquires more knowledge of the relevant values of each of the $p$ parameters.

## II

There are places in Hayek's methodological writings where he points to the fact that the case in which the scientist possesses an adequate theoretical understanding of the order under investigation (and, so, merely requires sufficient data in order to generate a precise prediction of particular events) is not the norm in the sciences of complex phenomena. As a matter of fact, there are many cases in which, relative to what would be required to generate a precise prediction, the scientist's theoretical knowledge is deficient, i.e., where the relevant "algebraic equation or set of such equations [that] defines […] a class of patterns" (Hayek 1967 [1964], 24) is itself underspecified—contains gaps or lacunae with regard to the parameters required of a full explanation of the order under investigation—so that a precise prediction of particular events could not be generated *even if the scientist possessed all of the relevant data*.[9]

That Hayek took this *theory problem* (as distinct from the *data problem* he so thoroughly explicated) to be common in sciences of complex phenomena is implicit in the argument of "The dilemma of specialization" (1967 [1956], 124),[10] an essay in which explicit methodological considerations take a backseat to concerns of best pedagogical practices in the social sciences. Hayek's claim that preeminence in these disciplines requires learning well beyond a narrow field of specialization is worth quoting at length as it gets to the heart of the theory problem implied by his methodology. "For almost any application of our knowledge to concrete instances", Hayek writes,

> […] the knowledge of one discipline, and even of all the scientific knowledge we can bring to bear on the topic will be only a small part of the foundations of our opinions. Let me speak first of the need of using the results of scientific disciplines other than our own, though this is far from all that is required. That concrete reality is not divisible into distinct objects corresponding to the various scientific disciplines is a commonplace, yet a commonplace which severely limits our competence to pronounce as scientists on any particular event. There is scarcely a phenomenon or event in society with which we can deal adequately without knowing a great deal of

[9] See, e.g., "A new look at economic theory", the first of four lectures delivered at the University of Virginia in 1961 (and now published as Hayek 2014 [1961]): "[*e*]*ven a true theory* will not enable us to make predictions of specific events unless we are able to ascertain all those relevant facts […] which govern the particular position" [italics added]. One is left to ponder the significance of predictions of theories that might be "true" as far as they go, but which do not go very far because their theoretical parameters are not fully specified.

[10] Hayek's concern for the theory problem is also implicit in his criticisms of Keynes's *Treatise on money* (1971 [1930]) in the early 1930s. The main point that Hayek persistently pushed in his multi-part review (1995 [1931a]; 1995 [1931b]; 1995 [1932]) of the *Treatise* was the absence of any theoretical account of capital in the book. In his rejoinder, Keynes (1973 [1931], 252-253) accepted that a treatment of capital would figure in a perfected theory of a money-using economy, but insisted that the theory presented in the *Treatise* was adequate for his purposes at the time. Hayek and Keynes obviously had different conceptions of the requisite "fullness" of a satisfactory explanation of a money-using economy (or misconceived each other's scientific purposes).

> several disciplines, *not to speak of the knowledge of particular facts that will be required* (Hayek 1967 [1956], 124; italics added).

In other words, a theory capable of generating a full explanation of some complex phenomena may well be a *composite system of theories*, spanning multiple disciplines each of which might investigate phenomena of greater or lesser complexity.[11] What is more, given that "concrete reality is not divisible into distinct objects corresponding to the various scientific disciplines", there may be phenomena that contribute to the emergence of a particular order and that must be accounted for if an explanation is to be "full", which fall under the heading of no extant scientific discipline. The relevant theoretical knowledge may not have been discovered (indeed, it may not even be discover*able*).

Hayek (1967 [1964], 40-42) argued in the "Postscript" that accompanied later versions of "The theory of complex phenomena" that the ordinary notion of a scientific law, i.e., of a relation between cause and effect, has a clear meaning (whatever it might be) only with respect to "those two-variable or perhaps three-variable problems to which the theory of simple phenomena can be reduced". However, with regard to more complex phenomena,

> [i]f we assume that all the other parameters of such a system of equations describing a complex structure are constant, we can of course still call the dependence of one of the latter on the other a 'law' and describe a change in the one as 'the cause' and the change in the other as 'the effect'. But such a 'law' would be valid only for one particular set of values of all the other parameters and would change with every change in any one of them. This would evidently not be a very useful conception of a 'law', and the only generally valid statement about the regularities of the structure in question is the whole set of simultaneous equations from which, if the values of the parameters are continuously variable, an infinite number of particular laws, showing the dependence of one variable upon another, could be derived (Hayek 1967 [1964], 41-42).

[11] The case in which a full explanation requires theoretical input from a number of disciplines illustrates the theory problem in all its ignominy, but the problem can just as well manifest in sciences where self-contained explanations are possible. All that is necessary (and sufficient) for the theory problem to arise is that, relative to the requirements of a full explanation of the phenomena under investigation, there be gaps in the specification of the parameters of the theory (or composite system of theories) meant to generate such an explanation.

In other words, the individual parameters are not, strictly speaking, causes in the sense in which we normally think of them in explanations of simpler phenomena. The "cause" of an order is the whole network of parameters and particular values from which it emerges, and to know the cause of some order is to possess knowledge of all of these theoretical and empirical considerations.

That a scientist's theoretical understanding of some complex phenomena may underwhelm the requirements of a precise prediction of particular events—regardless of the availability of the relevant data—implies that there are, on Hayek's methodology of sciences of complex phenomena, degrees of prediction less than the degree of mere pattern predictions. Consider a system of theories ($T$) fully specified in terms of $p$ theoretical parameters for which no data are available, i.e., a system of theories capable of generating only mere pattern predictions. Now, imagine removing one of the parameters. The result of this excision ($T'$), which consists of $p$-$1$ parameters, will, other things equal, imply a pattern that is missing at least one of the parts of the pattern implied by $T$. A system of theories that would be capable of generating mere pattern predictions *were its parameters fully specified* will describe only *partial* patterns. Moreover, if we remove another parameter from $T'$, the result ($T''$), consisting of $p$-$2$ parameters, will, *ceteris paribus*, imply a pattern that is missing at least one of the parts of the pattern implied by $T'$, and so on.

Indeed, a point will come in the process of eliminating parameters where the remaining variables will not imply anything like a substantive explanation that might be of interest to a scientist.[12] That is, for any complex phenomena the scientist might want to explain, there is a minimum positive number of theoretical parameters ($p$-$n$, where $p$ is positive and "large", and $n < p$) she must know in order for her theory to express even minimal empirical content. The predictive degree associated with the $T^n$ that includes $p$-$n$ parameters, the minimum number of parameters required for empirical content, sets a lower bound to the degree of associated partial pattern predictions. If the

[12] See Hayek (1952, 180): "In general, the possibility of forming a model which explains anything presupposes that we have at our disposal distinct elements whose action in different circumstances is known irrespective of the particular model in which we use them". Similarly, the possibility of even limited explanations requires that we have at our disposal *enough* of these elements. Hayek (1967 [1964], 26) refers to the "minimum number of distinct variables a formula or model must possess in order to reproduce the characteristic patterns of structures of different fields (or to exhibit the general laws which these structures obey)".

scientist knows fewer parameters than this, her theory cannot generate even partial pattern predictions, i.e., the theory's predictive degree is 0.

## III

It follows from the considerations adduced thus far that, *ceteris paribus*, the predictive degree of any theory ($T^m$, where $0 \leq m \leq n$) consisting of $p$-$m$ parameters, will be greater than or equal to that of $T^n$ and approach that of $T$ as more relevant parameters are added to the system. Of course, for any such $T^m$, predictive degree will, other things equal, increase as the $p$-$m$ parameters are increasingly populated with data. It also follows that, if we possess certain data points only some of which have clear theoretical interpretations with respect to the phenomena under investigation—that is, if we have some observations the theoretical significance of which is undetermined in the relevant context—then, as we are subsequently able to interpret these values in terms of (i.e., assign them to) particular parameters, the degree of associated predictions will increase (*ceteris paribus*).

However, an element of indeterminacy enters the frame when we consider the problem of comparing the predictive degrees of two (or more) conjunctions of theoretical and empirical knowledge, both concerning the same complex phenomena, one of which consists of a larger collection of parameters for which there are less data available and another that consists of a smaller collection of parameters for which there are more data available. To see this, consider the extreme case of comparing the predictive degree of a mere pattern prediction, i.e., the conjunction of $p$ parameters (which would suffice for a precise prediction if all of the data were available) and an empty data set, with the predictive degree of a conjunction of $p$-$n$ parameters (the minimum number of parameters required for empirical content) and a set of values each of which can be assigned to one of the $p$-$n$ parameters. Which of these conjunctions of theoretical and empirical knowledge has the higher predictive degree, i.e., which prohibits the occurrence of more events in the relevant domain? The situation here is that the first conjunction implies no details about a complete pattern of events while the second implies all of the details about a partial pattern of events. Whether this means that the predictive degree of the first conjunction is greater or lesser than that of the second is undecidable on the basis of a priori considerations alone.

In order to determine the predictive degrees of two rival conjunctions of theory and data, one consisting of both more parameters and less data than its competitor, we need *other* theories (more precisely, we need *meta-theories* or *second-order theories*) that tell us how, for each rival conjunction, theoretical and empirical knowledge interacts so as to influence the degree of associated predictions. Stated another way, these meta-theories would be multi-valued functions that take as inputs “measurements” of the extent of theoretical parameters and data, and return specific, quantitatively-precise, degrees of prediction. This is to say that, at the meta-level, the relevant data would be the rival conjunctions of (first-order) theories and their respective data. Our problem is precisely that, in the abstract, without these rival first-order theories and their respective data, nothing about the shapes of the relevant multi-valued functions can be established by way of philosophical analysis alone.

All of this is just a way of saying that the theory of predictive degree implied by Hayek’s methodology of sciences of complex phenomena is itself subject to both the data problem and the theory problem, and, therefore, cannot approach a full explanation of the epistemological aspects of predictions in the sciences of complex phenomena. However, this result is a strength rather than a weakness of both Hayek’s methodology and the present attempt to elucidate the latter further. It means, in essence, that the phenomena of predictive degree are themselves complex and, thus—in concert with Hayek’s methodology—that we are effectively limited by the complexity of these phenomena to an explanation of the principle of explanations of the principle. It is of course consistent for an explanation according to which explanations are limited to itself be limited. Indeed, the real threat to Hayek’s methodology would be a full explanation of the phenomena of predictive degree, for this would signify either that these phenomena are not complex in Hayek’s sense—a possibility which the slightest bit of reflection reveals to be unlikely—or that Hayek’s methodology is internally inconsistent, i.e., that full explanations of complex phenomena *are* possible.

## IV

To this point, I have relied on an implicit simplifying assumption to the effect that knowledge of any particular theoretical parameter or data point is equally important to a theory’s predictive degree. That is, I have assumed that if a full explanation of some complex phenomena requires knowledge of (a positive and "large") number $p$ of parameters and each of their values, then knowledge of *any p-1* of these parameters and their values would result in predictions of the same degree. However, there is nothing in Hayek's writings on complex phenomena that licenses this assumption. It may well be that knowledge of a certain collection of $p$-$m$ (where $0 < m \leq n$) parameters and their values results in predictions of greater (or lesser) degree than knowledge of some other group of $p$-$m$ parameters and their values. Indeed, it is even possible on Hayek's methodology that knowledge of some collection of $p$-$l$ (where $l > m$) parameters and their values results in predictions of *greater* degree than knowledge of a group of $p$-$m$ parameters and their values. In other words, it can matter for predictive degree not only *how many* items of theoretical and empirical knowledge the scientist possesses, but *which* pieces of knowledge she possesses, some items being more important for an adequate explanation of the emergent order than others. The scientist might possess fewer pieces of more important knowledge and generate predictions of greater degree than if she possessed more items of less important knowledge.

To see this latter point, consider again the quote above from "The dilemma of specialization". Presumably, what Hayek has in mind when he says that "the knowledge of one discipline, and even of all the scientific knowledge we can bring to bear on the topic will be only a small part of the foundations of our opinions" (1967 [1956], 124) is that a full explanation of, say, some economic phenomena, requires more knowledge than can be furnished by economic science alone—we would also need, e.g., historical, sociological, anthropological, and psychological knowledge, not to mention some knowledge provided by the physical and biological sciences. One would think—although, for the reasons stated in section III above, no proof can be offered—that the knowledge provided by economic science alone which, without any knowledge of these other disciplines, would represent a comparatively small proportion of the parameters required for a full explanation, would nonetheless yield a fuller explanation of the relevant economic phenomena than an explanation built entirely on knowledge of these other disciplines without any input from economics.

It should be obvious that dropping the simplifying assumption that knowledge of any particular theoretical parameter or data point is equally important leaves our theory yet further removed from a full explanation of the phenomena of predictive degree. Nonetheless, the considerations adduced thus far do license some very general statements about the effect of ignorance of relatively important theoretical parameters or their respective data on the degree of associated predictions. The partial pattern prediction implied by a system of theories missing relatively important parameters will, other things equal, describe less of the complex phenomena than an otherwise identical system missing the same number of less important parameters. To see this, return briefly to the example in the previous paragraph: it seems that, other things equal, the system missing *m* economic parameters must predict less of the relevant pattern of economic phenomena under investigation than another system missing *m* non-economic parameters. The former will imply lesser degree partial pattern predictions than the latter. Similarly, if the data associated with particularly important parameters are missing, then the resulting conjunction of parameters and data will imply the same pattern as an otherwise identical conjunction that happens to be missing the same number of data points associated with less important parameters; however, other things equal, the former will imply less of the more important details and more of the less important details of this pattern of events, while the latter will imply more of the more important details and less of the less important details of the pattern. In terms of our economic example, though they will both imply the same pattern of economic events, other things equal, the conjunction missing the same number of economic data points will imply less of the economic details than another conjunction missing the same number of non-economic data points. Thus, *ceteris paribus*, of two otherwise identical conjunctions of parameters and data, both missing the same number of data points, the system missing data for more important parameters will imply predictions of lesser degree than another system missing data for less important parameters.

## V

A bit of reflection, especially upon Hayek's definition of complex phenomena, reveals further dimensions with respect to which the scientist's knowledge may be limited to some extent or other. Moreover, these dimensions concern aspects of the phenomena of prediction that are not easily expressible in terms of *more or less*, or more or less *important*, knowledge, but instead concern the scientist's knowledge of purely qualitative matters relevant to some complex phenomena.

The theory problem begins with ignorance of at least one of the parameters relevant to the complex phenomena under investigation, but it does not suffice to resolve the problem that the scientist simply enumerate the pertinent variables and understand their relative importance to the phenomena; she must also know something of the *kinds* of values that each of the parameters can assume. It might seem natural to think that a particular parameter can only take values of the same kind, ontologically speaking—i.e., that one and the same parameter cannot assume, say, either the value *red* or *dog* or *the number seven*. In fact, nothing in Hayek's methodology of sciences of complex phenomena licenses this assumption. However, whether this assumption holds or not, the scientist must know something of the ontological properties of the relevant parameters—either that they each take values of the same kind or that some take values of different kinds. Of course, the scientist may be in a position to know both which of these latter conditions holds and the number of parameters, without knowing whether the values a particular variable assumes are all, e.g., of the kind *color*, *quadruped mammal*, *positive integer*, or *some combination thereof*. With regard to considerations of predictive degree, the most that can be established with regard to such circumstances is the rather trivial proposition that, other things equal, predictive degree increases as the scientist's theoretical knowledge improves in the relevant sense. That is, the scientist will be able to rule out more events as she improves her knowledge of the relevant properties of the phenomena represented by the parameters of her theory.

But, alas, the theory problem is not merely that of tallying, and identifying both the explanatory importance and ontological properties of, all of the parameters; it extends to that of specifying the interconnections of various subsets of these parameters both with each other and with the external environment. A full resolution of the theory problem requires knowledge of these latter circumstances. Moreover, given Hayek's definition of complex phenomena as those orders that emerge from the internal and external interconnections of a large number of elements, which possess "certain general or abstract features which will recur independently of the particular values of the individual data, so long as the general structure […] is preserved"

(Hayek 1967 [1964], 26), and his argument,[13] that the "cause" of a particular emergent order is the whole network of parameters and particular values from which it emerges, the theory problem includes determining the limits of the constancy of some order given changes in either the values of the data or the interconnections both between various subsets of parameters and between these subsets and the environment. That is, a full solution to the theory problem requires the scientist to consider the extent to which either the values or relevant interconnections of different subsets of parameters can change before the order is supplanted either by some altogether distinct order or by *disorder*.[14]

Again, philosophical analysis *in vacuo* is largely impotent to pronounce on the effects of these matters upon considerations of predictive degree beyond the rather trifling claim that, other things equal, the predictive degree of some theoretical system increases—i.e., the occurrence of more events can be ruled out—as the scientist acquires better knowledge of the interconnections both between different subsets of parameters and between these various subsets and the environment; and increases as well as better knowledge is acquired of the manner in which the emergence of some complex phenomena depends on the maintenance of particular interconnections and the values assumed by the parameters.

## VI

It remains to say a few brief words about the sense in which pattern predictions are falsifiable. If one is not careful to read Hayek charitably, it would be too easy to infer an inconsistency between Hayek's methodology and his numerous claims (e.g., at several points in both 1967 [1955] and 1967 [1964]) that pattern predictions are falsifiable. After all, the possibility of a *conclusive* falsification requires that the scientist of complex phenomena possess knowledge that she cannot possess according to Hayek's methodology, because she is always subject to some extent to either the "data problem" or the "theory problem" (and usually both).

[13] See discussion above, pp. 48-49.

[14] Hayek's methodology leaves open the possibility of "feedback" between an order and the phenomena from which it emerges. Thus, a full solution of the theory problem with respect to such an order would also require knowledge of the nature and extent of this feedback.

But, of course, this picture of falsification—which Imre Lakatos (1968-1969, 152-162) once dubbed the "naïve" conception—is untenable in virtually all philosophies of science. According to the naïve view, a single observation is sufficient to conclusively falsify a universal statement. The paradigmatic and oft-quoted example is the eighteenth-century discovery of black swans in Australia as a purportedly decisive falsification of the universal proposition that "All swans are white". But, as has been pointed out by Duhem (1954 [1906]), Popper (1959 [1934]), and Quine (1961 [1951]), a universal statement can always be saved from an apparent falsification if one is prepared to invoke some *ad hoc* hypothesis such as, in the case of the apparent observation of Australian black swans, that "The Australian climate causes spectral inversion in bird watchers". As Popper puts it,

> […] no conclusive disproof of a theory can ever be produced; for it is always possible to say that the experimental results are not reliable, or that the discrepancies which are asserted to exist between the experimental results and the theory are only apparent and that they will disappear with the advance of our understanding (Popper 1959 [1934], 50).

Because such *ad hoc* hypotheses are always available, naïve falsificationism is untenable.

So it is on Hayek's methodology of sciences of complex phenomena. The theory of predictive degree implies that only a precise prediction of a particular event—a prediction of degree 1—is fully falsifiable in the sense required by naïve falsificationism. Such a prediction rules out *all but one* of the multitudinous events relevant to the complex phenomena under investigation, i.e., it rules out all *ad hoc* hypotheses. If the scientist's knowledge were ever so comprehensive as to be able to exclude all but one event, then a falsification would indeed necessitate the rejection of the relevant theory. But, of course, given that Hayek (1967 [1964], 27) *defines* the sciences of complex phenomena in terms of the presence of (what I am calling) the "data problem", the impossibility of a scientist's knowledge being so total and, thus, the untenability of naïve falsificationism is part and parcel of Hayek's methodology. Naturally, the further possibility of the "theory problem" makes the naïve picture of falsifiability even less plausible. That is, if, as Hayek (1967 [1956], 124) says in "The dilemma of specialization", "all the scientific knowledge we can bring to bear […] will be only a small part of the foundations of our opinions", then this scientific—or, as Hayek (1967 [1964], 29) calls it elsewhere, "theoretical"—knowledge is not necessarily falsified when our opinions are controverted by experience.[15]

So, Hayek was at least not a "naïve" falsificationist. Unfortunately, considerations of brevity prevent me from providing a positive account of the precise nature and extent of Hayek's more "sophisticated" falsificationism in the present context. In particular, I cannot here explicate (my views concerning) the relationship of Hayek's methodology of sciences of complex phenomena to Popper's falsificationism. Suffice it to say here that so long as a conjunction of theoretical parameters and their respective values has a positive predictive degree, it rules out *some* events—perhaps, in the extreme, only one—and, if some such prohibited event is observed, then this would count as a falsification on Hayek's methodology.

## VII

Hayek used his methodology of sciences of complex phenomena—and, especially, his emphasis upon the limits of knowledge in the social sciences—as a weapon in his well-known debates with advocates of various social schemes that he believed to be epistemologically ill-conceived. Nonetheless, despite their significance for these and other aspects of his scientific and philosophical programs, he never fully explicated the consequences of his methodological arguments.

In an effort to make these implications plain, the present paper has applied what is in essence a method of decreasing abstraction to both Hayek's two-pronged epistemology and his definition of complex phenomena. That is, in section I, we considered the implications for predictive degree of the simplest case where the scientist possesses all of the relevant theoretical knowledge and none of the data. In section II, we dropped the former assumption and considered the degree of pattern predictions where the scientist's theoretical knowledge is also

[15] However, it must be said that Duhem-Quine considerations undermine the notion that a stark line can be drawn between sciences that investigate simple phenomena and sciences that investigate complex phenomena. That is, the presence of the data and theory problems that beset the latter sciences *demarcates* nothing: all sciences are subject to these problems to some extent or other. Of course, that Hayek was aware that no hard-and-fast distinction can ultimately be drawn between complex and simple sciences is reflected in the arguments of his original essays on "Degrees of explanation" (1967 [1955], especially 17-18) and "The theory of complex phenomena" (1967 [1964], especially 24-25).

limited. In section III, we established the impossibility of determining *in abstracto* the predictive degrees of multiple rival theories, which led to the observation that our theory of predictive degree is itself limited to pattern predictions. In section IV, we dropped the assumption that knowledge of each theoretical parameter (and its datum) is equally relevant to an explanation of some complex phenomena. In section V, we dropped the assumption that the scientist possesses knowledge of the ontological properties of the various parameters as well as knowledge of the interconnections both between the various subsets of parameters and between these subsets and the environment.

This approach highlights different respects, previously unnoticed in the literature on Hayek, in which relevant knowledge might be limited, that extend beyond the fact emphasized by Hayek that the scientist of complex phenomena is often ignorant of some or all of the data relevant to a particular analysis—a difficulty herein christened the “data problem”. In particular, the so-called “theory problem”—the possibility that the scientist’s ignorance might extend to one or more of the theoretical parameters relevant to a full explanation of some complex phenomena—has been especially emphasized here. A consequence of the theory problem is that predictions can be of lesser degree than that of Hayek’s (mere) pattern predictions: partial pattern predictions are possible. Furthermore, the extent of a scientist’s ability to fully explain or predict events with precision depends on her knowledge of parameters (and their associated data) that are particularly important to such an explanation. Hayek’s conception of complex phenomena also implies that a complete solution of the theory problem—and, thus, the possibility of precise predictions—requires knowledge of a more qualitative variety that is not easily expressible in terms of greater or lesser predictive degrees. But, more than this, the reflexivity of Hayek’s emphasis on fragmented and fallible knowledge has been established as a consequence of the possibility of only limited explanations of explanation. Hayek’s methodology implies the impossibility of either a complete explanation or a precise prediction of explanations and predictions.

## REFERENCES

Duhem, Pierre. 1954 [1906]. *The aim and structure of physical theory*. Trans. Philip P. Wiener. Princeton (NJ): Princeton University Press.

Hayek, F. A. 1948 [1935]. The present state of the debate. In *Individualism and economic order*. Chicago: University of Chicago Press, 148-180.

Hayek, F. A. 1948 [1940]. Socialist calculation III: the competitive ‘solution’. In *Individualism and economic order*. Chicago: University of Chicago Press, 181-208.

Hayek, F. A. 1948 [1945]. The use of knowledge in society. In *Individualism and economic order*. Chicago: University of Chicago Press, 77-91.

Hayek, F. A. 1952 [1942-1944]. *The counter-revolution of science: studies on the abuse of reason*. Glencoe (IL): Free Press. [Originally published in three parts as: Scientism and the study of society. *Economica*, 9 (35): 267-291; 10 (37): 34-63; 11 (41): 27-39]

Hayek, F. A. 1952. *The sensory order: an inquiry into the foundations of theoretical psychology*. Chicago: University of Chicago Press.

Hayek, F. A. 1967 [1955]. Degrees of explanation. In *Studies in philosophy, politics, and economics*. Chicago: University of Chicago Press, 3-21.

Hayek, F. A. 1967 [1956]. The dilemma of specialization. In *Studies in philosophy, politics, and economics*. Chicago: University of Chicago Press, 122-132.

Hayek, F. A. 1967 [1964]. The theory of complex phenomena. In *Studies in philosophy, politics, and economics*. Chicago: University of Chicago Press, 22-42.

Hayek, F. A. 1967. Notes on the evolution of systems of rules of conduct. In *Studies in philosophy, politics, and economics*. Chicago: University of Chicago Press, 66-81.

Hayek, F. A. 1978 [1975]. The pretence of knowledge. In *New studies in philosophy, politics, economics, and the history of ideas*. Chicago: University of Chicago Press, 23-34.

Hayek, F. A. 1991 [1920]. Contributions to a theory of how consciousness develops. Trans. Grete Heinz. Hoover Institution, Hayek Archives, box 92, folder 1.

Hayek, F. A. 1995 [1931a]. Reflections on the pure theory of money of Mr. J. M. Keynes. In *The collected works of F. A. Hayek*, ed. Stephen Kresge, vol. 9, *Contra Keynes and Cambridge: essays, correspondence*, ed. Bruce Caldwell. Chicago: University of Chicago Press, 121-146.

Hayek, F. A. 1995 [1931b]. A rejoinder to Mr. Keynes. In *The collected works of F. A. Hayek*, ed. Stephen Kresge, vol. 9, *Contra Keynes and Cambridge: essays, correspondence*, ed. Bruce Caldwell. Chicago: University of Chicago Press, 159-164.

Hayek, F. A. 1995 [1932]. Reflections on the pure theory of money of Mr. J. M. Keynes (continued). In *The collected works of F. A. Hayek*, ed. Stephen Kresge, vol. 9, *Contra Keynes and Cambridge: essays, correspondence*, ed. Bruce Caldwell. Chicago: University of Chicago Press, 174-197.

Hayek, F. A. 2014 [1961]. A new look at economic theory. In *The collected works of F. A. Hayek*, ed., Bruce Caldwell, vol. 15, *The market and other orders*. Chicago: University of Chicago Press, 373-426.

Keynes, John Maynard. 1971 [1930]. *A treatise on money: the pure theory of money*. In *The collected writings of John Maynard Keynes*, ed. Donald Moggridge, vol. 5. London: Macmillan.

Keynes, John Maynard. 1973 [1931]. The pure theory of money: a reply to Hayek. In *The collected writings of John Maynard Keynes*, ed. Donald Moggridge, vol. 13, *The general theory and after, part I: preparation*. London: Macmillan, 243-256.

Lakatos, Imre. 1968-1969. Criticism and the methodology of scientific research programmes. *Proceedings of the Aristotelian Society*, 69: 149-186.

Pareto, Vilfredo. 1927. *Manuel d’économie politique* [2nd ed.]. Paris: M. Giard.

Polanyi, Michael. 1966. *The tacit dimension*. Garden City (NY): Doubleday & Company, Inc.

Popper, Karl. 1959 [1934]. *The logic of scientific discovery*. London: Hutchinson.
Ryle, Gilbert. 1946. Knowing how and knowing that. *Proceedings of the Aristotelian Society*, 46: 1-16.
Quine, W. V. O. 1961 [1951]. Two dogmas of empiricism. In *From a logical point of view* [2nd revised ed.]. New York: Harper and Row. [Originally published in *The Philosophical Review*, 60: 20-43]

**Scott Scheall** teaches in the science, technology, and society department at Arizona State University. He received his PhD in philosophy from Arizona State in 2012. He is a former research fellow with Duke University's Center for the History of Political Economy and, during the current academic year (2014-2015), is a postdoctoral fellow with the F. A. Hayek program for advanced study in philosophy, politics, and economics at George Mason University. Scott is co-editor of *Research in the History of Economy Thought and Methodology*.
Contact e-mail: <scott.scheall@asu.edu>

*Erasmus Journal for Philosophy and Economics,*
*Volume 8, Issue 1,*
*Spring 2015, pp. 61-81.*
*http://ejpe.org/pdf/8-1-art-4.pdf*

# Orthodox and heterodox economics in recent economic methodology

D. WADE HANDS
*University of Puget Sound*

**Abstract:** This paper discusses the development of the field of economic methodology during the last few decades emphasizing the early influence of the "shelf" of Popperian philosophy and the division between neoclassical and heterodox economics. It argues that the field of methodology has recently adopted a more naturalistic approach focusing primarily on the "new pluralist" subfields of experimental economics, behavioral economics, neuroeconomics, and related subjects.

**Keywords:** orthodox, heterodox, neoclassical, economic theory, economic methodology

**JEL Classification:** A12, B41, B49, B50

> Myself when young did have ambition to contribute to the growth of social science. At the end, I am more interested in having less nonsense posing as knowledge (Frank Knight, 1956).

At the time I was finishing graduate school, there was no real "field" of economic methodology. There were of course methodological writings by influential economists (e.g., Robbins 1932, 1952; Friedman 1953; Samuelson 1964, 1965), but these works were seldom of the same intellectual quality as the research that had made these economists famous *as* economists. There were also brief discussions of economics in influential books on the philosophy of science (e.g., Hempel 1965,

AUTHOR'S NOTE: This paper began as a lecture delivered at the XVII Meeting on Epistemology of the Economic Sciences, School of Economic Sciences, University of Buenos Aires, Buenos Aires, Argentina, October 6-7, 2011. It was subsequently published in *Perspectives on epistemology of economics: essays on methodology of economics* (Lazzarini and Weisman 2012). It is reprinted here with the permission of the editors. I have made minor changes and updated the references, but I did not make any substantive changes to the argument in the original lecture. I would like to thank the many people who have made helpful comments on earlier versions of the paper and I would also like to thank the editors of the *Erasmus Journal for Philosophy and Economics*, particularly Luis Mireles-Flores, for suggesting it be reprinted here.

Nagel 1961), but they focused on general problems associated with the human and social sciences, rather than with specific issues concerning economics. There were two recently published case studies in the philosophy of economics written by philosophers—Hausman (1981) and Rosenberg (1976)—but in general the field was almost as unpopular among philosophers as it was among economists. Finally, and perhaps most importantly, there was beginning to be a collection of dedicated books on economic methodology—Blaug 1980a; Boland 1982; Caldwell 1982; Hutchison 1981; Latsis 1976; Wong 1978; and a few others—but it was a relatively assorted collection of texts with little to suggest that these books would end up being the foundational texts for the inchoate field of economic methodology. All in all, at that time there seemed to be very little to encourage a young scholar thinking about an academic career in economic methodology or the philosophy of economics.

However that was a long time ago, and I am happy to be able to report that the situation today is much improved. There are now dedicated journals such as *The Journal of Economic Methodology* and *Economics and Philosophy*, as well as numerous journals specializing in the history of economic thought that frequently publish methodological research. There are also a number of research institutes and professional societies dedicated to the intersection of economics and philosophy around the world. It is now possible for a young scholar to specialize in research connecting economics and philosophy without necessarily feeling like they are jeopardizing the possibility of a successful academic career. Of course, this does not mean that such careers are easy, or that all is well within the field—i.e., "better" certainly does not imply "good". Particularly in the United States, the economics profession still seems to have little or no interest in elevating economic methodology to the status of a legitimate field of inquiry within the discipline of economics. The financial crisis and the associated questioning of the methodological foundations of macroeconomic theory, seems to have initiated a momentary warming of the relationship between mainstream economics and economic methodology, but who knows how serious the overtures are or how long they will last. Also, it is probably not a good sign that the profession considers economic methodology to be an *inferior good* in the traditional microeconomic sense: that is, one that economists consume more of when incomes fall.

The last twenty or so years have also witnessed a significant change in the traditional relationship between "orthodox" and "heterodox" schools of thought within economics. For most of the second half of the 20th century the economic mainstream, the orthodoxy, consisted of neoclassical microeconomics combined with some version of macroeconomics (it was IS-LM Keynesian theory during the immediate post WW-II period, and new classical macroeconomics and real business cycle theory later). On the other hand, the periphery of the discipline was divided into a small number of self-consciously heterodox schools of thought: institutionalist, Marxist, Austrian, post-Keynesian, and others. There were two key features to this half-century long equilibrium in economic theorizing. First, there was a dominant orthodoxy based on neoclassical principles—prediction and/or explanation of economic phenomenon in terms of the coordinated equilibrium behavior of rational self-interested agents—and those principles were strictly enforced. If there were no maximizing agents in the model, then it was not mainstream, and for the majority of the profession, not scientific, economics.[1] And second, those outside of the mainstream tended to be self-conscious members of some particular heterodox school. It was not simply a matter of there being a dominant mainstream and a disparate group of outsiders—not just the discipline's "insiders" and the "others"—there was a dominant neoclassical school and a number of different, but distinct and self-consciously identified, heterodox schools in the periphery. Very few economists were engaged in theorizing that was outside of the mainstream and yet also outside of any of these clearly-labeled heterodox groups.

This relationship seems to have changed during the last few decades. On one hand, many of the most important recent developments within economics have occurred within fields such as

[1] The maximizing agents were explicit in microeconomics; in macroeconomics there were always ongoing efforts to find "microfoundations"—ways of grounding the macro-theoretical concepts on neoclassical principles. Although it is clearly recognized that the new classical macroeconomics that became dominant at the end of twentieth century was motivated by the desire for microfoundations, it is less well-recognized that even during the immediate post WW II period when Keynesian ideas dominated macroeconomics, there were also ongoing efforts to "ground" Keynesian ideas like the consumption function, liquidity preference, and the marginal efficiency of capital in individual maximizing behavior. The relevant "microfoundations" were defined more broadly during the Keynesian than the new-classical period, and perhaps the latter was more successful than the former in reaching its microfoundational goals, but the profession's preference for grounding macroeconomic concepts on neoclassical microeconomic principles was clearly revealed even during the Keynesian period.

experimental economics, behavioral economics, evolutionary economics, and neuroeconomics. These are fields that are not "orthodox" in the strict neoclassical sense—they often produce anomalous results that conflict with standard neoclassical theory and often characterized economic behavior in very non-neoclassical ways—but they are also not "heterodox" in the traditional sense either; they are not Marxist, or institutionalist, Austrian, and so on. For some of the economists working in these new research programs, their research provides a radical new (non-neoclassical) approach to the prediction and explanation of economic behavior, but even among those who are less radical—those who believe that some version of neoclassical theory will eventually be able to subsume these new developments—there still seems to be a consensus that the problems and anomalies these fields have identified are real and deserve the profession's attention. This is very different than had been the case for many of the criticisms traditionally raised by heterodox economists. The Marxian concern with the exploitation of the working class by the capitalist class, or the Veblenian distinction between business and industry, were for most mainstream economists, not real issues that deserved the attention of the discipline. This is very different from, say, the mainstream's response to the endowment effects, reference dependency, and other choice anomalies identified in the work of Daniel Kahneman, Amos Tversky, Richard Thaler, and others (see, e.g., Kahneman 2003; Kahneman, et al. 1991; Kahneman and Tversky 2000; Thaler 1980; Tversky and Kahneman 1991).[2] These concerns matter to mainstream economists in a way that most traditional heterodox concerns did not.[3]

There may also be changes underway within macroeconomics—changes initiated by what many see as the discipline's failure to predict, explain, or offer effective solutions for, the recent financial crisis—but I will focus primarily on microeconomic developments. There are a number of reasons for this. First, as I will argue later, microeconomics—

[2] One argument for the acceptance of these issues might be that some of these problems were recognized by the neoclassical economists of the ordinal revolution early in the 20th century. I have written in detail about this (Hands 2006, 2010, 2011), but it cannot be an argument for the recognition of these problems by the neoclassical mainstream, because there is essentially no recognition by contemporary economists that these same issues were also raised by economists during the ordinal revolution.

[3] One suggestion for why this has been the case is that while this literature has challenged the *descriptive-scientific* adequacy of mainstream theory, it accepts the mainstream view of rationality, i.e., the *normative* theory of what one *ought to do in order to be rational*. See Heukelom 2014 for a detailed historical discussion of this, and Sent 2004 for some other possible reasons for the mainstream attention.

individual choice theory in particular—is where much of the recent methodological research has been done—it is where the methodological action is, so to speak—and recent methodological research is the main focus here. Second, it is not at all clear at this point how, or if, macroeconomics will change. The changes taking place in microeconomics—whether they end up being revolutionary or reformist—have been ongoing for at least two decades and came mainly as a result of internal forces: the available laboratory and field evidence, new tools and ways of gathering data, and so forth. In the case of macroeconomics, the forces of change have been external—in the economy, not in economics—and have come quite quickly. The current crisis may end up having a profound impact on future macroeconomic theorizing in the way that the Great Depression did, but at this point that is not clear. Finally, given the particular features of the current crisis, *if* mainstream macroeconomics changes, it is possible that it will change back in the direction of Keynesian theory: not a new theory or a new methodological approach, but a revival of an earlier, and (at least on some readings of Keynes) once dominant, framework for macroeconomic analysis. This is quite different than in recent microeconomics where experimental and behavioral economists are now making it possible to do that which every influential methodological writer from John Stuart Mill, to John Cairnes, to Neville Keynes, to Lionel Robbins, to Milton Friedman, said was totally impossible—that is, experiments—and where neuroeconomics is adding new technology to render the previously immeasurable, now measurable.[4] It is useful also to note that this broadening of the base of acceptable approaches within mainstream microeconomics has occurred commensurate with a decline in the number of economists self-identifying with the traditional heterodox schools. This is not to say of course that institutionalist economics, or Marxist economics, or other heterodox schools have completely disappeared, but simply that while there are many economists critical of mainstream neoclassical practice, those who are, seem to be focused on particular problems, applications, and tools, rather than self-identifying with any general heterodox school of thought.[5]

---

[4] Although it is certainly possible to combine developments in experimental and behavioral economics with an analysis of the macroeconomic crisis. See, e.g., Heukelom and Sent 2010.

[5] See Dow 2010, or Lee 2009, for an alternative reading of the current situation in heterodox economics.

I want to explore this three-way relationship between orthodox economics, heterodox economics, and economic methodology during the last few decades. I will begin by characterizing how work in economic methodology related to orthodox and heterodox theory during (roughly) the period 1975-2000 and then turn to how this relationship has changed in recent years.

## ORTHODOX AND HETERODOX IN ECONOMIC METHODOLOGY: 1975-2000

Unlike most fields within economics, economic methodology does not have a standardized framework for inquiry; there are a wide range of approaches, styles, tools (from philosophy and elsewhere), as well as a wide range of goals (what it is the methodological research is supposed to "do"). Given this, how can I, in the space available, do justice to the methodological literature of the period 1975-2000? The truth is, I cannot, and for those interested in a detailed discussion of this literature I suggest a survey such as *Economic methodology: understanding economics as a science* (2010) by John Davis and Marcel Boumans, or my own *Reflection without rules* (2001). My focus here will be much more modest. I will focus on the relationship between orthodox and heterodox economics in the work of two influential economic methodologists during the second half of the 20th century: Mark Blaug and Terence Hutchison.[6] There were many others doing very different types of methodology during this period, but these two authors seem to be representative of the most influential work in the field (at least the work written by economists).

The first thing to notice about the methodological literature of this period is that it was based on what I have elsewhere called the "shelf of scientific philosophy" view of economic methodology (Hands 1994, 2001). Ideas from the (assumed given and stable) shelf of scientific philosophy were simply taken off the shelf and "applied" to the science of economics without reconfiguration or with much sensitivity to the peculiarities of the discipline. In the case of both Blaug and Hutchison, the relevant philosophical shelf was Popperian—based on Karl Popper's philosophy of science (1959, 1965, 1994)—and according to Popper in order to qualify as a real science a discipline needed to make bold

[6] A non-exhaustive list of their important contributions to the methodological literature includes: Blaug 1976, 1980a/1992, 1990, 1994, 2002, 2003; and Hutchison 1938, 1981, 1988, 1992, 2000, 2009.

(falsifiable, non ad hoc) conjectures and subject those conjectures to severe empirical tests.[7] Blaug and Hutchison both argued that while most economists claim to be engaging in this type of scientific activity, they in fact fail to do so: *economists do not practice what they preach*. Instead, economists are engaged in what Blaug called "innocuous falsificationism":

> I argue in favor of *falsificationism*, defined as a methodological standpoint that regards theories and hypotheses as scientific if and only if their predictions are at least in principle falsifiable, that is, if they forbid certain acts/states/events from occurring […] In addition, I claim that modern economists do in fact subscribe to the methodology of falsificationism: […] I also argue, however, that economists fail consistently to practice what they preach: their working philosophy of science is aptly characterized as "innocuous falsificationism" (Blaug 1992, xiii).

Such Popperianism offered tough standards—standards that Blaug and Hutchison argued economists could have, and should have, lived up to, but seldom actually did. It was an economic methodology that demanded that economists clean up their act.

There are of course many well-documented problems associated with Popperian falsificationism—in general, as well as when specifically applied to economics—but that is not my topic here.[8] The task here is not to evaluate these positions, but simply to try to characterize the general tone/attitude of the methodological discussion of this period (as represented by the work of Blaug and Hutchison) and relate it to orthodox and heterodox economics.

So what did the methodology of Blaug and Hutchison have to say about heterodox economics, or the relative scientific standing of orthodox and heterodox economics? On the face of it, quite a lot. Even a cursory examination of the methodological work of Blaug and Hutchison reveal that they directed a substantial amount of critical attention to

[7] Although it should be noted that neither Blaug nor Hutchison were entirely consistent about the substantive details of what a Popperian approach to economics would entail. For example, Blaug moved easily between advocacy of Popperian falsificationism and advocacy of Imre Lakatos's methodology of scientific research programs (MSRP). Although both approaches are broadly "Popperian", they are quite different in detail with Lakatos sharply differentiating his view from falsificationism, and Popper denying that MSRP was in any way Popperian. To be fair, it should also be noted that not all Popperians writing about economics (Larry Boland, for example) considered (or consider) falsificationism to be the proper interpretation of Popper's views.

[8] See Hands 2001, 275-304, or Hausman 1988.

heterodox theory of all persuasions: Marxian, institutionalism, post- and fundamentalist-Keynesianism, neo-Ricardian/Sraffian, Austrian, URPE-type late-1960s radical economics, and others.

Blaug began his career with a methodologically-inspired historical study of Ricardian economics (Blaug 1958) and he frequently criticized later Ricardians like John Stuart Mill for relying on introspection, ignoring the empirical facts of the mid 19th century British economy, and constructing various "immunizing strategies" to insulate Ricardian economics from empirical falsification (Blaug 1980a/1992). The Sraffa-based neo-Ricardians of the second half of the 20th century were also criticized on the same grounds, as well as for succumbing to "formalism" (Blaug 1990, 2009).[9] Blaug spent a substantial amount of time criticizing the labor theory of value and tendency laws (such as the falling rate of profit) in Marxian economics for not being falsifiable (Blaug 1980b, 1990) and noted Popper's own remarks about the unfalsifiability of the Marxian system (Popper 1976). Not to neglect the other side of the political spectrum, Blaug also had harsh methodological words for Austrian economists, particular Ludwig von Mises (Blaug 1980a/1992).

Similarly, Hutchison's first book (Hutchison 1938) was primarily a methodological critique of Lionel Robbins's *Nature and significance* (1932/1952), but it focused on the Austrian influence in Robbins's work. Hutchison continued to criticize Austrian economics throughout his life (Hutchison 1981) and while, like Blaug, the main methodological villain was von Mises, he included others such as Friedrich Hayek as well (Caldwell 2009). Hutchison criticized Marxian economics on grounds similar to Blaug's (Hutchison 1981) as well as the Cambridge-fundamentalist version of Keynesian economics (Hutchison 1981, 2009).

Based on all these criticisms, one might assume that Blaug and Hutchison used their Popperian methodology to defend the neoclassical mainstream against heterodox criticism. But that was not really the case. Both Blaug and Hutchison were just as critical of work in the neoclassical mainstream because it *also* was in conflict with the Popperian principles of bold conjectures and severe empirical tests. In particular, the formalist revolution which started during the 1950s and ended with the Arrow-Debreu abstract Walrasian general equilibrium theory that dominated microeconomics until quite recently, was harshly

[9] See Garegnani 2011, and Kurz and Salvadori 2011, for critical responses to Blaug on Sraffian economics.

criticized by both Blaug (1980/1992, 1997, 2002, 2003) and Hutchison (1992, 2000). For example, Blaug called the 1954 paper on the existence of competitive equilibrium by Kenneth Arrow and Gerard Debreu "a cancerous growth in the very centre of microeconomics" (Blaug 1997, 3) and Debreu's 1959 *Theory of value* "the most arid and pointless book in the entire literature of economics" (Blaug 2002, 27). Hutchison was only slightly more positive in his appraisal, calling general equilibrium theory the substitution of "fantasy content for realistic, or relevant, content" (Hutchison 2000, 18). But the criticism of neoclassical economics did not stop at the abstract Arrow-Debreu version of the theory. In fact, Blaug's survey of economic methodology (1980a/1992) was a veritable litany of criticisms of various aspects of the dominant neoclassical theory, with the eight chapters of Part III going topic by topic through standard theory from consumer choice, to production theory, to general equilibrium, to international trade, and so on, pointing out in each case how the theory failed to meet Popperian standards for scientific adequacy and/or progress. The only aspect of the mainstream theory of the day that Blaug seemed to give a positive nod was Keynesian economics, and even there he was critical of the "Mickey Mouse versions of Keynes in the 1950s" (1980a, p. 221) as well as the fundamentalist Cambridge versions of Keynesian theory. Hutchison was not quite as aggressive in his critical stance, but he too was critical of the formalism and lack of relevance of much of the dominant neoclassical theory (Hutchison 1981, 1992, 2000). Like Blaug, he was not very clear about exactly what kind of economics would meet the tough Popperian standards, but he was clear that both the neoclassical mainstream and heterodox theory were methodologically problematic.

The bottom line is that the Popperian "shelf of scientific philosophy" methodology of Blaug and Hutchison set the epistemic bar so high that essentially no economic theory could pass the scientific test. Although both Blaug and Hutchison probably favored the orthodox theory of the day—at least in its more applied, non-Arrow-Debreu, formulations—over various heterodox alternatives, it was a weak and frankly not very well-articulated preference since according to the methodological standards they endorsed, almost all economic theory was either unfalsifiable or false, and even the most serious empirical work was "like playing tennis with the net down" (Blaug 1980a, 256). The shelf of scientific philosophy approach was often defended as a "tough" approach to methodology, because it demanded compliance with a relatively strict set of methodological standards. For that reason it was often endorsed by those who sought to use it as a way to attack economic theories they did not support, but such a strategy was only effective as long as the critical fire was not turned back on one's own position (which, of course, it always could be). The toughness was explained as a kind of "tough love" because even though it was strict, it was ostensibly done in the interest of helping the economics profession be (epistemologically) all that it could be. Unfortunately, since no economic theory, orthodox or heterodox, really passed the test, the discipline was left without any guidance for how particular fields or models might be improved, or how the discipline's cognitive value could be increased at the margin.

The literature on economic methodology expanded significantly during the period 1975-2000—and for that we should be grateful since it helped establish economic methodology as a legitimate field—but it expanded in a way that prevented it from engaging in much constructive criticism, or in playing any significant role in the actual practice of economic theorizing, or in allowing orthodox theory to respond to the criticisms of heterodox economists (or vice versa) in any meaningful way.

## ORTHODOX AND HETERODOX IN ECONOMIC METHODOLOGY: THE RECENT LITERATURE

John Davis, my co-editor of *The Journal of Economic Methodology*, and others, have suggested that the mainstream of disciplinary economics is no longer neoclassical: that the once dominant neoclassical framework has been replaced by a new, more pluralistic, mainstream which is more open to psychology, less individualistic, accommodates various types of path-dependencies, and allows for a much broader class of modeling strategies and tools (Colander 2000; Colander, et al. 2008; Davis 2006, 2008, Santos 2011). As David Colander, Richard Holt, and Barkley Rosser put it: "Economics is moving away from a strict adherence to the holy trinity—rationality, selfishness, and equilibrium—to a more eclectic position of purposeful behavior, enlightened self-interest, and sustainability" (Colander, et al. 2008, 31). The most important piece of evidence for this change is the type of research that is currently being published in the most highly ranked economics journals: the *American Economic Review*, *Quarterly Journal of Economics*, *Economic Journal*, and even (although perhaps to a lesser extent) in the *Journal of Political Economy*. Another piece of evidence for this is that thirty years ago, most of the various specialty areas of research and teaching—labor economics, environmental economics, public finance, managerial economics, international economics, and the like—were simply particular "applications" of the standard neoclassical utility and profit maximizing framework. Now each of these fields is more likely to employ particular tools and conceptual frameworks that are indigenous, and in some cases endemic, to the particular subfield. International economics is now more than Walrasian general equilibrium theory with countries A and B replacing individuals A and B, environmental economists now need to actually know something about the relevant biological science, and so forth. Of course much of economic education—particularly undergraduate education—is still dominated by the neoclassical framework, but defenders of the "neoclassical is dead" thesis have tried to explain this in terms of lags and the institutional structure of the discipline (Davis 2006).

It is also important to note that the work identified with the new more pluralistic mainstream is not only not strictly neoclassical, it is also *not* heterodox either. Although many of the issues and anomalies identified in this recent literature have also long been identified by economists working within the heterodox tradition—think of the institutionalist critique of neoclassical choice theory or the institutionalist emphasis on evolutionary change, or the post-Keynesian or Austrian emphasis on path-dependency and hysteresis—the economists working in these new fields do not generally self-identify with heterodox schools of thought. For example, the histories of behavioral economics produced by practitioners (e.g., Camerer and Loewenstein 2004) often note Herbert Simon, James Dusenberry, and a few others from the middle of the 20th century, but do not generally cite any authors from the traditional heterodox literature. So too for earlier precursors. Behavioral ideas have been traced to Adam Smith (Ashrof, et al. 2005), David Hume (Sugden 2006), Jeremy Bentham (Kahneman, et al. 1997), and William Stanley Jevons and Francis Edgeworth (Bruni and Sugden 2007), but not to authors such as Karl Marx, Friedrich List, J. A. Hobson, or Thorstein Veblen. If there is a new more pluralist mainstream forming, it is neither neoclassical nor heterodox.

Although I am not quite as convinced as many others that the mainstream is no longer neoclassical, I do think the trend is clearly in that direction, and more importantly here, I definitely believe that a substantial change has taken place within *economic methodology*. In my book *Refection without rules* (2001) I argued that economic methodology was moving away from the "shelf of scientific philosophy" and more in the direction of naturalism, context-specific inquiries, and research that draws on a wider range of intellectual resources than just the philosophy of natural science. That process was ongoing at the time and has surely continued, but what was not clear a decade ago is how changes in *economics itself* have also initiated changes in the way that economic methodology is done. The bottom line is that almost all of the real "action" within contemporary economic methodology is in precisely the fields that Davis and others point to as elements of the new, more pluralistic, mainstream: neuroeconomics, experimental economics, behavioral economics, evolutionary economics; and the associated new tools such as computational economics, agent-based modeling, and various new empirical techniques. Neoclassicism may not be dead, but it is no longer the focus of the cutting edge of methodological research—but then nor is heterodox economics. Neither neoclassical nor heterodox economics are the main focus of recent methodological inquiry.

To provide some evidence for this claim about the recent methodological literature, let me just note a few of the methodological books published during the last few years that focus on a specific field, or small set of fields, within economics. A non-exhaustive list of such books would include those by Bardsley, et al. (2010), Guala (2005), Ross (2005), and Santos (2010). Notice that most of these books focus on experimental economics, but more importantly they all examine economic research in one or more of the new microeconomic fields. Also notice that they all focus on areas within economics that are neither heterodox nor strictly neoclassical. Finally, notice that these are also books with a normative philosophical focus—they are not (at least primarily) historical or sociological; they are philosophical—but again, it is a local or micro-philosophical focus, not the universal "one rule fits all science" approach of earlier methodological work like that of Blaug and Hutchison.[10]

[10] This emphasis on new more pluralistic fields is also reflected in recent methodological books with a broader focus such as: introductory textbooks (Reiss 2013), more general contributions to the philosophy of economics (Ross 2014), alternative methodological approaches to empirical research (Reiss 2007), or works concerned with philosophical ideas beyond epistemology and philosophy of science (Davis 2011). One exception might appear to be Hausman 2012—since it emphasizes questions about preference, choice, and welfare relevant to traditional neoclassical theory—but even here much of the discussion concerns behavioral and experimental economics.

As another, more personal, piece of evidence for this tendency, John Davis and I recently assembled a collection of papers by some of the most important contributors to the recent methodological literature: *The Elgar companion to recent economic methodology* (2011). The book has six sections: a section on methodological issues in contemporary choice theory, with papers on experimental economics, behavioral economics, and neuroeconomics; a second section on welfare economics, with many of the papers focusing on the economics of happiness and neo-hedonism; a third section on complexity, computational economics, and agent-based modeling; a fourth section on evolution and evolutionary economics; a fifth section on recent macroeconomics; and a final shorter section on the profession, the media, and the public. Notice that four sections out of six are dedicated to the areas of economics associated with the new pluralist mainstream in microeconomics. The last two sections are motivated in part by the recent macroeconomic and financial crisis and its impact on the profession (and the public's perception of the profession). The point is that when we attempted to put together a collection of papers that represented the best work in the most active research areas within recent economic methodology, we ended up with no papers on traditional neoclassical or heterodox topics.[11] This is not to say that none of the authors offered a methodological defense of neoclassical economics—a few did—but it was never the main subject. To me this is a nice example of the fact that not only has pluralism of intellectual resources replaced the once-dominant "shelf of scientific philosophy" within economic methodology, a new more pluralist mainstream has replaced the "neoclassical shelf of scientific economics" as the dominant domain of inquiry regarding the important questions and concerns for methodological inquiry.

As a final bit of evidence for these recent methodological trends, it is useful to look at what seems to be the most influential methodological research by economic practitioners, that is, economists who are not also contributors to the general methodological literature:[12]

[11] The possible exceptions, depending on how one defines orthodox and heterodox, are the four papers in the macroeconomics section.

[12] For example the various authors of Bardsley, et al. 2010 are all practitioners in experimental and behavioral economics, but since many of the authors are also regular contributors to the methodological literature, I listed this book as recent economic methodology (not practitioner's commentary).

Caplan and Schotter (2008).[13] Again, as with the methodological literature previously discussed, this book focuses on new pluralist areas like experimental economics, behavioral economics, and neuroeconomics. The volume contains the controversial "mindless economics" essay by Faruk Gul and Wolfgang Pesendorfer (2008) and a series of comments on that paper by economists who are practitioners in the relevant, or closely related, fields.[14] The Gull and Pesendorfer paper has been much discussed and elicits a wide range of responses, but it, and the commentaries on it, exhibit many of the same features as the recent literature from within the methodological community: the focus is on the new fields within microeconomics, it has a normative—but narrowly targeted—philosophical focus, and it exhibits a *pronounced disinterest* in most of the traditional methodological questions associated with either neoclassical or heterodox economics.

Two of the published responses from within the methodological community—Hausman (2008, 2012) and Ross (2011, 2014)—are quite different. Hausman is quite critical of not only Gul and Pesendorfer's methodological thesis, but also the revealed preference approach to choice theory on which it is based; while Ross is sympathetic to the revealed preference framework, but argues their methodological position needs to be strengthened in various ways.[15] Although the main subject of the Gul and Pesendorfer paper is behavioral and neuroeconomics, they end up defending what they call standard neoclassical economics (although they define neoclassical in a very idiosyncratic way). This said—and even though they are defending a view they consider to be neoclassical—their work, like the commentaries on it, and most of the recent research from within the methodological community, demonstrates that the "hot" methodological topics are in these relatively new microeconomic fields. The bottom line is that one does not need to be completely convinced that neoclassical economics has been displaced from its dominant position within the mainstream to recognize that the most interesting and important methodological questions are no longer about either traditional neoclassical or

[13] Another example is Smith 2009, but it explores a much wider range of topics.

[14] Only one of the contributors to the volume was a regular contributor to the methodological literature, the philosopher Daniel Hausman.

[15] My own view is that while contemporary revealed preference theory is an important tool in empirical demand analysis—and may prove to be useful in other areas of empirical economics as well—Gul and Pesendorfer's methodological use of this literature is extremely problematic. See Hands 2013a, and 2013b, for example.

heterodox economics, but rather, are about precisely the fields most often identified as representing a new more pluralistic mainstream.

This recent methodological literature is certainly less universalistic and more local, more naturalistic, and more sensitive to the particulars of the subfield within economics under investigation than the methodological literature of the period 1975-2000. Mark Blaug's book *The methodology of economics* (1980a/1992) provided a methodological assessment of various areas within economics, but the Popperian assessment tools were exactly the same for every single area. Do they make bold empirical conjectures and attempt to falsify them? If yes, then it is good science, and if no, then it is bad science (full stop). This is not the approach that is taken in most of the recent literature. A second point about this recent literature is that while it does exhibit the tendency to move away from the universalistic, and toward the particularistic, it is important that this movement does not imply an absence of philosophical rigor, a lack of normative assessment, or imply that anything goes. Not having a single narrow standard—what Deirdre McCloskey (1994) aptly called 3" x 5" card philosophy of science—does not mean having no philosophical standards at all. Again all of the works mentioned earlier are good examples of this.

## CONCLUSION

It is probably useful to conclude by summarizing the various parts of the argument I have presented. The earlier methodological literature like the work of Blaug and Hutchison was aggressively normative in its style, and negative in its assessment. The message was "this is what economists must do in order to produce scientific knowledge about the economy and economic behavior, and you (either neoclassical or heterodox) are not doing it". And yet the methodological rules it endorsed were offered at such an abstract and universalistic level, and so insensitive to the interests and concerns of the economists actually working in the various specific subfields within economic science, that it had essentially nothing to offer (either neoclassical or heterodox) practitioners about how disciplinary practice might be improved. There were very general injunctions to "test more" and "be more realistic", but there was no practical guidance to a group of economists working in a particular subfield struggling to extract as much knowledge as possible from the models and the data at their disposal while facing a wide range of subfield- and context-specific constraints. This is very different from the vast majority of the methodological literature of the last decade. For most of the recent research the domain of inquiry is neither neoclassical nor heterodox economics in general, but rather the many currently expanding subfields in microeconomics I have been discussing. In addition, it is not based on grand universalistic philosophy of science; it is applied philosophical inquiry aimed at the practical methodological issues of practitioners within specific subfields and sensitive to the issues, challenges, and constraints they face. It is important to note that while this more recent methodological work is local and close-focused, it is often critical—constructively critical—and it is *philosophy-based*. The argument that was often made in the earlier literature—Blaug 1994 is a good example—was that if one stepped down even a few steps from grand universalistic (and 3" x 5" card) rules for how all science must be done, one was necessarily on a slippery slope and will necessarily end up doing pure history, or sociology of science, or science studies, or some other type of inquiry that was not grounded in the (normative) philosophical justification of scientific knowledge and practice. Of course history, science studies, and sociological or anthropological studies of science (including economics) are interesting and important intellectual endeavors, but they do in fact have different goals, issues, and concerns than work grounded in normative philosophy. The point is that the recent literature in economic methodology clearly demonstrates that the entire slippery slope argument was an illusion. One can do local, subfield- and context-sensitive, studies in economic science that are philosophy-based and critical of current practice. Not only does one not need to give up on normative issues and philosophical justification, but one can produce work that actually offers the practicing economist some ideas about how knowledge production within specific subfields might be improved.

To conclude: there has been a lot of expansion and a lot of change within the field of economic methodology during the last few decades. During these years the field has changed its general philosophical focus from universal rules borrowed from the shelf of scientific philosophy to local practical advice grounded in the interests and concerns of particular sub-fields; and it has changed its domain of inquiry from neoclassical and heterodox economics in general to the more pluralistic microeconomic approaches at the edge of the current research frontier. Since interests always matter in the developmental path of any research program—within a particular science or within the study of a particular science—these changes will, and to some extent already have, contributed to the re-alignment of interests behind the field of economic methodology. My guess is that these changes will contribute to the steady growth and increased health of the field, but one never knows. Economic theorists have recently re-discovered path-dependency and the significance of context; we should not forget that these things matter to the future of economic methodology as well.

## REFERENCES

Ashrof, Nova, Colin F. Camerer, and George Loewenstein. 2005. Adam Smith's behavioral economics. *Journal of Economic Perspectives*, 19 (3): 131-145.

Bardsley, Nicholas, Robin Cubitt, Graham Loomes, Peter Moffatt, Chris Starmer, and Robert Sugden. 2010. *Experimental economics: rethinking the rules*. Princeton: Princeton University Press.

Blaug, Mark. 1958. *Ricardian economics: a historical study*. New Haven (CT): Yale University Press.

Blaug, Mark. 1976. Kuhn versus Lakatos, or paradigms versus research programmes in the history of economics. In *Method and appraisal in economics*, ed. S. J. Latsis. Cambridge: Cambridge University Press, 149-180.

Blaug, Mark. 1980a. *The methodology of economics: or how economists explain*. Cambridge: Cambridge University Press.

Blaug, Mark. 1980b. *Methodological appraisal of Marxian economics*. Amsterdam: Elsevier Science.

Blaug, Mark. 1990. *Economic theories, true or false?: essays in the history and methodology of economics*. Aldershot: Edward Elgar.

Blaug, Mark. 1992 [1980]. *The methodology of economics: or how economists explain* [2nd edition]. Cambridge: Cambridge University Press.

Blaug, Mark. 1994. Why I am not a constructivist: confessions of an unrepentant Popperian. In *New directions in economic methodology*, ed. R. E. Backhouse. London: Routledge, 109-136.

Blaug, Mark. 1997. Ugly currents in modern economics. *Policy Options*, 17 (7): 2-5.

Blaug, Mark. 2002. Is there really progress in economics? In *Is there progress in economics?*, eds. S. Boehm, C. Gehrke, H. D. Kurz, and R. Sturn. Cheltenham (UK): Edward Elgar, 21-41.

Blaug, Mark. 2003. The formalist revolution of the 1950s. *Journal of the History of Economic Thought*, 25 (2): 145-156.

Blaug, Mark. 2009. The trade-off between rigor and relevance: Sraffian economics as a case in point. *History of Political Economy*, 41 (2): 219-247.

Boland, Lawrence A. 1982. *The foundations of economic method*. London: George Allen & Unwin.

Bruni, Luigino, and Robert Sugden. 2007. The road not taken: how psychology was removed from economics, and how it might be bought back. *Economic Journal*, 117 (516): 146-173.

Caldwell, Bruce J. 1982. *Beyond positivism: economic methodology in the twentieth century*. London: George Allen & Unwin.

Caldwell, Bruce J. 2009. A skirmish in the Popper wars: Hutchison versus Caldwell on Hayek, Popper, and methodology. *Journal of Economic Methodology*, 16 (3): 315-324.

Camerer, Colin F., and George Loewenstein. 2004. Behavioral economics: past, present, and future. In *Advances in behavioral economics*, eds. C. F. Camerer, G. Loewenstein, and M. Rabin. Princeton: Princeton University Press, 3-51.

Caplan, Andrew, and Andrew Schotter (eds.). 2008. *The foundations of positive and normative economics*. Oxford: Oxford University Press.

Colander, David. 2000. The death of neoclassical economics. *Journal of the History of Economic Thought*, 22 (2): 127-144.

Colander, David, Richard P. F. Holt, and Barkley Rosser. 2008. The changing face of mainstream economics. *The Long Term View*, 7 (1): 31-42.

Davis, John B. 2006. The turn in economics: neoclassical dominance to mainstream pluralism. *Journal of Institutional Economics*, 2 (1): 1-20.

Davis, John B. 2008. The turn in recent economics and return of orthodoxy. *Cambridge Journal of Economics*, 32 (3): 349-366.

Davis, John B. 2011. *Individuals and identity in economics*. Cambridge: Cambridge University Press.

Davis, John B., and Marcel Boumans. 2010. *Economic methodology: understanding economics as a science*. London: Palgrave Macmillan.

Davis, John B., and D. Wade Hands. 2011. *The Elgar companion to recent economic methodology*. Cheltenham (UK): Edward Elgar Publishing.

Dow, Sheila C. 2011. Heterodox economics: history and prospects. *Cambridge Journal of Economics*, 35 (6): 1151-1165.

Friedman, Milton. 1953. The methodology of positive economics. In *Essays in positive economics*, Milton Friedman. Chicago: University of Chicago Press, 3-43.

Garegnani, Pierangelo. 2011. On Blaug ten years later. *History of Political Economy*, 43 (3): 591-605.

Guala, Francesco. 2005. *The methodology of experimental economics*. Cambridge: Cambridge University Press.

Gul, Faruk, and Wolfgang Pesendorfer. 2008. The case for mindless economics. In *The foundations of positive and normative economics: a handbook*, eds. A. Caplin, and A. Schotter. Oxford: Oxford University Press, 3-39.

Hands, D. Wade. 1994. Blurred boundaries: recent changes in the relationship between economics and the philosophy of natural science. *Studies in History and Philosophy of Science*, 25 (5): 751-772.

Hands, D. Wade. 2001. *Reflection without rules: economic methodology and contemporary science theory*. Cambridge: Cambridge University Press.

Hands, D. Wade. 2006. Integrability, rationalizability, and path-dependency in the history of demand theory. In *Agreement on demand: consumer theory in the twentieth century*, eds. P. Mirowski, and D. W. Hands. Durham (NC): Duke University Press [annual supplement to volume 38 of *History of Political Economy*], 153-185.

Hands, D. Wade. 2010. Economics, psychology, and the history of consumer choice theory. *Cambridge Journal of Economics*, 34 (4): 633-648.

Hands, D. Wade. 2011. Back to the ordinalist revolution: behavioral economic concerns in early modern consumer choice theory. *Metroeconomica*, 62 (2): 386-410.

Hands, D. Wade. 2013a. Foundations of contemporary revealed preference theory. *Erkenntnis*, 78 (5): 1081-1108.

Hands, D. Wade. 2013b. GP08 is the new F53: Gul and Pesendorfer's methodological essay from the viewpoint of Blaug's Popperian methodology. In *Mark Blaug: rebel with many causes*, eds. M. Boumans, and M. Klaes. Cheltenham (UK): Edward Elgar, 245-265.

Hausman, Daniel M. 1981. *Capital, profits, and prices: an essay in the philosophy of economics*. New York: Columbia University Press.

Hausman, Daniel M. 1988. An appraisal of Popperian methodology. In *The Popperian legacy in economics*, ed. N. De Marchi. Cambridge: Cambridge University Press, 65-85.

Hausman, Daniel. 2008. Mindless or mindful economics: a methodological evaluation. In *The foundations of positive and normative economics: a handbook*, eds. A. Caplin, and A. Schotter. Oxford: Oxford University Press, 125-151.

Hausman, Daniel. 2012. *Preference, value, choice, and welfare*. Cambridge: Cambridge University Press.

Hempel, Carl G. 1965. *Aspects of scientific explanation and other essays in the philosophy of science*. New York: Free Press.

Heukelom, Floris. 2014. *Behavioral economics: a history*. Cambridge: Cambridge University Press.

Heukelom, Floris, and Esther-Mirjam Sent. 2010. The economics of the crisis and the crisis of economics: lessons from behavioral economics. *Krisis*, (3): 26-38.

Hutchison, Terence W. 1938. *The significance and basic postulates of economic theory*. London: Macmillan.

Hutchison, Terence W. 1981. *The politics and philosophy of economics: Marxians, Keynesian, and Austrians*. New York: New York University Press.

Hutchison, Terence W. 1988. The case for falsificationism. In *The Popperian legacy in economics*, ed. N. De Marchi. Cambridge: Cambridge University Press, 169-181.

Hutchison, Terence W. 1992. *Changing aims in economics*. Oxford: Blackwell.

Hutchison, Terence W. 2000. *On the methodology of economics and the formalist revolution*. Cheltenham (UK): Edward Elgar.

Hutchison, Terence W. 2009. A formative decade: methodological controversy in the 1930s. *Journal of Economic Methodology*, 16 (3): 297-314.

Kahneman, Daniel. 2003. Maps of bounded rationality: psychology for behavioral economics. *American Economic Review*, 93 (5): 1449-1475.

Kahneman, Daniel, Jack L. Knetsch, and Richard Thaler. 1991. Anomalies: the endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives*, 5 (1): 193-206.

Kahneman, Daniel, and Amos Tversky (eds.). 2000. *Choices, values, and frames*. Cambridge: Cambridge University Press.

Kahneman, Daniel, Peter P. Wakker, and Rakesh Sarin. 1997. Back to Bentham?: explorations of experienced utility. *The Quarterly Journal of Economics*, 112 (2): 375-405.

Kurz, Heinz D., and Neri Salvadori. 2011. In favor of rigor and relevance: a reply to Mark Blaug. *History of Political Economy*, 43 (3): 608-616.

Latsis, Spiro J. (ed.). 1976. *Method and appraisal in economics*. Cambridge: Cambridge University Press.

Lazzarini, Andrés, and Diego Weisman (eds.). 2012. *Perspectives on epistemology of economics: essays on methodology of economics*. Buenos Aires: CIECE, UBA.

Lee, Frederic S. 2009. *A history of heterodox economics: challenging the mainstream in the twentieth century*. London: Routledge.

McCloskey, Deirdre M. 1994. *Knowledge and persuasion in economics*. Cambridge University Press.

Nagel, Ernest. 1961. *The structure of science*. New York: Harcourt, Brace & World.

Popper, Karl R. 1959. *The logic of scientific discovery*. New York: Basic Books.

Popper, Karl R. 1965 [1963]. *Conjectures and refutations*. New York: Harper & Row.

Popper, Karl R. 1976. *Unended quest: an intellectual autobiography*. LaSalle (IL): Open Court.

Popper, Karl R. 1994. *The myth of the framework: in defense of science and rationality*. London: Routledge.

Reiss, Julian. 2007. *Error in economics: towards an evidence based methodology*. New York: Routledge.

Reiss, Julian. 2013. *Philosophy of economics: a contemporary introduction*. New York: Routledge.

Robbins, Lionel. 1932. *An essay on the nature and significance of economic science*. London: Macmillan & Co.

Robbins, Lionel. 1952. *An essay on the nature and significance of economic science* [Reprint of 1935 2nd Edition]. London: Macmillan & Co.

Rosenberg, Alexander. 1976. *Microeconomic laws: a philosophical analysis*. Pittsburgh (PA): University of Pittsburgh Press.

Ross, Don. 2005. *Economic theory and cognitive science*. Cambridge (MA): MIT Press.

Ross, Don. 2011. Estranged parents and a schizophrenic child: choice in economics, psychology, and neuroeconomics. *Journal of Economic Methodology*, 18 (3): 217-231.

Ross, Don. 2014. *Philosophy of economics*. New York: Palgrave Macmillan.

Samuelson, Paul A. 1964. Theory and realism: a reply. *American Economic Review*, 54 (5): 736-739.

Samuelson, Paul A. 1965. Professor Samuelson on theory and realism: reply. *American Economic Review*, 55 (5): 1164-1172.

Santos, Ana. 2010. *The social epistemology of experimental economics*. London: Routledge.

Santos, Ana C. 2011. Behavioural and experimental economics: are they really transforming economics? *Cambridge Journal of Economics*, 35 (4): 705-728.

Sent, Esther-Mirjam. 2004. Behavioral economics: how psychology made its (limited) way back into economics. *History of Political Economy*, 36 (4): 735-760.

Smith, Vernon. 2008. *Rationality and economics: constructivist and ecological forms*. Cambridge: Cambridge University Press.

Sugden, Robert. 2006. Hume's non-instrumental and non-propositional decision theory. *Economics and Philosophy*, 22 (3): 365-391.

Thaler, Richard H. 1980. Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization*, 1 (1): 39-60. Reprinted in *Choices, values, and frames* [2000], eds. D. Kahneman, and A. Tversky. Cambridge: Cambridge University Press.

Tversky, Amos, and Daniel Kahneman. 1991. Loss aversion in riskless choice. *Quarterly Journal of Economics*, 106 (4): 1039-1061. Reprinted in *Choices, values, and frames* [2000], eds. D. Kahneman, and A. Tversky. Cambridge: Cambridge University Press.
Wong, Stanley. 1978. *The foundations of Paul Samuelson's revealed preference theory*. Boston: Routledge and Kegan Paul.

**D. Wade Hands** is distinguished professor of economics at the University of Puget Sound in Washington state and has taught history of economic thought for over thirty years. He has written on a wide range of topics in the history of economic thought and economic methodology. He is co-editor of *The Journal of Economic Methodology* and the author of *Reflection without rules: economic methodology and contemporary science theory* (Cambridge University Press, 2001). His *Agreement on demand: consumer choice theory in the 20th century*, edited with Philip Mirowski, was published in 2006 by Duke University Press and *The Elgar companion to recent economic methodology*, edited with John B. Davis, was published in 2011.
Contact e-mail: <hands@pugetsound.edu>

*Erasmus Journal for Philosophy and Economics,*
*Volume 8, Issue 1,*
*Spring 2015, pp. 82-97.*
*http://ejpe.org/pdf/8-1-int.pdf*

# Learning from the right neighbour: an interview with Jack Vromen

JACK J. VROMEN (Heerlen, 1958) is professor of theoretical philosophy, with a special emphasis on the philosophy of economics, dean of the philosophy faculty at the Erasmus University Rotterdam, and director (and co-founder) of the Erasmus Institute for Philosophy and Economics (EIPE). He earned master's degrees in economics and in philosophy of economics from the University of Tilburg, and a PhD in economics from the University of Amsterdam under the supervision of Neil De Marchi.

Vromen has a particular research interest in evolutionary thinking and economic methodology. He is the author of *Economic evolution: an enquiry into the foundations of 'new institutional economics'* (1995), and (co-)editor of numerous anthologies, including *Institutions and the evolution of capitalism; implications of evolutionary economics* (1999, with John Groenewegen), *The social institutions of capitalism: evolution and design of social contracts* (2003, with Hans van Oosterhout and Pursey Heugens), and most recently *The economics of economists* (2014, with Alessandro Lanteri).

EJPE interviewed Jack Vromen about becoming a philosopher of economics, his interest in evolution and its relation to economics, and the role he has played in the formation of EIPE, a major centre for the study of philosophy of economics. In this interview Vromen explains why he believes biology is a discipline much closer to economics than many economists realize, why the concept of evolution is important for understanding economic processes and for the economic discipline, and also why evolutionary economics never became mainstream when many believed that it would.

**EJPE:** ***You are an economist by training. How did you end up the dean of a philosophy faculty?***

JACK VROMEN: Well, I have to correct you there. I am not just an economist by training, I also did philosophy. In Tilburg University, there was the possibility of doing a double-degree program in philosophy of

**EJPE'S NOTE:** This interview was conducted by Willem van der Deijl and Vaios Koliofotis. Van der Deijl is a PhD candidate at the Erasmus Institute for Philosophy and Economics (EIPE) and co-editor of the *Erasmus Journal for Philosophy and Economics*. Koliofotis is a PhD candidate at EIPE, specializing in evolutionary explanations of economic behaviour.

economics next to your economics degree, which I did. So, I graduated in both. In fact, it was a little bit of a coincidence that I ended up doing my PhD in economics. Back then, it was required to choose a discipline in which your thesis was to be written in, a requirement that has now been dropped. If you look at my thesis, it is a little arbitrary that it ended up being a thesis in the field of economics. I think quite a few economists who read it at the time probably thought: is this really economics? In fact, one of the committee members at my defence asked me this very question. He told me that in a decent economics thesis there should be a model and an empirical test, and in my thesis there was neither a model nor a test. I think my work fell a little bit in between economics and philosophy. I think it probably would have qualified as a thesis in philosophy.

***So, you could have also ended up the dean of an economics faculty?***
[Vromen laughs] That does not follow. I did my PhD with Neil De Marchi, who was professor in economics in Amsterdam at the time. In those days, Amsterdam had a strong profile in philosophy and economics, starting with Johannes J. Klant, who preceded Mark Blaug's Popperian analysis of economics with his book *The rules of the game* (1984). The group was led by Mary Morgan, Mark Blaug, and later John Davis. I was part of that group. So, I have been part of an economics faculty, but not as a dean.

***Were there any particular thinkers, or texts, that influenced your early interest in philosophy and economics? When, how and why did you become interested in philosophy of economics?***
Starting with the last question, I started with doing a bachelor degree in econometrics actually, not in economics. My interest started with the building blocks, just the mathematics and statistics, without any applications. In the beginning I thought it was nice, but after a number of years it became too much for me. So, I switched to economics, which, in a basic sense, was just the lighter variant of econometrics. And then I decided to switch to philosophy of economics. Because there, I thought, I could find some answers to the questions I had when studying econometrics. While it was clear that the models used in econometrics rested on a number of assumptions, there was never any debate about the truth or reliability of the assumptions, neither in econometrics nor in economics. People just used them to derive useful applications. I

thought that there *should* be a debate about the assumptions. I took a few courses in philosophy and I thought it was *there* that these questions were being addressed.

When I did my undergraduate studies, I was intrigued by Habermas. I still think that for some purposes he has very interesting ideas. Not just historically—about how to think about the enlightenment—but also about the problems of contemporary societies. But when I moved on, and developed my own research projects that were related to philosophy of economics, I thought Habermas was not the right person to draw on. I do not think he really understood economics so well. He wrote about it from quite a large distance. The same applies to science in general.

At some point, I did develop an interest in the work of people who were trying to connect Habermas's ideas to science. One of them was Shaun Hargreaves Heap. He wrote a book in which he tried to connect Habermas to economics. I thought that was not bad at all. But I still think it was too distant from the discipline. If it comes to political or moral issues, I am still inspired by Habermas, but not for my work on philosophy of economics. So, I left it behind. But I was still very glad he visited the Erasmus University in October last year. I was still impressed by him. Not just by his brightness, but also by his overview on all kinds of things.

***So, you can say that your shift to philosophy was motivated by dissatisfaction with economics?***
Yes, definitely. Economics simply did not answer what I took to be very fundamental questions.

***Your specific research interest is evolution and economics. Your work can be divided into three parts. The first deals with industry behaviour, the second with the analysis of human behaviour in terms of evolutionary forces working on individuals and groups, and the third with types of explanation associated with evolutionary theorizing and modelling. What drew you to these topics and the theme of evolution?***
I was deeply interested in the issue of realism, or *realisticness*, of assumptions. The most central paper dealing with this in economic methodology is Friedman's The *methodology of positive economics* (1953). It occurred to me that there was already a large literature about

this paper at the time (in the mid or late 1980s). Most philosophers of economics had written something about it. What I took to be a very underdeveloped theme in Friedman was the evolutionary argument. While Mark Blaug had touched upon it, most methodologists neglected it completely. But I thought it was important, also as an important feature of the argument of the paper. At first I was very sceptical about the argument. It seemed to me that Friedman had only made it up for the occasion, to support the usage of unrealistic assumptions. I thought that if I wanted to assess the argument I should have a look at evolutionary theory itself in order to have a better grasp of it. I started to read books in natural selection and evolution, such as Elliott Sober's *The nature of selection* (1984). The funny thing is that the more I read about evolutionary theory, the more I thought: Friedman's argument is not at all that stupid; there is something to be said for it.

My supervisors advised me not to go in this direction because it was such an uncommon research topic. I ignored the advice for a while. I thought it was important and believed that it could become more important in the literature as well. And I was lucky, because when I finished there was quite some interest in it. My stubbornness had paid off.

So, this explains my first research interest—evolution in industry behaviour. I also started to explore other places where evolution and economics intersected. I discovered the work of Jack Hirshleifer (1977), who argued that it was strange that evolution entered economics at the level of industry behaviour. There is a direct unexplored link that could be drawn between evolution and individuals, and individual behaviour. Furthermore, as a philosopher with an interest in philosophy of science, I stumbled upon the work of Jon Elster (e.g., 1977) and Philippe van Parijs (e.g., 1981) on functional explanation, and to what extent functional explanation should be seen as a valid type of explanation outside of biology—in the social sciences. This opened up yet another area that I thought was interesting. So, I slowly found out that there is a variety of interesting links to be drawn between evolution and economics.

***You mentioned that you were originally motivated by a concern for the realism—or realisticness—of economists' assumptions. Is this still the link that connects these projects?***

Maybe a little, but it has moved to the background. At the time I was interested in the work of Uskali Mäki (who I did not know at the time) and Nancy Cartwright, and her book *How the laws of physics lie* (1983). The final chapter of my thesis was about the link between realism and evolution in economics. After my PhD I kept on working on this link, but it has now moved to the background.

***Do you think the link holds up? Does evolutionary thinking make economics more realistic?***
Yes, I still think so, but only if we understand evolutionary thinking in a specific way. Perhaps it is not biological evolution that is directly relevant for economics, but cultural evolution and related things. Friedman, and also Nelson and Winter (1982), believed that there was something like an evolutionary process going on in economic markets. And, to some extent, I still believe that this makes sense. Consider for example the idea that we should not assume that equilibria will be reached, or that they are already reached—as a working assumption—or that you should not assume that people are perfectly rational. These ideas are covered in evolutionary economics in an interesting way.

***Who is your work on the methodology of economics for? Do you write mostly for practicing economists or philosophers of science?***
Good question. I always hope, and try to make an effort, to engage with practising economists. In the work of evolutionary economists a lot of philosophical issues pop up and are sometimes explicitly addressed—but sometimes in an unsophisticated way. I would not be satisfied if only fellow philosophers paid attention to my work. I also participated in organizations like EAEPE (the European Association for Evolutionary Political Economy)—an umbrella for all sorts of heterodox schools of economics. Within EAEPE there have always been people—such as Geoff Hodgson—who have been interested in philosophical issues from the point of view of an economist. Unfortunately, this is quite exceptional. A lot of mainstream economists do not pay any attention to philosophy. They simply do not have an interest in it, thinking they can do without it. For evolutionary economists this is quite different. They think philosophical issues are important for their research practice. This makes it very interesting to address them in my academic work.

I also hope my work is interesting for philosophers of science. I have always thought that philosophy of science has a tendency to be very distant from actual scientific practice. It is generally very abstract, and does not address what practicing scientists actually think and do. I thought some improvements could be made there as well. I try to pay more attention to actual practice in economics, since this should be relevant for philosophy of science.

***Evolutionary economics in the Nelson and Winter tradition experienced rapid growth and presented itself as a radical alternative to neoclassical economics. However its ideas remain in the shadow of mainstream economics. Why do you think evolutionary economics never became mainstream?***

There are different stories to be told. One story is that Nelson and Winter type evolutionary economics was presented as a radical alternative to mainstream theorizing in economics, rather than an interesting addition. This was observed by William Baumol (1983), who wrote a review of Nelson and Winter's book. He wrote that it was interesting that the book was about things standard economics usually does not cover, and he believed it to be a contribution to the discipline. He did not see, though, why the authors felt the need to bash what they called 'orthodox economics'.

One problem is that both orthodox and heterodox economists did not completely feel at ease with evolutionary economics. It fell a little in between those camps. At the same time, I did meet a lot of economists in the 1990s who were very interested in evolutionary economics. Another contributing factor may be that this type of evolutionary theorizing in economics employed types of modelling that were not really *en vogue* with economists at the time, such as simulation modelling. Later on this became more commonly accepted in economics. In short, there are many different reasons, and it is still not very clear why evolutionary economics never really caught on.

***The scarcity of empirical work, and the lack of an overarching theoretical framework, could those be reasons?***

Yes, perhaps. Your latter suggestion goes in the direction of Geoff Hodgson and Thorbjørn Knudsen's (2010) attempt to formulate a theory of generalized Darwinism.

This might be an aspect, but I am not sure whether it is the full story, or even the main reason. I do not think it was so unclear what Nelson and Winter were arguing with regard to the general role of evolution in economics. At the same time, if it comes to the issue of attracting a critical mass of support among academic economists, Hodgson and Knudsen's attempt also fails. They present a generalized framework, but while there are many who find it interesting, these are mostly people who thought it was already there before their book came out.

What might also be relevant is that evolutionary economics was superseded by evolutionary game theory, which started to enter economics around roughly the same time—the late 1980s. Evolutionary game theory did really catch on. There were conferences and seminars in which the two were presented as alternative ways to make evolution relevant to economics. There was a clear preference for evolutionary game theory because it was much closer to the frame of mind of economists. So we can speak of a competition between the two research projects, and evolutionary game theory clearly won.

***So, there are a lot of different reasons. One is the ambivalent attitude of evolutionary economists towards mainstream economics. Is that something that you already saw at the time, or is it something you see now, looking back?***

I already observed it in my thesis. You could see that different people responded very differently to Nelson and Winter's book. There is a review by Philip Mirowski, who is very critical of standard economic theory, who criticized the book for not being radical enough. It was too close to standard economics.

If you look at the work of Richard Nelson, such as his work on innovation systems, it is very close to mainstream economics. In fact, this was already discussed explicitly in the book. They make a distinction between appreciative theorizing and formal theorizing. Formal theorizing comes in the form of equilibrium analysis and related practices. But appreciative theorizing is informal—sometimes in discussions that economists have or in working papers. It was the appreciative theorizing that Nelson and Winter wanted to formalize. This is closely related to work by Friedman, Alchian, and Machlup, who were all mainstream economists.

So, even in the book there was an ambivalence. Sometimes Nelson and Winter are very critical of the standard assumptions mainstream economists make, such as perfect rationality and the achievement of equilibrium—presenting themselves as radical reformers. At other times, they argue that the gist of their theory is already present in the appreciative theorizing of standard economics, and that they just aim to formalize this.

***Evolutionary economics has not converged in terms of theory. In a 2004 paper you noted that different authors have suggested a variety of ontological views, and that there is no ontological common ground in evolutionary economics—which seems to mean that they cannot agree on what they are talking about. Have things improved since the publication of your paper? Can the investigation of the foundations of evolutionary economics improve research in this field?***

As an answer to your first question: I failed miserably. It is indeed true that there was a debate about ontological issues in evolutionary economics. But, if you looked closely, they did not seem to agree on what ontology was supposed to be about.

There was a debate that came to be quite big within evolutionary economics. On one side there was the Hodgson camp, who argued that they presented *the* ontology for generalized Darwinism. On the other side, there were people such as Ulrich Witt, who argued for a different ontology based on continuity in evolutionary processes. The idea of continuity is that biological evolution produced intelligent creatures like us who can act deliberately but still have some remnants of the past in us, like tastes, preferences, etc. But the very fact that we are able to act deliberately shows, for Witt, that all these analogies to biological evolution are simply wrong: you cannot put human behaviour in a Darwinian framework of variation and selection.

What I noted is that both camps are really talking at cross-purposes. They are interested in different things. Hodgson, for instance, does not at all deny that there is such continuity. He even has a similar thesis in his book. What Hodgson has in mind are abstract principles. If you understand them in a very general way, they apply across the board, not only in biology but also in different systems. But this is not what Witt means when he uses the term ontology. In the article, I tried to show this. I analysed the different positions and different usages of the term ontology and see which positions are compatible and which ones are not. I tried to render them a service, but… mostly in vain.

***Recent empirical work on human behaviour has challenged the self-interest assumption of mainstream microeconomics while retaining***

***the rationality assumption. Can such experiments take us towards a unified theory of human behaviour incorporating insight from biology, economics, and psychology?***

I think it can in principle. But, it depends on what is meant by unified theory. It can mean that different theories are brought together in one framework. Herbert Gintis is known for doing this for the behavioural sciences (e.g., 2009). Another way of understanding unified theory is that it is a theory that many people working within the area of research accept. And the latter is something quite different. For example, Gintis thinks he can stick to a version of revealed-preference theory, but this is quite contested. Most behavioural economists want to get rid of this theory.

What you could have are different proposals for unified theories, but I do not see it happening sometime soon that there will be one that many people will accept.

There is also a large debate about whether rational choice theory and evolutionary theory amount to the same thing at some level of description. You could call this an attempt at unification as well, but of a very different type than Gintis's. There may be other sorts of attempts to arrive at a unified theory, but, given my experience, I am a little sceptical that one will succeed.

***The exchange of ideas and concepts between economics and biology is an important feature of your work. What is interesting for you in such an exchange between these two fields? What do you think economics can gain?***

A lot of methodologists of economics believe that economists try to imitate physicists. This started already with Adam Smith, who greatly admired Isaac Newton and wanted to be the Isaac Newton of economics. Philip Mirowski calls this 'physics envy'. Not only do economists imitate and emulate modelling techniques from physics, they even literally adopt exact equations from physics.

When I started to work on biology and evolutionary biology in particular, I started to think that the connection between economics and biology is much more natural, and much tighter, than the connection between economics and physics. I do not only mean that economists could learn more from biology, but also that economics and biology are very similar in the way the disciplines themselves evolved. To give an anecdote: we once invited Paul Krugman to an EAEPE conference with the specific question, "What can economists learn from biology?" He took a very serious look at biology and said: "What can economists learn from biology…? They are almost the same!"

The two disciplines are very similar indeed. Looking at certain modelling techniques in evolutionary biology, very similar questions arise about whether the processes converge to equilibrium or not. It looks very similar to economics. In a sense, economics has always been a 'population science'. As the debate about industry behaviour in Nelson and Winter illustrates, economics is mostly about aggregate behaviour rather than individual behaviour, just like biology. The similarities are so clear that Krugman concluded: economics *is* a biological science.

***So, economics should have less physics envy and more biology envy?***
I would not say it should be envious. But it is interesting to see that Darwin was inspired by Malthus. And the root notions of scarce resources and competition, which are, I think, the basis of the Darwinian idea of natural selection, can be found in Malthus's work. In many different respects there are really close similarities between biology and economics that should be acknowledged.

***In 1996 you founded the Erasmus Institute for Philosophy and Economics (EIPE) together with Uskali Mäki and Albert Jolink. Next year, EIPE will celebrate its 20th anniversary. How do you look back on this?***
I am very glad to have been part of it. The fact we started it is a little bit of a coincidence. At the time, there was a concentration of talents. Uskali Mäki, a strong figure in the debate about the realism of assumptions in economics, had just joined the philosophy department at Erasmus, and Arjo Klamer, who had been working on the rhetoric of economics with Deirdre McCloskey, had been hired by another faculty. At the time, these were the two hottest debates in the methodology of economics. So we were very fortunate to have these two scholars around. And there were more. Albert Jolink is a historian of economics who had written his thesis on Walras. Maarten Jansen is a game theorist trained in economics, econometrics, and philosophy, who was working as a methodologist of economics. John Groenewegen, an institutional economist, was also there, and so was Deirdre McCloskey—due to her close cooperation with Arjo Klamer—and later Mark Blaug. All these people were in the same place, and we thought it would be a waste not

to do something with this. We applied for some university subsidy and that is how we started. A large conference was organized around the inaugural lecture of Uskali Mäki. Nancy Cartwright, Mary Morgan, Kevin Hoover, Bruce Caldwell, Wade Hands and others all came, and that is when we launched EIPE.

There have been a lot of changes over the years. At first, EIPE was dominated by big names with strong views: Uskali Mäki, Deirdre McCloskey, and Mark Blaug. This has changed now. Also, EIPE is now less narrowly focused on methodology of economics. In general, philosophy of economics was, at the time, mostly about methodology; ethics had a much smaller part. So, there have been some interesting changes.

***Would you say that these influential thinkers—including also Roger Backhouse and Julian Reiss—have left their mark on EIPE?***

I think many people still associate EIPE with Uskali Mäki, Deirdre McCloskey, and Mark Blaug. Especially among those who have been in the field for a while. With Roger Backhouse it is a little different. He was only active at EIPE for a short while, before he started to write the biography of Paul Samuelson. Julian Reiss, though, has definitely left his mark on EIPE.

***What do you think about the role of EIPE in the philosophy of economics community?***

There are two well established journals in the philosophy of economics: the *Journal of Economic Methodology* (JEM) and *Economics and Philosophy*. The main competitor of EIPE as an institute is the philosophy and economics group at the London School of Economics and Political Science (LSE). The group at the LSE is mostly focused on decision theory, rational choice theory, game theory, social choice theory, etc. For that reason they have always had a close connection with *Economics and Philosophy*. EIPE, on the other hand, has always been associated with JEM.

At EIPE we had Blaug, who was a Popperian, and Klamer and McCloskey, who argued Popper should not be applied to economics. And Mäki's research interest lay with realism in economics. These are all central topics in the methodology of economics. Furthermore, EIPE has also been associated with sympathy for non-orthodox school of thought in economics, due to the work of John Groenewegen, as well as my own.

But this, I think, is now largely gone. At the time, there was a natural alliance between methodology of economics and criticism of mainstream economics. Most people considered it to be two sides of the same coin: if you went into methodology of economics, of course you were critical of mainstream economics.

***So there is a strong connection between EIPE and methodology of economics, but would you also say that there is a substantive position, something like an 'EIPE school'?***
Uskali Mäki's realism has been seen as the dominant approach at EIPE. Julian Reiss has somewhat different views, and that may have changed the image of EIPE as a realism centre. Although the differences may be smaller than some take them to be—I think Reiss also has some realist leanings.

***Looking at the future, do you have a vision about EIPE in 20 years?***
It will be flourishing, of course!

On a more serious note, the whole field of philosophy and economics has changed. There is a lot more ethics, which started with the work of Ingrid Robeyns. A lot of students are attracted to that. There is more formal work, related to decision theory and game theory, etc.—much more so than when we started. I could easily see that EIPE will be organized around a more diversified idea. There would be different areas that we try to cover with different people, much less focused only on methodology of economics.

I also think that there are nice opportunities to forge links again with the economists at Erasmus University, for example to work on the measurement of happiness, well-being, capabilities, and prosperity. Closer links can also be made with the behavioural economists, who are also prominent at Erasmus. So, there are all kinds of opportunities to connect more closely with economists.

***One of your published papers (Vromen 2009), the subject of your inaugural lecture and a later symposium which became a special issue of JEM, was about the 'economics made fun' genre. Could you explain how you came to be interested in this topic, and how it relates to your other interests? Should popular literature be an important part of philosophy and economics?***

The more general message of the inaugural lecture was that philosophers should try harder to understand what economists are doing and why they are doing it, rather than criticizing them from a distance. In this respect there is something interesting about the economics-made-fun genre. There are all kinds of jokes that economists make about their discipline.

I also showed a short clip of the stand-up economist Yoram Bauman, which I still think is very funny (see Vromen 2009, 74, n.5). The nice thing about him is that he is, to some extent, ridiculing parts of what economists do or think, but he does so with a thorough understanding of what they are saying and why they are doing it—without the condescending attitude that philosophers often have. This attitude—that economists must be crazy—is something I really dislike within philosophy of economics.

The economics-made-fun genre is interesting. If you just look at the name, you might think it is all about being funny. In some sense, though, the opposite is true. It is deadly serious. Often, authors writing in this genre argue that economics has a lot of interesting and worthwhile things to say about almost anything, and people had better pay attention to it. What struck me is that there was an interesting tension related to this project's timing. At this time, you could already see the financial crisis emerging. The popular image of economics was that it was not worth anything. No economists saw the crisis coming. They were all surprised—at least, that was the popular image. At the same time, Steven Levitt [the most popular economics-made-fun author] was telling the world that economists should be taken much more seriously than is commonly done, not only when it comes to standard economics topics, but also all sorts of other issues that are not typically part of the economics profession. This tension fascinated me.

It is interesting from a sociological point of view as well: why are economists doing this? How do fellow economists respond? It was interesting that Steven Levitt is a Chicago economist, and there were economists at Chicago who thought the economics-made-fun genre would attract exactly the wrong type of students. So, there was interesting debate within economics about this as well. This all made it very interesting.

***In light of the economic crisis, some economists have argued for more reflective courses in the curriculum and more space for philosophy***

***and history. Do you believe the influence of philosophy on the practice of economics has increased?***

This is hard to tell. There might be a little bit more openness. I think economists have become a little bit more self-critical. They have started to deal with philosophical issues. But what is striking is that they often do not turn to philosophers to address these issues. While I have argued that the term does not exactly apply, it is some sort of 'economics imperialism'. Itzhak Gilboa gave a talk at the last INEM conference [in Rotterdam, 2013], where he was dealing with the question "how do models relate to the world?". This is, of course, a key issue in the methodology of economics. When I saw an earlier version of his paper, I told him: "there are philosophers who are dealing with this issue as well", and his reaction was: "oh, really?".

Economists are dealing with philosophical issues themselves, but often do not know of the existence of our field, which I think is quite worrying. So, to answer your question, I think there is more interest among economists for philosophical issues, but this does not necessarily mean that philosophers of economics have a large say in economics. I think they should, but that is a battle still to be won. But, the first thing is that we should really engage with economists interested in these issues, and show them that we are around. Sometimes we only discuss these issues among ourselves, without reaching out. I think we should reach out more often.

***Could you tell us something about how philosophers of economics should do that? Your own work, for example, often seems motivated by the conceptual sloppiness of debates in economics.***

I would not like to say that this is the only way. But, indeed, what you often see with practicing economists is that they are typically very precise when it comes to mathematical or technical issues, but very sloppy indeed when it comes to conceptual issues. They simply do not care about this. And what often happens is that one economist means one thing with a term, while another means something else. Sometimes they do not even notice the difference. I really think they should pay more attention to these things. But, of course, this need not be the only way in which philosophers could enlighten economists.

There are two issues at stake. The first is what role philosophers could play in philosophy of economics, and the second is what role philosophers could play in economics. Philosophers can play both roles of course, but they are different. What I always rejected is the idea that philosophy of economics is applied philosophy—if applied philosophy means that you have general ideas in philosophy that can simply be applied to economics. I think that is completely the wrong way to go. One of the good things of the past decades is that within philosophy of science there has been a growing awareness about the peculiarities of different disciplines. Economics is really very different from sociology and other disciplines. I think it is really quite important for philosophers of economics to understand the economics that they are talking about. This often means that general ideas that philosophers have developed cannot simply be transferred to economics. You have to develop your own philosophical ideas to fit with practice in economics. The role of general philosophy in philosophy of economics should therefore be limited.

If it comes to the role of philosophy in economics, it is important not to have a condescending attitude. In my own work I try to avoid that. I certainly do not want to suggest that we should not take economists seriously because they are so conceptually sloppy. If you notice conceptual sloppiness you should also try to find out why this is the case, as I do in my work. Why do they not pay more attention to conceptual issues? Is it really fatal for their enterprise? Does it have disadvantages for what they want to accomplish? It is important to try to place yourself in the shoes of economists before you criticize them. So I think there is a role for philosophers of economics in economics, but hopefully without the attitude that we should teach them how to do economics.

## REFERENCES

Baumol, William. 1983. Book review of Nelson, R. R. and Winter, S. G. (1982). *Journal of Economic Literature*, 21 (2): 580-581.

Cartwright, Nancy. 1983. *How the laws of physics lie*. Oxford: Clarendon Press.

Elster, Jon. 1979. *Ulysses and the sirens: studies in rationality and irrationality*. Cambridge: Cambridge University Press.

Friedman, Milton. 1953. The methodology of positive economics. In *Essays in positive economics*. Chicago: University of Chicago Press, 3-43.

Groenewegen, John, and Jack J. Vromen. 1999. *Institutions and the evolution of capitalism: implications of evolutionary economics.* Cheltenham: Edward Elgar.

Hirshleifer, Jack. 1977. Economics from a biological viewpoint. *Journal of Law and Economics*, 20 (1): 1-52

Heugens, Pursey, Hans van Oosterhout, and Jack J. Vromen (eds.). 2003. *The social institutions of capitalism: evolution and design of social contracts.* Cheltenham: Edward Elgar Publishing.

Hodgson, Geoffrey M., and Thorbjørn Knudsen. 2010. *Darwin's conjecture: the search for general principles of social and economic evolution.* Chicago: University of Chicago Press.

Klant, Johannes J. 1984 [1979]. *The rules of the game: the logical structure of economic theories,* trans. Ina Swart. Cambridge: Cambridge University Press.

Lanteri, Alessandro, and Jack J. Vromen (eds.) 2014. *The economics of economists: institutional setting, individual incentives, and future prospects.* Cambridge: Cambridge University Press.

Levitt, Steven D., and Stephen J. Dubner. 2005. *Freakonomics: a rogue economist explores the hidden side of everything.* New York: William Morrow.

Nelson, Richard R., and Sidney G. Winter. 1982. *An evolutionary theory of economic change.* Cambridge (MA): Harvard Business School Press.

Van Parijs, Philippe. 1981. *Evolutionary explanation in the social sciences: an emerging paradigm.* Totowa (NJ): Rowman and Littlefield.

Sober, Elliott. 1984. *The nature of selection: evolutionary theory in philosophical focus.* Cambridge (MA): MIT Press.

Vromen, Jack J. 1995. *Economic evolution.* a*n inquiry into the foundations of new institutional economics.* London: Routledge.

Vromen, Jack J. 2004. Conjectural revisionary economic ontology: outline of an ambitious research agenda for evolutionary economics. *Journal of Economic Methodology*, 11 (2): 213-247.

Vromen, Jack J. 2009. The booming economics-made-fun genre: more than having fun, but less than economics imperialism. *Erasmus Journal for Philosophy and Economics*, 2 (1): 70-99.

*Erasmus Journal for Philosophy and Economics,*
*Volume 8, Issue 1,*
*Spring 2015, pp. 98-104.*
*http://ejpe.org/pdf/8-1-br-1.pdf*

# Review of Carsten Herrmann-Pillath and Ivan Boldyrev's *Hegel, institutions and economics: performing the social.* London: Routledge, 2014, 264 pp.

DON ROSS
*University of Waikato*
*University of Cape Town*
*Georgia State University*

The most fun an academic can have, at least on the job, comes from encountering a package of ideas one never expected to see that turn out to be deep and interesting. If, before I encountered Carsten Herrmann-Pillath's and Ivan Boldyrev's (henceforth, HPB) joint work at a conference two years ago, I had been asked to list canonical dead philosophers whose work might inspire fresh insights about current issues in economic methodology, I would have put Hegel near the bottom. Imagine anyone being so muddled about economic reasoning that that they could be set straight(er) by Marx! HPB's new book convinces me that this would have been a completely misjudged expectation.

By this comment I do not mean to endorse HPB's view that what both the academy and the policy world need now is a general embrace of Hegelian economics. Nor am I about to repeat the experience, which I recall with a shudder from much younger days, of actually reading Hegel's *Philosophy of right.* But I will go this far: the authors make a compelling case for the proposition that Hegel is the most important fountainhead for a coherent set of ideas about both economic behavior and political economy that, when expressed in an idiom closer to that of contemporary social science, deserve to be represented in both methodological and policy debates. Furthermore, as I will explain, going beyond anything to which HPB allude, if someone thinks (as I did) that the current German model of capitalism is largely a path-dependent consequence of Bismarck's cunning in designing a welfare state that needed paternalistic oversight by a Junker aristocracy, then HPB's book reveals that that is wrong too, or at least simplistic. German-style capitalism has deep intellectual roots, and they can be found in Hegel.

Before I get to economics, I will comment on HPB's basis for putting a convincing 21st-century gloss on Hegel's pure philosophy, which, as someone trained by analytic philosophers, I had previously found utterly archaic. Hegel, famously, goes on constantly about *spirit*, and, even more off-puttingly, about a kind of spirit he calls *objective*. This phrase smells like mysticism and, in its 19th-century context, nationalism to boot. Liberal cosmopolitan economists like me can hardly imagine a more repugnant mixture than that. But HPB make a convincing case that objective spirit is in fact Hegel's pre-Darwinian name for an element of the ontological furniture that so-called *externalist* philosophers of mind, following Tyler Burge (1986), Daniel Dennett (1987), Edwin Hutchins (1995), Ron McClamrock (1995), Andy Clark (1997), Radu Bogdan (1997, 2000, 2009, 2010, 2013), and Tad Zawidzki (2013), have established as central to an adequate science of human behavior: the socially scaffolded but primarily self-narrated *person*. I have argued repeatedly—but see especially Ross (2005, 2014)—that the standard story economists tell about the philosophical foundations of their most important theoretical concept, economic agency as inferred from revealed preference, snaps in a satisfying gestalt switch from incoherence to rich profundity if only one distinguishes such socially scaffolded but richly individuated people from the neural computers studied by psychologists and neuroscientists. Only then can we understand how preference consistency is stabilized as an achievement that simultaneously embeds normative individualism *and* identification with points in complex vector spaces of social markers. The relevant concept to replace the mind-as-internal-computer has unfortunately been established in the literature under the label of the 'extended mind' (Menary 2010). Since this suggests that the mind begins as an internal computer that is then accessorized, in terms of connotations the label is not much of an improvement on 'objective spirit'. So, HPB convince me, Hegel also anticipated the important philosophers listed just above in discovering a scientifically crucial conceptual insight and then botching its branding.

HPB update Hegel by reference to three main reference points that emerged subsequent to his death. First, his unrestricted teleology is replaced by the non-teleological but nevertheless developmental dynamics of Darwinian evolution, as generalized to apply to culture. Second, his highly abstract account of the formal stabilization of the social relations of free people is set into the context of contemporary industrial and post-industrial society, with the business corporation as the appropriate central exemplar, by filtering it through the work of

Masahiko Aoki (2001, 2010). Finally, Hegel's political economy, which according to HPB was a refinement of Adam Smith's theory of moral sentiments in light of a critical response to Kant's accounts of reason and morality, is mapped onto the concept space of contemporary economics according to the template that is fully worked out in Herrmann-Pillath's own recently published magnum opus, *Foundations of economic evolution* (2013). In light of all this updating, the reader might wonder whether the resulting comprehensive picture is 'Hegelian' merely by courtesy. HPB argue for Hegel's substantive primacy by reference to his originality with respect to all the picture's core elements. I find this case convincing for a reason they do not mention, and to which I will return below: HPB have effectively conjured the deep intellectual roots of the German model of capitalism; and whereas those roots can reasonably be associated with Hegel, it would make no sense to say that Aoki, or the leading modelers of cultural evolution, have elliptically been describing Germany.

On HPB's exegesis, Hegelian persons—that is, socially scaffolded ones—collectively participate in *institutions* that are subject to three primary constraints: *continuity*, *performativity*, and (reciprocal) *recognition*. Continuity is the denial of ontological dualism: people and groups of people are embodied in physical complexes that include their brains, and their potential actions and thoughts are constrained by the limits of these systems. Both of the systems in question, and therefore also the coupled system that arises from their continuity, are complex, in the full sense modeled by contemporary theorists of such systems. Performativity refers to the idea that theoretical models of society, including economic models, are realized by concrete actions—performances—which feed back upon the dynamics of theory articulation and change. Finally, people can only participate in institutions to the extent that they recognize others, and are recognized by others, as entities who enjoy subjective points of view on the basis of which they frame *choices* and are influenced by (partly) idiosyncratic preferences. Recognition reflects continuity, in that people can only construct themselves as individuated persons—in Hegel's language, can only achieve objective freedom—in response to recognition of their personhood by others.

This recognition is mediated and stabilized by institutions, in the absence of which mutual recognitions would be too ephemeral to serve as the basis for building concentrations of human, cultural and financial capital. The most important of these institutions, languages and money,[1] are created by people but not deliberately designed by them. The case of money is worth specific elaboration, particularly given its special interest to economists. Hegel, according to HPB, anticipated and inspired Georg Simmel's (1978 [1907]) conception of money as collectively structured scaffolding that makes valuations among many people, including strangers, commensurable. This function permits individuals to partly transcend their parochial social contexts, as it allows them, at least in principle, to build their extended selves out of any assemblage of available materials that (some) others can understand. Thus money is a collectively constructed institution that fosters modern normative and performed individualism.

This example nicely illustrates the central *apparent* paradox in Hegel's thought, namely, that people expand their freedom by forging their identities within the normative structuring of institutions. Hegel has often been ridiculed for this idea by thinkers who are not only *normative* individualists, but also descriptive and methodological individualists. For example, Bertrand Russell (1946, 701-715), writing about Hegel during the Second World War, saw the seeds of national socialism in the latter's philosophy and duly turned on Hegel the same mocking tone he took when writing about Nazis. Of course this was unfair, but appreciating Russell's motivations can assist us in taking the measure of Hegel's contemporary significance and relevance to economics. Institutions, and especially political institutions, serve liberating functions and they are *also* very often oppressive. The Nazi state surely tipped the net balance in favor of the latter, while democracies filled with ideological individualists, like Russell's England, create more opportunities than they foreclose. This raises an interesting possibility. Suppose that Hegel is right—as I join HPB in thinking he is—about the essential role of institutions in creating constitutive conditions for the development of modern (and, for that matter, post-modern) individuals. This allows that we might observe significantly different social forms—and, in particular, different forms of capitalism—depending on whether most people living in a society idealize the Hegelian character of their political economic dynamics, or

[1] In this list Hegel would have included the family and the state; but in contemporary rich societies the former is being crowded out in importance by professional and cultural networks, and the latter may also be shrinking in significance as it competes with rival forms of political organization.

invent implausibly individualistic mythologies about these dynamics and use this as a basis for normative institutional criticism.

To further frame the point at which I am driving here, let us consider one of HPB's examples of an economic institution that can be better understood in light of the Hegelian insight. Economists have often complained that international trade bargaining that proceeds by way of reciprocal tariff reductions is irrational. Since every country would usually (though not always) improve the average welfare of its citizens by unilaterally eliminating protectionist barriers, reciprocal commitments to lower tariffs resemble a situation in which one person agrees to stop punching herself in the face if another does likewise. But HPB use their Hegelian perspective to remind us that although tariffs are of course instruments for protecting comparatively disadvantaged domestic producers, they are *also* the principal mechanism by which market access rights are institutionalized and countries recognize one another as, among other things, agencies responsible for promoting development at national scales. They are therefore the natural focus of any institution, such as the WTO or a national trade ministry, that regards international trade as a *managed* system of relationships.

Free traders, of course, might wish that international trade relationships were *not* institutionalized in this way; but it is a fact of global political life that they *are*. On the other hand, by direct contrasting analogy, in a few countries, particularly the United States, individual people are largely left to sort out their access to labor markets on their own terms (with some ethical/institutional restrictions such as the ban on selling oneself into slavery or even selling one's labor below a mandated minimum wage). In Germany, by comparison, labor market access is strongly institutionalized through the role that unions and artisanal associations play in corporate governance. This model goes back to Bismarck's time and so has been highly resilient, especially in light of the upheavals in political structures that have occurred in Germany over that stretch of history. Thus, although the populations of both the USA and Germany benefit (extravagantly) from (culturally) evolved Hegelian freedom, the extent to which prevailing ideologies are consistent with this inheritance differs in ways that make for significant divergences in economic performance. (Each country, I would argue, tends to reliably out-perform the other on predictable and familiar dimensions of assessment, with the American economy being more dynamic and the German economy being less vulnerable to business cycle volatility.) If HPB's general interpretation of the philosophical history of economics is correct—and I indeed find it very persuasive—then there is valuable insight to be had from considering the German form of capitalism as reflecting a more pervasively Hegelian economic sensibility than the capitalism of the USA. (One could arrange other national capitalisms along a continuum stretched between them.) I find that to be an illuminating insight, potentially fecund with others, which could not have been obtained in the absence of HPB's highly original and rigorously constructed contribution to the philosophy of economics.

## REFERENCES

Aoki, Masahiko. 2001. *Toward a comparative institutional analysis.* Stanford: Stanford University Press.

Aoki, Masahiko 2010. *Corporations in evolving diversity.* Oxford: Oxford University Press.

Bogdan, Radu. 1997. *Interpreting minds*. Cambridge (MA): MIT Press.

Bogdan, Radu. 2000. *Minding minds*. Cambridge (MA): MIT Press.

Bogdan, Radu. 2009. *Predicative minds*. Cambridge (MA): MIT Press.

Bogdan, Radu. 2010. *Our own minds*. Cambridge (MA): MIT Press.

Bogdan, Radu. 2013. *Mindvaults*. Cambridge (MA): MIT Press.

Burge, Tyler. 1986. Individualism and psychology. *Philosophical Review*, 95 (1): 3-45.

Clark, Andy. 1997. *Being there*. Cambridge (MA): MIT Press.

Dennett, Daniel. 1987. *The intentional stance*. Cambridge (MA): MIT Press.

Herrmann-Pillath, Carsten. 2013. *Foundations of economic evolution*. Oxford: Edward Elgar.

Hutchins, Edwin. 1995. *Cognition in the wild*. Cambridge (MA): MIT Press.

McClamrock, Ron. 1995. *Existential cognition*. Chicago: University of Chicago Press.

Menary, Richard (ed.). 2010. *The extended mind*. Cambridge (MA): MIT Press.

Ross, Don. 2005. *Economic theory and cognitive science: microexplanation*. Cambridge (MA): MIT Press.

Ross, Don. 2014. *Philosophy of economics*. Basingstoke: Palgrave Macmillan.

Russell, Bertrand. 1946. *A history of western philosophy*. London: Unwin.

Simmel, Georg. 1978 [1907]. *The philosophy of money*, trans. Tom Bottomore and David Frisby. London: Routledge.

Zawidzki, Tad. 2013. *Mindshaping*. Cambridge (MA): MIT Press.

**Don Ross** is professor of economics and dean of management at the University of Waikato, professor of economics at the University of Cape Town, and program director for methodology at the Center for the Economic Analysis of Risk at Georgia State University. His areas of recent research include economic methodology; experimental economics of risk and time preferences in vulnerable populations; strategic foundations of human sociality; and scientific metaphysics. His many

publications include *Economic theory and cognitive science: microexplanation* (2005), *Every thing must go: metaphysics naturalized* (with James Ladyman) (2007), *Midbrain mutiny: the picoeconomics and neuroeconomics of disordered gambling* (with Carla Sharp, Rudolph Vuchinich, and David Spurrett) (2008), and *Philosophy of economics* (2014).
Contact e-mail: <don.ross931@gmail.com>

*Erasmus Journal for Philosophy and Economics,*
*Volume 8, Issue 1,*
*Spring 2015, pp. 105-109.*
*http://ejpe.org/pdf/8-1-br-2.pdf*

# Review of *The economics of economists: institutional setting, individual incentives, and future prospects*, edited by Alessandro Lanteri, and Jack Vromen. Cambridge: Cambridge University Press, 2014, 374 pp.

PHILIP MIROWSKI
*University of Notre Dame*

It is distressing these days to be a book reviewer. And no, I am not bemoaning the fact that no one under the age of 40 reads books anymore, due to having the attention span of a cocker spaniel after having spent their life surfing the Web. Rather, it seems that publishers do not care anymore to constrain the titles of books to have any bearing on the actual topics covered therein. Here—case in point—one might reasonably expect to find a gaggle of economists applying microeconomics to the behavior of economists, perhaps to praise the rational virtues of that most sagacious of agents, the model of a modern economist. That is what I expected when I agreed to review it. But no: what we have here is a jumble of disjoint exercises attempting to conduct a scattershot armchair *sociology* of economics, almost exclusively carried out by economists who have little time for or background in real sociology. If I wanted to sample a random selection of people spouting off on what is wrong with economics, I could always just turn to Google or the *Real world economics review* or blogs such as *Nakedcapitalism*. It is not clear to me why these particular papers warranted being collected together between these covers, lumbered with its misleading title, especially since so many of them had been published elsewhere previously.

Let me try to make the point about sociology in a terse manner. The only article in the book which really sets out to explain the shape of the modern economics profession is the superb chapter by a real sociologist, Marion Fourcade. That chapter, along with some more recent work, presents a plausible account of the global rise to power of the modern economics profession, in conjunction with the extraordinary intellectual contempt for other social sciences and parochial standards of argument so characteristic of its members. Fourcade adopts a comparative approach to the structures of epistemic authority and power, founded on data relevant to her thesis, as captured by this table from her paper "The superiority of economists" (Fourcade, et al. 2014).

Table 2 Response to the statement "In general, interdisciplinary knowledge is better than knowledge obtained by a single discipline" (in percent)

| American university professors in | Agree, strongly agree | Disagree, strongly disagree | No answer, don't know |
|---|---|---|---|
| Economics | 42.1 | 57.3 | 0.6 |
| Sociology | 72.9 | 25.3 | 1.8 |
| Political science | 59.8 | 28.0 | 12.2 |
| Psychology | 78.7 | 9.4 | 11.9 |
| Finance | 86.6 | 9.6 | 3.8 |
| History | 68.2 | 31.7 | 0.1 |

As one can observe, there is something different about American economists: they are more likely to disparage neighboring fields, and much more likely to enjoy an overweening confidence in their own epistemic legitimacy. After the global crisis, this really is quite extraordinary, and deserves to be a subject of inquiry. But what we find instead in the current book is further exemplification of that very hubris and insularity: various economists light upon some random aspect of their profession, and proceed to 'explain' it using the folk sociology of the natives. They ignore the work of people like Fourcade, as if she were not right there, in the book. Standards of evidence are lax; the big interesting questions are mostly evaded. And by this, I do not mean that they do not conform to a few stylized standards of hypothesis testing; rather, in most cases they write as though the economics profession were not embedded in the larger predicament of the modern university, nor reveal any interest in the more general sociology of knowledge.

I do not usually do this, but I am going to list the authors and topics, just to provide some feeling for how tone-deaf many of these contributions are. Arjo Klamer suggests that economists do not behave like self-interested agents, because they are really engaged in some sort of Habermasian ideal speech community. Perhaps things are different where he lives. Margit Osterloh and Bruno Frey list some 'disadvantages' of academic rankings in economics, hinting that it is encouraged by some modern regime of New Public Management. They seem oblivious to the large literature which traces the imposition of metrics in the modern university to the neoliberal imposition of stunted notions of 'competition' and 'quality' as a prelude to a marketplace of ideas, often spearheaded by economists. David Colander bemoans the imposition of

American metrics and standards in so-called 'reforms' of European economics departments, fearing the spread of mediocrity under the banner of standardization and competition. He should read some Fourcade, to begin to comprehend how power and epistemic authority have been operating since the 1960s; maybe even a little Foucault on neoliberalism. Wendy Stock and John Siegfried report on a survey of attitudes of 207 economists who earned their PhDs in 1996/1997. No surprise: they are a pretty smug and self-satisfied bunch. Wade Hands asks why economists do not experience the kind of priority fights that one observes in the natural sciences; reaching back to Mertonian sociology, he suggests that such flare-ups are rooted in emotions of moral indignation, and economists seem somehow devoid of collective notions of professional morality. This thesis is countermanded by the chapter by Deirdre McCloskey, who ascends the pulpit to condemn the various 'sins' of the economics profession from within the catechism of the seemingly amoral (but deeply neoliberal) stance of one D. McCloskey. Donna Ginther and Shulamit Kahn document that women economists fare worse in their academic careers than in other disciplines. Here identity politics blocks any consideration that the bigger immunities enjoyed by orthodox economists might also extend to immunities from social movements towards gender equality. Alessandro Lanteri and Salvatore Rizzello attempt to argue away the experimental literature which has found that undergraduate majors and graduate economists are more selfish and experience diminished solidarity with others. The authors suggest that students in such experiments are just telling their economist overlords what they think they want to hear. I think they should let some real psychologists have at the question, or at least read some Leon Festinger.

There are two interesting papers included herein that do not really seem to fit into even a broad conception of what this book purports to be about. Robert Frank reprises his theme song about what it means to teach the unwashed about how to 'think like an economist': mostly, it involves shoehorning some sort of cost-benefit analysis into the most unlikely of everyday situations. Jack Vromen takes the time to read Frank very carefully, and, in one of the better takedowns of the whole 'economics is just common sense' literature, he insists that, "the cost-benefit principle lacks specific content, [and] can be interpreted and applied in widely different ways, and therefore the first cost-benefit explanation that a student can think of need not be the only possible cost-benefit explanation that can be given" (p. 283). In other words, most of applied microeconomics is effectively empty, which is why it can be extended to any human experience. This is a salutary philosophical lesson; but I still fail to see how it contributes to an understanding of the modern economics *profession*.

I was thinking of ending the review by proposing that the next time Cambridge wants to put together an anthology of work about the shape of the economics profession, perhaps they should place it in the hands of a sociologist. But then I remembered that the latest hot thing in economic sociology has been Michel Callon's notion of 'performativity', which, crudely stated, says that economists reshape the world in the image of their theories, which explains their epistemic power in modern life. That whole program has turned out a bust, primarily because it retailed economists' own stories about their purported close coherence of theory and empiricism as if it were a 'radical' thesis, when in fact the target economic theory had rarely described how the constructed markets actually functioned 'in the wild'.

Interdisciplinarity in and of itself is no free-standing virtue. It is not that sociologists would naturally see the economists more clearly than (by the evidence of this volume) they see themselves; it is rather that a powerful sociology of knowledge requires the analyst to toggle back and forth between a number of competing perspectives. It would seek to render strange and precarious what the denizens believe is cozy and unexceptional.

The panoply of images and beliefs that support the dominance of the economic orthodoxy runs much deeper than simplistic notions of 'rationality' and 'cost-benefit analysis'; it is a regimen extending across many disciplines. What might be required is a weaving together of forms of knowledge, relations of power, and techniques of the discipline of the self (Foucault 2011). Where is Foucault when you really need him?

## REFERENCES

Foucault, Michel. 2011. *The courage of truth (Lectures at the Collège de France)*, ed. Frédéric Gros, trans. Graham Burchell. New York: Palgrave Macmillan.

Fourcade, Marion, Etienne Ollion, and Yann Algan. 2014. The superiority of economists. *MaxPo Discussion Paper* 14 (3). Max Planck Sciences Po Center on Coping with Instability in Market Societies, Berkeley, CA.

**Philip Mirowski** is Carl Koch chair of economics and the history and philosophy of science, and fellow of the Reilly Center, University of Notre Dame. He is author of, among others, *Machine dreams* (2002), *The effortless economy of science?* (2004), *More heat than light* (1989), *Never let a serious crisis go to waste* (2013), and *ScienceMart: privatizing American science* (2011). He is co-editor of *Agreement on demand* (2006), and *The Road from Mont Pèlerin: the making of the neoliberal thought collective* (2009), among other works. His next book, with co-author Edward Nik-Khah, will be *The knowledge we have lost in information: a history of the economics of information*, which will be accompanied by a video lecture series from INET.
Contact e-mail: <mirowski.1@nd.edu>

*Erasmus Journal for Philosophy and Economics,*
*Volume 8, Issue 1,*
*Spring 2015, pp. 110-115.*
*http://ejpe.org/pdf/8-1-br-3.pdf*

# Review of Till Düppe and E. Roy Weintraub's *Finding equilibrium: Arrow, Debreu, McKenzie and the problem of scientific credit.* Princeton: Princeton University Press, 2014, 304 pp.

S. ABU TURAB RIZVI
*Lafayette College*

In an all-night session in early 1697 Isaac Newton solved two problems posed by Johann Bernoulli and Gottfried Leibniz as a challenge to the mathematical community. Newton sent the solutions to the Royal Society for anonymous publication. Despite Newton's self-effacing gesture, Bernoulli recognized the author from the proofs, and proclaimed, "tanquam ex ungue leonem": we know the lion by his claw. In Till Düppe and Roy Weintraub's engaging and illuminating history of the mid-twentieth-century proofs of competitive general equilibrium and the three authors associated with them, the issue of the personality behind the proofs appears again and again. In this case, authorship is well known to us. Yet the role of personalities is hardly clear. The authors have raised an interesting problem, and enlighten us with their investigation.

Düppe and Weintraub's clearest contribution is to reveal the richness of the recently available archival material that surrounds the proofs. They provide evidence for what many have conjectured: that Arrow and Debreu, despite writing together one of the most famous papers in economic theory, had very different aims. Weaving together archival and interview material, the authors convincingly demonstrate that Arrow's goal in his joint work with Debreu was to stress 'economic meaning' and, at that point, 'verifiability', while Debreu stressed mathematical formalism, technique, and generality. They quote Debreu as later saying that such theorizing is not a "statement about the real world" but is only an evaluation of a model. These are commitments the two Nobel Prize winners carried throughout their careers. McKenzie's story is far less known and revelations about him and his work are particularly welcome. Düppe and Weintraub do a real service by showing the hesitations and setbacks McKenzie had to overcome to publish his paper, and the struggle for recognition that he engaged in only partly successfully. Düppe and Weintraub relate numerous pertinent and illuminating facts, events, and conversations that surround these authors and the circles in which they traveled.

The authors' methodological points will be more controversial than their expansion of the historical record. I will focus on three claims regarding the personality of authors and their work.

The first derives from the Mertonian observation that science aims for impersonal authoritativeness, yet credit for scientific discoveries is perforce personal. This sets up a tension that our protagonists have to address. Should they maneuver to gain credit or simply be pleased by their contributions to the advancement of an overarching enterprise? Arrow is the clearest exponent of the anonymity and sociality of the enterprise. He writes that economic research is "more and more a cooperative matter, requiring teams of individuals trained along similar lines". He says of the proof, "If I had not done it, somebody else would have" (p. 235). Düppe and Weintraub show that while Arrow tended toward this self-effacing position, Debreu and McKenzie, confidentially but persistently, hungered for the recognition and respect that comes from priority in publication—credit for work, appointments and recognitions at prestigious institutions, and prizes for a lifetime of achievement, most notably the Nobel Prize.

Düppe and Weintraub are successful at showing us the tensions that arise between the norms of science and personal ambition. As with Newton, personality seems hard to stamp out. If a claw can be seen, it denotes a lion, not a mere mortal. The authors take Arrow's attitude at face value. They write that in contrast to the backbiting of Debreu in his priority battle with McKenzie, "Arrow, instead, remained generous" (p. 212). Arrow may well have displayed such characteristics regardless of circumstance, but it is fair to point out that he got recognitions that others may have deserved, and got them more easily and earlier than the other two. From this more doubting perspective, it is possible to see Arrow's magnanimity as something he could afford to display. Similarly, McKenzie and Debreu's circumstances may have propelled them into more decisive action to advance their cases. Debreu was awarded the Nobel Prize eleven years after Arrow and had been denied tenure at Yale to boot; McKenzie suffered through a failed D.Phil. at Oxford, a sense of isolation and provincialism, and no Nobel Prize even though he had publication priority. That is, it may be that the record shows more what psychologists would call 'state' rather than 'trait'.

Nonetheless, Düppe and Weintraub do choose a side and make the claim that, pace Merton, it is primarily personality factors rather than the norms of science—'social forces' is their term—that result in battles for priority and credit (p. 240). Yet it is hard to disengage the two. They themselves point to institutional factors as being dominant in McKenzie's case, who was "acutely aware of his professional marginalization" and argue that he got less of a hearing because of his intellectual pedigree and institutional affiliation, for being an "academic outsider" (pp. 241-242). This conservatism of the economics profession contrasts with Martin Shubik's description of the atmosphere of the Princeton mathematics department, in which "[i]f a stray ten-year-old with bare feet, no tie, torn blue jeans, and an interesting theorem had walked into Fine Hall at tea time, someone would have listened" (pp. 94-95). McKenzie, and Debreu earlier on, had a much harder time getting someone to listen. Thus Düppe and Weintraub's claim that as a general rule it is more personality and less social norms at work seems hard to sustain. In fact, read more generously than their statement about Merton would suggest, we can see the authors arguing for a more complex causality or, as they term it, a mutual or reciprocal stabilization of the antagonistic demands of personality and work. Arrow, Debreu, and McKenzie struggled to produce work that was valued by the communities in which they found themselves without contravening important personal considerations, which were themselves molded by circumstance. This is why Düppe and Weintraub spend considerable space on the contexts of the work, exploring the activities of the RAND Corporation, the Cowles Commission, mathematics departments at Princeton and Chicago, and the like.

A second aim of Düppe and Weintraub is to repersonalize (p. xv) and depersonalize (p. xiii) economic theory at the same time. If this sounds complicated, it is, and the authors note the "apparent paradox" (p. xiii). But they have a coherent position. While they do want to repersonalize the equilibrium proofs, which means that we see the personalities behind them, they want to maintain the overall impersonality aspired to by economic theory as a science. They try to have it both ways.

Here is how they are able to argue in this manner. They "seek to understand the peculiar human practice of economic theory by viewing it less as a way of representing the world than as a way of dealing with the world" (p. xxi). What they mean is that economic theory is not about the world, about the economy, but is instead about models and the academic communities in which they make sense. Thus personal commitments about the world or the economy do not enter into consideration. On the other hand, personality and career allow the theorist to deal with the life of being an economist: thus matters like credit, prestige, priority, acceptance, preferment, influence, and the like figure very strongly in their account. These are worldly, social matters the theorist has to navigate. To repersonalize economics means to bring to the fore such career concerns.

Here, too, the situation is more complicated. They several times mention Debreu's view that the theorist deals with models, not with the economy. They also point out that modern economic theory differs from personality-laden schools, such as those associated with Schumpeter, Keynes, or Hayek, each of whom had a particular economic vision. Thus the economic theorist cares about a career, not about the economy, cares about the narrow world of personal ambition in a closed community, not about influence in the real world. The problem with this view is not that it is insular, verging on gossip and anecdote, but that it does not jibe with the view of one of the main protagonists. Arrow always felt that an economic theorist should deal with the world. After all, Arrow's case is a continuing contrast to Debreu that Düppe and Weintraub illuminate so well (e.g., pp. 196-203). (McKenzie does not say nearly as much as the other two about method and approach.) Arrow seems to have thought (mistakenly) that the models resulting in the proofs of the existence of competitive equilibrium could be modified to bring them into closer relation to the real world and that such developments would allow for the model's use and relevance. This is true as of the 1971 writing of *General competitive analysis*, Arrow's treatise with Frank Hahn, which tried to extend the basic proofs to debt, bankruptcies, imperfect competition, uniqueness and stability analysis, econometric identification, comparative statics, and Keynesian economics. It is fair to say that Arrow's hopes were dashed by the Sonnenschein-Mantel- Debreu results published in the following several years that showed that the competitive model had no particular implications at the aggregate level.

Thus, Arrow had particular ambitions for the proofs of the 1950s to be relevant in ways that it turned out they could not be: he says in 1986 that the hypothesis of rationality had few implications at the aggregate level (Rizvi 2006, 232). His personal commitments notwithstanding, he had not understood that the deep mathematical structure of the model would not allow for its elaboration along hoped-for lines, something that always seemed to have been clear to Debreu, as Düppe and Weintraub point out well (and was implied by the unheralded Hirofumi Uzawa 1962). While the authors seek to understand economic theory as less “representing the world than as a way of dealing with” it, Arrow, in his own words, wants to be a more “complete economist” and sees his economics as deriving from a “social conscience”, since “when we talk about studying people, we also talk about advising them” (p. 201). Thus he wants not just to deal with the world, but to represent it and intervene in it.

A third position follows from Düppe and Weintraub’s view that economic theory deals with models, not with personal (non-career) goals and visions for the use of economics in the world. This allows the claim that economic theory is universal, ‘authoritative’ in Düppe and Weintraub’s phrase, and not tied to particular visions. The analogy is with natural science. If science is discovery about nature, the personality or individuality of the discoverer is irrelevant. The authors contrast this view with the “differentiated collection of Marxian, Keynesian, neoclassical, Marshallian, Ricardian, Institutionalist, Austrian (and so forth) economists” who are particular and whose views bear an originator’s ‘personal stamp’ (p. xiii). They support their view by quoting Debreu who says that, “Even though a mathematical economist may write a great deal, it usually remains impossible to make, from his works, a reliable conjecture about his personality” (p. xiii).

But, as we have seen, personality means much more than behavioral traits. When it entails political and economic commitments, these might be discernible in mathematical economic theory. Düppe and Weintraub’s argument for the contrary is that it made sense to depoliticize economics in order to escape McCarthyite scrutiny (pp. 81-82). This stratagem does not explain the depoliticized nature of economics thereafter. There has to be more to the story. One wishes that Düppe and Weintraub had supported their view further. After all, it is quite a trick for one particular brand of economics to end up having the anonymity and universality of science, while its competitors are seen as personal and particular.

## REFERENCES

Rizvi, S. Abu Turab. 2006. The Sonnenschein-Mantel-Debreu results after thirty years. *History of Political Economy*, 38 (1): 228-245.

Uzawa, H. 1962. Walras' existence theorem and Brouwer's fixed point theorem. *Economic Studies Quarterly*, 13 (1): 59-62.

**Abu Rizvi** studies the methods and history of microeconomics. He has written in particular about the theories of games and general equilibrium. A book, co-authored with David Levine, *Poverty, work and freedom: political economy and the moral order*, explores another area of interest. He is provost and professor of economics at Lafayette College.
Contact e-mail: <rizvia@lafayette.edu>

*Erasmus Journal for Philosophy and Economics,*
*Volume 8, Issue 1,*
*Spring 2015, pp. 116-122.*
*http://ejpe.org/pdf/8-1-br-4.pdf*

# Review of Joseph Heath's *Morality, competition, and the firm: the market failures approach to business ethics*. Oxford: Oxford University Press, 2014, 424 pp.

CARSON YOUNG
*Wharton School of Business, University of Pennsylvania*

Anyone who has taught or taken a business ethics course will be familiar with the tired debate between shareholder theory and stakeholder theory. Shareholder (or stockholder) theory is almost always represented by a Milton Friedman opinion piece from a 1970 issue of *The New York Times Magazine* that traditionally plays a role in the business ethics classroom comparable to that of the 'heel' in a professional wrestling match. It announces in its title the view it purports to defend: "The social responsibility of business is to increase its profits" (Friedman 1970). (Friedman's piece actually contains a variety of arguments that are difficult to reconcile with each other, which makes his ultimate views on the social responsibility of business both more nuanced and harder to pin down than the title suggests.) Swooping in to rescue business from Friedmanite moral laxity is usually an article expounding stakeholder theory by its founding father, R. Edward Freeman, who claims that the firm's managers should advance the interests of *all* of a firm's stakeholders, not merely those of shareholders, as an ultimate goal.

Shareholder theory and stakeholder theory have both generated enormous literatures. When reading through the three decades' worth of contributions to this literature, one gets the sense that there is little left to say. The shareholder-stakeholder debate has grown stale, but it never reached a particularly satisfying resolution. The big questions that originally set off the debate remain open: What are the ethical responsibilities of the corporate manager? What factors must the corporate manager take into account, and how, in order to run the corporation ethically?

Over the past decade, University of Toronto philosopher Joseph Heath has written a series of papers that put forward a new way of thinking about these central questions of business ethics. Heath's work features extensive use of economic theory. This is relatively common in

**AUTHOR'S NOTE:** I thank Wayne Norman for his guidance in writing this review. I also thank Eric Orts and Nicolas Cornell for their helpful comments.

business ethics. Indeed, as I have mentioned, one of the articles that gave rise to business ethics as an academic discipline was written by none other than Milton Friedman. But usually, economic approaches to business ethics are used to argue for extremely minimalist views about the ethical obligations that apply to corporate managers. What separates Heath's work from much of the previous business ethics literature is his extensive use of economic theory to justify a much more demanding set of ethical norms for business. *Morality, competition, and the firm* is a collection of essays on his novel alternative to stakeholder and shareholder theories, which he dubs the 'market failures' approach to business ethics. (Six of the chapters in the book develop and defend core elements of this approach, and the remaining eight chapters address related subjects relevant to the evaluation of markets, firms, and market agents.)

Heath's book is essential reading for scholars and students interested in new ways of thinking about the foundations of business ethics. But the themes in the book are also likely to be relevant to scholars working at the intersection between ethics, political philosophy, political economy, and economics more broadly. There is tension, to put it mildly, between mainstream views in political philosophy and mainstream views in economics. What is distinctive about Heath's work is that it links mainstream egalitarian views about justice in political philosophy to certain aspects of mainstream thinking about economics. Indeed, one of the most important chapters in the book, "Efficiency as the implicit morality of the market" (which I will briefly discuss later), gives an explicit and detailed account of how norms of economic efficiency are compatible with a commitment to a strict egalitarian theory of justice.

Before continuing, it is worth mentioning a few of the book's minor flaws, all of which stem from the fact that it is a collection of essays rather than a unified monograph. First, some of the chapters overlap with each other enough that they feel repetitive. Second, nine of the book's fourteen chapters were previously published elsewhere, so those already familiar with Heath's work may find it somewhat redundant. Third, the book's main themes might be easier to follow if the connections between the chapters were more explicit. That said, this third quibble is largely mitigated by the book's excellent introduction, in which Heath gives an intellectual history of his project as well as a survey of the field of business ethics as he sees it. Heath is a master at distilling the main literature on a topic into a gripping intellectual narrative in order to set up his own substantive arguments, and the introduction provides a fine example of Heath's expository savvy.

The first of the book's three sections contains the 'greatest hits' of Heath's previous work on business ethics. These are the articles in which Heath developed the fundamental ideas of the market failures approach. "A market failures approach to business ethics", originally published in 2004, was Heath's first foray into the field. Stakeholder and shareholder theorists had been arguing for decades about whether a corporate manager's obligations extend beyond maximizing profit. Heath's insight was that adherents to both of these approaches, with few exceptions, attempted to defend views on the ethical status of profit maximization without considering how the profit motive's role in the broader economic system is (or can be) justified in the first place.

> It is in seeking to justify the profit motive that we discover that the appropriate form of managerial responsibility is not to maximize profits using any available strategy, but rather to take advantage of certain specific opportunities for profit (p. 26).

This starting point led Heath to argue for a more subtle version of Milton Friedman's defense of profit maximization. Heath gives his own reconstruction of Friedman's argument, claiming that in order to be consistent with the underlying economic logic on which he relies, Friedman cannot defend all forms of profit seeking, since

> managers have no right to take advantage of market imperfections in order to increase corporate profits. The set of permissible profit-maximizing strategies is limited to those strategies that would be permissible under conditions of perfect competition (p. 34).

This line of argument is further developed in two chapters that were both originally published as journal articles in 2006. In "Business ethics without stakeholders" Heath writes,

> [P]rofit is not intrinsically good. The profit-seeking orientation of the private firm is valued only because of the role that it plays in sustaining the price system, and thus the contribution that it makes to the efficiency properties of the market economy as a whole. Ideally, the only way that a firm could make a profit would be by employing one of the preferred [non-market-failure-exacerbating] strategies. However, for strictly practical reasons, it is often

> impossible to create a system of laws that prohibits the non-preferred ones. Thus according to the market failures perspective [...] the ethical firm does not seek to profit from market failure (p. 89).

Heath makes a similar argument with somewhat different points of emphasis in "An adversarial ethic for business: or, when Sun-Tzu met the stakeholder", which stresses the importance of accounting for the adversarial structure of the market when developing a theory of business ethics.

In addition to the above three chapters that primarily focus on ethical obligations in an extra-firm context, the book's first section also has two chapters on corporate governance. "Stakeholder theory, corporate governance, and public management" (co-authored with Wayne Norman) discusses the governance problems that arise when firm management has a strong duty to serve different stakeholder groups beyond shareholders. "Business ethics and the 'End of History' in corporate law" includes a sympathetic discussion of Henry Hansmann and Reinier Kraakman's defense of shareholder primacy, which Hansmann and Kraakman argue is best understood as a special case of *owner* primacy, and as such would apply equally to worker-owned co-ops, or to tenant-owned condominiums (Hansmann and Kraakman 2003). The chapter concludes with an argument that Hansmann and Kraakman's endorsement of shareholder primacy is too strong, given their premises. Heath claims, "if there is a conflict between the interests of various constituency groups, management should assign priority to the interests of shareholders", but that when "the conflict is one between the interests of shareholders and the principle that managers should refrain from taking advantage of market power in dealing with other constituencies, then the principle trumps the interests" (p. 141).

In the background of all of these articles is Heath's commitment to the idea "that the market is essentially a staged competition, designed to promote Pareto efficiency" (p. 5). This will strike most readers as an implausibly minimalist normative principle. What is so great about Pareto efficiency? In perhaps the most important non-previously published chapter in the book, "Efficiency as the implicit morality of the market", Heath explains why, even if one were starting from a G. A. Cohen-style strict egalitarian theory of justice, there are good reasons to conclude that "the guiding idea in business ethics should be the principle of Pareto efficiency" (p. 173). Heath's explanation for this is technical and much too complex to explain adequately here, but the basic idea should be familiar to anyone who has read Rawls. The reason Rawls adopts the Difference Principle rather than strict egalitarianism is that, because of incentive problems, a principle that allows for certain economic inequalities will be better for the least well-off than a stricter egalitarian principle (Rawls 1971). Likewise, Heath thinks that given the incentive and information problems that plague egalitarian and prioritarian principles, we should adopt the Pareto principle for evaluating market behavior.

The book's remaining chapters cover topics somewhat outside the book's central themes, but they are worth reading for those with relevant interests. In "The benefits of cooperation", Heath argues that there are five main mechanisms through which cooperation yields benefits: economies of scale, gains from trade, risk pooling, self-binding, and information transmission. Clearly distinguishing these mechanisms, Heath claims, is vital for understanding the normative foundations of the welfare state. "Contractualism: micro and macro" argues that there is a tension between versions of contractualist ("social contract") theories of justice that take small group interactions as an analytical point of departure (such as David Gauthier's, see Gauthier 1986) and those that focus instead on society as a whole (such as John Rawls's, see Rawls 1971). Microcontractualist theories are unable to provide principles that ensure justice at the society-wide level, while macrocontractualist theories lack the resources to generate principles that ensure justice in small-scale, particular interactions. Heath proposes a contractualist framework that he claims can resolve this puzzle. "The history of the invisible hand" examines the evolution of the invisible hand argument from Adam Smith to Friedrich Hayek. As its title suggests, "The uses and abuses of agency theory" argues against some prominent conceptions and applications of agency theory and suggests how agency theory should be understood and employed. In "Business ethics and moral motivation", Heath criticizes certain folk theories of moral motivation that have been popular among business ethicists and suggests criminological literature as a promising resource for insights into the causes of unethical behavior in organizations. "Business ethics after virtue" urges business ethicists to "put virtue theory behind us once and for all" (p. 323). Finally, in "Reasonable restrictions on underwriting", Heath looks at insurance markets, which he argues are different in significant ways from other markets. He shows how these differences can lead to otherwise counterintuitive conclusions. For example, he defends the position that (with some caveats) it is morally permissible for private insurers to charge individuals different insurance premiums based on statistical predictions about how expensive those individuals will be to insure.

As Heath recognizes, the market failures approach to business ethics remains very much a work in progress. So before concluding, I will mention one deficiency that adherents to the view need to address. In the book's first chapter, Heath writes that "[t]he firm should behave as though market conditions were perfectly competitive, even though they may not in fact be" (p. 37). This principle entails a variety of restrictions on how firms may pursue profit. For example, ethical firms must "[m]inimize negative externalities" and "[t]reat price levels as exogenously determined" (p. 37). Heath recognizes that we cannot hold firms to these ethical standards in our non-ideal world, since any firm that abided by them would be unable to survive in a competitive marketplace. However, as Heath acknowledges, we cannot even claim that firms should abide by these constraints as best they can given the competitive pressures they face because of the 'second-best theorem' (Lipsey and Lancaster 1956). The second-best theorem implies that if there is a distortion in the market, the most efficient possible outcome may require introducing other market distortions. Therefore, one cannot straightforwardly apply principles derived under ideal assumptions to a non-ideal context in which those assumptions do not hold. Heath writes in the book's introduction that he "would someday like to address more thoroughly […] the non-ideal aspect of the theory" (p. 20). I hope he does, because as long as its non-ideal aspect remains unresolved, the market failures approach risks being strictly academic. A good approach to business ethics should be able to give concrete, practical guidance to a firm manager who wants to do business ethically.

One could quibble about *Morality, competition, and the firm* being a collection of articles more so than a coherent, unified book. But it contains such a wealth of challenging and thought-provoking ideas about the ethics of business and economics that I believe those interested in these fields will find it worthy of their attention.

## REFERENCES

Friedman, Milton. 1970. The social responsibility of business is to increase its profits. *New York Times Magazine*, September 13.

Gauthier, David. 1986. *Morals by agreement*. Oxford: Oxford University Press.
Hansmann, Henry, and Reinier Kraakman. 2003. The end of history for corporate law. In *Convergence and persistence in corporate governance*, eds. Jeffrey Gordon, and Mark J. Roe. Cambridge: Cambridge University Press, 33-68.
Lipsey, Richard, and Kelvin Lancaster. 1956. The general theory of the second best. *The Review of Economic Studies,* 24 (1): 11-32.
Rawls, John. 1971. *A theory of justice*. Cambridge (MA): Harvard University Press.

**Carson Young** is a PhD student in the Ethics and Legal Studies department at the Wharton School of the University of Pennsylvania.
Contact e-mail: <carsony@wharton.upenn.edu>

*Erasmus Journal for Philosophy and Economics,*
*Volume 8, Issue 1,*
*Spring 2015, pp. 123-126.*
*http://ejpe.org/pdf/8-1-ts-1.pdf*

# PHD THESIS SUMMARY: Using case studies in the social sciences: methods, inferences, purposes

ATTILIA RUZZENE
*PhD in philosophy and economics, November 2014*
*EIPE, Erasmus University Rotterdam*

Despite fads and fashions in the academic culture, case-based reasoning has proved to be a persistent form of analysis in the social sciences, in the humanities, and even in moral thinking. Broadly understood, case-based reasoning locates the ultimate source of our epistemic and moral intuitions in the concreteness and idiosyncrasy of particulars. Even though they can be traced back to a common root, different traditions of reasoning with cases and of using case studies coexist in the academic landscape. This thesis focuses primarily on the use of case studies in the social sciences as an epistemic strategy to formulate, establish, and generalize causal hypotheses. A secondary goal is an investigation into the use of causal findings generated within and by means of case studies to inform policy making in the social realm.

The thesis is organized in four chapters. In chapter 1, I characterize what can be regarded as two alternative views of case studies and the understanding of science in which they are embedded. The first approach flourished in the 1970s and looked at case studies as a special, and typically weaker, form of the experimental, statistical, or comparative methods. Since this approach tends to evaluate case studies by criteria belonging to other methodological traditions, it can be said to present a *heteronomous* paradigm. The second, alternative view, which was developed in the last decades, is taking shape gradually and is still far from being fully articulated. This approach strives for an understanding of case studies liberated from the narrow mindset that caricatures case studies as the method of last resort. In particular, it sees case studies as an *autonomous epistemic genre* (Morgan 2012).

In chapter 2, I address internal validity in historical narratives. Historical narratives are case studies that aim to formulate and substantiate causal hypotheses by articulating descriptions of the sequences of events leading to the outcome of interest. They typically make use of process-tracing to draw causal inference, and often rely on the additional use of the methods of comparison. Despite the important role of historical narratives in the social sciences, how process-tracing operates in the narratives is still poorly understood. The debate on process-tracing in fact, even though it is growing thanks to a number of recent contributions, is still muddy and under-developed. In particular, there are no shared criteria to assess its epistemic contribution; moreover, the conditions proposed so far tend to tie the validity of the findings to the use of specific kinds of evidence and are thus unhelpful when this specific evidence is not available.

I argue that the proposed conditions are unduly restrictive and fail to acknowledge the actual contributions process-tracing can offer to valid causal inference. I formulate new conditions to assess process-tracing performance in cases in which the favourable evidential circumstances do not occur and existing criteria fail to apply.

In chapter 3, I address the problem of generalizability. I provide an outline of what I define as the traditional view on external validity. This approach is conditioned by a statistical viewpoint on case study research (CSR) and reduces external validity to issues of mere representativeness. In so doing it leads the debate on the generalizability of case-study results to a dead end as it quickly dismisses external validity as the downside of CSR. At the same time, it suggests that CSR is comparatively stronger in providing internally valid results. On this ground this approach recommends the use of case studies when internal validity is the main research goal of interest, while turning to other methods when one pursues generalizations instead. This outcome is unfortunate because, as a matter of fact, case studies are often performed with the explicit or implicit purpose of drawing lessons from the studied case to be carried over to new contexts yet unstudied.

I attempt to solve this tension by examining the assumptions behind the traditional view on the external validity of CSR. Some of these assumptions have already been addressed, and actually disputed, in the current debate. In chapter 3, I focus instead on those assumptions that, to the best of my knowledge, have not been addressed yet and seem to be responsible for the dead end in which the discussion among social scientists seems to be trapped now. In particular, I suggest that the debate should focus on how make case studies comparable rather than how select the typical case. Typicality and comparability are concepts closely related but distinct. The traditional view conflates the two and thus run into confusion about what external validity is really about and how it can actually be confronted in a fruitful manner. I surmise that by enhancing the comparability of studies unnoticed room for improvement is made for formulating more reliable assessment of the external validity of results obtained in case studies.

In chapter 4, I discuss issues of relevance when policy making purposes are at stake. In particular, I focus on the debate on the use and usefulness of randomized controlled trials (RCTs) to find the key to economic and social development. The participants to this debate agree that RCTs are affected by limited external validity, and that this impinges on their usefulness for policy making. They diverge, however, on the strategies to overcome this problem. I analyze three alternatives that are found in the economic literature: replication of RCTs, which has been proposed by the promoters of RCTs; cross-country regressions, which have been typically endorsed by RCT-sceptics; and the causal models proposed by James Heckman. I argue that these strategies succeed in their attempt to a different, and limited, extent.

Proponents of the first two strategies fail to take into adequate consideration the distinction between external validity and relevance, and treat the latter as a spill-over of the former. Their strategies, in fact, aim to improve the external validity of causal effects on the assumption that relevance will automatically follow. I argue that this is not the case because external validity and relevance are distinct concerns and should thus be confronted separately. The proposal by Heckman succeeds in delivering causal effects that are, as a matter of fact, more relevant to policy makers' purposes. I argue, however, that his model cannot adequately address the type of problems policy makers are likely to confront in developing contexts. Whereas Heckman's model is equipped to face problems of *prediction*, in developing contexts policy makers face problems of *planning*. Planning is a complex procedure that depends on various pieces of evidence and raises several concerns. Causal effects are but one epistemic input in this procedure; case-study evidence is also relevant to the crucial phases of planning.

## REFERENCES

Morgan, Mary. 2012. Case studies: one observation or many? Justification or discovery? *Philosophy of Science*, 79 (5): 667-677.

**Attilia Ruzzene** is a lecturer at the Faculty of Philosophy of Erasmus University Rotterdam. She studied international and diplomatic science at the University of Torino where she obtained an MSc (2006) and a PhD (2010) in economics. In the meantime she developed an interest in the philosophy and methodology of the social sciences and moved to Erasmus University Rotterdam. At the Erasmus Institute for Philosophy and Economics (EIPE), she obtained a research master in philosophy of economics (2009) and embarked on her second PhD project under the supervision of Julian Reiss and Jack Vromen.
Contact e-mail: <ruzzene@fwb.eur.nl>

*Erasmus Journal for Philosophy and Economics,*
*Volume 8, Issue 1,*
*Spring 2015, pp. 127-130.*
*http://ejpe.org/pdf/8-1-ts-2.pdf*

# PHD THESIS SUMMARY: The role of negotiations in achieving Pareto optimality in multi-dimensional cooperation games: implications for the ethical conduct of business

RICHARD STOMPER
*DBA/PhD in business administration, October 2014*
*SMC University and Central University of Nicaragua*

Numerous excellent macroeconomic studies have been done concerning the recent and, to some extent, still ongoing financial crisis and the painstaking recovery which has followed it. Not a small number of experts have expressed their puzzlement that such a phenomenon could occur in apparent contradiction to efficient market theory. Through it all, however, orthodox microeconomic theory has remained intact and, seemingly, has not even been seriously challenged. Yet the need to reexamine the fundamental propositions of this theory is obvious as the recent financial earthquake found its epicenter in a consumer debt crisis in the United States, the result of innumerable microeconomic decisions of consumer households.

The thesis proposes to formulate a microeconomic explanation for a contemporary macro-economic phenomenon that forms the essential backdrop of the consumer debt crisis, the failure of the growth of real wages to keep pace with productivity increases in the overall economy (Madland and Walter 2009). This fact, in itself, violates a fundamental premise of orthodox economic theory and necessarily implies that wages do not equal marginal productivity. At this point, it is obvious that the entire matter cannot be analyzed in terms of aggregate demand, but rather requires a disaggregation of that demand and an examination of the inter-sectoral mechanisms producing the consequent structure of demand.

Despite the marginal revolution and marginal utility theory, a subjective theory of value has been only partially embraced by orthodox economic theory. Austrian economic theory gives a fuller development to subjective value in its analysis of the market, while, simultaneously, almost entirely neglecting both economic exchange and the role of value within institutions. The largely unresolved problem of marginal cost posed by Ronald Coase (1946) is a problem only because orthodox theory largely relies, in the final analysis, on a cost of production basis to economic exchange. The pervasive existence of varying opportunity and conflict costs, and of highly differentiated markets have little impact, at least in theory, on the price point at which exchange takes place, giving to subjective value only a truncated expression in the final result.

Furthermore, orthodox economic theory has traditionally distinguished two very general areas of exchange, those that are external to organizations, i.e. market exchange, and those that occur within organizations and which are governed by hierarchical structure. The latter is generally the subject of a quasi-independent field of study known as institutional economics. Neo-classical economic theory, relying almost entirely on the market approach, historically has treated the firm as a *black box* (Coase 2008; Demsetz 1993) whose internal workings are largely separate from the broader market. Central to the theoretical propositions developed in the thesis is a model of economic exchange which seeks to identify the common elements of both market and hierarchical organization and, abstracting from these, to present a completely general model which is applicable to the entire range of transactions involved in economic exchange, and which sees both pure market and pure hierarchy as idealized abstractions comprising opposite poles of a spectrum.

In developing a general approach to economic exchange, the thesis takes as its starting point social choice theory (SCT) and particularly Arrow's general possibility theorem (GPT). The latter, in particular, in concluding that no social ordering is possible which can satisfy the ordinal preferences of the members of a society, essentially demands modification of its fundamental axioms to establish a stable societal equilibrium. The fact that SCT and the GPT derive from ordinal non-comparability of individual preferences makes them compatible with a theory of subjective value and methodological individualism. The theory developed in the thesis begins by introducing a dyadic model of exchange (Brennan and Buchanan 1985) into social choice theory. This dyadic model necessarily excludes political processes which are beyond its scope as is public choice theory in general.

From this starting point, the thesis performs three modifications to Arrow's (1963) axioms. The first is that the domain of preferences is restricted to those applicable to each dyadic exchange. The second is that the independence axiom is modified to permit exchange over future stochastic events and, therefore, over risk. The third modification is that, while the assumption that value preferences remain as both ordinal and non-comparable at the societal level, in each individual exchange, the parties develop sui generis evaluations which are comparable. Cooperation games are used to model the general theory of exchange. Cooperation games differ from non-cooperative games in that the payoffs are not ordained by the rules of the game but may be modified by the parties. The resulting models involve an exchange of information and value and are labeled *negotiations.* A definition of bargaining power for each party is derived.

The central theoretical conclusion at which the thesis arrives is that the distributive and allocative results of negotiations under conditions of equal bargaining power are equivalent to those produced by a hypothetically perfectly efficient market. Ethically, therefore, from the point of view of both efficiency and fairness, ex-ante attempts to equalize bargaining power are preferred over ex-post attempts at redistribution. Moreover, economic, social, and political changes can alter bargaining power and, consequently, move outcomes along a Pareto frontier. Since, at the societal level, no measure of value is fully applicable, both a significant degree of federalism and decentralization are called for to achieve negotiations of the optimal scope. Broad policy solutions should be limited to those few issues on which genuine broad agreement is assured. In general, individuals and dedicated groups, e.g., firms, labor organizations, etc., can better negotiate for their own interests.

## REFERENCES

Arrow, Kenneth. 1963. *Social choice and individual values.* New York: John Wiley and Sons.

Brennan, Geoffrey, and James Buchanan. 2000 [1985]. *The reason of rules: constitutional political economy.* Indianapolis: Liberty Fund.

Coase, Ronald. 1946. The marginal cost controversy. *Economica*, 13 (51): 169-182.

Coase, Ronald. 2008. The institutional structure of production. In *Handbook of new institutional economics*, eds. Claude Ménard, and Mary Shirley. Berlin: Springer-Verlag, 31-39.

Demsetz, Harold. 1993. George J. Stigler: midcentury neoclassicalist with a passion to quantify. *The Journal of Political Economy*, 101 (5): 793-808.

Madland, David, and Karla Walter. 2009. Unions are good for the American economy. Center for American Progress Action Fund website.

http://www.americanprogressaction.org/issues/labor/news/2009/02/18/5597/unions-are-good-for-the-american-economy-2 (accessed May 2015).

**Richard Stomper** received his BA degree in physics from the University of Chicago and his MBA from Loyola University of Chicago. He holds the dual degree of doctorate of business administration from SMC University and PhD from Central University of Nicaragua. He works as a field representative for the Illinois Fraternal Order of Police Labor Council. He has extensive experience negotiating and administrating labor contracts.
Contact e-mail: <richardstomper@hotmail.com>

*Erasmus Journal for Philosophy and Economics,*
*Volume 8, Issue 1,*
*Spring 2015, pp. 131-135.*
*http://ejpe.org/pdf/8-1-ts-3.pdf*

# PHD THESIS SUMMARY: The performativity of economics: a conventionalist approach

NICOLAS BRISSET
*PhD in economics, December 2014*
*Université de Lausanne and Université Paris 1 Panthéon-Sorbonne*

The performativity of economics has been the focal point of attention for the last decade in economic methodology and economic sociology. Initiated by Michel Callon, the performativity *thesis*, the idea that economics, more than merely describing an external social world, shapes it in its own image, directly challenges the scientificity of a social science that pushed to the extreme its desire to become an objective science. Since this thesis claims to destroy the whole battlefield of economic controversy it is not surprising that we find critics (let us recall that, for Popper, criticism is the basis of objectivity) of the performativity thesis in mainstream as well in more heterodox economics fields: if the social world takes the form of any theory which describes it, there is no longer such a thing as an external world in the name of which we can judge the validity of those theories' claims.

This point seems odd to the reader of J. L. Austin, the father of the idea of performativity in the philosophy of language. Indeed, for Austin, a performative utterance is above all an utterance that cannot fail to mean something, but can fail to do what it calls for. In his masterpiece, *How to do things with words* (1962, 5), Austin gives famous examples of this kind of utterance: "I name this ship the *Queen Elizabeth* […] I give and bequeath my watch to my brother […] I bet you sixpence it will rain tomorrow". Here, the state of the world can *or cannot* be changed; the performative utterance can be *happy* or *unhappy*. The happiness of an utterance rests, in Austin's theory, on a set of felicity conditions. As we can see, the potential failure of performativity is consubstantial of the definition of the performativity. I would even say that the limits of performativity are the heart of Austin's definition of this concept. The idea of the performativity of economics proposed by Callon seems to be missing this perspective, which I wanted to bring back in my dissertation: a definition of the performativity of economics according to its possibility of failure.

In the first part of the dissertation, I argue that much of the criticism toward Callon's performativity thesis stems from its lack of consideration of the potential *unhappiness* of economic theory. For instance, Mirowski and Nik-Khah (2007), and also Miller (2002), emphasize the conservative way in which performativity theory is engaged when it claims, contrary to most critics of rationality assumptions, that *homo economicus does exist*, because of the *embeddedness of the economy in economics*. If economics shape the world, there are no longer true or false theories *per se*, and there is no longer the possibility of factual challenge. Some other authors (Ferraro, et al. 2005; 2009; Felin and Foss 2009a; 2009b) argue that if economics matters in the construction of reality, we need to explain how and why. It is necessary to understand why, in a situation of competition between several economic theories, one is adopted and not the other. Felin and Foss (2009b, 676) call for a 'reality check' and argue that only true theories impact the social reality because agents choose and keep a theory in mind after its confrontation with reality. This is a contestation of the direction of causality. In Callon's view of performativity, a theory *T* is applied and thereby becomes pertinent regarding agents' expectations. Ferraro, Felin and Foss reverse this chain: A theory *T is* a good description of the social reality; that is why it becomes more used by agents.

After having pointed out what Callon left out of Austin's account, the main purpose of the second part of the dissertation focuses on elaborating a new approach to performativity centered around a set of felicity conditions. Since Austin's theory comes from the philosophy of language, the conditions he emphasizes cannot be used directly for the question I ask: what are the conditions for a theory to perform the social world? I therefore develop new felicity conditions, resting on David Lewis's theory of conventions. Following Lewis (1969, 76; 1975, 5-6), a regularity *R* in the behavior of members of a population *P* in a recurrent situation *S* is a convention if and only if:

(1) Everyone conforms to *R*
(2) Everyone expects everyone else to conform to *R*
(3) Expectation (2) gives everyone a good reason to conform to *R* themselves

(4) Everyone prefers a general conformity to *R* rather than a slightly-less-than general conformity
(5) Everyone would prefer that everyone conform to *R'*, on condition that *R'* meets the last two conditions
(6) Conditions (1) to (5) are common knowledge

My claim is that a theory becomes performative if it becomes a convention *à la* Lewis. From points 1 to 6 of David Lewis's definition of convention, I develop several conditions a theory has to fulfill in order to be performative. To say that economics performs the economy by becoming a convention emphasizes two major points. First, to perform in the social world, a theory has to potentially be a convention. This implies, according to Lewis's definition, that it has to be *empirical* and *self-fulfilling*.

1.1 *Empiricity*: a theory is said to be empirical if it permits the identification of and discrimination among at least two coordination points. If there is no choice between *R* and *R'*, there is no need for a convention: "this condition provides for the arbitrariness of conventions" (Lewis 1975, 6).

1.2 *Self-fulfilling:* people conform to *R* because the fact that everyone conforms to *R* makes it a fixed point, and thus a self-fulfilling prophecy. Everybody conforms to *R* because everybody thinks that everybody conforms to *R* (it is common knowledge), and *R* is therefore efficient in the sense that it permits people to coordinate with each other. As Lewis argues "reasons for conforming to a convention by believing something […] are believed premises tending to confirm the truth of the belief in question" (Lewis 1975, 5).

The second point is external to the strict definition of convention. To perform the social world, a theory must become a new convention in an existing social world made of conventions. As a consequence, performativity is closely linked to the degree of coherence between new and existing conventions.

2. *Coherency:* to perform the world, a convention derived from an economic theory has to fit with existing conventions.

In the third part of my dissertation, I analyse three limits of performativity: the form of the theory, the necessity to be self-fulfilling and coherency with the conventional world. Each limit is the object of a specific case study, namely, the theory of rationality, financial markets and market design. In each case study, I follow the different transformations a theory had to undergo, some of which permitted these theories to perform the social world.

To give a sample of the potential use of my theoretical framework, the 'financial market' case focuses on the famous study by Donald MacKenzie and Yuvan Millov (2003) of the performativity of the Black-Scholes-Merton model (BSM) on the Chicago Board Options Exchange (CBOE) from 1973-1987. MacKenzie and Millov argued that BSM's empirical success resulted not so much from discovering pre-existing price regularities, but because traders used it to anticipate each others' pricing of options. As a result, actual options prices came to correspond with the prices predicted by BSM. I point out in Chapter 7 that this conclusion is partly incorrect since the BSM model never became a *self-fulfilling* model (condition 1.2). The stock market crash of October 1987 is, I defend, empirical proof that the financial world never fitted with the economic theory contained in BSM.

## REFERENCES

Austin, J. L. 1962. *How to do things with words*. Cambridge (MA): Harvard University Press.

Felin, Teppo, and Nicolai Foss. 2009a. Social reality, the boundaries of self-fulfilling prophecy, and economics. *Organization Science*, 20 (3): 654-668.

Felin, Teppo, and Nicolai Foss. 2009b. Performativity of theory, arbitrary conventions, and possible worlds: a reality check. *Organization Science*, 20 (3): 676-678.

Ferraro, Fabrizio, Jeffrey Pfeffer, and Robert Sutton. 2005. Economics language and assumptions: how theories can become self-fulfilling. *Academy of Management Review*, 30 (1): 8-24.

Ferraro, Fabrizio, Jeffrey Pfeffer, and Robert Sutton. 2009. How and why theories matter: a comment on Felin and Foss. *Organization Science*, 20 (3): 669-675.

Lewis, David. 1969. *Convention*. Oxford: Blackwell Publishers.

Ferraro, Fabrizio, Jeffrey Pfeffer, and Robert Sutton. 1975. Languages and language. In *Minnesota studies in the philosophy of science*, Volume VII, ed. Keith Gunderson. Minneapolis: University of Minnesota Press, 3-35.

Miller, Daniel. 2002. Turning Callon the right way up. *Economy and Society*, 32 (2): 218-233.

MacKenzie, Donald, and Yuval Millo. 2003. Constructing a market, performing theory: the historical sociology of a financial derivatives exchange. *American Journal of Sociology*, 109 (1): 107-145.

Mirowski, Philip, and Edward Nik-Khah. 2007. Markets made flesh: performativity, and a problem in science studies, augmented with consideration of the FCC auctions. In *Do economists make markets?,* eds. Donald MacKenzie, Fabian Muniesa, and Lucia Siu. Princeton: Princeton University Press, 190-224.

**Nicolas Brisset** obtained his PhD in economics from both the University of Lausanne and the Université Paris 1 Panthéon-Sorbonne. He is currently a postdoctoral fellow of the Foundation House of Human Sciences (Paris). His areas of specialization are the philosophy of economics, history of economic thought, and economic sociology. He has published papers in the *Revue de Philosophie Économique*, *Revue Économique*, *l'Année Sociologique*, and *Œconomia*.
Contact e-mail: <brissetni@gmail.com>

*Erasmus Journal for Philosophy and Economics,*
*Volume 8, Issue 1,*
*Spring 2015, pp. 136-138.*
*http://ejpe.org/pdf/8-1-ts-4.pdf*

# PHD THESIS SUMMARY: The ontology of money: institutions, power and collective intentionality

GEORGIOS PAPADOPOULOS
*PhD in philosophy and economics, January 2015*
*Erasmus University Rotterdam*

The aim of the thesis is to revisit elementary questions about the nature and the existence of money and to propose an alternative framework to the textbook description of money according to three functions, i.e., as a means of exchange, unit of account and store of value. The main task that the thesis sets for itself is to investigate and to present how individual attitudes, social institutions, and technological contingencies ascribe to money its social significance, its functions and its value, in an effort to understand how the monetary system can be studied in the current socio-technological juncture. The motivation of the project is the dissatisfaction with the dominant commodity theory of money and its inability to contribute to the conversation on the recent economic crisis or on the technological transformation of money through digital payment systems.

The framework of analysis is developed through a comparison between the two major scientific research programs on money, the commodity and the state theories. In order to compare the two theories, three fundamental questions are raised: "What is money? How does it get or lose its value? Where does it come from or how does it get into society?" (Ingham 2004, 10) The two research programs offer different answers to the aforementioned questions because they adopt different methodological and ontological starting points. The commodity theory describes the economy as an all-encompassing market characterized by rationality, individualism, complete information and free choice. In this universe there is no place for power or the state, while the relations and the rules that regulate social interaction are minimal. The state theory of money is developed in a different, historically informed, theoretical framework, where state authority, rules, and norms are acknowledged and money is defined as an abstract standard of value.

The analysis of money according to the state theory is compatible with the reality of state-sponsored nonconvertible paper currency, but the appeal to state authority alone cannot provide a full explanation for the existence of money as an institution that regulates individual and collective behavior. In the thesis, social ontology and institutional economics supplement the state theory by providing a comprehensive framework for the analysis of the existence, the operation and the evolution of the institution of money. State money is supported by an account of social existence built upon the notions of collective intentionality and constitutive rules. Intentionality is a philosophical notion that defines the relation of the mind to the world. Collective intentions express a 'we-mode' rather than the 'I-mode' that characterizes individual intentionality. Constitutive rules establish the shared meaning of institutional facts and provide desire-independent reasons for action (Searle 2005, 5). An account of social existence, based on collective intentionality and constitutive rules, can provide the basis for an institutional analysis of money, delineating a form of collective acceptance that is both able to carry the ontology of money and that is consistent with an institutional analysis of its identity and evolution.

The definition of money as an institution is necessary, according to the ontological framework of the thesis, because the functions of money and its social significance depend on a system of constitutive and normative rules. Normative rules have the form of 'do *X* in context *C*' and constitutive rules communicate the status-functions of money through the structure '*X* counts as *Y* in the context *C*'. The combination of normative and constitutive rules establish money and explain how it gets invested with a specific social significance—as an abstract standard of value—in virtue of which it assumes a specific institutional status and can perform its social function. The normative and constitutive rules that create money are selected for their ability to facilitate the functions of money and in consequence to instantiate its status. The identity-constituting function of money remains unchanged and defines money, but the meaning and the fulfillment of this function within the specific social context of its constitution depends on the constitutive and normative rules, at the same time as their persistence depends on their ability to support the fulfillment of the function of money in the same context. The interplay between institutional rules and status suggest a relationship of mutual dependence and a mechanism for social evolution.

Technology is the motor of change in the process of social evolution, with financial innovation leading to institutional change. The interplay between the functions of money and the technological devices that are used to support its operation, including the regulatory framework that constitutes them, provides the mechanism for the evolution of money. The function of money remains unchanged, but the technology for its fulfillment evolves through time following innovations and changes of the institutional rules that establish the institution of money. The antagonism between ceremonial and instrumental values sets the pace of the integration of technological innovations in the institutional structure of the monetary system, regulating institutional adjustment (Waller 1982, 757). Ceremonial values account for the conservative inertia to social development, describing how the privileges, power and rents define the conditions for the social constitution of innovative technologies. Instrumental values are directed towards the application of new technological knowledge for the solution of specific social problems and refer to the progressive influence of technology on social attitudes and institutions.

The main contribution of the thesis is the proposed framework for the evolutionary analysis of money and economic value that combines the state theory of money with an account of social ontology developed from the concepts of collective intentionality, constitutive rules and social status, and with original institutional economics and its theory of institutional change.

## REFERENCES

Ingahm, Geoffrey. 2004. *The nature of money*. London: Polity.

Searle, John. 2005. What is an institution? *Journal of Institutional Economics*, 1 (1): 1-22.

Waller, William T. 1982. The concept of habit in economic analysis. *Journal of Economic Issues*, 22 (1): 757-771.

**Georgios Papadopoulos** combines economics, philosophical analysis, and aesthetics with an exploratory artistic practice. His research gravitates around money and its socioeconomic functions. He studied philosophy and economics at the London School of Economics and then at Erasmus University Rotterdam under the supervision of Prof. Dr. Arjo Klamer. In 2008 and 2009 he was a researcher at the theory department of Jan Van Eyck Academy in Maastricht, and in 2012 was granted the Vilém Flusser Award for Artistic Research from the School of Fine Arts and the transmediale festival in Berlin.
Contact e-mail: <languagegames@gmail.com>