ERASMUS JOURNAL FOR PHILOSOPHY AND ECONOMICS
VOLUME 5, ISSUE 1, SPRING 2012

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>

EDITORS

François Claveau
C. Tyler DesRoches
Joost W. Hengstmengel
Luis Mireles-Flores
Thomas Wells

EDITORIAL ADVISOR

Julian Reiss

ADVISORY BOARD

Erik Angner, Kenneth L. Avio, Roger Backhouse, Marcel Boumans, Richard Bradley, Nancy Cartwright, David Colander, Job Daemen, John B. Davis, Sheila Dow, Till Grüne-Yanoff, D. Wade Hands, Conrad Heilmann, Frank Hindriks, Clemens Hirsch, Geoffrey Hodgson, Elias L. Khalil, Arjo Klamer, Alessandro Lanteri, Aki Lehtinen, Uskali Mäki, Caterina Marchionni, Deirdre N. McCloskey, Mozaffar Qizilbash, Ingrid Robeyns, Malcolm Rutherford, Margaret Schabas, Eric Schliesser, Esther-Mirjam Sent, Robert Sugden, Jack Vromen.

ACKNOWLEDGMENTS

EJPE WOULD LIKE TO THANK ALL THOSE WHO HAVE ASSISTED IN PREPARING THE PRESENT ISSUE:
Christopher Clarke, David Colander, Sharon Crasnow, Bjørn-Ivar Davidsen, George DeMartino, Fredrik Hansen, Sven Ove Hansson, Darian Heim, Catherine Herfeld, Richard B. Howarth, Uskali Mäki, John Maloney, Maria Cristina Marcuzzo, Judith Mehta, Michiru Nagatsu, Deren Olgun, Kathryn Plaisance, Stephen Pratten, Mauro Rossi, Joseph Rouse, Attilia Ruzzene, Ben Saunders, Guido Schröder, R. David Simpson, Bradley Turner, Jeroen Van Bouwel.

SPECIAL THANKS GO TO THE MEMBERS OF THE JUDGING PANEL FOR THE MARK BLAUG PRIZE 2012:
Conrad Heilmann, Kevin D. Hoover, Caterina Marchionni, Deirdre N. McCloskey, Margaret Schabas.

ERASMUS JOURNAL FOR PHILOSOPHY AND ECONOMICS
VOLUME 5, ISSUE 1, SPRING 2012

TABLE OF CONTENTS

ARTICLES

Valuing environmental costs and benefits
in an uncertain future: risk aversion and discounting
FABIEN MEDVECKY [pp. 1-23]

Ethics in economics: lessons from human subjects research
MEGAN BLOMFIELD [pp. 24-44]

Doing the best one can:
a new justification for the use of lotteries
ITTAY NISSAN-ROZEN [pp. 45-72]

SPECIAL CONTRIBUTION

The potentials and limitations of rational choice theory:
an interview with ***GARY BECKER*** [pp. 73-86]

BOOK REVIEWS

Hilary Putnam and Vivian Walsh's
The end of value-free economics
DANIEL LITTLE [pp. 87-92]

Malcolm Rutherford's *The institutionalist movement*
in American economics, 1918-1947: science and social control
DAVID GINDIS [pp. 93-99]

Peter J. Boettke's
Handbook on contemporary Austrian economics
DAVID HOWDEN [pp. 100-105]

Iara Vigo de Lima's
Foucault's archaeology of political economy
RYAN WALTER [pp. 106-111]

Mark D. White's *Kantian ethics and economics:*
autonomy, dignity, and character
NICOLAS GRAVEL [pp. 112-120]

Philippe Steiner's
Durkheim and the birth of economic sociology
JARED L. PEIFER [pp. 121-127]

James E. Alvey's
A short history of ethics and economics: the Greeks
SPENCER J. PACK [pp. 128-130]

PHD THESIS SUMMARIES

Essays in the economics of knowledge
SAMULI LEPPÄLÄ [pp. 131-134]

Erasmus Journal for Philosophy and Economics,
Volume 5, Issue 1,
Spring 2012, pp. 1-23.
http://ejpe.org/pdf/5-1-art-1.pdf

Valuing environmental costs and benefits in an uncertain future: risk aversion and discounting

FABIEN MEDVECKY
University of Sydney

Abstract: A central point of debate over environmental policies concerns how future costs and benefits should be assessed. The most commonly used method for assessing the value of future costs and benefits is economic discounting. One often-cited justification for discounting is uncertainty. More specifically, it is risk aversion coupled with the expectation that future prospects are more risky. In this paper I argue that there are at least two reasons for disputing the use of risk aversion as a justification for discounting when dealing with long-term decisions, one technical and one ethical. Firstly, I argue that technically, it implies an inconsistency between theory and practice. And secondly, I argue that discounting for uncertainty relies on a form of individualism which, while reasonable in standard microeconomic theory where an agent chooses how to spread her own consumption over her own lifetime, is inappropriate in the context of inter-generational social decisions.

Keywords: discount rate, risk aversion, uncertainty, inter-generational justice, economics, environmental decisions

JEL Classification: B41, D61, D63, D81, Q51

Long-term environmental policy decisions, such as those concerning climate change and nuclear energy, raise many important scientific, social, and economic issues. A central point of debate over such decisions concerns how future economic costs and benefits should be assessed. The methods chosen to assess future costs and benefits, and the justifications we have for choosing such methods have deep social,

AUTHOR'S NOTE: I am grateful to the Centre of Excellence for Environmental Decisions and in particular Hugh Possingham. This paper was presented at the Harvard Climate: Science + Humanities conference, and the Eindhoven University of Technology workshop Cost-Benefit Analysis: Uncertainty, Discounting and the Sustainable Future where it benefitted greatly from insightful comments. I am also very thankful to Simon Caney, Katie Steele, Kim Sterelny, Mark Colyvan, and two anonymous referees for further helpful comments.

political and ethical implications (Gardiner 2004, 572; Caney 2009, 163). Arguably the most commonly used method for assessing the value of future costs and benefits in both scientific and economic models is discounting, a cost-benefit analysis tool that decreases future costs and benefits by a yearly rate, the discount rate. The discount rate is an aggregate of various parameters, each representing a different motivation and justification for discounting, and these motivations and justifications are each a point of debate amongst experts in the field (Stern 2007, 41; Nordhaus 2007, 689; Quiggin 2008, 200).

One such justification for discounting is uncertainty. The rationale for this view is that all things being equal, we are better able to assess what the state of affairs will be like tomorrow than next year, and we are better able to assess what the state of affairs will be like next year than the following year, and so on for any subsequent year. Given this increasing uncertainty over time, we are justified in placing greater value on present consumption than on future consumption, as present consumption is more certain. However, it is not uncertainty in any and all of its forms that is used to justify discounting; it is uncertainty over the likelihood of hitting the expected value of a project, which is used to justify discounting. It is a preference for a guaranteed outcome over a gamble that has the same expected value as the guaranteed return, or simply risk aversion, which is usually presented as a justification for discounting (Smith 2011, 4; Howarth 2009, 24; Brent 1996, 168; Department of Infrastructure and Planning 2007, 9).

I will argue that there are at least two reasons for disputing the inclusion of risk aversion in our justifications for discounting when dealing with long-term decisions. Firstly, appealing to uncertainty and risk aversion as a justification for discounting is technically problematic because it implies an inconsistency between theory and practice. This has implications for cost-benefit analysis which have not been noted before. Secondly, the premise that we are justified in preferring the more certain present to the less certain future is ethically problematic as this justification relies on a form of individualism. This form of individualism might be reasonable in standard microeconomic theory where an agent chooses how to spread her own consumption over her own lifetime. However, when faced with long-term environmental decisions, one moves from the individual decisions to the social decisions and from intra-generational decision to inter-generational decisions.

Unlike discounting in an intra-generational setting, discounting in an inter-generational setting presents us with a discontinuity between the beneficiaries of the decision (the present generation) and the subsidisers of the decision (the future generations). This discontinuity—which has played an important role in the ethical discourse concerning climate change—challenges the microeconomic rationale for using discounting when dealing with uncertainty. Before arguing my case, I will briefly discuss discounting, and how uncertainty in the form of risk aversion is used to justify discounting.

1. DISCOUNTING AND UNCERTAINTY

Most, if not all of our economic decisions involve comparing commodities. We might compare one commodity at a price against another commodity at a price, such as an apple at $1 against a pear at $1.05. Or we might compare one commodity against the same commodity, but at different points in time; such as an apple at $1 now against an apple at 90c next year. In cost-benefit analysis (CBA), discounting is a way of assessing the value of consumption—either as costs or as benefits—at different points in time. It is a way of assessing how much next year's apple is worth to me relative to today's apple. As part of this paper is concerned with the technical application of discounting, let me briefly explain the mechanics of discounting. Consider the following pair-wise comparison:

1) $200,000 now and further $100,000 in 10 years time or

2) $280,000 now and nothing further.

Assuming both options cost the same, in order to work out which option offers the highest return, we need to assess whether $100,000 in 10 years time is greater than, less than, or equal in value to $80,000 now. Working out the Net Present Value (*NPV*) of $100,000 in 10 years is just what discounting does. We can think of discounting as the reverse of interest; how much money would I have to invest today at a given interest rate r such that in 10 years time I have $100,000? Put in a general form, the equation for the net present value of a good $NPV(G)$, with future value $fv(G)$ in t years, at discounting rate r is:

$$NPV(G)=fv(G)/(1+r)^t$$

At a discount rate of 2.8% p.a., the rate suggested by the Office of Management and Budget, $100,000 in 10 years time has a *NPV* of $75,870 (OMB 2011). Thus option 2) is better, paying $280,000, compared to option 1) paying $275,870. But how do we work out the discount rate? A common way to arrive at a discount rate is to consider the various motivations we have for discount independently, and to work out how much discounting each motivation justifies. The amount of discounting justified by the various motivations is then aggregate into a single rate. These motivations are usually given as impatience or pure time preference, economic growth, opportunity cost, and risk and uncertainty.

In this paper, I will only be concerned with how much discounting can be justified by uncertainty. More specifically, I will be concerned with uncertainty spelled out in terms of risk aversion, as this is a common model. For example, Robert Brent, in his textbook on applied CBA, explains, "the risk premium is added to the discount rate to correct for the uncertainty characteristic of the benefit being in the future" (Brent 1996, 168). In regard to long-term environmental decisions, such as the climate change debate, Kathryn Smith states that one of the roles of η, a parameters of the social discount rate, "is to measure relative risk aversion. In a stochastic model, higher η implies more disutility from exposure to risk" (Smith 2011, 4). One point to note before proceeding is that there has been substantial discussion on whether the discount rate should be constant for through time or not. In this paper, I will be assuming constant rates as this is the practice advocated by the OMB amongst others (OMB 2011).

Risk aversion and risk premiums

The economic definition of a decision under uncertainty is a decision in which the probability of each outcome is known, but in which no single outcome is certain (Perloff 2004, 574). According to proponents of discounting for uncertainty, uncertainty leads us to discount because we prefer more certain returns to less certain returns, and the present is more certain than the future. As codified in the expression "a bird in the hand is worth two in the bush", we place greater value on what we know we have over what we think we might get.[1]

[1] In fact this preference for the certain option over the uncertain (even if the certain option is of lesser value) is the economic definition of risk aversion; see Perloff 2004, 582.

Consider the returns of a forestry project and deciding whether to harvest now, or in five years, assuming both options have the same expected return. We might be quite certain of the current market price of the wood and the current harvesting rate, but both of these are liable to change over time. Those changes in price and harvesting rate may be either to our benefit, or to our detriment. The prices might go up or down, and the harvested quantities might increase or decrease. But as far as discounting is concerned, it does not matter that uncertainty cuts both ways; that events might equally turn out to be better or worse than expected. What concerns us is that it might turn out worse. In effect, we want to give more weight to the possibility that the returns will be lower than the expected returns. To express our concern over the negative possibility, we turn to expected utility theory (EU) and include some risk aversion.

Risk aversion is the preference for investments which offer returns with smaller deviation from the mean over investments which offer returns with larger deviation from the mean, where the mean is the expected returns. For example, consider choosing between two coin-toss games with pay-offs as described in table 1 (assuming all games cost the same to play and the coin is fair). Both games are uncertain and both have an expected return of $50, but game *B* is less risky in that the payoffs do not diverge as far from the expected return as they do in game *A*.

Table 1

	Heads (*Pr*=0.5)	Tails (*Pr*=0.5)	Expected returns
Game *A*	$100×0.5 = $50	0×0.5 = 0	$50+$0 = $50
Game *B*	$60×0.5 = $30	$40×0.5 = 20	$30+$20 = $50

The worst case scenario of game *B* is that the player wins $40, which is $10 less than the expected return, whereas the worst case scenario of game *A* is that the player wins nothing, which is $50 less than the expected return. A risk-averse individual would prefer to play game *B* to game *A*. Such an individual would value game *B* more highly than game *A*, and would be willing to pay a premium to play game *B*. This premium, the risk premium, varies according to the situation and the decision maker's risk aversion utility of wealth function. An individual's utility of wealth function tells us how much good (broadly construed) an

individual derives from a given return. Using a text-book risk-averse utility of wealth function such as $U(W)=\sqrt{W}$ for the above coin bet, we get the following risk-averse table:

Table 2

	Head (*Pr*=0.5)	Tail (*Pr*=0.5)	EU
Game *A*	$U=\sqrt{\$100}\times 0.5 = 5$	$U=\sqrt{\$0}\times 0.5 = 0$	5
Game *B*	$U=\sqrt{\$60}\times 0.5 = 3.872$	$U=\sqrt{\$40}\times 0.5 = 3.162$	7.034

While both games have the same expected return, for risk-averse individuals, the riskiness of game *A* reduces its utility compared to game *B*. In order to perform a CBA on this game, we turn the expected utility back in monetary value by calculating the *risk-free* (or *certainty*) *equivalent* of the expected utility. This is done by converting the utility back into dollar values using the reverse of the utility of wealth function, in our case *risk-free equivalent (W)*=U^2. In the coin-toss games above, the risk-free equivalent of game *B*'s payoff is \$49.47, and game *A*'s risk-free equivalent payoff is \$25. In other words, an individual exhibiting the risk-averse utility of wealth function $U(W)=\sqrt{W}$ would derive the same amount of utility out game *B* as she would out of a guaranteed \$49.47, and should be indifferent between game *B* and a guaranteed \$49.47. Given that game *B* has an expected monetary value of \$50 and the agent is indifferent between game *B* and \$49.47, the risk premium she is willing to pay in this instance is \$0.53 (\$50 - \$49.47).

While different degrees and structures of risk aversion can be represented by altering the utility of wealth function, all of them result in a certainty equivalent amount smaller than the expected monetary value. Moreover, for the same expected value the greater the uncertainty of the returns in terms of variance, the lower the expected utility of the investment. This decrease in utility leads to a decrease in the risk-free equivalent value of the return, and the more the risk-free equivalent decreases, the greater the risk premium becomes. If uncertainty increases over time, then the risk premium a risk-averse individual is willing to pay also increases over time. This increase in the risk premium over time leads to progressively decreasing the value of future returns over time. It is this progressive decrease in the value of future returns that discounting for uncertainty is supposed to represent.

2. DISCOUNTING FOR UNCERTAINTY: THE TECHNICAL PROBLEM

Recall the forestry example above. I know what I get if I harvest the wood now, but much can happen in five years (for better or for worse). The longer I delay my harvesting, the greater the possibility of deviation from my expected return. Being risk-averse, I would be willing to pay a premium to guarantee my returns, and discounting the expected return of the harvest in five years just reflects my willingness to pay a risk premium. However, in this forestry example, I was solely concerned with the benefits or returns on my investment. But when we discount, we discount both costs and benefits.

If we discount future costs and benefits because we are risk-averse, then the future costs demand a different treatment to the future benefits. This is contrary to the standard practice in CBA, where the discount rate is insensitive as to whether it is applied to costs or to benefits (Department of Finance and Administration 2006, 41). For example, assume that we are assessing a forestry project with expected returns of $20M a year and expected costs of $15M a year for the next ten years. The standard practice in CBA is to discount both the expected returns and the expected costs at the same rate before balancing them, or (equivalently) to first subtract the costs from the returns, and then to discount the result. The problem is that discounting costs with the same positive rate used to discount benefits in fact involves a risk-loving preference, not a risk-averse preference. To accurately reflect risk aversion, the part of the discount rate which reflects risk aversion in regards to future costs should be *negative*.

For risk-averse individuals, discounting reflects a willingness to pay a premium for certainty. According to risk aversion, the more uncertain we are, the more we are willing to pay a premium to compensate for the uncertainty. And to pay a premium is to incur a cost. In the case of benefits, placing a cost on the returns reduces the magnitude of the benefits, so a positive discount rate accurately reflects our willingness to pay a premium. However, in the case of costs, placing a cost (for the premium) on existing costs *increases* the magnitude of the existing costs. This is contrary to discounting costs, which leads to a *decrease* in the magnitude of the costs.

Consider a case where a risk-averse individual is unsure about the amount of a cost, say a forthcoming tax bill. Assume the tax bill could be $1000, $1500, or $2000, each with equal probability 1/3.

The expected monetary value of the forthcoming tax bill is clearly $1500, but a risk-averse individual who is willing to pay a premium for certainty would be indifferent between a tax bill *T* which is greater than $1500 but certain, and the present scenario (which has lower expected monetary value than *T* but greater variance). A risk-averse individual would be willing to pay *T* - 1500 premium for certainty, a premium which *increases* the magnitude of the costs.

As is explained in the Office of Management and Budgets (OMB) Circular No A-94, "a risk-averse individual may have a certainty-equivalent for an uncertain set of costs that is larger in magnitude than the mathematical expectation of costs" (OMB 1992). Yet, the standard use of a positive discount rate leads to a decrease in the value of the costs. If we were to discount the expected return of -$1500 for one year, at say 7%, the *NPV* would be -$1400. If we want to reflect an agent's willingness to pay a premium for risk aversion in regards to costs, we should include a negative rate in the agent's discount rate.

The fact that uncertainty leads to a different treatment of costs and benefits opens a veritable can of worms. There are four issues I will consider in the following subsections, namely:

1) Is risk aversion the right description of why we discount for uncertainty?

2) If costs and benefits require different treatment, should we start by discounting them separately, and then subtracting the costs from the benefits, or should we first subtract the costs from the benefits, and then discount the result accordingly?

3) Do all motivations for discounting lead to costs and benefits requiring different discounting rates, or is this unique to discounting for uncertainty?

4) Does uncertainty always lead to different discount rates for costs and benefits?

My answers will be that 1) risk aversion may not be descriptively accurate of empirical evidence regarding individual choices under uncertainty, but trying to be descriptively accurate does not help, 2) we should start by discounting the costs and benefits, and then balancing them out, 3) the need for different discount rates for costs and benefits is unique to uncertainty, and 4) uncertainty does not always lead to different rates for costs and benefits, but even in the cases where we

could use the same rate for both costs and benefits, we ought to reject uncertainty as a motivation for discounting on pragmatic grounds.

Is risk aversion the right description of why uncertainty leads to discounting?

Having a uniform risk-averse value function for both costs and benefits requires that we use a different discount rate for each, which is contrary to standard practice. In order to rectify this problem, two options are open to us. On the one hand we could accept the implication of the above discussion and endorse a 'two rates' approach to discounting, with a lower rate for costs to reflect the negative risk-averse component. Alternatively, we could hold on to the one discount rate for both costs and benefits and argue that risk aversion fails to accurately represent how agents reason under uncertainty. Indeed, empirical evidence suggests that agents making decisions under uncertainty do not behave according to an expected utility model as is normally assumed in economics (Loewenstein and Prelec 1992, 573; Rabin and Thaler 2001, 220). If we want discounting to accurately reflect how people behave when making decisions under uncertainty, we should consider the empirical evidence about how people make such decisions.[2]

The best known evidence for how people behave when making decisions under uncertainty comes from Kahneman and Tversky (Kahneman and Tversky 1979, 263). Participants in Kahneman and Tversky's experiments were offered hypothetical choices between pairs of options, with different values and probabilities in each case. The results of the experiments showed that individuals making decisions under uncertainty were not only sensitive to payoffs and probabilities, which is what would result from adhering to expected utility theory, but individuals are also sensitive to other factors such as the direction of the payoffs. For example, one of the choices faced by participants was a choice between either A: \$6,000, with P=0.45, or B: \$3,000, with P=0.90. In that case, the overwhelming majority of participants (86%) exhibited a risk-averse attitude and preferred option B (Kahneman and Tversky 1979, 267). When offered a negative version of the above pairwise choice, namely either A: -\$6,000, with P=0.45, or B: -\$3,000, with P=0.90, the overwhelming majority of participants, 92%,

[2] Whether we do in fact want this is an interesting question, but sadly it lies outside the scope of this paper.

preferred the riskier option A (Kahneman and Tversky 1979, 267; Loomes and Sugden 1982, 805).

Experimental results show that participants' utility of wealth function was asymmetric across gains and losses; individuals value uncertainty over costs differently to uncertainty over benefits, which could be good news for those wanting to hold on to discounting unaltered. If individuals making decisions under uncertainty do not have a uniform risk aversion for gains and losses, then we may not need to resort to positive discount rates for benefits and negative discount rates for costs. We could hold on to a uniform rate for costs and benefits so long as individuals making decisions under uncertainty displayed 1) a risk-averse attitude in regard to gains, and 2) a risk-loving attitude in regard to losses, and 3) an equal degree of risk aversion in regard to gains as the degree of risk love displayed in regard to losses.

As it turns out, individuals making economic decisions under uncertainty did usually display a risk-averse attitude in regard to gains (all things being equal), so a positive discount rate for benefits seems appropriate. Moreover, the empirical evidence also shows that in regard to economic losses, individuals do display the risk-loving attitude which is required to justify a positive discount rate for costs. But while individuals might be risk-averse when considering gains and risk-loving when considering losses, the empirical evidence shows "that agents are more sensitive to losses than to gains, resulting in a utility function that is steeper for losses than for gains" (Köbberling and Wakker 2005, 120). This has come to be called loss-aversion (Rabin and Thaler 2001, 226). As Kahneman and Tversky explain:

> A salient characteristic of attitudes to changes in welfare is that losses loom larger than gains. The aggravation that one experiences in losing a sum of money appears to be greater than the pleasure associated with gaining the same amount (Kahneman and Tversky 1979, 279).

If the exhibited rate of risk aversion in regard to gains does not match the exhibited rate of loss aversion in regard to losses, then we again require a different discount rate for costs and benefits. To accurately reflect the empirical evidence about individuals' risk preferences, we should use a positive discount rate for both gains and losses, but the discount rate for losses should be of a greater magnitude. Thus, while it is true that the standard risk-averse expected

utility model is an inaccurate (or at least incomplete) description of how agents behave under uncertainty, appealing to descriptive accuracy would not help justify the present standard practice of using the same positive discount rate for both costs and benefits. Moreover, even if using the same positive discount rate for both costs and benefits was descriptively accurate, it does not follow that it is ethically defensible.[3]

Should we start by discounting costs and benefits separately, and then balancing them, or should we first balance the costs and benefits, and then discount the result?

One advantage of using the same discount rate for both costs and benefits is that it simplifies the process of CBA by making it insensitive to the order in which we proceed. It makes no difference whether we balance the costs and benefit first and then discount, or whether we discount first, and then balance the costs and benefits. However the introduction of different discount rates for costs and benefits complicates this. Now, the order in which we proceed does make a difference. Recall the above forestry example with expected returns of $20M, and expected costs of $15M. If we first balance the costs and benefits, and then discount the result at 5% for one year, we have a *NPV* of $4.75M. If, on the other hand, we first discount the benefits and the costs at 5% and -5% respectively, and then balance them, we get a *NPV* of $3.25M. Clearly, it cannot be the case that both these methods are right. Either we first balance the costs and benefits, and then discount, or we discount the costs and benefits first, and then balance them.

I argue that if we want to discount for uncertainty we must first discount all costs and benefits separately, and then balance them out. Ordering the process the other way around by first balancing and then discounting is untenable; here is why. Let us assume for a moment that we do in fact balance the costs and benefits first and discount second. The balancing of the costs and benefits merges the two figures into single number: the balance. Presumably, if the balance is positive it means we anticipate a profit or benefit, so we discount the result with a positive discount rate. And if the balance is negative, we discount it with a negative discount rate because we anticipate a loss or cost from our project. Yet this can only be reasonable if the costs and benefits occur at approximately the same time. If costs and benefits occur at different times, the 'balance first' option becomes problematic.

[3] I will consider this ethical problem in section 3, below.

Again consider the forestry example above, but with the following variations: a) all costs occur in 6 months from now, and all benefits occur in 1 year, b) all benefits occur in 6 months from now, and all costs occur in 1 year. If we first balance out the costs and benefits and then discount them, these work out the same. If, on the other hand, we begin by discounting the costs and benefits, and then balance them (at -5% and 5% respectively), option a) has a *NPV* of $3.65M, while option b) has a *NPV* of $3.75M. This may not be a great difference, but nor is it negligible, and it is a difference which is lost if we balance out the costs and benefits before we discount them.

While the question of how often we should discount (every year, every six month, every quarter) is a vexing issue for discounting *even* if we use the same discount rate for both costs and benefits (OMB 1992), the inclusion of different discount rates for costs and benefits make this issue even more vexing. Discounting the costs and benefits first (with their respective discount rate) before balancing them out has the great advantage of not being susceptible to this problem. Since costs and benefits are discounted independently, we can accurately reflect the time at which the costs or benefits occur.

Do all reasons for discounting lead to different discounting rates for costs and benefits, or is this unique to uncertainty?

While uncertainty requires that we treat costs and benefits differently, other motivations for discounting do not raise the same demands. The reason why uncertainty is unique in requiring a positive discount rate for benefits and a negative rate for costs is that uncertainty is the only reason for discounting which leads to a preference for *both* costs and benefits to be as close to the present as possible. Whether we appeal to opportunity costs, economic growth or pure time preference (impatience) as justifications for discounting, all of these lead us to prefer benefits as early as possible and costs as delayed possible. Put simply, in all cases other than uncertainty, we have a preference for early returns over delayed returns. Benefits are returns, so we prefer them as early as possible; costs on the other hand, are the opposite of returns, and hence are preferred as late as possible.

As an example consider opportunity costs as a motivation for discounting. Opportunity costs lead us to discount costs and benefits because of the financial returns we could have had if we had invested the resources under consideration (Torgerson and Raftery 1999, 914).

Recall the forestry example above with $20M returns and $15M costs. Assuming I could get 5% p.a. return on an investment, if I had the $20M benefits now, by next year that would be worth $21M. Conversely, all I need to have today to generate $20M next year at 5% is $19M—namely the discounted or *NPV* of the $20M, so the sooner the returns, the better. As for the costs, the process is reversed; the later the costs the better. If I had to pay the costs today, I would have to pay $15M. However, if I do not have to pay the $15M till next year, then I can invest the $15M such that by next year, I have $15.75M. And if I do not have to pay the $15M till next year, but I can get 5% return on my investment, all I need to invest today to generate $15M for next year is $14.3M—again the discounted or NPV of the $15M.

With opportunity costs, the nearer (temporally) the benefits are, the better off I am, hence the positive discount rate. And with the costs, the further in the future the costs are, the better off I am, which again requires a positive discount rate. As stated above, all motivations for discounting other than uncertainty behave like opportunity costs. They all lead us to prefer benefits to be in the near future and costs to be as far into the future as possible. This set of opposing preferences justifies a positive discount rate for both costs and benefits. Uncertainty over returns, on the other hand, leads to parallel preference across costs and benefits. The more certain a cost or benefit is, the more desirable it is; the earlier a cost or benefit is, the more certain it is; therefore both costs and benefits are preferred as early as possible.

Does discounting for uncertainty always lead to different rates for costs and benefits?

So far I have proceeded under the (unstated) assumption that the reasons for the uncertainty over the costs and benefits were independent of each other. Now let us consider what happens if costs and benefits are correlated. Recall one more time the forestry project with expected returns of $20M and expected costs of $15M. Let us assume that there are only two uncertain variables: 1) how much timber will be sold—these are the benefits—and 2) how many new seedlings will be required—which are the costs. Let us further assume that trees are only felled to fill the timber orders and seedlings are only bought to replace felled trees. In this example, the costs and benefits are correlated; if we sell more timber than expected (and hence have a greater turnover than expected), we will need to fell more trees than

expected. If we fell more trees than expected, we will need to buy more seedlings to replace those trees (and hence have a greater cost than expected). Conversely, if we sell less timber than expected (and hence fell fewer trees than expected) we will need to buy fewer seedlings (and hence have a lower cost than expected).

In such a case, using different discount rates for costs and benefits seems unreasonable because any increase or decrease in benefits relative to the expected benefits will be mirrored by a proportionally equivalent change in the costs. Therefore, it cannot be that we will have lower than expected benefits and higher than expected costs. In cases where the uncertainty over the costs and benefits exactly correlates (when the uncertainty in regard to cost equals the uncertainty in regard to benefits), it seems reasonable to use the same discount rate for both. When the uncertainty in regard to cost does not exactly correlate with the uncertainty in regard to benefit, we are no longer justified in using the same rate for both (for the reasons given at the beginning of this section). However, I will argue that even in such cases where costs and benefits exactly correlate, we have good pragmatic reasons to reject discounting for uncertainty.

Correlated costs and benefits

In cases where costs and benefits are exactly correlated we are permitted to use the same discount rate for both costs and benefits. But in such cases we are faced with a new decision in regards to the discounting for uncertainty, namely: what form of risk aversion should our discount rate reflect? The discount rate could reflect an aversion towards uncertainty in regards to benefits or an aversion towards uncertainty in regards to costs. Put differently, we could use a positive or a negative discount rate to account for uncertainty. While the standard practice in CBA is to use a positive discount rate, this does not need to be the case.

I will refer to a preference for certainty in regards to benefits over a preference for certainty in regards to costs simply as *risk aversion*. *Loss aversion* will be used to denote a preference for certainty in regards to costs over a preference for certainty in regards to benefits. Recall that all we require when costs and benefits are correlated is that we use the same rate across both costs and benefits. We can use a positive discount rate, a negative discount rate, or a zero discount rate depending on

whether we want to reflect risk aversion, loss aversion, or no preference either way.

Consider a 'break-even' environmental programme with correlated expected costs and benefits of $4 billions each. Using a positive rate to discount for uncertainty would reflect risk aversion; the thought would be something like "we might not get the benefits we expect, so we better under value our returns and let the costs follow suit". We could use a negative rate to discount for uncertainty, which would reflect loss aversion; the reasoning would be "we may have higher costs than we expect, so we better over value our costs and let the benefits follow suit". Alternatively, we may choose to use a zero discount rate, and not discount at all: "costs and benefits might not be what we expect them to be, but they will move together, so we may as well be neutral towards risk". All we required to be economically consistent is that the discount rate be the same across costs and benefits, whether it be positive, negative, or neutral. I will argue that while we have no economic reasons to sway one way or the other, we have pragmatic reasons to use a zero discount rate for uncertainty.

Consider what the consequences might be of applying each of these attitudes to risk in CBA to see whether this can help guide our decision. In many and possibly most cases—when costs and benefits are not correlated—discounting for uncertainty is a technical minefield. Until we either revise our theory of discounting (allow for two rates—one for costs and one for benefits—and work through the effects of this revision) or revise our theory of risk aversion (and revise it in such a way that would warrant a uniform discount rate), we ought not to include a rate for uncertainty in our discount rate. I will argue that not including a rate for uncertainty in the discount rate is the best option in all cases, as it is the only rate that reconciles cases when costs and benefits are not correlated with cases when they are correlated.

Most projects involve many economic measures, all of which are to be discounted. In order to remain consistent throughout a project, one needs to use the same discount rate for all measures. But this can only be justified if all of the measures in the project are correlated. Furthermore, projects often interact with other projects. Often, one project relies on the data and outcomes of previous project, making uniformity a desirable attribute of the discount rate. Consider a project, *P*, in which all the internal costs and benefits are correlated. The inclusion of some non risk-neutral discount rate in the assessment

of *P* rules out including the outcomes of *P* into any project which has costs and benefits that are not correlated with *P.* This gives us a reason to reject both risk aversion and loss aversion, at least on some occasions. Given that we have no reason to endorse either risk aversion or loss aversion over risk neutrality, but that we sometimes have reasons to reject both risk aversion and loss aversion, risk neutrality stands as the consistently least troublesome contender, including the cases where costs and benefits are correlated.

In this section, I have shown that in order to correctly reflect risk aversion and the willingness of risk-averse agents to pay a risk premium in the discount rate, the part of the discount rate which represents risk aversion should be positive when discounting benefits, but negative when discounting costs. The current practice, however, is to use the same, usually positive rate to account for risk aversion in the discount rate for both costs and benefits. I have argued that we should therefore be risk-neutral and not include a rate for risk aversion in the discount rate. Proponents of discounting for uncertainty might counter that this is an unnecessarily rash move and that we should just use different rates for costs and benefits to correctly reflect risk premiums. But, aside from the technical difficulties using different rates would create, I believe there is also an ethical problem with discounting for uncertainty in inter-generational social decisions.

3. DISCOUNTING FOR UNCERTAINTY: THE ETHICAL PROBLEM

One important difference between individual and social choices is that with individual decisions, decision theorists are only concerned with the demands of rationality. In social decisions, however, decision theorists also have to consider "the relationship between the demands of rationality and those of justice or fairness" (Resnik 1987, 177).[4] While I will not consider the various definitions of "justice" and "fairness", I will argue that discounting for uncertainty runs contrary to our intuitions of what is just or fair under almost any definition of these terms (Marglin's view aside).[5]

[4] This is not to say that there are no ethical considerations to individual decisions; rather, the suggestion is that individual decision theory does not take into account notions of justice or other ethical considerations. Social decision making does.

[5] Marglin has famously (or infamously) argued that "I want the government's social welfare function to reflect only the preferences of present individuals. Whatever else democratic theory may or may not imply, I consider it axiomatic that a democratic government reflects only the preferences of the individuals who are presently

Standard CBA discounts for uncertainty under the assumption that individuals prefer more certain returns to less certain returns and that this preference justifies decreasing the value of less certain returns relative to more certain returns. While I will accept these assumptions, I believe this is a poor justification for discounting in long-term social decisions. A central factor of the ethical dimension of inter-generational decisions is the discontinuity between decision makers and those that will be affected by the decision. It is this discontinuity which leads me to reject discounting for uncertainty in long-term decisions.

Discounting for uncertainty is simply a form of risk aversion and can be compared to single-time period decisions with different probability distribution, as I will do below. Indeed, comparing risk aversion in inter-temporal decisions as is the case in discounting to single-time period decisions is a fruitful approach as it shows explicitly what we are doing (Parfit 1983, 33).

One problem with uncertainty as a justification for discounting in long term social decisions is that this justification is founded on individual choice theory, and this translates poorly into a justification for inter-generational social decisions. With regard to individual choices, the argument is that individuals have certain attitudes to risk, and this attitude justifies their preferences. The standard microeconomic explanation of discounting is a representation of an individual's degree of risk aversion in regard to *her* returns; whether I prefer *A* with probability distribution *D* or *A* with probability distribution *D',* where *D'* has the same expected value as *D,* but greater deviation from the mean than *D*. This justification relies on a form of individualism based on an agent's right to chooses how to spread her own consumption over her own lifetime.

However, an inter-generational social decision involves a preference over uncertainty *and* across persons; whether I prefer *A* with probability distribution *D* for me or *A* with probability distribution *D'* for some future individual, again where *D'* has the same expected value as *D,* but greater deviation from the mean than *D*. The discontinuity between decision makers and those that will be affected by the decision is reflected in the fact that discounting in inter-generational choices also involves preferences across persons. It is this dimension of discounting

members of the body politic" (Marglin 1963, 97). Almost all other economists and philosophers believe that social distribution decisions should extend further than simply reflecting the views of the electorate. For some other views, see Scanlon 1982; Sibley 1953; Rawls 1993; Dobson 1996; Ekeli 2005.

for uncertainty in inter-generational decisions which I believe is of concern.

Since the temporal dimension, *when A* occurs, is only relevant in so far as it determines the probability distribution, we can replace the temporal dimension with a change in probability distribution (Parfit 1983, 32). A preference for $1000 now over $1000 at some future time *t* can be represented as a preference for $1000 guaranteed over some gamble with an expected monetary value of $1000, for example, a 50/50 chance at $1500 or $500. Now if we translate this into a social decision and make our preference over uncertainty *and* across persons, we get a preference for $1000 guaranteed for some individual *A* over a 50/50 chance at $1500 or $500 for some other individual *B*. This already seems problematic; why should the fact that there is uncertainty over the value of the outcome affect *who* gets the outcome.

It is worse still if we translate our decision into an inter-generational decision, where we make our preference over uncertainty and across persons, and in which we (the decision makers) are often some of the persons involved. We get a preference for $1000 guaranteed for some individuals, often ourselves, over a 50/50 chance at $1500 or $500 for someone else, namely future generations. In the cases where we are some of the persons involved, we will prefer the guaranteed money for us to the possible money for someone else. But because we are considering choices which distribute goods across persons including ourselves, we cannot differentiate between the preference for certainty (which leads to discounting) and the preference for self-interest (which does not necessarily lead to discounting). Indeed, we will prefer money for us with any probability distribution (all things being equal) over money for someone else. This can lead us to run against our preferences for the more certain over the less certain. We would likely prefer a 50/50 chance at $1500 or $500 for us over $1000 guaranteed for someone else, even though there is greater variance in our preferred option. This is comparable to preferring future goods for us over present goods for someone else, which runs contrary to the theory of discounting.

Our incapacity to differentiate between our preference for certainty and our preference for self-interest poses an ethical challenge. A preference for self-interest has nothing to do with uncertainty. It can, however, be a reason to prefer one outcome over another. We might prefer the present because *we*, the current generation, want the

consumption. But such an approach to social decision making seems irreconcilable with most conceptions of fairness and justice. Indeed, most philosophers and economists explicitly object to the use of such self-interest in social decision making (Ramsey 1928, 543; Smart 1973, 63; Smith 1976 [1776-1789], III-2; Stern 2007, 31).

For the majority not willing to endorse such self-interest, discounting for risk aversion must be distinguishable from discounting for self-interest. But it is not possible to distinguish between the two. This creates a dilemma for those wanting to discount for uncertainty but not for self-interest: either they must accept that self-interest might be "smuggled in" with uncertainty; or they must reject discounting for uncertainty. As I am not willing to accept self-interest as a motivation for discounting in social decisions, I side with those who reject discounting for uncertainty.

Discounting as compensation

Proponents of discounting for uncertainty might reply to my ethical objection by arguing that future individuals might accept the imposition of risk if they were to be compensated for this imposition. According to this line of reasoning, to impose a risk on another individual is morally permissible if that other individual is willing to accept the risk for an agreed amount of compensation. Now the issue turns not on whether we can discount for risk, but on what the discount rate should be, such that it incorporates the agreed compensation.

While this argument might have some weight in intra-generational decisions, it has little traction when applied to inter-generational issues. The challenge this argument faces, when applied to inter-generational decisions, stems from our incapacity to assess what counts as compensation. As the individuals to be compensated (future generations) are not part of the decision making framework, they cannot state what they would accept as a just compensation, thus, there is no agreed upon compensation. At best, we can allow for what we believe would be a just compensation for the imposition of risk. But, as was stated previously, in many inter-generational distribution decisions, we—the current generation—are not only decision makers, but are also potential beneficiaries of the decision. This self-interest skews our capacity for detached and fair judgement, much as it did in regard to uncertainty. The 'compensation argument' does not solve the problem, it simply shifts the problem. Instead of imposing our view of acceptable

risks, we impose our view of acceptable compensation for acceptable risks. This approach, in my view, creates more problems than it solves.

Consider uncertainty over the cost of decommissioning a nuclear power station. Currently the average cost of decommissioning a nuclear plant is around US$400M (Nuclear Energy Institute 2011). If we were to discount this value for risk aversion, we would reduce its magnitude and have a possible shortfall. According to the argument above, we could compensate for the risk by altering the discount rate. But I simply would not know how much to compensate people in 80 or 100 years from now for the possible shortfall they might face when dealing with the decommissioning of nuclear power stations. In fact, it has been argued that the uncertainty over issues such as nuclear power should lead us to be cautious about our assessment of future costs and benefits and, if anything, we should err on the side of over-allocation of funds, not under-allocation of funds (Caney 2009, 176; Gardiner 2006, 3; Steele 2006, 19).[6]

Whichever way we go, we impose views of acceptable compensation for acceptable risks. While some imposition of our view of what sits as acceptable compensation is an inevitable part of inter-generational decisions, the imposition of compensation for our risk aversion is not. We do not need to impose our view of acceptable risks onto future generations, thus we do not need to impose our view of acceptable compensation for acceptable risks. All we need to do is remain risk-neutral when choosing a discount rate.

CONCLUSIONS

In the economic debate surrounding environmental decision making, discounting stands as one of the most controversial issues. In this article, I have focussed on one of the justifications for discounting, uncertainty. I argued that uncertainty spelled out in terms of risk aversion cannot justify discounting in the case of inter-generational social decisions. I argued that there are two reasons to reject discounting for uncertainty in such cases.

Firstly, I argued that on technical grounds, discounting costs and benefits by the same rate inaccurately reflects risk aversion. I noted an exception to this case, namely when costs and benefits are correlated and have argued that while in such cases we have no economic reasons to outright reject discounting for uncertainty, we do have a pragmatic

[6] This issue is closely related to the precautionary principle.

reason: that only a risk-neutral discount rate (or no discounting) can lead to a uniform discount rate across projects.

Secondly, I argued that on ethical grounds, discounting for uncertainty relies on a form of individualism. This individualism might be reasonable in standard microeconomic theory, where an agent chooses how to spread her own consumption over her own lifetime. However, when applied to long-term environmental decisions, the discontinuity between the decision makers (the present generation) and those who will bear the consequences of the decision (the future generations) undermines the rationale for discounting for uncertainty.

Since discounting for uncertainty is both technically and ethically problematic in its current form, I believe we ought to reject uncertainty as a justification for discounting in cases of long-term environmental decisions and not include a rate for uncertainty in our discount rate.

REFERENCES

Brent, Robert. 1996. *Applied cost-benefit analysis*. Cheltenham: Edward Elgar Publishing.

Caney, Simon. 2009. Climate change and the future: discounting for time, wealth, and risk. *Journal of Social Philosophy*, 40 (2): 163-186.

Department of Finance and Administration. 2006. *Introduction to cost-benefit analysis and alternative evaluation methodologies*. Canberra (ACT): Commonwealth of Australia.

Department of Infrastructure and Planning. 2007. *Project assurance framework: cost-benefit analysis*. Brisbane (QLD): Queensland Government.

Dobson, Andrew. 1996. Representative democracy and the environment. In *Democracy and the environment: problems and prospects*, eds. William M. Lafferty, and James Meadowcroft. Cheltenham: Edward Elgar, 124-139.

Ekeli, Kristian Skagen. 2005. Giving a voice to posterity: deliberative democracy and representation of future people. *Journal of Agricultural and Environmental Ethics*, 18 (5): 429-450.

Gardiner, Stephen. 2004. Ethics and global climate change. *Ethics*, 114 (3): 555-600.

Gardiner, Stephen 2006. A core precautionary principle. *Journal of Political Philosophy*, 14 (1): 33-60.

Howarth, Richard B. 2009. Discounting, uncertainty, and revealed time preference. *Land Economics*, 85 (1): 24-40.

Kahneman, Daniel, and Amos Tversky. 1979. Prospect theory: an analysis of decision under risk. *Econometrica*, 47 (2): 263-291.

Köbberling, Veronika, and Wakker, Peter P. 2005. An index of loss aversion. *Journal of Economic Theory*, 122 (1): 119-131.

Loewenstein, George, and Drazen Prelec. 1992. Anomalies in intertemporal choice: evidence and an interpretation. *The Quarterly Journal of Economics*, 107 (2): 573-597.

Loomes, Graham, and Robert Sugden. 1982. Regret theory: an alternative theory of rational choice under uncertainty. *The Economic Journal*, 92 (368): 805-824.

Marglin, Stephen. 1963. The Social rate of discount and the optimal rate of investment. *The Quarterly Journal of Economics*, 77 (1): 95-111.

Nordhaus, William D. 2007. A review of the Stern review on the economics of climate change. *Journal of Economic Literature*, 45 (3): 686-702.

Nuclear Energy Institute. 2011. Costs: fuel, operation and waste disposal. NEI Website. http://www.nei.org/resourcesandstats/nuclear_statistics/costs/ (accessed April 2012).

Office of Management and Budget (OMB). 1992. Guidelines and discount rates for benefit-cost analysis of federal programs, Circular No. A-94. U.S.A. Executive Office of the President Website. http://www.whitehouse.gov/omb/circulars_a094 (accessed April 2012).

Office of Management and Budget (OMB). 2011. Discount rates for cost-effectiveness, lease purchase, and related analyses, Circular A-94, Appendix C (Revised December 2011). U.S.A. Executive Office of the President Website. http://www.whitehouse.gov/omb/circulars_a094/a94_appx-c/ (accessed April 2012).

Parfit, Derek. 1983. Energy policy and the further future: the social discount rate. In *Energy and the future*, eds. Douglas Maclean, and Peter Brown. Totowa (NJ): Rowman and Littlefield, 31-37.

Perloff, Jeffrey. 2004. *Microeconomics*. Boston: Pearson Addison Wesley.

Quiggin, John. 2008. Stern and his critics on discounting and climate change: an editorial essay. *Climatic Change*, 89 (3): 195-205.

Rabin, Matthew, and Richard Thaler. 2001. Anomalies: risk aversion. *The Journal of Economic Perspectives*, 15 (1): 219-232.

Ramsey, Frank. 1928. A mathematical theory of saving. *The Economic Journal*, 38 (152): 543-559.

Rawls, John. 1993. *Political liberalism*. New York: Columbia University Press.

Resnik, Michael. 1987. *Choices: an introduction to decision theory*. Minneapolis: University of Minnesota Press.

Scanlon, Thomas. 1982. Contractualism and utilitarianism. In *Utilitarianism and beyond*, eds. Amartya Sen, and Bernard Williams. Cambridge: Cambridge University Press, 103-128.

Sibley, W. M. 1953. The rational versus the reasonable. *The Philosophical Review*, 62 (4): 554-560.

Smart, John J. C. 1973. An outline of a system of utilitarian ethics. In *Utilitarianism: for and against*, John J. C. Smart, and Bernard Williams. Cambridge: Cambridge University Press, 3-74.

Smith, Adam. 1976 [1776-1789]. *An inquiry into the nature and causes of the wealth of nations* (WN), eds. R. H. Campbell, A. S. Skinner, and W. B. Todd. Indianapolis: Liberty Fund.

Smith, Kathryn. 2011. Discounting, risk and uncertainty in economic appraisals of climate change policy: comparing Nordhaus, Garnaut and Stern. *Commissioned paper for the Garnaut Review update 2011*. Department of Climate Change and Energy Efficiency, Canberra, AU.

Steele, Katie. 2006. The precautionary principle: a new approach to public decision-making? *Law, Probability and Risk*, 5 (1): 19-31.

Stern, Nicholas. 2007. *The economics of climate change: the Stern review*. Cambridge: Cambridge University Press.

Torgerson, David, and James Raftery. 1999. Economics notes: discounting. *The British Medical Journal*, 319 (7214): 914.

Fabien Medvecky submitted his PhD at University of Sydney on the justifications for the social discount rate in inter-generational decisions. He is currently working as a research fellow at the University of Queensland and teaches in philosophy and economics. His interests are philosophy of economics, environmental philosophy, and ethics.
Contact e-mail: <fabienmedvecky@gmail.com>

Erasmus Journal for Philosophy and Economics,
Volume 5, Issue 1,
Spring 2012, pp. 24-44.
http://ejpe.org/pdf/5-1-art-2.pdf

Ethics in economics: lessons from human subjects research

MEGAN BLOMFIELD
University of Bristol

Abstract: Many economists, it is said, "are inclined to deny that moral philosophy has anything to do with economics" (Hausman and McPherson 2006, 291). In this paper I challenge such inclinations by drawing an analogy between economic interventions and human subjects research. It is undeniable that investigators engaged in the latter should adhere to specific ethical principles. I argue that analogous features of economic interventions should lead us to recognise that similar ethical concerns actually arise in both activities, and thus that economic interventions should also be conducted in accordance with ethical principles. By exploring the analogy further I formulate some ethical guidelines for economic practice, which in turn imply that ethical responsibilities will extend to all members of the economics profession.

Keywords: experimental economics, positive-normative distinction, research ethics

JEL Classification: A13, B40, B41, C90

It is said that "many economists are inclined to deny that moral philosophy has anything to do with economics" (Hausman and McPherson 2006, 291). Whether this accusation is entirely fair can be questioned,[1] but there is some agreement that the belief that economics

[1] DeMartino—an economist who is very concerned with ethics—gives evidence that economists are often acutely aware of the moral quandaries that their profession may place them in (DeMartino 2011, chapter 3). The position DeMartino defends in his book is similar to mine in many respects, but our approaches differ somewhat. His work builds a strong and extensive case for a professional code of ethics for economists, grounded in the influence that economists enjoy, their institutional power, the harm that economic interventions can do, and the uncertainty under which

AUTHOR'S NOTE: This article was originally written for a course I took at the University of Toronto with Jossi Berkovitz, for whose instruction I am very grateful. I would also like to thank George DeMartino, one anonymous referee of this journal, and editor Thomas Wells for extensive and constructive comments. They helped me to address a fundamental flaw that appeared in the original version and identified various other places where important revisions were needed.

is an objective science, "insulated from ethics", is widely held in the profession and "continues to inform economic training and practice" (DeMartino 2011, 70, 62). This understanding of economics enables economists to argue that, even when they are called on to give policy advice, their scientific input can be value-free "provided the means-ends distinction is rigidly maintained" (Blaug 1992, 129); in other words, provided that it is governments or policymakers that decide on the *goals* of the policy, and the economist's involvement is restricted to the provision of objective information about options for achieving those goals. As Per Pinstrup-Andersen—an economics professor at Cornell University—relates: "I was trained to believe that economists are supposed to do value-free, positive policy analysis and leave the normative part to the policymaker" (Pinstrup-Andersen 2005, 1097).

In this paper I will be challenging this understanding of economics by arguing that there is an important class of normative questions that should be the concern of all economists—questions of how to practice ethically within their field. Positive and normative economists alike seek to understand or influence real-world economic systems.[2] There is immediate potential for ethical concerns to arise here because human beings are an essential part of these real-world economic systems. To see where such concerns materialise, I start by discussing one area of economic research where it is indisputable that economists need to follow certain ethical guidelines—research involving human subjects. I argue that anyone who accepts the significance of ethical concerns in human subjects research must acknowledge that analogous ethical concerns arise in all economic interventions. I then discuss what responsibilities economists, in particular, will bear in relation to such ethical concerns; for this I use a brief discussion of the lessons that can be learned from clinical research ethics to investigate how far within the economics profession the need to adhere to ethical standards extends. I conclude that economics cannot and should not be considered value-free.

---

economists operate (DeMartino 2011, 116). I also identify some of these aspects of economic practice as being important. However, instead of employing them directly as a *ground* for economic ethics, I use them to develop an analogy between economic practice and human subjects research. This analogy is intended to persuade those who acknowledge the existence of ethical concerns in the latter activity that they cannot consistently deny their presence in the former. Insights that I have gained from DeMartino's book have undoubtedly helped to clarify and strengthen my arguments.

[2] I am assuming that economists are in fact interested in real-world economic systems. Those who are not seem more appropriately classed as mathematicians or philosophers.

EXPERIMENTAL ECONOMICS

Economists have increasingly been turning to experimental methods in their pursuit of knowledge. Laboratory experiments have been used in economics since at least the 1930s (Roth 1993, 184-185). What I will be concentrating on in this section, however, is the growing use of field experiments in economic research.

A field experiment is an experiment conducted on a system of interest in its natural environment. In economics such experiments are generally thought to have been pioneered by the 1967 New Jersey Income Maintenance experiment (Ferber and Hirsch 1978, 1380; Greenberg, et al. 1999, 158). Field experiments are designed to measure the effects of changes in economic policy "by applying these changes to human populations under conditions of controlled experimentation similar to that used in the physical and biological sciences". Effectively, this involves using "the real world as a laboratory" (Ferber and Hirsch 1978, 1379-1380).

Economic field experiments have been particularly popular in the United States, where economists play "a major role in designing, operating, and evaluating social experiments" (Greenberg, et al. 1999, 170). More recently, development economists have been turning to such experimentation and as a result there is "a growing methodology for analysing micropolicy questions in randomised controlled trials (also known as Randomized Evaluations, REs)" (Cohen and Easterly 2009, 1; see also Banerjee and Duflo 2009, 152).

Social experiments, involving as they do "the manipulation of people's resources and even life styles", clearly raise ethical questions (Ferber and Hirsch 1978, 1395). To see just how important these questions can be, consider an example of one of the REs that are becoming popular in development economics. This study, conducted by Jessica Cohen and Pascaline Dupas, was designed to test the hypothesis that "cost-sharing—charging a subsidized, positive price—for a health product is necessary to avoid wasting resources on those who will not use or do not need the product" (Cohen and Dupas 2010, 1). This hypothesis is based on economic theories implying three potential effects of positive pricing. Firstly, positive pricing may give rise to a selection effect, allocating the product to those who value—or need—it the most. Additionally, it is claimed that paying for a product encourages an individual to use it, because they do not want to consider

the money wasted ('sunk cost' effects), or because they perceive the product as being of higher quality (Cohen and Dupas 2010, 2).

Cohen and Dupas tested this hypothesis for the provision of insecticide-treated bed nets (ITNs), used to prevent malarial infection. They selected 20 prenatal clinics in Western Kenya. Four clinics were assigned as the control group which would continue charging the standard subsidised price for ITNs—a price higher than any of the other groups. The rest of the clinics were randomly assigned a price at which they could sell ITNs to pregnant women, including five clinics that gave them out for free. Cohen and Dupas then monitored the uptake and usage of the ITNs to see if this was in agreement or disagreement with the hypothesis (Cohen and Dupas 2010, 3, 13).

This intervention thus manipulated the environment of the pregnant women under study in such a way that there was a perceived risk (based on the economic theories mentioned above) that they would become less likely to use an ITN—and less likely to ensure that their children used an ITN—as a result of being randomly assigned to the group where ITNs were given out for free, and thus more likely to catch malaria. These women and children were thereby exposed to a perceived risk of severe harm. In addition, Cohen and Dupas's study employed the invasive procedure of measuring haemoglobin levels in the women who attended the clinics in order to assess which of them in fact needed the ITNs the most. Such influence on the lives of the women involved in the study clearly necessitated serious ethical consideration regarding the way the study was set up and reviewed; for example, the implementation of safeguards to ensure that the study would not continue if it proved to be discouraging ITN use in some groups, and that the women involved consented to having blood samples taken for research purposes.

Essentially, anyone who engages in human subjects research has a responsibility to proceed according to certain ethical guidelines, and this includes economists pursuing such research in the field or the laboratory. Luckily for them, biomedical scientists and psychologists have been engaging in human subjects research for long enough to have developed various rules of practice. For academics, guidelines are laid down by their university.[3] Economists working for private companies on

[3] The University of Toronto, for example, has an Ethics Review Office which provides information regarding guidelines and practices for human subjects research. There are several Research Ethics Boards which are responsible for the ethical review of all such research, although primary responsibility for ensuring that research is carried out in an ethical manner falls to the researcher (Ethics Review Office 2007).

the other hand—who have traditionally implemented and evaluated a large proportion of social experiments in the United States (Greenberg, et al. 1999, 166)—have a responsibility to ensure that those companies have similar ethical rules in place.[4]

The claim that economists engaging in human subjects research have a responsibility to adhere to certain ethical standards does not seem to be one that many would dispute. In what follows, however, I will argue that accepting this claim should also lead one to accept that analogous ethical responsibilities extend across the economics profession.

ECONOMIC INTERVENTIONS AND EXPERIMENTS ON HUMAN SUBJECTS

One reason why it appears to be inconsistent to accept that economic experiments involving human subjects should be conducted in accordance with ethical standards, but at the same time deny that this holds for other economic interventions, is that many economic interventions arguably *constitute* experiments. If economic interventions are indeed experiments, then they are experiments involving the human beings that are part of the economy in question. Thus, *prima facie*, those engaging in such interventions are experimenting on a human population, and should therefore follow analogous ethical procedures to those that guide smaller scale human subjects research. As Iain Chalmers puts it, "selectively designating some [social] interventions as "experiments" […] ignores the reality that policy makers and practitioners are experimenting on other people most of the time" (Chalmers 2003, 30).[5]

Support for the idea that some economic interventions can rightly be considered experiments can be found in the writing of a number of economists: Alfred Eichner, for example, suggests that *whenever* economic theories are turned into policies we can consider the implementation of the policy a test of the theory, "but with society itself as the test subject" (Eichner 1983, 210); Milton Friedman too believed that policy implementation can constitute an experiment on an economy (sometimes, even, a crucial experiment) (Friedman and Schwartz 1963, 687-688); George DeMartino claims that during the economic

[4] In some countries the government will have a policy statement on human subjects research that should help privately employed economists to do this; see, for example, Panel on Research Ethics 2010.
[5] Chalmers in turn attributes this idea to Campbell (1969).

transition of formerly socialist Central and Eastern Europe, the rapid reforms enacted by economists "subjected countries to economic experimentation" (DeMartino 2011, 9); and Dani Rodrik suggests that macro-development economists are coming to realise the importance of an experimental approach, taking China as an example of a country that has implemented macroeconomic interventions in an explicitly experimental manner (Rodrik 2009, 40, 43).

The plausibility or significance of my initial claim that some large-scale economic interventions constitute experiments on a human population can be undermined, however, by arguments that most—or even all—economic interventions lack certain essential features of experimentation. Firstly, one might argue that economic interventions *cannot* in fact ever constitute experiments because they appear to lack controlled experimental conditions. Secondly, one might argue that even if Rodrik and the other economists discussed above are correct that *some* large-scale economic interventions constitute experiments, this will not be the case for the vast majority of cases. This is because whilst experiments are essentially *epistemically-orientated*—i.e., interventions on a system with the specific aim of improving knowledge of that system—economic interventions are usually primarily designed to produce a desired change in an economy.

I will address these two objections in turn. I will argue that the first is incorrect because the elements of control lacking in economic interventions are not in fact defining features of experimentation, and that, even if they were, their lack does not obviate the need for ethical concern in this domain. I will support this latter claim with a more detailed analysis of the features of human subjects research that engender an obligation to adhere to ethical standards. This more detailed analysis will in turn help me to deal with the second objection.

Experimental control

Unlike most laboratory or field experiments, when economy-wide interventions are implemented one cannot observe a control system to which the new policies are not applied. This in turn means that it is difficult to conclude that any other changes in the economy are the result of the intervention, rather than the influence of background effects (which will always be many and varied). Neither, of course, is it possible to have a random assignment of different policies in the economy-wide case. The absence of these methods of experimental

control might lead some to believe that such interventions cannot constitute experiments, particularly because random allocation often appears to be taken as a defining feature of social experiments (see, for example, Greenberg, et al. 1999, 157; Orr 1999, 10).

However, the use of a control group or randomised allocation are features that make an experiment a *good* experiment, rather than necessary features for something to count as an experiment *simpliciter*. Experimental control is a matter of degree. In any science it will be easier to secure adequate experimental control in some cases than others, and "the experimentalist can never, not even in principle, exhaustively demonstrate that no disturbing effects are present" (Galison 1987, 3). Looking more closely at randomisation, we should note that this technique is actually only used in experiments in order to avoid the disturbing effects of 'selection bias' (see Chalmers 2003, 29).[6] Thus, the absence of random allocation does not prevent something from being an experiment; rather, it makes an experiment's results vulnerable to selection bias. Experimental use of a control group, on the other hand, is not the only way an experimentalist can try to isolate background effects.

Peter Galison's exploration of Cavendish's experiments to measure the effects of gravity provides a useful point of comparison here. Just as it is impossible to have a control system when conducting an economy-wide intervention, it was not possible for Cavendish to observe a control system where the effects of gravity were not in play. Cavendish instead pursued experimental control in two ways: firstly, by constructing his apparatus in such a way that the disturbing effect of any change in temperature was blocked (by shutting his device in a room and operating it by remote); and secondly, by measuring and calculating features of the background to subtract from his observations as necessary (Galison 1987, 2-3).

When implementing economic policies, analogous attempts at experimental control can be made: firstly, by implementing one new policy at a time rather than, say, five; secondly, by measuring or calculating any background features of the economy that have a known effect on the system. These measures can help economists to assess economy-wide interventions by comparing them to counterfactual

[6] Selection bias results when the different behaviours of two experimental systems that are being compared result from systematic differences between them, rather than the different experimental interventions made.

scenarios rather than actual control groups (see Ravallion 2008, 25). This method will obviously be far less able to secure a level of control adequate to ensure that experimental results are informative, but this is a difference of degree, not of kind. Furthermore, even if each individual economic intervention is insufficient to prove or disprove a hypothesis, the accumulation of evidence from many such experiments from many countries over many years can be used by econometricians as a more or less convincing test of economic theory.

I do not believe, therefore, that the impossibility of employing a control group or randomisation gives us reason to hold that large-scale economic interventions cannot count as experiments on a population. Rather these missing features just give us reason to believe that economy-wide interventions can only ever constitute experiments with very significant and unavoidable flaws in terms of internal validity—experiments for which we should be very cautious in drawing any conclusions.

In any case, the absence of these elements of control in economic interventions is irrelevant to whether such work should be held to similar ethical standards as human subjects research. This is because such controls are not the only features of human subjects research that necessitate ethical concern, and thus their absence is not something that will obviate the need for such concern in economic interventions. In order to explain this claim in more detail, I will now discuss the features of human subjects research that give rise to ethical concerns.

Ethically problematic features of human subjects research

The use of certain types of experimental control *can* engender specific ethical concerns in human subjects research. The introduction of randomisation in medical trials, for example, resulted in worries that investigators were violating their duty to act in the best interests of the patients involved and led to the development of the principle of 'clinical equipoise' to indicate when the use of randomised allocation is ethically acceptable (see Freedman 1987). The use of a control group in human subjects research, on the other hand, gives rise to particular questions of *justice* if participants may be benefitted or disadvantaged in morally important ways depending on which group they are assigned to. Therefore the use of control groups calls for particular forms of monitoring, designed to ensure that the experiment does not continue

beyond the stage where it can be reasonably concluded that one group is receiving significantly greater benefits than the other.

The absence of a control group or randomisation in an economic intervention does, therefore, give us reason to believe that *some* ethical concerns analogous to those found in human subjects research will probably not arise here. However, such experimental controls are not the only, or most obvious, features of human subjects research that give rise to ethical questions, as I will now attempt to detail.[7]

Perhaps the strongest cause for concern in human subjects research is that such experiments may potentially—or even certainly—harm the participants. Compounding this risk is the fact that experiments are likely to involve scientific theory that is unproven, making their effects on the human subjects involved (whether beneficial or deleterious) uncertain. Whether or not harm is perceived to be likely, experiments have an impact on participants and researchers must therefore show respect for the autonomy of potential subjects by seeking their consent. For consent to be meaningful it must be informed (subjects must understand the goals of the study, what participation involves, and what the risks are), voluntary (free from manipulation and domination), and ongoing (see, e.g., Panel on Research Ethics 2010, 27). Acquisition of voluntary consent can clearly be problematic when financial incentives are offered for participation. An additional complicating factor in the acquisition of meaningful consent is the asymmetry of power between researcher and subject. Researchers will be at an advantage in terms of their knowledge of the area of research, and are likely to rely on their claimed expertise in the field to gain the trust of participants.

Now, whether or not economic interventions instantiate standard features of experimental control, they clearly share these other ethically problematic features of human subjects research—briefly: possible harm, uncertain effects, potential for undermining human autonomy, and asymmetries of power—and thus give rise to analogous concerns. To start with, large-scale economic interventions may potentially—or even, often, certainly—harm the human beings that will be affected by them. Economic policies can lead people to lose their jobs, their homes, and their ability to provide for themselves and their families; they can result in deprivation, starvation, poor health and death.

[7] This list—which is by no means intended to be exhaustive—builds on and adds to briefer comments made by Norman Daniels in his discussion of clinical research ethics and health sector reform (2006, 447). Daniels's work has had a significant influence on this paper, as will become apparent in later sections.

In some cases economists will be able to calculate the probability of such outcomes with some degree of confidence. However, because the applicability of economic theory to real-world economies and the exact effects of economic policies are in general unproven, in many cases it will not even be possible to assess the likelihood that particular harms will result from an economic intervention, or to predict the magnitude of any harms or benefits that are expected to result.

Even putting the issue of such obvious harms to one side, large-scale economic interventions can affect individuals in ways that undermine or neglect their autonomy: economists, through their policies, "introduce and restrict liberties and freedoms, incentives, rewards, punishments, and risk; they affect incomes, careers, entitlements" (DeMartino 2011, 4). Autonomy can be undermined further by the offer of economic incentives for acquiescence. For example, when international financial institutions "provide not just advice but also the material resources that the client state requires to meet basic needs, retain or restore credit, and achieve other vital objectives" (DeMartino 2011, 51), this can undermine individual autonomy by weakening a society's capacity for collective self-determination. And finally, economic policies are formulated and proposed by professional economists whose claimed expertise in a field that is opaque to most of those affected engenders trust in their recommendations. Those affected by policies will in general be at a disadvantage relative to the economists who formulate them—both in terms of their knowledge of economic theory and their influence on policymakers—which means that economic advisors occupy a position of power that can potentially be exploited.

The goals of experiment and economic intervention

With this analysis of the ethically problematic features of human subjects research in hand, I shall now return to the second objection identified above: that most economic interventions are essentially concerned with producing change, not knowledge. This objection is in a sense more successful than the first, because epistemic-orientation is undeniably an essential aspect of experimentation. Because most economic interventions do not have knowledge generation as their primary purpose they will not constitute experiments after all and, one might conclude, my analogy will therefore only have succeeded in suggesting that ethical concerns arise in the minority of economic interventions that are explicitly experimental.

However, the detailed analysis of the previous section shows that the analogy between human subjects research and economic intervention stands despite this difference in goals. Even economic interventions that are not primarily concerned with knowledge generation will share at least some of the ethically problematic features identified in the previous section, and thus will still give rise to analogous ethical concerns. I will illustrate this with an example.

Imagine a hypothetical scenario in which Cohen and Dupas do not conduct their study and instead the government of Kenya decides—on the advice of economists—to adopt a new policy whereby ITNs will now be available for free, in the firm belief that this will encourage uptake. Economists and policymakers are so convinced that this intervention will be successful that they do not put any systems in place to monitor the actual effects of the policy, and have no intention of using the intervention to improve their knowledge.

In this hypothetical case—as in Cohen and Dupas's study—it is possible that the 100% subsidy would undermine the objective of ensuring that ITNs are used by those who need them the most, because economists have underestimated the importance of positive pricing's selection and sunk cost effects, and public perceptions of price as an indicator of quality. The hypothetical national policy would, in addition, place thousands more individuals in a situation of potentially greater risk than the field study because the goal of producing change rather than knowledge discourages policymakers from engaging in monitoring that would indicate if the policy is causing harm (and being prepared to intervene if so). This hypothetical economic intervention therefore calls for a number of ethical guidelines analogous to, but if anything even stronger than, those that would apply to Cohen and Dupas's field experiment, despite the fact that the imagined intervention fails to constitute an experiment.

WHAT DOES THIS MEAN FOR ECONOMISTS?

Even if it is accepted that large-scale economic interventions share many of the features that engender ethical concerns in human subjects research, and therefore acknowledged that such interventions should be implemented according to analogous ethical guidelines, must one accept that it is *economists* who have a responsibility to ensure that this is the case? Can one still argue that economists only need provide objective, value-free advice on how to attain an end chosen by the policymaker,

and that it is the responsibility of the latter to ensure that ethical guidelines are followed? And what about economists who are engaged solely in research or education, and who do not advise any policymakers—do they need to concern themselves with any standards of ethical practice?

I think that the above discussion undermines the claim that economic policy advice can be value-free, because the means-ends distinction can no longer be seen to divide positive questions regarding economic policy formulation from value-laden ones. In Cohen and Dupas's experiment, it was not just the end that they were pursuing (experimental evidence regarding an economic hypothesis) that could raise ethical questions, but also the means (i.e., the experimental method) by which they pursued that end. Analogously, it is not only the end of economic intervention (the achievement of a given economic outcome) which should be subject to ethical assessment, but also the means by which the intervention is conducted. Economists, as the ones called upon by policymakers to suggest the means, therefore need to be involved in ensuring that those means are ethical.

In the next section I shall try to motivate this claim by considering what form ethical guidelines for economic interventions should take and what role economists should play in implementing them. I will do this by drawing on Norman Daniels's recent attempt to apply principles from clinical research ethics to health sector reforms. My discussion will, in addition, suggest that economists engaged in solely research or educational (rather than advisory) roles will also have a responsibility to adhere to certain ethical standards in their work, due to its influence on economic policy.

CLINICAL RESEARCH ETHICS

In his recent discussions of health system transformation, Daniels argues that health sector reforms "constitute social experiments on a population". Therefore, "the rationale for proactively evaluating clinical experimentation on human subjects also applies to these social experiments". This proactive evaluation will involve "ethical and scientific review before [the reforms] are implemented and ethical and scientific monitoring and evaluation afterwards" (Daniels 2006, 447). If clinical research ethics can teach us a lesson in health sector reform,

then it seems that it may also be informative in the realm of large-scale economic intervention.[8]

Daniels identifies three main elements of the ethical and scientific review of clinical research which should also apply to health sector reforms: assessment of objectives, assessment of the suitability of the measures to the pursuit of the objectives, and proper governance (Daniels 2006, 447). He suggests that an interdisciplinary team including policy makers, *academics* and civil society groups should be involved in this review (Daniels 2006, 449; emphasis added). I will discuss each of these elements in turn to see what they suggest about the role economists should play in the ethical formulation and implementation of economic policies.

Objectives

As is often pointed out, economic objectives (ends) are usually selected by policymakers and politicians, with economists called on to offer advice regarding the *means* of pursuing those ends. The assessment of those objectives thus seems to be the element of ethical review which is most likely to fall outside the purview of economists.

This is not entirely so, however. For a start, one can dispute the claim that economists are rarely asked to advise policymakers on the selection of goals (Blaug 1992, 129). Refusing to offer such advice is one way for economists to escape some of the ethical responsibilities that arise in this domain, but even then there will be cases where economists have a responsibility to condemn the authentic objectives of governments or policymakers as unethical. In cases where economic policies are sought in the pursuit of unquestionably immoral objectives—say a government was looking for an efficient way to starve the opposition and enrich their family members—an economist surely has an ethical and professional obligation not to offer advice on ways to achieve this. Thus, the assessment of objectives can be the ethical responsibility of economists after all.

[8] In fact, the examples that Daniels gives of unsuccessful health reforms that should have been, but were not, subject to ethical review—the introduction of user fees and expansion of the private health sector in developing countries (Daniels 2006, 447)—also constitute economic reforms, and were presumably influenced by economic arguments as much as biomedical ones.

Assessment of means

Once the (ethically acceptable) goals of economic intervention have been agreed upon, means for pursuing those ends need to be chosen. In choosing these means, the evidence that any proposed measures will lead to the desired objective must be examined; practical plans for implementing policies must be devised (Daniels 2008, 270-271); and consideration must be given as to whether there are other (desirable or undesirable) consequences that could result from pursuing the goals of the intervention in any given way. It is here that economists' involvement in the ethical review of interventions will be much more significant.

Economic research is vital for providing evidence that particular economic policies are likely to promote certain economic results. The economic advisors in each case have primary ethical responsibility for ensuring that the evidence they draw on is as reliable as possible, and is not presented as being more conclusive than it really is. Foreseeable risks should be clearly identified. In some cases, when evidence regarding the efficacy of policies for attaining a given objective is lacking, economists could request that they be allowed to first conduct laboratory or field experiments. Such experiments, as I have discussed above, also raise ethical concerns. However, when there are perceived risks from implementing a given economic policy it is surely preferable to first implement it on a smaller rather than a larger scale if possible, and the presence of a control group means that harms and benefits arising from that policy may be identified more easily (and faster).

To return to the example I have been using, the results of Cohen and Dupas's study in fact suggested that women who received free ITNs were no less likely to use them than women who paid a positive price. Furthermore, charging a positive price for ITNs significantly reduced uptake (Cohen and Dupas 2010, 1). Although the small scale of this field experiment creates some problems regarding the external validity of the results, they can still be taken as *some* evidence that implementing a free distribution policy across Kenya may not put citizens at risk by discouraging use after all, and that failing to implement this policy may well be harming citizens by discouraging the uptake of ITNs.

It is in the assessment of means that purely research economists have ethical responsibilities too. Our body of economic knowledge and theory is a product of the economics profession as a whole, and the

entire community therefore has a responsibility to take heed of certain ethical requirements. Any economic research which is available in the public domain has the potential to influence policy makers, whether directly or through the advice of other economists. The tentative and unproven nature of most economic hypotheses should therefore be clearly stated, with economists ensuring that they do not make unwarranted claims regarding the predicted consequences of implementing certain economic policies, or the applicability of their economic models to real-world economies.[9]

These requirements can be seen as addressing the possibility of harm, uncertainty and imbalance of power concerns that I identified earlier. As David Colander points out, those not familiar with economic modelling will often have "an elevated view" of the insights it provides, which means that such models can be used inappropriately. Colander suggests that economic modellers should therefore "see themselves as having an ethical responsibility to make the limitations of their models clear to others when they see their models being misused by others in ways that could cause harm to those other people or to society" (Colander 2010, 421-422). Economists involved in teaching have a duty to ensure that their students are made aware of such ethical responsibilities.

Martin Ravallion too emphasises the influence of economic research on policy, in particular in development settings where policymaking "draws on accumulated knowledge built up in large part from published research findings". The realm of publishing comes with its own ethical dilemmas. Difficulty in acquiring funding can mean that relevant research does not get published, conservative publication biases favour papers that confirm received views, and there is a tendency to focus on the internal rather than external validity of research findings. All of these factors can affect the achievement of development goals (Ravallion 2008, 25-26). Here again, academic economists who are involved in the publishing sphere have a responsibility to address such ethical problems, whether or not they personally engage in policy advice.

Another ethical standard that should be adhered to in publishing is to clearly state sources of funding in order to make potential biases

[9] Eichner argues that we should go so far as to insist that "any argument, before it is passed on to others as part of the cumulative body of economic knowledge, be empirically validated—and that any argument which cannot meet this test be viewed as at best only a tentative hypothesis, if not rejected outright" (Eichner 1983, 240).

identifiable. Academic research as a whole is becoming increasingly dependent on funding from commercial organisations, and there is some reason to believe that this can affect the outcomes of studies. In the biomedical sciences, for example, there is evidence suggesting that clinical researchers are more likely to find positive results when their studies are funded by the pharmaceutical industry (e.g., Lexchin, et al. 2003; Bhandari, et al. 2004). Such concerns led medical journals to require that authors declare any potential conflicts of interest. It is important that economic researchers do likewise.[10]

All the above considerations should help to ensure that economic research and advice is—as far as possible—honest, transparent, and evidence-based. Nevertheless, due to the intrinsic difficulty of acquiring conclusive economic knowledge, economic policies must still be subjected to continuous monitoring and re-evaluation once implemented. Any potential risks that were identified should be carefully watched for. A commitment from economists to such monitoring and evaluation would also help to ensure that the intervention is properly governed. It is governance to which I will now turn.

Governance

In clinical research, Daniels explains, good governance involves making sure that subjects have given informed consent to their participation in the study and that concern is shown for their welfare. Informed consent is important to ensure that "subjects can affirm the goals of the research and avoid manipulation, deception, or exploitation". Subjects should also be assured that "adequate surveillance of the outcomes, including risks, is done so that harms can be minimised or benefits optimised" (Daniels 2008, 271). Ethical review of health sector reforms, Daniels argues, should have analogous aspects of governance.

Actually getting every affected individual to give informed consent to health sector or economic reform would simply be impossible. Daniels suggests that when health sector reforms are implemented by democratically accountable agencies, the analogous role to informed

[10] In an interesting parallel, Mark Friedberg, Bernard Saffran, Tammy Stinson, Wendy Nelson, and Charles Bennett, find evidence that pharmaceutical industry sponsored economic analyses of cost-effectiveness regarding oncology drugs are less likely to report unfavourable findings (Friedberg, et al. 1989, 1453). Since I wrote this piece, the American Economic Association has adopted more stringent principles of disclosure for conflicts of interest, and now urges economists to apply them in all their activities: www.aeaweb.org/PDF_files/PR/AEA_Adopts_Extensions_to_Principles_for_Author_Disclosure_01-05-12.pdf

consent can be played by "democratic oversight of the reform process". When the institution imposing the reform is not democratically elected, securing anything resembling informed consent will be difficult. This is a particular risk in cases where "powerful external agencies offer large incentives and are not themselves held accountable for the reforms they impose". In such cases, efforts must be made to improve accountability and "empower civic society" (Daniels 2008, 271-272).

Daniels's comments again suggest that policies forced upon struggling economies by non-democratic agencies such as the IMF and the World Bank in return for aid are highly morally problematic because they undermine a society's powers of collective self-determination (even when the countries in question are democratically governed). Increasing the local public accountability of such organisations is thus a particularly important way to protect individual autonomy and provide recourse if the economic policies they proscribe prove harmful. Mark Weisbrot and Dean Baker argue that improving the accountability of International Financial Institutions will have the added bonus of increasing their effectiveness and efficiency, and identify four principles of accountability that can be used to monitor the economic policies they propose: well specified goals prior to implementation; evaluations during implementation to see if the country is on course to satisfy these goals (and, where it is not, a clear indication of why this is the case); wide public availability of any reports to ensure citizens of the country in question are kept fully informed; and identification of the economists and their supervisors who are responsible for the policy design and recommendation (Weisbrot and Baker 2004, 2).

Amanda Wolf's work on the principle of prior informed consent—which she examines through an investigation of the role it plays in two international trade treaties—might also help us figure out what consent could involve on an economy-wide scale. Informed consent, according to Wolf, is "procedurally principled by disclosure, comprehension, non-coercion, competence, and consent" (Wolf 2000, 503). These elements will be interpreted differently depending on context. In the economic case they likely require economists to ensure that a proper risk assessment—detailing hazards, expected results, and the probability of unintended consequences—of the intervention is completed before it is carried out; that the expected implications of the policy are publicised and explained to those who stand to be affected; and that some form of consultation is carried out in which citizens have the opportunity to

influence the decision without incurring significant costs, and without being subject to manipulation. One clear implication is that economists have an ethical and professional responsibility not to give advice in favour of controversial or unproven economic policies when those interventions will be implemented by authoritarian regimes. As well as being imposed without any plausible form of consent, such interventions have the potential to do extreme harm to a powerless population, with no guarantee that their effects will be monitored or risks assessed.[11]

Even when informed consent is given it is necessary to continue to manage risks through the course of a social intervention if harm is to be minimised (Wolf 2000, 499). In order to show adequate ethical concern for the population being affected by an economic policy, we should follow Daniels's suggestion that ethical review "be both proactive and ongoing, relying on monitoring and evaluation to make sure that the risks to a population are understood and can be minimised by the timely modification of reforms" (Daniels 2008, 271). We should recognise that reforms in pursuit of worthy societal goals are complicated, that it can take time for changes to take effect, and that things may get worse before they get better. In all of this, though, there should be an understanding of what risks and harms to the interests of the population involved can be considered acceptable.

Plans should be in place to implement alternative policies should the monitoring of interventions suggest that the results of the policy are unacceptable. In this respect, economists would do well to follow Colander's suggestion and acknowledge that rather than involving direct application of economic theory to real-world problems—what he terms the 'economics of control' model—economic policy making really is (and ought to be) a case of 'muddling through'. This latter trial and error approach—in which economic theory is just one of many inputs into decision making—seems much more likely to engender the flexibility and humility that will ensure unsuccessful policies are modified before they do too much harm (see Colander 2003, 197, 202, 208).

Economists can also show concern for those affected by economic policies through a willingness to learn from mistakes. Ignorance, incompetence, or reluctance to properly assess the results of economic

[11] Unfortunately this is an ethical imperative that some economists have ignored. See Friedman's letter to General Pinochet (reprinted in Friedman and Friedman 1998, 591-594). I will not here attempt to answer the question of whether it could *ever* be morally acceptable for economists to give advice to authoritarian regimes.

interventions could well be partly responsible for the trend in development economics for "trying the same thing over and over again, despite a long record of previous failures" (Cohen and Easterly 2009, 21). Repeated implementation of economic policies that past experience has shown may do more harm than good shows a distinct lack of concern for those who stand to be affected.

Ravallion, in a paper on the importance of evaluation in development economics, agrees that "too little evaluative research on development effectiveness gets done" (Ravallion 2008, 3). In part this is due to significant practical and logistical difficulties in conducting evaluations of what will often be long term and diffuse impacts of economic interventions. There is also, however, the problem that such evaluations are a public good and current researchers may not take the benefits that will accrue to future practitioners into account when deciding how much time and effort to spend on them. Ravallion suggests that institutional support for evaluations is therefore important (Ravallion 2008, 5-6); economic researchers will presumably play a crucial role in pressing for such support.

CONCLUSION

Those who hold economics to be an objective science devoid of ethical concerns, and who claim that economic research and advice can therefore be value-free provided that the means-ends distinction is maintained, are wrong. I have discussed a number of ways in which ethical considerations should enter into economic practice. Economists themselves have a responsibility to further develop the necessary ethical guidelines for their profession, and to encourage the relevant values through active discussion and training (particularly if they do not want to risk having their academic freedom limited by the powers of an external agency that is instead assigned with such a duty). This is another area in which academic economists will find that they possess ethical responsibilities, to ensure that their students—our future economists—are made to consider the ethical questions faced by the profession, such as have been discussed in this paper.[12] To paraphrase a line that Nobel Prize winning economist Trygve Haavelmo wrote over sixty years ago: Should we expect any less of economists, if their work is

[12] Perhaps, for example, professors should ensure that the economics department of which they are a member has mandatory courses in economic ethics.

to be the basis for economic policy upon which may depend the general economic welfare of billions of people? (see Haavelmo 1944, 115).

REFERENCES

Banerjee, Abhijit, and Esther Duflo. 2009. The experimental approach to development economics. *Annual Review of Economics*, 1: 151-178.

Bhandari, Mohit, Jason W. Busse, Dianne Jackowski, Victor M. Montori, Holger Shünemann, Shiela Sprague, Derek Mears, Emil H. Schemtisch, Dianne Heels-Ansdell, and P. J. Devereaux. 2004. Association between industry funding and statistically significant pro-industry findings in medical and surgical randomised trials. *Canadian Medical Association Journal*, 170 (4): 477-480.

Blaug, Mark. 1992. *The methodology of economics: or how economists explain*. Cambridge: Cambridge University Press.

Campbell, Donald T. 1969. Reforms as experiments. *American Psychologist*, 24 (4): 409-429.

Chalmers, Iain. 2003. Trying to do more good than harm in policy and practice: the role of rigorous, transparent, up-to-date evaluations. *The Annals of the American Academy of Political and Social Science*, 589 (1): 22-40.

Cohen, Jessica, and Pascaline Dupas. 2010. Free distribution or cost sharing? Evidence from a randomised malaria prevention experiment. *Quarterly Journal of Economics*, 125 (1): 1-45.

Cohen, Jessica, and William Easterly. 2009. Introduction: Thinking big versus thinking small. In *What works in development? Thinking big and thinking small*, eds. Jessica Cohen, and William Easterly. Washington (DC): Brookings Institution Press, 1-23.

Colander, David. 2003. Muddling through and policy analysis. *New Zealand Economic Papers*, 37 (2): 197-215.

Colander, David. 2010. The economics profession, the financial crisis, and method. *Journal of Economic Methodology*, 17 (4): 419-427.

Daniels, Norman. 2006. Toward ethical review of health system transformations. *American Journal of Public Health*, 96 (3): 447-451.

Daniels, Norman. 2008. *Just health: meeting health needs fairly*. Cambridge: Cambridge University Press.

DeMartino, George F. 2011. *The economist's oath*. Oxford: Oxford University Press.

Eichner, Alfred. 1983. Why economics is not yet a science. In *Why economics is not yet a science*, ed. Alfred Eichner. Armonk (NY): M. E. Sharpe Inc., 205-241.

Ethics Review Office. 2007. Guidelines and practices manual for research involving human subjects, Version 1.0. University of Toronto Website. http://www.research.utoronto.ca/wp-content/uploads/2009/03/ERO_Guidelines_Manual-2007.pdf (accessed April 2012).

Ferber, Robert, and Werner Z. Hirsch. 1978. Social experimentation and economic policy: a survey. *Journal of Economic Literature*, 16 (4): 1379-1414.

Freedman, Benjamin. 1987. Equipoise and the ethics of clinical research. *The New England Journal of Medicine*, 317 (3): 141-145.

Friedberg, Mark, Bernard Saffran, Tammy J. Stinson, Wendy Nelson, and Charles L. Bennett. 1999. Evaluation of conflict of interest in economic analyses of new drugs used in oncology. *Journal of the American Medical Association*, 282 (15): 1453-1457.

Friedman, Milton, and Rose Friedman. 1998. *Two lucky people: memoirs*. Chicago: University of Chicago Press.

Friedman, Milton, and Anna Schwartz. 1963. *A monetary history of the United States 1867-1960*. Princeton: Princeton University Press.

Galison, Peter. 1987. *How experiments end*. Chicago: University of Chicago Press.

Greenberg, David, Mark Shroder, and Matthew Onstott. 1999. The social experiment market. *The Journal of Economic Perspectives*, 13 (3): 157-172.

Haavelmo, Trygve. 1944. The probability approach in econometrics. *Econometrica*, 12 (supplement): iii-vi, 1-115.

Hausman, Daniel M., and Michael S. McPherson. 2006. *Economic analysis, moral philosophy, and public policy*. Cambridge: Cambridge University Press.

Lexchin, Joel, Lisa A. Bero, Benjamin Djulbegovic, and Otavio Clark. 2003. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *British Medical Journal*, 326: 1167-1170.

Orr, Larry. 1999. *Social experiments: evaluating public programs with experimental methods*. Thousand Oaks (CA): Sage Publications Inc.

Panel on Research Ethics. 2010. Tri-council policy statement: ethical conduct for research involving humans (TCPS 2). Government of Canada Website. http://www.pre.ethics.gc.ca/pdf/eng/tcps2/TCPS_2_FINAL_Web.pdf (accessed April 2012).

Pinstrup-Andersen, Per. 2005. Ethics and economic policy for the food system. *American Journal of Agricultural Economics*, 87 (5): 1097-1112.

Ravallion, Martin. 2008. Evaluation in the practice of development. *Policy Research Working Paper*, No. 4547. World Bank Development Research Group, Washington, DC.

Rodrik, Dani. 2009. The new development economics: we shall experiment, but how shall we learn? In *What works in development? Thinking big and thinking small*, eds. Jessica Cohen, and William Easterly. Washington (DC): Brookings Institution Press, 24-47.

Roth, Alvin E. 1993. The early history of experimental economics. *Journal of the History of Economic Thought*, 15 (2): 184-209.

Weisbrot, Mark, and Dean Baker. 2004. Applying economics to economists: good governance at the international financial institutions. *CEPR Briefing Paper*. Center for Economic and Policy Research, Washington, DC. http://www.cepr.net/documents/publications/ifi_2004_07.pdf (accessed April 2012).

Wolf, Amanda. 2000. Informed consent: a negotiated formula for trade in risky organisms and chemicals. *International Negotiation*, 5 (3): 485-521.

Megan Blomfield is a PhD candidate at the University of Bristol (UK). Her current research focuses on questions of global justice that are raised by the problem of climate change. Her paper "Global common resources and the just distribution of emission shares" is forthcoming in *The Journal of Political Philosophy*.

Contact e-mail: <mb3586@bris.ac.uk>

Erasmus Journal for Philosophy and Economics,
Volume 5, Issue 1,
Spring 2012, pp. 45-72.
http://ejpe.org/pdf/5-1-art-3.pdf

Doing the best one can: a new justification for the use of lotteries

ITTAY NISSAN-ROZEN
Hebrew University in Jerusalem

Abstract: In some cases in which rational and moral agents experience moral uncertainty, they are unable to assign exact degrees of moral value—in a non-arbitrary way—to some of the different acts available to them, and so are unable to choose with certainty the best act. This article presents a new justification for the use of lotteries in this kind of situation. It is argued that sometimes the only rational thing for a morally motivated agent to do here is to use a lottery.

Keywords: moral uncertainty, lotteries, indivisible goods

JEL Classification: B41, D01, D03, D80

Many people share the intuition that in some choice situations using a lottery among (some of) the acts available to an agent is the morally right thing to do. In the philosophical literature several justifications for this intuition are presented. The most famous is John Broome's justification, which is based on the idea that what makes using a lottery the morally right thing to do (when it is the morally right thing to do) is that it is fairer than any of the definite choices available to the agent.[1] Thus, Broome's explanation of what makes a lottery right has two parts: first he presents an account of the fairness of lotteries, and second he argues that in some situations the fairness consideration is strong enough to make the fair act the right act.

In this paper I will present a new justification for the rightness (in a sense specified below) of lotteries. According to my justification a lottery is justified in some situations where an agent suffers from moral uncertainty, i.e., in some situations when an agent is unsure what the morally right thing to do is (however, not in every situation in which this

[1] See, e.g., Broome 1990, 1991, 1994. Other discussions of the questions include Hooker 2005; Sher 1980; Saunders 2009; Rescher 1969; Glover 1977.

AUTHOR'S NOTE: I would like to thank Richard Bradley, Alex Voorhoeve, and three anonymous referees for their very helpful comments on previous versions of this text.

is the case). I will argue that in these situations using a lottery is the best one can do, given one's moral uncertainty. I will also characterize the set of situations in which a lottery is justified according to my account and present an explication for the term "the best one can do".

However, unlike Broome, I will not argue that using a lottery, when it is the right thing to do according to my account, is also the fair thing to do. One could take a further step and try to argue that what makes a lottery right according to my account is also what makes it fair. Hence, one could argue that being fair is just doing the best one can to do the right thing. I think there might be good reasons to take this further step,[2] but I will not argue for it here. Here I only present a justification for the use of lotteries, not an account of fairness.

Is my account a rival to Broome's account? Not necessarily and actually there are reasons to think that the two are never rival accounts, as they seem to address different issues. One can at the same time hold the position that some lotteries are morally right for the reasons Broome presents and that some are right (in a different sense of "rightness") for the reasons I present (and that some may be right for other reasons).

The rest of the paper will be organized in the following way. In the first section, using Broome's discussion of the tension between the sure-thing principle and the rightness of lotteries, I will present some background issues that will be of later use (I will not, however, present Broome's account). In the second section, I will discuss the idea of moral uncertainty; and in the third section, I will present my account for the rightness of lotteries using a formal framework. In the fourth section, I will discuss the recommendations that my account gives in some cases and will argue that these provide independent support for my account.

THE FAIRNESS OF LOTTERIES

Broome's starting point is the intuition that: "Sometimes a lottery is the fairest way of distributing a good" (Broome 1990, 87). Broome also holds that because of this fact "there will certainly be some

[2] Hooker (2005) acknowledges (and refers to others who acknowledge) that "fair is often used with a very broad meaning. A 'fair decision', in this very broad sense of 'fair', means a decision that appropriately accommodates all applicable moral distinctions and reasons" (Hooker 2005, 331). This understanding of the term "fair" is in line with the "being fair as doing the best one can" thesis, only that under the explication presented here for "doing the best one can" such an understanding of fairness can also explain why lotteries are sometimes fair.

circumstances where it is better to hold a lottery than to choose the best candidate deliberately" (Broome 1990, 99).

This latter claim, poses a problem for Broome that he has to deal with even before presenting his justification for the intuition he started with: it seems that any moral preferences ordering that ranks a lottery between two actions above both of these actions must violate an intuitive principle of rationality called the sure-thing principle (SP). The sure-thing principle requires that when an agent is uncertain what the consequences of some of the actions available to him will be, then when he evaluates these actions he can disregard any state of the world in which all of them bring about the same outcome, as shown in Table 1:

Table 1

	ω_1	ω_2
L	A	B
A	A	A
B	B	B

The SP requires that if the agent prefers act A to act B then he should prefer act A to act L and act L to act B. Thus, it is easy to see that a lottery between two alternatives should never be preferred to both of them.

One way to deal with this problem is to reject the SP in moral contexts.[3] However, this is not the strategy Broome adopts and he (as well as others) has presented very convincing arguments against it (see Broome 1984, section 2). To deal with this problem Broome suggests that in cases in which a lottery seems to be morally preferable to any of the alternatives over which it is defined, we have to include the fairness achieved by using the lottery in the description of the outcomes.[4] By following this suggestion, the SP is not violated because it does not apply. This is illustrated in Table 2:

[3] This is the position adopted, for example, by Diamond (1967), who first introduced this problem.
[4] See Karni 1996, for a similar suggestion.

Table 2

	ω_1	ω_2
L	A achieved by a lottery.	B achieved by a lottery.
A	A achieved by a definite choice.	A achieved by a definite choice.
B	B achieved by a definite choice.	B achieved by a definite choice.

Since, under the new interpretation of the situation, the two possible outcomes that act L might bring are different from the outcomes acts A and B bring, the SP does not apply to the decision problem and so is not violated.[5]

However, regardless of the question of whether the SP must be violated in cases where a lottery is ranked above all of the acts over which it is defined, the following claim does hold: whenever an agent prefers a lottery to all the definite acts over which it is defined, the agent does not maximize the expectation of a quantity that we can call goodness.[6] He may be maximizing expected moral value—i.e., the expectation of the overall value of an act, given both the act's expected goodness and its fairness—but not expected goodness.

The last observation is the reason why I have lingered on the discussion above. The important distinction I want to make is between accounts—like Broome's—that recommend a lottery even in cases in which it is possible for the agent to choose a definite act with higher expected goodness, and accounts that recommend a lottery only when it is impossible for the agent to do that. Accounts of the latter type can, for example, recommend a lottery in cases in which the agent is morally indifferent between two acts. In such cases, any lottery between the two

[5] In the literature there are several objections to Broome's use of the "redescribing the outcomes" strategy (see Steele 2006, for a good overview). Broome (1991) discusses one of them: the violation of the "rectangular field assumption". I will not discuss these objections here. See Bradley 2007, however, for a formal framework that resolves Broome's worries.

[6] "Goodness", as I use the term here, refers to the moral value of an outcome (i.e., an act in a specific state), *not including the* moral value added to the act that brings this outcome in virtue of it being a fair lottery. "Expected goodness" is to be understood in an analogous way. I think that Broome uses the term in the same way most of the time (at least in his earlier work), but as there are some places in which he makes comments that can be understood as implying the contrary, I do not want to argue that he does. In any case, this is the way I am going to use the term here.

acts has the same expected goodness as each of the two acts and so it is impossible for the agent to choose an act with higher expected goodness. However, such accounts cannot recommend a lottery in cases in which the agent morally prefers one definite act to another.

The account that I will put forward here can do—in an important sense—both: it never recommends a lottery when it is possible for the agent to choose a definite act with higher expected goodness, and it does sometimes recommend a lottery over all the definite acts over which it is defined, even when the agent is not morally indifferent between all of them. How can this be the case? The answer lies in my use of the idea of moral uncertainty. I will discuss the idea in more detail in the next section, but even before doing that, it is easy to see how using this idea can make such an account possible.

When an agent is certain that one of the acts available to him has higher expected goodness than the other(s), but is not certain which one it is, then although it is metaphysically possible for him to choose the act with the highest expected goodness (it is, after all, one of the acts available to him), it is not epistemically possible for him to do so. In some such cases (but not in all of them), my account will recommend a lottery. Thus, in those cases, a lottery will be recommended even though the agent is certain that there is another act available to him with higher expected goodness, and yet the requirement to always choose the act with the highest expected goodness (when it is epistemically possible to do that) is kept.

So my account recommends choosing a lottery—when it does recommend that—not because this is the "objectively" right thing to do, but rather because, given the agent's epistemic state, this is the only "subjectively" right thing for him to do.

To avoid confusions, a terminological comment is necessary here. As mentioned—and as will be discussed at more length in the last section—my account is consistent both with the position according to which being a fair lottery adds some positive moral value to an act, and with the position according to which this is never the case. In order to avoid taking a stand in this debate, I need to use three different terms that describe different types of moral features of acts. I have already introduced the first one—"goodness"—explicitly: the "goodness" of an act is the moral value of the outcome that an act will actually deliver (not including the moral value added to it in virtue of it being a fair lottery).

Now we see that there are two more moral features of acts that will play a role in my discussion. The first one is the overall moral value of an act (including the possible moral value added to the act in virtue of it being a fair lottery). I will use the term "overall moral value" to refer to this feature. The second one is the possible moral value added to an act in virtue of it being a lottery. I will use the term "fairness" to refer to this feature.

When an agent does not suffer from moral uncertainty he ought to choose the act with the highest overall moral value. I have just used the term "the objectively right act" in order to refer to this act and will continue to do that. However, when an agent does suffer from moral uncertainty he cannot choose the objectively right act with certainty. Which act should the agent choose in such a case? Whatever the answer to this question is, we need a term in order to refer to this act. I have just used the term "the subjectively right act" in order to refer to it and I will continue to do that.

Notice that in light of the distinction between the best act and the right act, we can already distinguish between two types of moral uncertainty: regarding which act is the objectively right act to choose and regarding which act is the best act (when the "best act" is the one with the highest expected goodness). My discussion of moral uncertainty, in the next two sections, will only refer to the latter type of moral uncertainty, i.e., to uncertainty regarding which act is the morally best act (and for convenience I will just use the term "best act").

It should be noted, though, that by limiting my discussion only to this type of moral uncertainty, I do not mean to suggest that Broome's account—and others like it—are either false or cannot be applied to the cases that I discuss. I offer an alternative account for the use of lotteries that can justify choosing a lottery even when being a fair lottery does not add any moral value to an act (or adds some moral value but not enough to make the fair act, the right act), but this account can live peacefully with other accounts. In the last section, I will go back to the above distinction and will discuss the advantages that my account has in this context. I move now to discuss the idea of moral uncertainty.

MORAL UNCERTAINTY

The idea of being uncertain regarding what is the morally best thing to do is the key element that will help me develop my account. It is possible to make a distinction between three different types of

uncertainty regarding which act is the morally best act to choose in a given situation, only one of which will be dealt with here. The first type is that of such uncertainty that results exclusively from uncertainty regarding the truth of some non-normative claims. For example, an agent might be uncertain whether one act, A, is better than another act, B, because he is uncertain what the consequences of the acts would be. If the agent knew for sure what their consequences would be, he would not be uncertain regarding which act is better. There are some discussions in the literature (see, e.g., Diamond 1967) regarding the question of how a rational moral agent should make his decisions in the face of such uncertainty, but these will not concern us here.

The second type of uncertainty is that which cannot be reduced to uncertainty regarding the truth of some non-moral claim, but still can be reduced to uncertainty regarding the truth of some general moral claim. For example, an agent might be uncertain whether one act, A, is better than another act, B, only because he is uncertain whether consequentialism is true or not. If he was sure that consequentialism is true (or not) he would not be uncertain regarding which act is the best act to choose. This type of uncertainty is the one which is usually discussed in the literature under the title "moral uncertainty" (see, e.g., Lockhart 2000; and Sepielli 2009).

Most of the accounts that belong to this literature treat moral uncertainty much in the same way that decision theory treats uncertainty regarding states of the world, i.e., by requiring that in face of moral uncertainty one should maximise expected goodness (relative to the moral uncertainty one suffers from). Such a requirement, however, is based on two assumptions: 1) one is able to tell how good or how bad every possible act is, according to each of the moral theories (or general moral claims) one believes might be true; and 2) one is able to compare these values across theories.

Much of the discussion in the literature is dedicated to an examination of the second assumption, i.e., to the question of which principles should govern the inter-theoretical comparison of moral value. However, in some cases, prior to dealing with the second assumption, one has to deal with the first one. Before one can present a prescription for how to compare the degrees of moral value different moral theories assign to an act, one has to explain how it is possible to get those degrees according to a single theory.

While in some choice situations it seems that the theories themselves might offer an answer to this question, in other choice situations this cannot be the case. Situations of the latter kind are usually (but maybe not always) those in which different moral considerations (according to the same theory) push in different directions and thus, in order to assign exact levels of goodness to the acts, the theory must assign weights to the different considerations.

We can judge, for example, that saving the life of another person is better than slightly improving his wellbeing, but it is really hard for us to say exactly how much better it is. Thus, when we have to decide between saving the life of one person and slightly improving the wellbeing of many people we may become uncertain regarding what we ought to do in a specific case (i.e., when the number of people whose wellbeing we can improve is high enough). The same is true for the moral judgements we make, under the assumption that a specific moral theory (or a general moral claim) is the right one. One can accept, for example, utilitarianism, or even a specific type of utilitarianism, but still be uncertain regarding how different aspects that increase utility ought to be weighed against each other.

Notice that in this kind of case the moral uncertainty arises *because* we do not have direct access to degrees of moral value. If we had such access we would know how to weigh the different moral dimensions against each other and would just form our moral preferences according to the levels of expected goodness of the different acts.

Now, one might argue that although an agent does not have direct access to the level of goodness of the different acts available to him; he does have direct access to the level of goodness of the different acts conditional on some betterness relation holding between them. In other words, it might be that an agent who is uncertain whether act B is better than act C, but is certain that act A is better than both act B and act C, is also certain, *regarding every possible lottery between* A *and* C, that if it is the case that act B is better than act C, this lottery is either better or worse than act B. This is what it means, for a rational agent, to have direct access to the level of goodness of the acts conditional on some betterness relation holding between them.

I do not want to argue that such cases never happen. I certainly believe that in many cases, people have partial information regarding degrees of goodness conditional on some betterness relation holding.

However, I also believe that there are cases in which this type of information is unavailable.

Consider an agent who tries to follow the strategy described above when coming to make a moral decision between two acts, of which she is uncertain which is better than which. The agent formulates a number of hypotheses regarding the degrees of moral value of the acts available to her, conditional on either of the two betterness relations holding. Then she has to assign a probability value to each one of these hypotheses. On what basis can this be done? Surely we do not want to argue that she should do this arbitrarily. Arguing this is like arguing that she should choose an act arbitrarily as by assigning different probabilities to the hypotheses, she can make either of the acts the one that maximises expected goodness.

A more plausible answer is that she should do this according to her actual degrees of beliefs in these hypotheses, which should probably be based on what she takes to be moral evidence. However, it is not clear what can constitute evidence for a specific hypothesis regarding the exact degree of moral value of an act, other than the kind of betterness judgements that assigning such a degree to the act (together with assigning other degrees to other acts) leads to. Actual moral theories do not give us exact degrees of moral values in the kind of situations I am referring to and people do not generally have intuitions regarding such levels. What people do when they have to assign such levels is to implicitly judge which one is better among many possible acts.

Thus, in such choice situations, since moral agents do not have direct access to the degrees of moral value that different theories assign to different acts, the need arises to present an account for decision making under conditions of moral uncertainty that does not make any reference to these degrees of moral value. This is not because the requirement to maximize expected moral value is not a principle of rationality. I believe it is. Rather, it is because when one does not have access to these degrees, one cannot possibly obey this requirement and thus the need arises for an alternative decision rule.

This last type of moral uncertainty, i.e., moral uncertainty that cannot be reduced either to uncertainty regarding non-moral claims or to uncertainty regarding general moral claims, is the one which I am going to discuss in the next section.

MORAL UNCERTAINTY AND LOTTERIES

When an agent is uncertain regarding the morally best thing to do in a specific situation because she does not have access to the degrees of moral value of the different acts available to her, but still must make a decision, what should she do? The immediate answer is, I think, that she should try to minimize this uncertainty as much as she can: she should spend some time reflecting on the matter, she should consult with people whose opinions she values, she should read some books, and so on. But when she is done with this process, when she has used any sources of moral information available to her, then if she is still uncertain regarding what is the best thing to do, she has no plausible alternative but to go with the judgements she is more certain about. So if she believes that some act A is better than another act B more strongly than she believes that B is better than A, she should choose A over B. Let us call this requirement the Likelihood of Betterness Constraint (LBC).

Such a prescription clearly falls short of the ideal of maximizing expected moral value (which the agent cannot obey because she does not have access to degrees of moral value), but it does require that the agent make use of the information she does have, i.e., her degrees of belief regarding the moral betterness relations that hold between different acts.[7]

Notice, however, that if one accepts the LBC, but still wants one's moral choices to be transitive, one commits oneself to the requirement that for any three alternatives, A, B and C, if one believes it is more likely than not that A is better than B and that it is more likely than not that B is better than C, one must believe that it is more likely than

[7] It might be argued that there are some situations in which the agent does not even have enough information to allow her to assign in a non-arbitrary way degrees of belief to the different possible betterness relations that might hold between different acts. Maybe there are such cases. However, there are certainly cases in which we do feel that we have enough moral evidence to assign in a non-arbitrary way degrees of belief to the different possible betterness relations, while we do not have enough moral evidence to assign in a non-arbitrary way degrees of beliefs to the different possible hypotheses regarding the exact level of moral value of each possible act. The reason for that is the one discussed in the previous section: in situations in which more than one morally significant aspect is involved, and thus the need to weigh the different aspects against each other arises, we usually do that using the betterness relations that we take to hold between different acts. See footnote 11, for a formal illustration of this point: the information that we need in order to assign degrees of belief to different betterness relations holding is strictly weaker than the information that we need in order to assign degrees of belief to different hypotheses regarding the degrees of moral value of different acts.

not that A is better than C. This condition does not follow, however, from the requirement that the agent believes with probability 1 that the betterness relation is transitive, as there are many probability distributions over the set of all possible orderings of alternatives that do not respect it.[8]

When this happens and one finds oneself in a situation in which in order to obey the LBC one must violate transitivity, at least one of these two requirements has to go. It is tempting to argue that this should be the LBC. This is so since it is clear that in such situations, by obeying the LBC, the agent will necessarily find himself violating a second-order moral judgement he should hold: the moral judgement that he ought to choose consistently when making moral choices.[9] He will be violating this judgement by violating transitivity. Does this consideration give us a reason to reject the LBC? Not on its own, I will argue now.

Implicit in the objection to the LBC presented in the last paragraph is the assumption that the moral judgement that one ought to always choose consistently ought always to have priority over any other moral judgement. However, this assumption is dubious. The rationality requirements get their normative force from our belief that rationality is a guide for choices that will best serve the agent's interests (in our case, moral interests: the interests of the agent when acting as a moral agent). The rationality here is instrumental rationality: there is no substantive moral value in obeying its requirements. The moral value of obeying its requirements comes from the further belief that doing so will best serve other purposes that do have intrinsic moral value.

[8] Think of any "Condorcet paradox" style situation. For example, consider the following table:

1/3	1/3	1/3
A	C	B
B	A	C
C	B	A

Here, the agent believes with probability 2/3 that A is better than B, that B is better than C, and that C is better than A, and still gives a positive probability only to transitive rankings.

[9] I do not argue that this is a moral judgement the agent should hold. I assume this for the sake of argument as I believe it is not an unreasonable position and is indeed a position to which many scholars are committed. John Harsanyi argued, for example, that "an individual making a moral value judgment must follow, if possible, even higher standards of rationality than an individual merely pursuing his personal interests" (Harsanyi 1978, 226). The argument that follows aims at showing that, even if one accepts the second-order moral judgement that one's moral judgements ought to be consistent, in the kind of cases that I discuss here this moral judgment is defeated by other considerations.

However, when an agent believes it is more likely that one act is better than another rather than *vice versa*, it is clear that what will best serve the agent's moral interests—in the absence of sufficient information about degrees of moral value—is to choose this act over the other. Requiring that such an agent does otherwise, in the name of transitivity, amounts to putting the cart before the horse. It amounts to requiring that the agent gives priority to a moral judgement that gets its moral force from more fundamental moral judgements over one of those more fundamental moral judgements.

One might argue that, from a wider perspective, giving such a priority is justified, since by choosing in an intransitive way, the agent exposes himself to "money pumps", or in the moral context, to "positive moral value pumps". However, this argument misses the point. If the agent has good reasons, *in a particular case,* to suspect that by choosing intransitively, he will be drawn into a money pump, then this consideration ought already to be taken into account through his assessment of the possible consequences of the acts available to him. However, the mere possibility of being money pumped, *without having any reason to suspect that this possibility will actually be realised*, should not matter much to an agent who must make a specific decision.

So I think it is not the LBC that has to go in such situations. However, by relaxing transitivity we are left without a decision rule. Which act should the agent choose in a case where he has intransitive moral preferences over three acts and all three acts are available to him?

Here is one possible answer: if we allow the agent to use mixed strategies, i.e., if we require that the set of acts available to the agent is convex then—under a natural condition that describes the connection between the agent's factual beliefs and the agent's moral beliefs—there always exists an act that the agent believes is more likely or equally likely better than any other act available to him. In other words, there exists an act such that the agent believes that no other act is better than it. It seems reasonable to require from the agent to choose such an act.

This requirement can be seen as a generalisation of the requirement of not to choose an act to which another act is preferred, which is usually used to justify the transitivity axiom (e.g., Davidson, et al. 1955). When it is impossible for the agent to have transitive preferences, and when there is no act that is preferred to all the acts over which the intransitivity occurs, then this requirement cannot be respected. However, I will now show that the generalization of this requirement

(i.e., *the requirement not to choose an act if there is another act that one believes it is, more likely than not, better than the first*) can always be satisfied if the set of acts is convex, provided that one natural condition holds.[10]

Here is the condition. First informally: the agent's degree of belief that one mixed act is better than another equals his expected degree of belief that this act is better than the other. In order to express this condition formally, we need some more structure.

Let $\Omega = \{\omega_1 \ldots \omega_n\}$ be a finite set of possible states. Let p be a probability distribution over Ω. Let D = {A, B, C…} be a set of outcomes and let $E = \{a_1 \ldots a_k\}$ be a set of acts, where an act is a function from Ω to D. Let ≥ be the agent's moral preference relation over E. In addition let >* denote the moral betterness relation between pairs of acts, i.e., >* is a binary relation over elements of E. For simplicity, we will assume that for any two elements, a_i and a_j, $a_i >^* a_j$ or $a_j >^* a_i$. By assuming this, I am ignoring here the possibility that the agent gives a positive probability to the possibility that two acts are equally good, i.e., that neither one of them is better than the other. This assumption will make the discussion simpler and nothing is dependent on it.

Since we want to allow the agent to have beliefs regarding the betterness relation, we will usually need to refer to the betterness relation as a variable. In these cases we will just use the notation ">". Finally, let q be a probability distribution over all possible >*s. To be clear, the expression $q(a_i > a_j)$ denotes the sum of the probabilities q gives to all >* such that $a_i >^* a_j$.[11]

---

[10] There are certainly other decision rules one might consider as plausible candidates in this respect. See, for example, Loome and Sugden's (1987) discussion of a similar question arising in the context of regret theory and SSB utility theory, and some of the articles in Arrow, Sen, and Suzumura's (2002) edited volume, which discuss a similar question in the context of social choice. I am not going to argue against any of these possibilities. Even if rationality does not require one to choose according to the decision rule I have offered above—which is in the spirit of Laffond, Laslier, and Le Breton's (1993) solution in the context of social choice—it would be hard to deny that rationality allows for using it. The discussion that follows in this section and the next one provides further support for this decision rule: by accepting it one gains a strong justification for the use of lotteries that does better—in terms of its ability to predict our moral intuitions regarding specific lotteries—than any other justification around.

[11] It is important to stress that by taking q to be a probability distribution over the set of all possible betterness relations, I do not commit myself, and do not intend to suggest, that either ordinary people or ideal rational moral agents deduce their beliefs regarding the betterness relations that hold between different pairs of acts from their beliefs over the set of all possible rankings of all the possible acts available to them.

As Leonard Savage (1972) does, we can define each element of D as the constant act (i.e., an act that gives the same outcome in every state) whose value is this element and require that E includes all the possible—constant and not constant—acts. With this we can treat the agent's beliefs regarding the betterness relation between constant acts as his beliefs about the betterness relation between outcomes, and the agent's preferences over constant acts as his preferences over outcomes. For convenience we will use the notation q(A>B) to refer to $q(a_A>a_B)$ when a_A is the constant act that gives A and a_B is the constant act that gives B.

In the interpretation, p represents the agent's degrees of belief about factual matters in the world, while q represents the agent's degrees of belief about the betterness relation between different acts. Now, we have the conceptual resources to formally express both the LBC (informally introduced above) and another constraint, the EBC.

Likelihood of Betterness Constraint (LBC):

1. $q(a_i > a_j) > q(a_j > a_i)$ iff $a_i > a_j$ and

2. $q(a_i > a_j) = q(a_j > a_i)$ iff $a_i = a_j$.

Expectation of Betterness Constraint (EBC):

For every two acts, a_i and a_j,

$$q(a_i > a_j) = \frac{\sum_{\omega_k : a_i(\omega_k) \neq a_j(\omega_k)} p(\omega_k) q(a_i(\omega_k) > a_j(\omega_k))}{\sum_{\omega_k : a_i(\omega_k) \neq a_j(\omega_k)} p(\omega_k)}$$

The agents might form their beliefs in such a way (although I find it psychologically implausible and normatively unappealing), but nothing in the model requires them to do so. This is because I do not assume anything about conditional probabilities; that is the probability of one act being better than another conditional on other betterness relations holding between other acts. Thus, I do not use any information that one gains from access to a specific probability distribution over the set of all possible rankings of the acts and that one does not have if one only has access to the probability of one act being better than another, for all pairs of acts.

Now we can see that the distinction made in the previous section between three types of moral uncertainty on a conceptual level—that is, the distinction between 1) moral uncertainty that can be reduced to uncertainty about non-normative propositions, 2) moral uncertainty that can be reduced to uncertainty about which moral theory is the correct one, and 3) "primitive" moral uncertainty—can be represented formally in a straightforward way: The first kind of moral uncertainty happens when there is no uncertainty regarding the agent's own preferences, the second happens when there is such uncertainty but all the probabilities, including the conditional probabilities are known to the agent, and the third happens when only non-conditional probabilities are known (or in other words, when the probabilities of conjunctions are not known).

Intuitively, the EBC says that one's degree of belief that one act is better than another should be equal to one's expected degree of belief that this act is better than the other, in case one of the two acts is better than the other. In other words it requires that the agent's degree of belief that one act is better than another is equal to the agent's degree of belief that the world is such that this act is better than the other.

It will be useful to demonstrate how the EBC works, using an example. Consider the following table.

Table 3

	$p(\omega_1) = 0.2$	$p(\omega_2) = 0.3$	$p(\omega_3) = 0.4$	$p(\omega_4) = 0.1$
a_i	A	B	C	B
a_j	B	C	A	B
a_A	A	A	A	A
a_B	B	B	B	B
a_C	C	C	C	C

Suppose the agent's degree of belief that outcome A is better than outcome B (that is that act a_A is better than act a_B) is 0.7, that his degree of belief that B is better than C is 0.8 and that his degree of belief that A is better than C is 0.9. What should his degree of belief be that a_i is better than a_j? According to the EBC it should be $(0.2 \times 0.7 + 0.3 \times 0.8 + 0.4 \times 0.1) / 0.9 = 0.4666$.

Here is how the calculation goes: firstly the agent should check in which states the two acts give the same outcome and ignore these states. In our example this only happens in state ω_4. Next, the agent should give each of the remaining states a weight which is equal to its probability and add up his weighted degrees of belief that act a_i is better than act a_j.[12] Lastly, he should normalise this sum by dividing it by the sum of the probabilities of all the states he did not rule out in the first stage. This last move is necessary in order for the agent's degrees of belief to be probabilistic.

To see why, given the EBC, there always exists a mixed act such that the agent believes regarding any other act that it is less likely or equally

[12] Notice that here I used the assumption that two acts cannot be equally good. It is easy to see that if we relax this assumption, the EBC will have to be slightly adjusted, but nothing significant will change.

likely better than this mixed act, let us start with the case of only three acts with regard to which the agent has intransitive preferences, if he obeys the LBC. We can do this by using the following example:

An agent has to choose between three acts that can bring about, in different states of the world, three possible outcomes: that all the 100 inhabitants of village A will die, that all 200 inhabitants of village B will die, or that all 400 inhabitants of village C will die. Assume that the agent is absolutely confident that it is better to save more people than fewer people, thus, q(A>C)=q(A>B)=q(B>C)=1. However, the choice he has to make is not between sure outcomes, but between the following three acts:

Table 4

	$p(\omega_1) = 4/9$	$p(\omega_2) = 3/9$	$p(\omega_3) = 2/9$
a_i	B	B	B
a_j	A	C	C
a_k	B	A	C

The agent is following the two conditions mentioned above:

1. EBC: for every two acts, a_i and a_j,

$$q(a_i > a_j) = \frac{\sum_{\omega_k : a_i(\omega_k) \neq a_j(\omega_k)} p(\omega_k) q(a_i(\omega_k) > a_j(\omega_k))}{\sum_{\omega_k : a_i(\omega_k) \neq a_j(\omega_k)} p(\omega_k)}$$

2. LBC: for every two acts a_i, a_j, $a_i \geq a_j$ iff $q(a_i > a_j) \geq q(a_j > a_i)$.

Now, since $p(\omega_2)+p(\omega_3)>p(\omega_1)$, he believes a_i is better than a_j to degree 5/9. Since $p(\omega_1) >p(\omega_2)$, he believes that a_j is better than a_k to degree 4/7, but since $p(\omega_2) >p(\omega_3)$, he also believes that a_k is better than a_i to degree 3/5 and thus he has intransitive preferences.

We are looking now for a mixed strategy, M, over the three acts such that the agent will believe that M is better than or equal to each one them. We can look at this in the following way. When the agent is using a mixed strategy, he adds some uncertainty to the uncertainty he already suffers from: he transforms any world ω_i to which he gives a positive probability into three worlds, the probability of each one of these being the multiplication of the probability of the original world by the probability that the mixed strategy the agent uses gives to one of the original acts. Here is how this is done in our example:

Table 5

	$p(\omega_1)$* $M(a_i)$	$p(\omega_1)$* $M(a_j)$	$p(\omega_1)$* $M(a_k)$	$p(\omega_2)$* $M(a_i)$	$p(\omega_2)$* $M(a_j)$	$p(\omega_2)$* $M(a_k)$	$p(\omega_3)$* $M(a_i)$	$p(\omega_3)$ *$M(a_j)$	$p(\omega_3)$* $M(a_k)$
M	B	A	B	B	C	A	B	C	C
a_i	B	B	B	B	B	B	B	B	B
a_j	A	A	A	C	C	C	C	C	C
a_k	B	B	B	A	A	A	C	C	C

Now, M is preferred or equal to a_i only when the agent believes it is more likely or equally likely that M is better than a_i, i.e., when the sum of the degrees of beliefs that the outcomes that M brings about in every possible world in which M and a_i bring about different outcomes, weighted by the probabilities of these worlds, is higher than this sum for a_i, i.e., since we assumed that the agent's degrees of beliefs regarding the betterness relations among pure outcomes are all equal to 1, when:

$$p(\omega_1)*M(a_j) + p(\omega_2)*M(a_k) \geq p(\omega_2)*M(a_j) + p(\omega_3)*M(a_j) + p(\omega_3)*M(a_k)$$

We can do the same for M in relation to a_j and a_k, and we get three inequalities with three variables. Each inequality can be derived from the other two, but we also know that $M(a_i) + M(a_j) + M(a_k) = 1$. It is easy to see that there is a unique solution to this system in which the equality relation holds for all inequalities. For the values in the example, this solution is when $M(a_i) = M(a_j) = M(a_k) = 1/3$, and in the general case:[13]

$$M(a_i) = (2q(a_j>a_k) - 1) / ((2q(a_j>a_k) - 1) + (2q(a_i>a_j) - 1) + (2q(a_k>a_i) - 1))$$

$$M(a_j) = (2q(a_k>a_i) - 1) / ((2q(a_j>a_k) - 1) + (2q(a_i>a_j) - 1) + (2q(a_k>a_i) - 1))$$

$$M(a_k) = (2q(a_i>a_j) - 1) / ((2q(a_j>a_k) - 1) + (2q(a_i>a_j) - 1) + (2q(a_k>a_i) - 1))$$

[13] By "the general case" I do not mean only that this solution holds for any p(.), but also that it holds for any q(.), and for any finite number of outcomes (over which the three acts are defined). Finding solutions for more than 3 acts is more difficult. I will show, however, that such a solution always exists.

These values also have an intuitive interpretation, which will be discussed in the next section. The story, however, does not end here, as it is easy to see that for every mixed strategy, such as M, there exist two other acts such that the agent has intransitive preferences over M and these two acts. In our example, this can be done in the following way:

Table 6

	$p(\omega_1)$* $M(a_i)$	$p(\omega_1)$* $M(a_j)$	$p(\omega_1)$* $M(a_k)$	$p(\omega_2)$* $M(a_i)$	$p(\omega_2)$* $M(a_j)$	$p(\omega_2)$* $M(a_k)$	$p(\omega_3)$* $M(a_i)$	$p(\omega_3)$* $M(a_j)$	$p(\omega_3)$* $M(a_k)$
M	B	A	B	B	C	A	B	C	C
N	A	A	B	C	C	A	C	C	C
L	B	A	B	A	C	A	C	C	C

The reasons are identical to the reasons for the intransitivity in the original example. However, notice that N and L are not mixed strategies over the three original acts. Given the set of the original acts and every mixed strategy over them, there is a unique mixed strategy that respects the condition that the agent should never choose a strategy when there exists another strategy available to him that he believes is, more likely than not, better. It seems, then, that in this kind of case the only rational choice for the agent is this mixed strategy.

What happens, though, when the set of available strategies contains more acts? For example, what happens if this set contains the three acts from our example, acts N and L, and every mixed strategy over these five acts? Is it still true that there exists a unique mixed strategy, M, over this set, such that there is no strategy in this set that the agent believes is, more likely than not, better?

The answer to the existence question is yes (I will get back to the uniqueness question soon). To see that, we can think of the agent as playing a game against himself in which the payoffs for every combination of strategies are the agent's degrees of belief that one of these strategies is better than the other: the intuition is that when the agent has to make a choice, my demand from him is that, given what he chooses, there is no other strategy he could have chosen that he believes will be better. So we can think of it in the following way: the agent looks at the strategies available to him and asks himself—for each

one of them—given that I choose this strategy, will there be a better strategy for me to have chosen? If the answer is yes he should not choose that strategy. It is easy to see that this condition holds for the two players in the game only when they play Nash equilibrium strategies.

Now, since the agent plays against himself, the game is symmetric: the strategies and the payoffs for each combination of strategies for the two players are identical. In the same way, since the two players represent the same agent, the equilibrium must be a symmetric one, since the agent can choose only one strategy. So what we have is a two-player symmetric game and every symmetric game has a symmetric Nash equilibrium (see Nash 1951).

To see things more clearly, let us construct such a game, using our original example. Each player has three pure strategies, a_i a_j and a_k and the payoff every player gets from choosing an act a, while the other agent chooses act b, is just his degree of belief that a is better than b. Since we assume that the agent ignores, in his reasoning, worlds in which the two acts give the same outcome, we can assign a payoff of ½ to every result in which the two players choose the same pure strategy. So here is the game:

Table 7

	a_i	a_j	a_k
a_i	½ , ½	$q(a_i>a_j)$, $q(a_j>a_i)$	$q(a_i>a_k)$, $q(a_k>a_i)$
a_j	$q(a_j>a_i)$, $q(a_i>a_j)$	½ , ½	$q(a_j>a_k)$, $q(a_k>a_j)$
a_k	$q(a_k>a_i)$, $q(a_i>a_k)$	$q(a_k>a_j)$, $q(a_j>a_k)$	½ , ½

Notice that if the agent has transitive preferences, i.e., if $q(a_i>a_j) \geq ½$, $q(a_j>a_k) \geq ½$, and $q(a_i>a_k) \geq ½$, the only Nash equilibrium is that both players play the pure strategy a_i. However, when the agent has intransitive preferences (which is the case we are interested in), i.e., when $q(a_i>a_j) \geq ½$, $q(a_j>a_k) \geq ½$, but $q(a_k>a_i) \geq ½$, there is no pure strategies Nash equilibrium. However, there is a mixed strategies equilibrium and in this case it is unique.

Now, although this kind of equilibrium is not always unique, the following still holds: given the set of all symmetric Nash equilibrium mixed strategies, the agent always has transitive preferences among them. The reason is simple. It is a well-known fact that in a mixed strategies Nash equilibrium, given the strategy played by player 1, player 2 is indifferent between all the mixed strategies available to him which are defined over the set of all rationalizable strategies. In particular, he is indifferent between all mixed strategies which belong to the set of symmetric Nash equilibrium mixed strategies. Thus, the transitivity requirement is satisfied in a trivial way: the agent is indifferent between all the mixed strategies in the relevant set. So what we have now is a choice rule that respects the requirement that the agent should never choose a strategy if he believes there exists another strategy available to him which is better, which sometimes recommends (i.e., whenever the agent has intransitive preferences) the use of a mixed strategy.[14]

To conclude, what we have shown is that if an agent respects the LBC and the EBC, then—even if he holds intransitive moral preferences— if the agent is allowed to use lotteries, there always exists a lottery which he believes is more likely than not better than all other definite acts or lotteries. Thus, for such an agent it seems that the only rational choice will be to choose this lottery (or, if this lottery is not unique, one of the lotteries included in the set).

Recall now the discussion in the first section. One of our lessons from that discussion was that any account—like Broome's account—that recommends lotteries also in cases where the agent is not morally indifferent between the definite acts which are available to him, must be committed to the claim that sometimes the objectively right thing to do is not to maximize expected goodness. By following the account presented here, we can see that the agent (subjectively) ought to choose a lottery exactly in those cases in which he cannot maximize any quantity anyway, i.e., when his preferences are intransitive.

To be more precise, what I am arguing is that whenever the agent does not suffer from moral uncertainty he should simply choose the

[14] Note that this result does not depend on the LBC and the EBC. Many other decision theories that allow for intransitive preferences can serve. For example, if instead of using the degrees of belief in the betterness relations as the payoffs of the game, we use expected regret levels, the situation will be the same. More generally, Peter Fishburn (1984) has proved that whenever intransitive preferences can be represented by an SSB utility function, this will be the case.

objectively right act. However, when the agent does suffer from some moral uncertainty and does not have an access to degrees of moral value, then if he obeys the LBC he might find himself having intransitive moral preferences. This does not mean that he believes the moral betterness relation is intransitive. We can assume that the agent believes it is transitive. However, since all he can rely on are his beliefs about this relation—in the kind of situations I have pointed to—he has no way to avoid intransitivity. Thus, in the cases where the intransitivity arises, it seems that *the only rational thing for him to do is to choose a lottery.*

So, in my account, choosing a lottery is not an irrational thing to do, but rather—whenever it is justified to choose a lottery—the only rational thing to do. It is clear that in this account there is no need to claim that sometimes the objectively right thing to do is not to choose the best act: one can (but not "must") argue that one ought always to choose the best act, but that when one is uncertain about which act that is, the only rational thing to do is to use a lottery. Is it also the best thing to do? Well, yes and no. No, in the sense that when choosing a lottery the agent knows for sure that there is another act available to him that brings a higher amount of expected goodness (but he does not know which act that is). Yes, in the sense that—given his uncertainty—this is the only rational thing for him to do, and if we accept that one ought to be rational in one's moral choices (which we should) then choosing the lottery is the only morally justified act (what I have called "the subjectively right" thing to do).

It turns out that this account also has some nice predictions regarding the kinds of lotteries we ought to use. Some of these will be discussed in the next section.

WHICH LOTTERIES ARE JUSTIFIED?

In this section, I will consider some of the predictions of my account regarding when, and which, lotteries are justified. I will aim to show that by accepting my account we gain an explanation for some judgements that, I think, are intuitive. This, in turn, serves as independent evidence in favour of my account.

Case 1

Let us begin with a paradigmatic case: there are three individuals, i, j, and k, all in need of a kidney. There is only one kidney available and the moral evaluator is uncertain regarding who should get the kidney.

His degrees of beliefs are such, though, that he believes it is more likely than not that i should get the kidney rather than j, it is more likely than not that j should get the kidney rather than k and it is more likely than not that k should get the kidney rather than i. In such a case, it is easy to confirm that my account will recommend the following lottery among i, j, and k:[15]

$$M(a_i) = (2q(a_j>a_k) - 1) / ((2q(a_j>a_k) - 1) + (2q(a_i>a_j) - 1) + (2q(a_k>a_i) - 1))$$

$$M(a_j) = (2q(a_k>a_i) - 1) / ((2q(a_j>a_k) - 1) + (2q(a_i>a_j) - 1) + (2q(a_k>a_i) - 1))$$

$$M(a_k) = (2q(a_i>a_j) - 1) / ((2q(a_j>a_k) - 1) + (2q(a_i>a_j) - 1) + (2q(a_k>a_i) - 1))$$

In other words, the weight individual i gets in the lottery, that is the chance that he will get the kidney (denoted $M(a_i)$), should be proportional to the moral evaluator's degree of belief that giving the kidney to j is better than giving it to k. Note that another sensitivity does not hold: it is not the case that the chance that individual i gets in the lottery is proportional to the moral evaluator's degree of belief that giving the kidney to i is better than giving it to j. On the face of it, I find this phenomenon surprising.

Now, this is simply a result of the assumptions presented in the previous section. However, here is one way to make this requirement intuitive. The moral evaluator believes that if k does not get the kidney, i should get it (since he believes that giving the kidney to i is, more likely than not, better than giving it to j). The only reason the evaluator thinks i should not get the kidney is that he believes it is more likely than not that it is better to give it to k than to i. Thus, to the extent that the evaluator believes the kidney should not go to k, he should give it to i. The extent that the evaluator believes the kidney should not go to k is his degree of belief that it is better to give the kidney to j than to give it to k. Thus, it makes sense that the evaluator should give the kidney to i with a probability that is proportional to his degree of belief that k should not get it, i.e., his degree of belief that it is better to give the kidney to j than to k.

[15] This is the case when the agent assigns probability 0 to the possibility of every two acts being morally equivalent. Relaxing this assumption does not change anything substantial, but it does make the mathematical expressions a little bit more complex. For the sake of clarity, then, I chose to use the assumption.

Case 2

Let us consider, now, the simplest case, which is also the one most discussed in the literature. This is the case in which there is no moral uncertainty and the moral evaluator is morally indifferent between two possible acts.[16] For example, consider a kidney case in which there is one available kidney and two candidates, identical in every respect that the evaluator takes to be morally relevant. In this case, although my account allows the use of a lottery, it does not make it strictly (subjectively) superior to either one of the two definite acts (i.e., of giving the kidney to one of the candidates). Broome's account (and any other account like it—in the sense discussed in the first section of this article) does make the lottery that gives equal chances to the two candidates (objectively) morally superior to both either of the two possible definite acts or any other lottery.

Although this is surely an advantage of Broome's account, as intuitively the lottery that Broome's account recommends in this case *is* strictly morally superior to any other possible act, it is not a weakness of my account. This is so because, as explained, my account is not a rival to Broome's account. When my account justifies using a lottery, it is because using a lottery is the subjectively right thing to do. When Broome's account justifies using a lottery it is because doing so is the objectively right thing to do. In the case considered here, there is no moral uncertainty involved and thus my account does not apply. It is not inconsistent with my account, however, to accept the claim that a different account—like Broome's—does apply in this case and does justify using a lottery (in the same way, it is not inconsistent with my account to deny that).

It is important to note, however, that there is a price that Broome must pay here. If the fairness consideration adds some moral value to the lottery in case there is no moral uncertainty and the evaluator is morally indifferent between the two candidates, it should do so also in the case where there is no moral uncertainty, but the evaluator is not morally indifferent between the two candidates.

For example, consider yet another kidney case involving only two candidates, but this time the candidates are identical in everything, apart from the fact that one has a slightly higher chance of a successful operation. According to Broome's account, there must be some cases in which a lottery between the two definite acts would be morally superior

[16] See, for example, Diamond 1967.

to the act of giving the kidney to the candidate with a slightly higher chance of a successful operation.

In order to generate a lottery under Broome's account you can reduce the difference in the chances of a successful operation between the two candidates as much as you want. At some point—if Broome's account is not empty—you will reach a difference in chances such that choosing a lottery between the two candidates will become morally preferred to simply giving the kidney to the one with the (slightly) higher chance of success.

However, if you are consistent in your choices, you will always make the same choice. Thus, if you face a similar choice over and over again you will always prefer the lottery to the option of simply giving the kidney to the candidate with the slightly higher chances. But no matter how small the difference is between the two candidates' chances of a successful operation, after making this decision enough times this will result in preferring a policy that generates more loss of life to one that generates less.

The trade-off has now become clear: if one is willing to accept that in the indifference case, a lottery is not strictly morally superior to the definite acts, one can deny that the fairness consideration is strong enough to lead to morally preferring a policy that generates more loss of life to one that generates less. If, on the other hand, one is willing to accept that sometimes a policy that generates more loss of life is morally superior to a policy that generates less, one can argue that in the indifference case, the lottery is strictly morally superior to any other act.

A third option is to retain both the judgement that in the indifference case the lottery is strictly morally superior to any other act and the judgement that when the evaluator is not morally indifferent and there is no moral uncertainty involved, a lottery is never justified. One can do this by limiting (in a somewhat artificial way) Broome's account (or any other account like it) to cases of indifference, or by arguing that, for some reason, Broome's account does not apply to the case we consider. This move is unattractive for obvious reasons, but these reasons are theoretical, not ethical.

Here, I am not going to argue in favour of any one of the three possible positions I have just presented. The point I want to make is that my account is consistent with all three of them and thus—no matter what your position regarding the above trilemma is—it should

not stop you from accepting my account as valid. The reason for that is, again, that my account is silent regarding the question of the objective rightness of lotteries, and this is exactly what is at stake here.

Case 3

Consider again a single kidney case, but this time there are ten people, i, j, k, and l1... l7, waiting for the kidney. Assume that the evaluator, after thinking about the decision for a while and gathering relevant information, summarises his judgements using the following table:

Table 8

Age	Chances of success	Any other relevant consideration
i	k	j
j	i	k
k	j	i
l1… l7	l1… l7	l1… l7

In other words, the evaluator believes that, from the point of view of the age of the candidates, i is more suited to get the kidney than j, j is more suited than k, and k is more suited than any of l1… l7. However, from the point of view of the chances for a successful operation, k is ranked above i, who is ranked above j, who is ranked above l1… l7. Finally, when the evaluator thinks of any other relevant moral consideration he ranks j above k, k above i, and i above l1… l7.

What should the moral evaluator do? One thing he can do is to try to give a relative weight to each one of the categories and, using these weights, derive a combined ordering. If he manages to do this and get a transitive ordering, I believe he should simply give the kidney to the person ranked at the top, which will be, of course, either i, j, or k.

The problem, though, is that this kind of case is exactly the kind in which the agent might become uncertain regarding which act is the best choice but does not have access to the degrees of moral value of the different acts (this is why the moral uncertainty arises in the first place) and so it might happen that—by following the LBC—he will find that he has intransitive preferences among i, j, and k. In such a case, my

account will suggest a lottery, but this lottery will give a positive chance only to i, j, and k and no chance at all to l1… l7.

To see why this is the case, recall the analogy with a game that I used in the previous section to show why there always exists a lottery that is weakly preferred to any other act. It was demonstrated that, when the agent chooses such a lottery, his choice must constitute a Nash equilibrium in the game he plays against himself.

Now, it is well known that a mixed strategies Nash equilibrium must give a positive chance only to rationalisable strategies, i.e., strategies that can survive the process of iterated elimination of dominated strategies. It is clear that giving the kidney to any of l1… l7 is not a rationalisable strategy because it is dominated by giving the kidney to either i, j, or k. Thus, according to my account, if the agent should use a lottery (which might or might not be the case depending on the agent's beliefs) this lottery must give a positive chance only to i, j, and k.

This seems to me very intuitive. Giving a positive chance to all of the candidates reduces the chances of i, j, and k, and this is so even though the evaluator is sure that it would be wrong to give the kidney to anybody but i, j, or k. So my account, again, gives the "right" recommendation in this case, although it was not designed in any way to do that.

CONCLUSION

I have presented an account of why choosing a lottery over a definite act is sometimes the (subjectively) right thing to do. According to this account, one ought always to choose the best act available when one can. When one cannot, one should use a lottery, and this is because using a lottery is the only rational thing to do in such a situation. So my account succeeds in satisfying both the requirement that moral preferences be rational and the requirement that one ought always to choose the best act available. Moreover, I have argued that the lotteries suggested by my account are the right ones.

One can accept the account presented here for the rightness of lotteries and reject other accounts, but one can also accept my account alongside other accounts as different valid justifications for the use of lotteries. One can also take the account presented here not only as an account of the rightness of lotteries, but also as an account of the fairness of lotteries, but one does not have to do so. If one does, then one can think of being fair as trying the best one can to do the right

thing. If one does not, than this is ok too, as long as one still believes one *ought* to try the best one can to do the right thing.

REFERENCES

Arrow, Kenneth J., Amartya K. Sen, and Kōtarō Suzumura (eds). 2002. *Handbook of social choice and welfare*. Amsterdam: Elsevier.

Bradley, Richard. 2007. A unified Bayesian decision theory. *Theory and Decision*, 63 (3): 233-263.

Broome, John. 1984. Uncertainty and fairness. *Economic Journal*, 94 (375): 624-632.

Broome, John. 1990. Fairness. *Proceedings of the Aristotelian Society*, 91: 87-102.

Broome, John. 1991. *Weighing goods.* Oxford: Blackwell Publisher.

Broome, John. 1994. Fairness versus doing the most good. *The Hasting Center Report*, 24 (4): 36-39.

Davidson, Donald, J. C. C. McKinsey, and Patrick Suppes. 1955. Outlines of a formal theory of value, I. *Philosophy of Science*, 22 (2): 140-160.

Diamond, Peter A. 1967. Cardinal welfare, individualistic ethics, and interpersonal comparison of utility: comment. *The Journal of Political Economy*, 75 (5): 765-766.

Fishburn, Peter C. 1984. SSB utility theory: an economic perspective. *Mathematical Social Sciences*, 8 (1): 63-94.

Glover, Jonathan. 1977. *Causing death and saving lives.* New York: Penguin.

Harsanyi, John C. 1978. Bayesian decision theory and utilitarian ethics. *American Economic Review*, 68 (2): 223-228.

Hooker, Brad. 2005. Fairness. *Ethical Theory and Moral Practice*, 8 (4): 329-352.

Karni, Edi. 1996. Social welfare functions and fairness. *Social Choice and Welfare*, 13 (4): 487-496.

Laffond, Gilbert, Jean-François Laslier, and Michel Le Breton. 1993. The bipartisan set of a tournament game. *Games and Economic Behavior*, 5 (1): 182-201.

Lockhart, Ted. 2000. *Moral uncertainty and its consequences.* New York: Oxford University Press.

Loomes, Graham, and Robert Sugden. 1987. Some implications of a more general form of regret theory. *Journal of Economic Theory*, 41 (2): 270-287.

Nash, John. 1951. Non-cooperative games. *Annals of Mathematics*, 54 (2): 286-295.

Rescher, Nicholas. 1969. The allocation of exotic life saving therapy. *Ethics*, 79 (3): 173-186.

Saunders, Ben. 2009. A defence of weighted lotteries in life saving cases. *Ethical Theory and Moral Practice*, 12 (3): 279-290.

Savage, Leonard J. 1972. *The foundations of statistics.* New York: Dover Publications.

Sepielli, Andrew. 2009. What to do when you do not know what to do? In *Oxford studies in metaethics*, vol. 4, ed. Russ Shafer-Landau. New York: Oxford University Press, 5-28.

Sher, George. 1980. What makes a lottery fair? *Noûs*, 14 (2): 203-216.

Steele, Katie. 2006. What can we rationally value? Manuscript presented at the *Third Annual Austin-Berkeley-CMU Formal Epistemology Workshop*. http://fitelson.org/few/few_06/steele.pdf (accessed April 2012).

Ittay Nissan-Rozen is post-doctoral fellow at the Edelstein centre for the history and philosophy of science at the Hebrew University in Jerusalem. His main research interests are in the areas of philosophy of economics, philosophy of probability, Bayesian epistemology, and formal ethics.
Contact e-mail: <I.nissan@lse.ac.uk>

Erasmus Journal for Philosophy and Economics,
Volume 5, Issue 1,
Spring 2012, pp. 73-86.
http://ejpe.org/pdf/5-1-int.pdf

The potentials and limitations of rational choice theory: an interview with Gary Becker

CATHERINE HERFELD
Witten/Herdecke University

Gary S. Becker (Pennsylvania, 1930) is a university professor at the Departments of Economics, Sociology, and the Graduate School of Business at the University of Chicago, Illinois. Becker earned his undergraduate degree from Princeton University and was awarded a PhD by the University of Chicago in 1955 for a thesis on the economics of discrimination, under the supervision of Milton Friedman. After teaching at Columbia University from 1957 to 1969, he returned to the University of Chicago where he has been based ever since.

Becker's work and research interests encompass a wide range of topics, unified by what he calls *The economic approach to human behavior* (Becker 1976). He considers this refined version of the neoclassical theory of consumer behavior as a method that can be applied to analyzing individual choices beyond the boundaries of traditional economics domains, including discrimination, education (human capital), crime, addiction, the family (marriage, divorce, fertility), and altruism. Becker's path-breaking work has been recognized with numerous honors, including the John Bates Clark Medal (1967), and the Presidential Medal of Freedom (2007). In 1992, he was awarded the Nobel Memorial Prize in Economic Sciences "for having extended the domain of microeconomic analysis to a wide range of human behavior and interaction, including nonmarket behavior" (Nobel Prize press release).

Professor Becker was interviewed by Catherine Herfeld at his office on the Campus of the University of Chicago on December 8th, 2010. The discussion ranged over a number of issues including the consequences of the recent financial crisis for the economics profession, the role of mathematics in economic modeling and the role of modeling in economics, the significance of the rationality-principle, and the

NOTE: Catherine Herfeld is currently a PhD candidate at the chair of economics and philosophy at Witten/Herdecke University (Witten), and a fellow at the Max-Planck-Institute for the History of Science (Berlin). Her thesis is on the history and philosophical foundations of rational choice theory. At the time of this interview, she was a visiting scholar at the Department of Philosophy of Columbia University (NY). Contact email: <Catherine.Herfeld@uni-wh.de>

development of Becker's 'economic approach' and its distinctiveness from behavioral economics.

CATHERINE HERFELD: ***Professor Becker, looking back on the recent economic crisis and the failure of the majority of economists to forecast what happened, do you think the economics profession faces a crisis?***

GARY BECKER: No, I do not think so. I think that the profession will be affected in the sense that people will be working on problems in order to understand the financial crisis—and people are doing this already. But forecasting major events like that is very hard to do in any field. And I think to hold that as a standard of what one can do in a field is not the right standard. It is true that specialists, maybe in fields such as 'asset pricing' and 'assets and banking', should have seen that banks had high ratios of assets to capital and that only a few people in the profession forecasted that that might be a problem, some of them maybe even only by accident. But I do not think that we can observe a crisis in economics.

What we can indeed observe is much more a tension, particularly in the macroeconomic literature, in trying to understand what actually happened. Already before the crisis, there was a literature on the financial sectors that was concerned to understand business cycles by making use of the so-called real business cycle theories. Those theories are on the defensive now. I think, as we go forward, theories of business cycles will have to give much more attention to the financial sector. In that sense the crisis has taught economists an important lesson, but it will not radically change what most economists are doing.

Following the crisis, many economists and methodologists have argued that more realistic behavioral underpinnings of economic theory would have made forecasts more accurate. Do you think that one of the things the recent crisis has shown us is that people just do not behave rationally? Or did the crisis rather show exactly the opposite—that people did in fact react to incentives and that the consequences of introducing new financial instruments were just not foreseeable?

I think it is mainly the latter. There were incentives, both on the borrower and on the lender side, that these subprime loans would be

made available at the lowest interest rates; and there was pressure from the government to do so; and probably those involved did not understand the financial instruments. Now, is it that we have to change our theories radically with respect to their behavioral structure or even switch to a new behavioral framework? There is very little evidence that would support such a move.

There is a whole field of behavioral economics that I follow pretty closely, and parts of it I have even contributed to. But did the behavioral economists predict the crisis any better? When taking a look at the literature, one does not find better results. The rational choice model is an abstraction and as is the case with all abstractions and all theories from whatever discipline, say physics, you abstract from some things that sometimes may be important. And this is also true of the rational choice model. In terms of understanding the crisis, I do not think that more realistic behavioral assumptions would solve the problem. It has always been difficult in rational choice models to adequately account for the coordination of people's expectations. To some extent, the crisis involved the coordination of irrational expectations. This might be something we should think about and improve.

With respect to how the crisis affects our models in terms of being based on a more realistic assumption structure: what will occur is that models become refined to help us understand what happened. But I do not see a fundamental change in the models with respect to the underlying structure of human behavior, nor do I see a need for such a change.

So was the crisis more a source for a critique against the rational expectations hypothesis, rather than towards the behavioral core of economic theory, i.e., rational choice theory?

Well, to some extent it was a critique of that hypothesis. The expectations turned out to be, to some extent, not rational; there is no question about that. Price increases were for example expected to continue. But the theory of rational expectations always said that people make a forecast and could coordinate on a bad forecast. That has always been part of the theory, however there is more attention being paid to that phenomenon now.

In the post war period, mathematization was (and still is) considered a prime virtue of economic theory, and important to improving the

scientific status of economics. Finance, especially, embraced highly technical models that produced precise calculations and predictions. However, it is claimed that it is exactly this extensive use of mathematical models that ultimately weakens the scientific status of the discipline; the failure of economists to predict the crisis being taken as evidence. What role should mathematics play in economics? And in which ways and to what extent can these mathematical models inform us about the complexity of the social world and of the economy?

Let me give you an example. The great depression was a far more serious crisis than this financial crisis we are currently facing. Economists made no use of mathematical models then. Did they predict that crisis very well? No, they did not. Going back and analyzing this failure is a good lesson to take. The economists back then did just as badly as the economists do now in terms of predicting. So, I do not think that the problem lies in the use of mathematics.

There is a lot of critique against mathematics in economics, from non-economists, from Austrian economists and from other groups, and I think it is misplaced. Mathematics can be a very useful servant; when it becomes the master, we are not in a good situation. However, I do not think it has become a master in economics. I think we made mistakes in understanding how economies move forward, even in understanding the pricing of derivatives. But one can make these mistakes, and plenty of mistakes have been made, without using any mathematics. Sociologists make a lot of mistakes without using mathematics. So I do not think that the problem is the use of mathematics per se.

The discipline will continue to be heavily mathematical but hopefully will learn from this crisis. I always say that mathematics is useful but you have to have good economic content. If you do not have good economic content then, whether you do it mathematically, verbally, or with a graph, you are doing bad analysis. I am not one of the most mathematical of the economists; a lot of people use much more mathematics than I do. But I have never thought that the use of mathematics is the problem. Only bad use of mathematics is a problem and will continue to be a problem. If we did it all verbally, would that improve our science? Economics was a verbal science until the 1940s and I would say we are now doing much better than the economists back then.

How do you think economics should be done?

The way I like to put it is that we have to have a dialogue between the theory or model and the data. Theory informs us about what data to look at and how to interpret the data. But data also informs the theory. So if you have theoretical predictions that continue to turn out wrong, you have to change the theory. As I said, the real business cycle theory ignored the financial sector. This crisis showed us that the financial sector is really important. Economists who are working in that area are going to change that now. And that is how I think it should be.

A discipline where the theory is isolated from what is going on in the world will become a sterile discipline. And a discipline that is only looking empirically without any modeling will also become sterile. I think the disciplines that are active, that are productive, are those that have an active dialogue between the two aspects. That does not mean that everybody has to be doing both, but I have always believed that the ideal economist works with theory, looks at the data, gets a data feedback on the theory and vice versa. Some people just work on theoretical issues and that is fine. Some people just put data together in a useful way, which is fine too. But the bulk of them should be looking at this combination.

Coming to your own work, you had a major influence on 20th century economics by introducing a broad range of human motives into economic theory, something that could be considered as going into the same direction as what is today known as behavioral economics. Yet, unlike behavioral economists, you retained "an irrational passion for dispassionate rationality", as you once expressed it.[1] Why did you decide to stick with the rationality-principle in explaining and predicting human behavior?

I felt that the rationality-principle was a powerful tool that was useful for explaining behavior. In all my work, even if it is purely theoretical, I am looking at data, talking about observations, sometimes even gathering data. For example, when I used a rational model of altruism to look at the family, it seemed to help me to understand why parents do support their kids and under which conditions they do so, to analyze relations between spouses, and so on (e.g., Becker 1981). When I went to crime, I thought it helped us to understand crime and deterrence and

[1] Oral remark at a luncheon seminar at Tulane University in 1986 (Choi 1993, 19 fn.).

the effects of education (e.g., Becker 1968). In areas where the rational choice model does not work so well, one has to modify it, but I have been persuaded, at least by my own thinking and by looking at the world and the actual data, that it does a very good job, and that there is no other comparable approach in the social sciences with the same degree of explanatory power, or even anywhere near.

And you can look at a lot of behavior, not only to bring in aspects such as altruism or envy, which are now part of behavioral economics, but also to include the idea that people discount the future hyperbolically instead of exponentially. David Laibson, who introduced the idea of hyperbolic discounting into economics, makes use of axiomatization (see, for example, Laibson 1997). Everything else he uses is very much the same as what I use. Maybe people are hyperbolic; if that is true, we have to alter our theories, but that is where we would need feedback from the actual data. Yet we still have a basic framework in use, which is some version of the rational choice model. If you would start to abandon this framework, you would end up with a loose group of findings. If you want to abandon rational choice theory altogether, you have to substitute it with a new framework, and I do not see any new framework available at the moment—neither in the behavioral economics literature nor anywhere else—that has comparable explanatory and predictive power. That is the test.

It is an old saying that you need a theory to beat a theory. That does not mean that you cannot extend the existing theory or modify it—you can and you should. As we learn more, we will modify rational choice theory. Maybe fifty years from now it will not be like rational choice theory anymore, because by then it will have been modified and changed in so many ways. That is how things evolve. Einstein modified Newtonian mechanics, but Newtonian mechanics is still applied to a wide range of phenomena.

Would you consider behavioral economics to be a revolution in economics analogous to the transition from Newtonian mechanics to Einsteinian relativity theory, in that, although the rationality-principle has been questioned fundamentally, the rational choice model remains in use in economics as a parallel framework for explaining a wide range of behavior?

Well, of course I would not say that behavioral economics has been as important as Einstein's revolution; I would not compare them with one

another. In fact, I do not think that behavioral economics is a revolution. However, it has added some insights into human behavior and those insights, to the extent that they are verifiable, will be absorbed into the rational choice model. They will not lead to a radical change of the model. The real issues are how important are those insights and where do they apply?

So, for example, the explanation that consumers were somehow misled in the credit market, and that this in turn contributed to the financial crisis: I think there is very little empirical support for that. A lot of consumers were making pretty rational decisions, even those who were taking out mortgages with low interest rates and low down payments. Maybe they were going to default. But they did not default on their own capital. They defaulted on the lender's capital. So I see very little evidence from this that consumers are not rational, in the sense that the rational choice model cannot explain most of what they did.

How would you assess the epistemic value of these recent developments, such as behavioral and experimental economics, in terms of providing us with new knowledge about how the economy works?

I think they have been stimulating in terms of leading to further lab and field experiments. They have shown that people can be fooled by how a question is framed. And they have shown that similar people in different contexts may behave differently. I think that all this is valuable. But at the bottom line, economists deal with markets and group responses, and there it is hard to see major modifications coming out of behavioral economics as yet.

But has there not been a shift in the last eighty years towards economics becoming concerned with explaining individual choices rather than aggregate behavior?

Well there was a shift in putting much more reliance on individual choices in terms of modeling behavior. But when you look at the data economists have mainly used, the ultimate goal of most of the studies is to look at how people respond to incentives. Economists can make use of individual data panels and other data based on observations about the individual. However, what we are interested in are aggregates and market relations. For example, if the political aim is to subsidize education, economists do not care about how you respond or I respond

in particular. Maybe there are differences in how Germans respond or Americans respond, or how people who study at the University of Chicago respond in comparison to how students from Columbia University respond, and I guess that we would care about that. But not about how the individual responds. In my opinion, this is a fundamental difference between psychology and economics.

During the 1970s and 1980s various psychologists, such as Herbert Simon, Daniel Kahneman, and Amos Tversky, showed that people systematically violate the rationality principle and argued that economics could be improved by making the underlying psychological assumptions of economic theory more realistic. How seriously did you take these new developments at the time? Did they influence you or did they make you question your own approach? Did they inspire your later work—for example on endogenous preferences in your Accounting for tastes (Becker 1996)?

Well, it is hard to know where the influences come from. How I proceed in my work is that I try to keep up with what is being researched on in the discipline and then I think about potential contributions. Kahneman and Tversky made contributions that were very influential and highly cited in economics. So, for example, looking at the utility function as hinging around some usual position—that there is a reference point and you are highly risk averse towards losses, and so forth. I do not think it is completely clear from market evidence that its effects are significant, but I do believe the reference point analysis is itself important and maybe some of that literature influenced the work I did with Luis Rayo, very formal work on evolutionary theory where we show how to derive reference points and some other properties (Rayo and Becker 2007).

Their work has been important to psychologists and it has had some influence on the economics profession. And I would say it has had some influence on my work, although it is hard for me to know exactly how much. But I like to believe that my work has evolved and that what I believe today is not the same as what I believed in the 1980s and 1970s. I learn from what other people are doing; that is what intellectual interaction does for you. So I do not think their work has radically changed my approach, but I have been affected by it.

Turning to a somewhat different issue: rationality is a concept that originated in philosophy and its various economic formulations and

uses have been discussed extensively in the philosophical literature on the methodology of economics, such as by Alexander Rosenberg, Philip Mirowski, D. Wade Hands, and Mark Blaug. Were you ever interested in that literature? Or where did you get inspiration from when thinking about improving how rationality is conceived of in economics?

Primarily, I get inspiration from my own discipline, economics. For example, I wrote my doctoral dissertation on racial discrimination.[2] I wrote three papers before I did my work on discrimination. One was with Milton Friedman on Keynesian models (Friedman and Becker 1957), one was on monetary trade (Becker and Baumol 1952), and one was on international trade (Becker 1952). I would say that the last two were rather traditional papers; the one with Friedman was not traditional as we were very critical of Keynesians, but I got a lot of that from Friedman, so it was more him than me.

The work on discrimination I would say was my own work. As an undergraduate, I always felt that economics was too narrow. I thought of being a sociologist, but I found sociology difficult and so I was not satisfied with it. Friedman really taught me—although his own way did not take that path—that economics could be a powerful tool, and I began to think about racial discrimination and how economists were not discussing such an important topic that affects so many people. That is how I got into it. It was not from the methodological literature. I read some of that literature. I read Karl Popper and I studied Rudolf Carnap when I was a graduate student here at the University of Chicago, so I did read a lot of philosophical literature that was relevant to economics. But I cannot say that it directed me towards the topics I dealt with.

Friedman's article on methodology of economics (Friedman 1953) was very important—that I did read very carefully. I knew that article very well and it influenced me to a certain extent. Maybe not on the topics that I chose, but more in how I approached these topics. That value theory does not have to be realistic in any dimension was the part influenced by Friedman, and he got it from people like Karl Popper and others.

[2] Later published under the title *The economics of discrimination* (Becker 1957).

In an interview, the economist Leonard Rapping said that "many Chicago people would argue that the world is in fact competitive. They tend to believe their own pragmatic myth" (Klamer 1984, 221), i.e., that people in fact maximize profits and utility. In your own work, you have often denied a commitment to a realist interpretation of the rationality-principle. In your Nobel Lecture for example you state that the economic approach is a "method of analysis, not an assumption about particular motivations" (Becker 1993, 385). However, to make your approach work, the rationality-principle seems to figure as a behavioral assumption. It has to be at least approximately true to provide meaningful explanations and predictions; one cannot derive a true conclusion from false premises. Could you comment on this seeming contradiction? Does this for example reflect the influence on your work of the strong version of instrumentalism propagated by Milton Friedman?

The way I restated Milton Friedman's view in my own thinking is that one cannot evaluate a set of assumptions individually. You have to evaluate the whole set of assumptions collectively, because that is what a model is: a collective set of assumptions about behavior that is predicting behavior. And how do you evaluate a collective set of assumptions? It is very difficult to say "this assumption does not work" a priori, because it is the collective set what is relevant. The only way to evaluate assumptions is to ask whether this collective set of assumptions is in fact explaining behavior. Are you doing well in predicting and understanding how people respond to a tax cut, tariffs, globalization, returns to education, and the like? So that is my methodology.

Now, it is true that I like to believe that the individual assumptions are in some sense reasonable, but you have to look at them together. And I think that this is a problem with the behavioral economists. They take an assumption, for example that people cannot calculate probabilities very well and that there are other people on the other side who will try to exploit that weakness in them and will offer them various deals. If there is competition on the other side of the market that will mean that they will be offered some compensation. That is what competition does. For example if we play any kind of gamble, let us say we throw a dice and you think that the most likely outcome is ten, well I can exploit that. I do not even have to cheat, let us play. But if a lot of people are going to want to exploit you, we are going

to have to compensate you. So you have to ask what the market equilibrium looks like. You are a fool, but the market is competing to take advantage of that. That is how I would analyze that problem, and that has been one of my critiques of some of the behavioral economists: they do not embed their insights in a complete model of behavior.

Are you thereby implying that equilibrium analysis is a worthwhile undertaking?

You do not need to use the concept of a complete equilibrium. You can do it with modified equilibrium: as long as there are other people who are recognizing that I am a fool, they would compete. If you were the only person with that information, you could exploit it. A monopolist for example could exploit me. One of the great advantages of competition is that it prevents such exploitation, and you do not need perfect competition to have a strong effect in that direction.

So how does equilibrium analysis feature in your 'economic approach'?

Economists from the Austrian school hate equilibrium analysis in some sense, but I never understood their criticism. What do philosophers not like about equilibrium analysis?

Philosophers raise several objections against this way of analyzing the economy, one being that the application of the concept of equilibrium to an environment which is actually never in equilibrium is meaningless and does not provide us with any understanding about the real world. Take, for example, comparative statics.

But what do they substitute for it?

Well, philosophers do take a critical perspective; they often tend to evaluate the shortcomings of a theoretical framework first.

Yes, but as I said before, you need a theory to beat a theory. I think the equilibrium concept in economics is very subtle. It could take into account and does often take into account dynamic issues, changes—it is not static and it is not stationary. You have dynamic models of behavior that incorporate the concept, so they are still equilibrium models but it is dynamic equilibrium. It is a broad issue.

I have read some of the literature on the critique of equilibrium, not so much by philosophers but by the Austrian school of economics, and I could just never make sense out of it, because I do not see what they are substituting for it. Even Friedrich Hayek, who is listed as one of the top Austrians, if you read his analysis, you see that he is using equilibrium analysis.

But Hayek suggested the concept of a 'reflective equilibrium', which has however so far not been formalized.

Well, dynamic equilibriums can be formalized, as dynamic general equilibrium analysis does. I agree with you that a lot of analysis needs to be dynamic and comparative statics is not the right analysis for every issue. In economics, we are of course trying to improve, but I think we can do so with the tools we have available. I do not think that there is in principle any philosophical barrier to doing so. We could do it with a rational choice model or any other model. I do not think that would destroy the concept of equilibrium, and I do not think we should try to destroy it, because I think it is a very valuable concept.

In your work, you mainly look at aggregate demand and supply curves. In your textbook Economic theory (1971), for example, you look at a model of the irrational behavior of households to show that "the basic demand relations are derived fundamentally from scarcity alone rather than from an assumption that behavior is 'rational' and that the main conclusions of demand analysis [i.e., negatively inclined market demand curves] stem from a much more general principle than rational behavior—the scarcity of resources that defines an economic problem. Accordingly, we are able to derive the usual demand functions even when households behave 'irrationally'" (Becker 1971, 11 fn.). Yet, you go on to use the concept of rationality because of the power of the implication "consumers prefer more to less" (which is empirically questionable) and you say that a model that implies such behavior is to be preferred. Why is that to be preferred? And why exactly do you need rationality if scarcity is what is fundamental? Or, to put it differently, do you think that the rationality-principle is a necessary ingredient of economic theory and, if so, why?

Do you think it is empirically questionable that people prefer more to less? I do not think so. I do not see people giving away, except to charity

but that is another good we would introduce. Yet, looking at the world, do you see many people who prefer less to more?

Well people become satiated or even reject striving for endless material improvement.

Are people satiated with regard to leisure? Anyway, let us get back to your question. I argued that the rationality assumption is required to introduce the aspect that people prefer more to less, which in turn helps us to understand market outcomes and explain prices. In the analysis you cited I assumed that prices are given to consumers and producers, but I think you need to have some analysis to answer the question, where do prices come from? Maybe there are other theories you could use but you need to amend or add to the probabilistic type models some theoretical framework that tells you what types of prices are finally picked out. And for that, rational choice analysis really is very useful.

Are you after truth?

Absolutely! I think there is a truth out there. We are only approximating it but we are getting better. I think that the goal is to find the truth and I think there is something like that. I know that there is a lot of philosophical discussion about what truth is, which I however do not really find useful. I think there is a truth, and I think that economists have found a significant amount of the truth in economic behavior. There are a lot of things we do not know, but there are also a lot of things we do know, which the non-economist gets completely wrong. A simple idea like showing that when gasoline is substituted, people are going to buy more gasoline elicits some truth about people's behavior. And these are important truths. This is what I call the truth in a particular case and this is what I want to find out and analyze. So, yes, I am after the truth.

REFERENCES

Becker, Gary S. 1952. A note on multi-country trade. *American Economic Review*, 42 (4): 558-568.

Becker, Gary S. 1957. *The economics of discrimination*. Chicago (IL): Chicago University Press.

Becker, Gary S. 1968. Crime and punishment: an economic approach. *Journal of Political Economy*, 76 (2): 169-217.

Becker, Gary S. 1971. *Economic theory*. New York: Alfred A. Knopf.

Becker, Gary S. 1976. *The economic approach to human behavior*. Chicago (IL): University of Chicago Press.

Becker, Gary S. 1981. *A treatise on the family*. Cambridge (MA): Harvard University Press.

Becker, Gary S. 1993. Nobel lecture: the economic way of looking at behavior. *Journal of Political Economy*, 101 (3): 385-409.

Becker, Gary S. 1996. *Accounting for tastes*. Cambridge (MA): Harvard University Press.

Becker, Gary S., and William J. Baumol. 1952. The classical monetary theory: the outcome of the discussion. *Economica*, 19 (76): 355-376.

Choi, Young Back. 1993. *Paradigms and conventions: uncertainty, decision making, and entrepreneurship*. Michigan: University of Michigan Press.

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

Friedman, Milton, and Gary S. Becker. 1957. A statistical illusion in judging Keynesian models. *Journal of Political Economy*, 65 (1): 64-75.

Klamer, Arjo (editor). 1984. *Conversations with economists: new classical economists and opponents speak out on the current controversy in macroeconomics*. Totowa: Rowman & Littlefield Publishers.

Laibson, David. 1997. Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, 62 (2): 443-478.

Rayo, Luis, and Gary S. Becker. 2007. Evolutionary efficiency and happiness. *Journal of Political Economy*, 115 (2): 302-337.

The Nobel Prize in economics 1992: press release. Nobelprize.org Website. http://nobelprize.org/nobel_prizes/economics/laureates/1992/press.html (accessed January 30, 2012).

Gary S. Becker's Webpage: http://home.uchicago.edu/gbecker/

Erasmus Journal for Philosophy and Economics,
Volume 5, Issue 1,
Spring 2012, pp. 87-92.
http://ejpe.org/pdf/5-1-br-1.pdf

Review of *The end of value-free economics*, edited by Hilary Putnam and Vivian Walsh. Routledge, 2011, 230pp.

DANIEL LITTLE
University of Michigan-Dearborn

This recent volume edited by Vivian Walsh and Hilary Putnam, collating contributions to a debate in the pages of the *Review of Political Economy* between 2000 and 2009, brings to a fine point a line of argument that has been brewing for fifteen years: is the logical positivist insistence on separating "fact-based" science from "value-based" ethics any longer tenable? Most particularly, are there now compelling reasons for declaring that mainstream economics needs to recognize that the distinction is wholly untenable? Is the zeal for insisting on "positive" economics now unsupportable? Should economists at last recognize that Lionel Robbins's strong exclusion of normative language from the science of economics (1932) was both unjustified and unwise? Walsh and Putnam argue that the answer to each of these questions is definitive: the strict dichotomy between fact and value in economics can no longer be supported.

The contemporary issue of facts and values in the sciences was constructed from a number of sources within the empiricist tradition. There is Hume's claim that we cannot derive "ought" from "is"; or, in other words, that moral judgments are logically independent from empirical beliefs. There is the positivists' criterion of significance, according to which the meaning of an utterance reduces to the empirical experiences that would demonstrate its truth or falsity. (The two propositions together imply that moral sentences are meaningless or "non-cognitive", since the first proposition holds that no empirical experience can demonstrate the truth or falsity of a normative statement.) And there is the positivists' view that science is exclusively concerned with "facts"; and, since the first two propositions consign moral statements to the category of "value" rather than "fact", therefore science cannot contain normative vocabulary. Another source was internal to debates within neoclassical economics itself: Lionel Robbins's arguments against interpersonal comparisons of utilities, based on the

idea that making such comparisons unavoidably involves taking an evaluative stance towards the individuals in question (Robbins 1932).

The key idea advanced in *The end of value-free economics* is that none of these philosophical ideas have survived the critique of positivism offered within philosophy of science and philosophy of language over the past fifty years. The attempt to draw a sharp line between "fact" and "value" turns out to be impossible. And this is equally so in economics.

Consider an example. The concept of Pareto efficiency is defined in value-neutral terms: a distribution is Pareto-efficient if there is no other distribution that benefits some individuals without harming at least one individual. Whether or not a distribution is Pareto-efficient can be determined without making any ethical or value-based judgments. (There is, to be sure, a corresponding normative principle: "If a distribution is Pareto-efficient then it is unjust to alter the distribution". This principle is indeed value laden, but it is not part of the economic concept of Pareto efficiency.) The concept of distributive justice is not value-neutral; it invokes the idea that some distributions are better because they are fairer or more just than others. The positive economist holds that such distinctions are legitimate to make—in some other arena. But within economics the language of justice and equity has no place. The economist, according to this view, can work out the technical characteristics of various economic arrangements; but it is up to the political process or the policy decision-maker to arrive at a governing set of normative standards. Walsh and Putnam (as well as Amartya Sen) dispute this view on logical grounds; and thereby show that the discipline is free to have a rational and reasoned discussion of the pros and cons of various principles of distributive justice.

Raising the issue of value-neutrality for economics is a frontal assault on the uncritical positivism that neoclassical economics incorporated from the 1930s onwards. But it is also an attack on something else—the no-longer acceptable idea that economists can only tell us how things are, not how they should be. Is famine worse than food sufficiency? Is literacy better than illiteracy? Is good health an improvement in wellbeing? If we take the view that "positive economics" cannot contain normative judgments, then none of these questions could be answered by an economist. "It depends on what you value". What Walsh, Putnam, Sen, and other contributors to this volume want to say is that this response is idiotic, and there is no basis in logic,

science, or methodology that would support it. Of course economics, and economists, can find that starvation is a bad thing. Instead, they maintain that the best philosophy of language and philosophy of science supports the idea that value concepts and descriptive concepts are intermingled or "entangled", and that we can offer good reasons and evidence for evaluating claims involving both.

Why, some readers will ask, has Hilary Putnam become a central figure in this emerging debate? Putnam is known as a technically astute philosopher of mathematics, logic, and physics, and a philosopher of language; he is known for a sometimes wavering adherence to several versions of scientific realism; and he has made contributions of the greatest importance to each of these fields. But how did he come to get deeply immersed in the issue of the role of values in economics?

Vivian Walsh is one important part of the answer. Walsh undertook a series of publications in the 1980s and 1990s that were critical of the logical positivist assumptions that have lingered within the methodology of neoclassical economics (Walsh 1987; 1994; 1996). He took encouragement from the writings of Amartya Sen on welfare economics that confidently dismissed these positivist assumptions—for example, the idea that science could not incorporate values or that statements about values were meaningless. (Lionel Robbins was offered as a particularly clear advocate of these views.) And Putnam after reading Walsh's (2000) reconsideration of Adam Smith worked up his reactions to these ideas into a novel book in 2002, *The collapse of the fact/value dichotomy and other essays* (Putnam 2002).

A key construct in the collaborative thinking that Putnam and Walsh have done together is the idea of the "second phase of classical theory". (Harvey Gram discusses this construction in detail in his contribution.) Walsh introduces the idea and Putnam follows up in his essay. What this refers to is the fact that classical political economy, as founded by Smith and Ricardo, underwent a major intellectual revival in the 1960s when thinkers like Piero Sraffa proposed reappropriating some of their key analytical ideas. Sraffa's *Production of commodities by means of commodities: prelude to a critique of economic theory* (1960) was a key product of this rethinking. The rethinking itself came about because of an uneasiness about the premises of neoclassical economics. According to Walsh, this phase restricted itself to providing a rigorous reformulating of the technical core of classical economics rather than reintroducing the broader context of argument offered by Adam Smith.

The first phase of the revival focused on Ricardo, but the second phase, Walsh argues, has given a much more nuanced interpretation of Smith himself. Walsh finds that this reconsideration has been led by Amartya Sen and is more wide-ranging. Here is why Walsh thinks this reconsideration of Smith is important:

> This is because Smith embedded a remarkable understanding of the core concepts of a political economy whose implications for moral philosophy he understood and explored. The Smith texts as a whole offer a rich tapestry, interweaving threads of classical analysis, moral philosophy, jurisprudence, and history (p. 7).

And here is how Putnam summarizes Sen's contribution to this reconsideration of classical political economy:

> If we are to understand Sen's place in history, the reintroduction of ethical concerns and concepts into economic discourse must not be thought of as an abandonment of "classical economics"; rather it is a reintroduction of something that was everywhere present in the writings of Adam Smith, and that went hand-in-hand with Smith's technical analyses. This is something that Sen himself stresses (quoted by Walsh, 29).

Amartya Sen has argued robustly throughout his career for the feasibility of philosophical and analytical reasoning about value issues—in economics and elsewhere—a very early place where Sen takes up this topic is in "The nature and classes of prescriptive judgements" (1967). Much of what Sen brings to this debate within economics, according to Walsh and Putnam, is found in his capability approach as a foundation for a theory of welfare or wellbeing. It is based on the idea of human 'functionings' (what people can be and do), in which there is a plain intermingling of factual and evaluative ideas. We need to know what kind of lives people can lead before we can say how well off they are. And this means bringing human values into the discussion of economics at the beginning rather than at the end. Putnam draws attention to Martha Nussbaum's list of central human capabilities (Nussbaum 2000). Anyone reading Nussbaum's descriptions would agree that they presuppose human values. And Nussbaum (as well as Sen and Putnam) believes that we can rationally discuss and evaluate these. But if welfare economics is to incorporate a substantive notion of human wellbeing, then it plainly cannot be maintained that it is "value-free".

Another important locus for Sen's reintroduction of ethical concepts into economics is his critique of the narrow conception of individual economic rationality. As Sen puts the point in "Rational fools" (1977):

> A person thus described may be "rational" in the limited sense of revealing no inconsistencies in his choice behavior, but if he has no use for these distinctions between quite different concepts, he must be a bit of a fool. The purely economic man is indeed close to being a social moron. Economic theory has been much preoccupied with this rational fool decked in the glory of his one all-purpose preference ordering. To make room for the different concepts related to his behavior we need a more elaborate structure (p. 336).

Sen introduces the idea of "commitments" directly into the concept of economic rationality. Individuals choose among preference rankings based on their commitments—to each other, to political ideas, to groups with whom they have decided to affiliate. And this brings normative ideas directly into economic reasoning at the level of the actor—and therefore into the domain of economics.

Walsh and Putnam insist on a point that seems very important to me as well: it is the *dichotomy* based on strict separability between facts and values, or between positive and normative analysis, that they reject. They do not reject the idea that there are facts and there are values. But they believe in important respects these categories are intertwined and inseparable. They argue for "entanglement" and "rich description". So we can distinguish between value aspects and factual aspects of a situation; but we cannot strictly separate these aspects in our reasoning about the situation. Putnam writes in his contribution to the volume, "I try to show that value judgment and factual judgment are entangled in many ways, not just one. But one of the most important ways is this: there are facts […] which only come into view through the lenses of an evaluative outlook" (p. 112). Both Putnam and Walsh believe that it is fully possible and acceptable to engage in rational debates over the best theory of justice, or human nature, or human freedom; and to do so within economics as well as outside of economics. And they believe that science can pursue its goals without this sharp dichotomy.

This is a valuable contribution, offering the most sustained presentation to date of the reasons why the fact-value dichotomy cannot stand as a criterion of adequacy of the science of economics. It is a stimulating and rigorous conversation among a set of highly gifted philosophers and economists who have engaged deeply with the

underlying issues in the philosophy of science and the philosophy of economics. Anyone who wants to see the development of a discipline of economics that is better able to confront the economic and social challenges of the twenty-first century will certainly want to read it.

REFERENCES

Nussbaum, Martha. 2000. *Women and human development: the capabilities approach*. Cambridge: Cambridge University Press.

Putnam, Hilary. 2002. *The collapse of the fact/value dichotomy and other essays*. Cambridge: Harvard University Press.

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

Sen, Amartya. 1967. The nature and classes of prescriptive judgements. *The Philosophical Quarterly*, 17 (66): 46-62.

Sen, Amartya. 1977. Rational fools: a critique of the behavioral foundations of economic theory. *Philosophy and Public Affairs*, 6 (4): 317-344.

Sraffa, Piero. 1960. *Production of commodities by means of commodities: prelude to a critique of economic theory*. Cambridge: Cambridge University Press.

Walsh, Vivian. 1987. Philosophy and economics. In *The New Palgrave: a dictionary of economics*, eds. John Eatwell, Murray Milgate, and Peter Newman. London: Macmillan, 861-869.

Walsh, Vivian. 1994. Rationality as self-interest versus rationality as present aims. *American Economic Review*, 84 (2): 401-405.

Walsh, Vivian. 1996. *Rationality, allocation, and reproduction*. Oxford: Oxford University Press.

Walsh, Vivian. Smith after Sen. *Review of Political Economy*, 12 (1): 5-25.

Daniel Little is professor of philosophy and chancellor of the University of Michigan-Dearborn. His research interests concern the philosophy and methodology of the social sciences. His most recent books include *The future of diversity* (Palgrave, 2010; edited with Satya Mohanty) and *New contributions to the philosophy of history* (Springer, 2010). He publishes a web based monograph, *Understanding society*, on the philosophy of social science and the nature of the social world.
Contact e-mail: <delittle@umd.umich.edu>
Blog: <http://understandingsociety.blogspot.com/>

Erasmus Journal for Philosophy and Economics,
Volume 5, Issue 1,
Spring 2012, pp. 93-99.
http://ejpe.org/pdf/5-1-br-2.pdf

Review of Malcolm Rutherford's *The institutionalist movement in American economics, 1918-1947: science and social control*. Cambridge (UK): Cambridge University Press, 2011, 410pp.

DAVID GINDIS
University of Hertfordshire Business School

In publications spanning just over three decades, Malcolm Rutherford has been an important contributor to the study of American institutionalism as both a theoretical framework and an intellectual movement. His earlier research appraised the ideas of key figures of American institutionalism, namely Thorstein Veblen, John R. Commons, Wesley C. Mitchell, and Clarence Ayres. Following the rise of the so-called "new institutional economics" in the 1980s, associated with future Nobel laureates Ronald H. Coase, Douglass C. North, and Oliver E. Williamson, many of Rutherford's subsequent publications involved the comparative assessment of this "new" institutionalism and what had become by implication the "old". This line of inquiry gave us the earlier monograph, *Institutions in economics: the old and the new institutionalism* (Rutherford 1994).

For the last fifteen years or so, Rutherford's efforts have centered on developing a systematic study of American institutionalism in the interwar period, thereby meeting the challenge spelled out in his Presidential address to the History of Economics Society (Rutherford 1997). The result, based on meticulous archival labor, is this goldmine of a book. Although much of the material Rutherford invites us to consider has been previously published in a series of journal articles, this book, published in the "Historical Perspectives on Modern Economics" series edited by Craufurd D. Goodwin, brings together what we had only seen glimpses of in a well-written and tightly-knit narrative.[1] This is the definitive history of American institutionalism, and surely the benchmark for any research to come.

Like previous book-length accounts of American institutionalism (Yonay 1998; Hodgson 2004), part one ("Introduction") of Rutherford's book demonstrates that institutionalism was anything but the somewhat

[1] A *précis* can be found in the Autumn 2010 issue of this journal (Rutherford 2010).

juvenile, anti-theoretical dissent from neoclassicism that later critics such as George J. Stigler and others depicted it to have been. Contrary to this "standard view", to which Coase and Williamson also subscribe, Rutherford's "revisionist view" (p. 5ff.) stems from the observation that institutionalism in the interwar period cannot be characterized simply as dissent against neoclassicism since neoclassicism's rise to dominance in American academia coincided with the rise of Keynesianism, that is, mainly after World War II.

In the period under consideration, American economics was pluralistic in terms of the variety of ideologies, methods, and policy prescriptions (see also Morgan and Rutherford 1998). Institutionalists published in leading journals, held positions in major research universities, served as presidents of major scientific associations, were active in the creation of research and educational organizations, were recipients of substantial funds from all the important research-sponsoring foundations, and were heavily involved in policy making, particularly during the New Deal. Institutionalism, in other words, was part of the mainstream.

Institutionalism, however, was not a well-defined "school" of economic thought but was rather what Rutherford calls a "movement", that is, a network of people who actively promoted the development of an "institutional approach" to economics in line with Walton H. Hamilton's foundational institutionalist manifesto, presented at a special session of the American Economic Association meetings in 1918 (Hamilton 1919). Rutherford's argument is that the group of explicitly self-identified members of the institutionalist movement, led by Hamilton, Mitchell, and John M. Clark, were held together by a common conception of "science" and the desire to use science as an instrument of "social control" (chapter 2).

In a nutshell, the institutionalists were critical of the abstract theorizing associated with neoclassical economics, and rejected the universal depiction of man as a "lightning calculator of pleasures and pains", to use Veblen's celebrated expression. They believed that economic theory had to be based on assumptions that conform to real-world conditions, and took this to require not only a strong empirical, investigative and problem-centered approach but also a strong degree of consistency with scientific knowledge in contiguous fields, such as psychology, sociology, and law (p. 8). Without these features, they argued, economic theory would never be scientific and relevant to

the pressing social problems of the day (e.g., labor unrest, business cycles, unemployment, poverty, externalities, monopoly, and so forth).

Hence, while Veblen's concepts and concerns, such as his stress on pecuniary institutions, were retained by institutionalists, they generally did not adopt Veblen's evolutionary approach, viewing it as ill-suited for deliberative social guidance and legislative reform (pp. 38-39). In consonance with the earlier generation of American progressive economists—particularly Richard T. Ely and Henry C. Adams—the institutionalists were pragmatists of Deweyan persuasion, emphasizing reform through legislative and legal change, and many of them, including Hamilton, Clark, Commons, and Robert Lee Hale, had close connections with the legal realist movement that similarly rejected abstract jurisprudence and advocated an empirically-grounded legal science (see also White 1976 [1949]; and Schlegel 1995).

Rutherford's argument is skillfully substantiated in the most detailed analysis to date of the writings, projects, and careers of all the self-proclaimed institutionalists. In part two ("Institutionalist careers") Rutherford concentrates on Hamilton (chapter 3), the chief promoter of the institutional approach, and illustrates Hamilton's influence by offering a rare discussion of Morris A. Copeland, a conspicuous representative of the second generation of institutionalists who worked in both academia and government after studying under Hamilton, Clark, and others (chapter 4).

The narrative goes on to present the network of people, research programs and curricula to be found in the *hauts lieux* of interwar institutionalism in part three ("Centers of institutional economics"). Rutherford supplies helpful tables of selected faculty and graduate students at each of these centers, allowing us to get a real sense of the mobility within the network by following the trails of individual careers. Rutherford begins with the University of Chicago that under Veblen's influence was undeniably pivotal in the formation and early history of the institutionalist movement prior to 1918 (chapter 5). Rutherford then guides us through Hamilton's two main educational experiments, Amherst College in Massachusetts and the Brookings Graduate School in Washington D.C., which produced many members of the second generation, including Copeland (chapter 6).

Our attention is next directed to the University of Wisconsin where, following the path broken by Ely, Commons and his numerous students developed a distinct variety of institutionalism (chapter 7).

From Wisconsin we turn to the other major institutionalist stronghold, Columbia University in New York, then one of America's top PhD-granting universities, where the institutionalist cohort included Mitchell, Clark, John Dewey, Adolf A. Berle, and Gardiner C. Means (chapter 8). Finally, we learn of the vital links between institutionalists and key executives of the Carnegie and Rockefeller Foundations who, based on a remarkable convergence of beliefs about science and social control, backed the creation in 1920 of the National Bureau of Economic Research (NBER) in New York, where Mitchell and many others engaged in ground-breaking statistical research on business cycles (chapter 9).

The oft-debated factors contributing to the decline of American institutionalism in the post-1945 period are discussed in part four ("Challenges and changes"). Although institutionalism cannot be properly characterized as merely dissent against neoclassicism in the interwar period, after World War II it certainly did become a dissenting heterodoxy lying outside of the mainstream of American economics that became dominated by both neoclassical and Keynesian economics. Rutherford argues that the relationship between institutionalism and what became Keynesian economics is more complex than is usually acknowledged (chapter 10),[2] much more complex, in any case, than the view that institutionalists were left helpless by the Great Depression and fell victim to its Keynesian remedy (e.g., Ross 1991, 419).

As Rutherford explains, the institutionalists' interest in under-consumption goes back to Veblen. In the 1930s many of them, particularly Clark at Columbia, Mitchell at the NBER, and Rexford G. Tugwell (a prominent member of President Franklin D. Roosevelt's original "Brains Trust") in government, were explicitly proposing counter-cyclical public expenditures to deal with unemployment (p. 292ff.). Indeed, Mitchell and others viewed their empirical work on cycles of various sorts as informing decisions of this kind. A few institutionalists converted partly or wholly to Keynesian economics, but many others, in particular Mitchell, Clark, and Copeland, remained highly critical of Keynesian macroeconomics and the associated new econometric modeling, repeatedly underlining the factors omitted in the models and the "Ricardian vice" involved. This apprehension, however, was not shared by the new post-1945 generations of economists. Keynesianism offered them the possibility of social control based on the latest and most scientific approach to economics.

[2] This chapter is based on an article co-authored by Tyler DesRoches.

By the end of what G. L. S. Shackle (1967) called the "years of high theory", the neoclassical economist's toolbox had significantly expanded from its Marshallian foundations, posing a serious challenge to institutionalism (chapter 11). When combined with Arthur C. Pigou's analysis of externalities and social costs, the pioneering work on imperfect competition by Edward H. Chamberlin and Joan Robinson allowed neoclassical economists to address questions that were previously in the institutionalists' domain. In the meantime, key developments in utility theory, associated with John R. Hicks and Paul A. Samuelson, began to slowly but irreversibly separate economic theory from psychology (pp. 316-317). By the time of Samuelson's (1947) *Foundations of economic analysis*, the institutionalists' concern with realistic assumptions was increasingly ignored by neoclassical economists, and within years Milton Friedman's positivist defense of an "as if" methodology of model-building seemed to have settled the matter for the profession.

Rutherford's explanation of the relatively rapid decline of American institutionalism during what Mark Blaug (2003) described as the "formalist revolution" hinges on the same elements that allowed him to define institutionalism as a movement and to document its internal dynamics between the wars. Important changes in the American academic environment transformed the standards of scientific economics, weakening the movement's claim to science. The institutionalists' general reluctance to embrace the new econometric modeling techniques led to the lasting, albeit misleading, reputation of institutionalism as "measurement without theory" (Koopmans 1947). As mainstream economics became increasingly separated from psychology, sociology, and law, institutionalism's interdisciplinary orientation lost its appeal. In this context, the institutionalist network could no longer retain its position at the pinnacles of American education and research. America's top universities stopped hiring institutionalists, and the movement grew smaller and more scattered, both geographically and in terms of social cohesion, until it became marginalized within the profession (pp. 340-341).

There is much to commend in this book. It performs a great service to the community of historians of American economics, who will relish the details of the flow of people and ideas within the institutionalists' network during the interwar period, assembled here thanks to many years of painstaking efforts in the archives. Readers with an interest in

institutional economics will appreciate that caricatures of American institutionalism as anti-theoretical dissent fail to come to terms with the historical facts, and learn from Rutherford's disagreements with Hodgson (2004) regarding the relative importance of Veblen's evolutionary economics for interwar institutionalism (p. 349), and their contrasting interpretations of Frank H. Knight's position (p. 146). Finally, readers with a methodology background will welcome Rutherford's discussions of the institutionalists' views of science and scientific methods, and his illustration that in economics, as in other disciplines, changes in what is widely perceived as normal science may help paradigms rise but also ensure their fall.

Rutherford's book also raises a number of questions that future research will need to address. An important issue that fits well with the definition of institutionalism as a movement concerned with science and social control is the matter of the institutionalists' ontological commitments. I suspect that the institutionalists' social ontology can be partially revealed by an examination of their engagement with the corporate personality controversy that dominated the legal literature between roughly 1900 and 1930 (Gindis 2009). Rutherford mentions the controversy only in passing, mainly in connection with Dewey's dismissive remarks (p. 244), and with Hamilton's later disapproval of legal fictions as barriers to the domestication of the "corporate ghost" (p. 92). Given that other members of the institutionalist network, particularly Commons and the British political scientist Harold Laski, were concerned with the nature of corporate personality, and by extension with the nature of corporations and human associations more generally, the connection deserves further investigation.

REFERENCES

Blaug, Mark. 2003. The formalist revolution of the 1950s. In *A companion to the history of economic thought*, eds. Warren J. Samuels, Jeff E. Biddle, and John B. Davis. Oxford: Blackwell, 395-410.

Gindis, David. 2009. From fictions and aggregates to real entities in the theory of the firm. *Journal of Institutional Economics*, 5 (1): 25-46.

Hamilton, Walton H. 1919. The institutional approach to economic theory. *American Economic Review*, 9 (1): 309-318.

Hodgson, Geoffrey M. 2004. *On the evolution of institutional economics: agency, structure and Darwinism in American institutionalism.* London: Routledge.

Koopmans, Tjalling C. 1947. Measurement without theory. *Review of Economic Statistics*, 29 (3): 161-172.

Morgan, Mary S., and Malcolm Rutherford. 1998. American economics: the character of the transformation. *History of Political Economy*, 30 (1): 1-26.

Ross, Dorothy. 1991. *The origins of American social science*. Cambridge: Cambridge University Press.

Rutherford, Malcolm. 1994. *Institutions in economics: the old and the new institutionalism*. Cambridge: Cambridge University Press.

Rutherford, Malcolm. 1997. American institutionalism and the history of economics. *Journal of the History of Economic Thought*, 19 (2): 178-195.

Rutherford, Malcolm. 2010. Science and social control: the institutionalist movement in American economics, 1918-1947. *Erasmus Journal for Philosophy and Economics*, 3 (2): 47-71.

Samuelson, Paul A. 1947. *Foundations of economic analysis*. Cambridge (MA): Harvard University Press.

Schlegel, John H. 1995. *American legal realism and empirical social science*. Chapel Hill (NC): University of North Carolina Press.

Shackle, George L. S. 1967. *The years of high theory: invention and tradition in economic thought, 1926-1939*. Cambridge: Cambridge University Press.

White, Morton G. 1976 [1949]. *Social thought in America: the revolt against formalism*. New York: Oxford University Press.

Yonay, Yuval. 1998. *The struggle over the soul of economics: institutionalist and neoclassical economists in America between the wars*. Princeton (NJ): Princeton University Press.

David Gindis is lecturer of economics at the University of Hertfordshire Business School (UK), where he is a member of the Group for Research in Organizational Evolution (GROE). His research interests include institutional economics, evolutionary economics, law and economics, economic methodology, and social ontology. His specific focus is on the institutional aspects of the emergence, persistence, and evolution of firms in market economies. In 2008 he was awarded the Herbert Simon Young Scholar Prize by the European Association for Evolutionary Political Economy (EAEPE).
Contact e-mail: <d.gindis@herts.ac.uk>

Erasmus Journal for Philosophy and Economics,
Volume 5, Issue 1,
Spring 2012, pp. 100-105.
http://ejpe.org/pdf/5-1-br-3.pdf

Review of *Handbook on contemporary Austrian economics*, edited by Peter J. Boettke. Cheltenham (UK): Edward Elgar, 2010, 174pp.

DAVID HOWDEN
St. Louis University, Madrid Campus

Peter J. Boettke brings together ten third-generation "modern" Austrian economists to outline ten core propositions of the Austrian school of economics.[1] The edited collection is a further development of Boettke's 2008 entry to the *Concise encyclopedia of economics* entitled "Austrian school of economics". While the core tenets of Austrian economics are certainly open to debate, Boettke meets the challenge and the propositions provide the reader with his particular view of the discipline.

The Austrian school of economics exists largely as a heterodox school outside of the mainstream of economic thought. It has its share of controversies, both internally and with its more orthodox peers. The most visible defining characteristic of the school—and what is likely the aspect most interesting to the majority of *EJPE* readers—is its methodological base. Shunning positivist and empiricist approaches, the Austrian school generally applies deductive reasoning (mostly of the verbal form) to construct theory. From this basis spring further distinctions of varying degrees between Austrian economists and the mainstream, which provide the primary fodder for this book.

The book is structured similarly to other recent overviews of Austrian economics, and complement texts such as Boettke (1994) and Huerta de Soto (2010). It sets itself apart by being a little more systematic, focused and, in the end, geared at a different audience. It is systematic in the sense that the reader does not need to work too hard to understand exactly what the authors are trying to stress—the

[1] A complete listing of the ten propositions is as follows: 1) only individuals chose, 2) the study of the market order is about exchange behavior, 3) the "facts" of the social sciences are what people believe, 4) utility and costs are subjective, 5) prices economize on information people need for decision making, 6) private property in the means of production is a necessary condition for rational economic calculation, 7) the competitive market is a process of entrepreneurial discovery, 8) money is non-neutral, 9) the capital structure consists of multispecific and heterogeneous goods, and 10) social institutions are the result of human action, but not of human design.

ten propositions clearly form the basis of the body of knowledge being imparted. The use of ten propositions reminds this reviewer of N. Gregory Mankiw's "10 principles of economics" that open the pages of his widely read introductory textbook to neoclassical economics. The uninitiated to Austrian economics will be at ease in these pages, as the overview sticks to the basic points. The contributing authors also provide enough depth to hold the attention of intermediate-level Austrian economists.

While methodological differences differentiate the Austrian from the mainstream economist, there is also considerable debate within the school. What is the proper use of empirical facts? Should theory be constrained to strictly a priori statements, or are even these tainted by empirical influences? For the most part this book wisely sidesteps such issues, and instead focuses on the central features of the methodological divide with mainstream economics.

Without getting embroiled in specific debates as to what defines the Austrian school, Boettke aims to delineate what sets it apart from its more mainstream peers. Most of the ten entries are noncontroversial and unique to the Austrian school (or at the least, are outside of the mainstream of economic reasoning). In particular this reviewer is thinking of proposition 4: utility and costs are subjective, and proposition 8: money is non-neutral. Other propositions, however, could find themselves at home in the mainstream economist's toolbox. Proposition 1: only individuals choose, as a case in point, is unobjectionable to many microeconomists (perhaps behaviorists aside), regardless of their preferred school. Likewise, proposition 5: the price system economizes on information people need to make decisions, is well incorporated in the corpus of mainstream economic theory (indeed, it could be the contribution of Friedrich Hayek best incorporated by the mainstream, though there is more to say on that later).

In what follows, I wish to overview some of the more stimulating contributions, before turning to the real strengths (and weakness) of the book.

Anthony J. Evans takes the contemporary understanding of the term "methodological individualism" to task as a vulgar version of its original use. Modern usage has the economist reducing all complex phenomena to their most elemental (individual) components for further analysis. Evans advocates instead an "institutional individualism" approach, whereby one can still reduce actions to an atomistic level yet remain

cognizant of more institutional (and less individual) influences. While reading in many ways like a redefinition of existing terminology, the chapter is more of a rallying call to return to the original use of methodological individualism. While reducing social phenomenon to their primal elements clarifies many ambiguities, there is no denying that the "group" influences many actions. Evans's middle ground approach allows for just the right amount of attention to the individual while paying heed to institutional influences.

In two complementary chapters, Virgil Henry Storr and Edward P. Stringham discuss the problem of subjectivity in economic science. Storr takes a more theoretical approach, explaining why any interpretation of events must be subjective. He ably sidesteps controversies over how "subjective" different analyses are, while succinctly making the case for a subjective core for economics. Stringham addresses the question of how "subjective" one is by compiling a "Subjectivism purity test" for the reader to establish to what degree they believe the facts of social sciences are subjective. Interesting results abound.

One of interest to this reviewer is Stringham's conclusion that "[f]rom this perspective the famous mathematical economist Kenneth Arrow was more of an economic subjectivist than Mises!" (p. 54). Stringham refers to Mises's explicit preference for a certain state of the world, versus Arrow's acceptance that an optimal state is indefinable. This reviewer thinks that Stringham overstates his case. One can prefer, after all, a certain outcome (and advocate for it) and still remain subjectivist provided one does not attempt to impose this state onto others.

Stephen C. Miller contrasts two views of prices and their role in the market. The neoclassical economist views a price as a summary of an existing equilibrium situation. The Austrian views a price as an active protagonist in the entrepreneurial discovery process, spreading information effortlessly to various agents. Indeed the difference is subtle, and Miller serves the reader well to shed light on the different approaches. While the neoclassical focus emphasizes the process that prices undergo to get to a new equilibrium, the Austrian counterpart emphasizes the process by which demanders of a good respond to price changes. While the mainstream has superficially incorporated Hayek's stress on the role of prices in the market economy, Miller guides the reader through the intricate role that prices actually play.

Frederic Sautet outlines the intimate relationship between competition and entrepreneurship. In one interesting schema, Sautet lists and differentiates between three definitions of a "market": the old classical view of a market as a place; the Austrian view of a market as a process; and the now more mainstream conceptualization of a market as a metaphor. This reviewer was struck by how useful this trichotomy was in identifying where some aspects of economic theory have gotten off-track. The shift of the most generic institution of economic theory—the "market"—from a place (a spatial relation) to a process (a temporal relation) to a metaphor (an abstract relation) has resulted in a purging of some essential facets from economic theory. As Sautet outlines, chief among these facets are the entrepreneurial function and its competitive nature.

In many respects Austrian economists differ from their more mainstream peers the most in the area of monetary economics. J. Robert Subrick sheds light on the reasons for non-neutral money and its implications. Bringing attention to the six main rationales for money's non-neutrality, Subrick clearly shows where Austrian economists have established themselves, and where more work needs to be done. In comparison to their mainstream colleagues, Austrians have paid plenty of attention to the negative consequences of Cantillon effects and forced saving; certain strands of Austrian economics have focused on the money illusion and sticky prices; while Mundell-Tobin effects remain relatively underexplored. Subrick could have paid closer attention to the distinction between commodity and fiat money, and its implications for price stickiness, but on the whole the chapter gives the reader a refreshing primer on the current state of understanding of the concept of non-neutral money among Austrian economists.

In the book's final chapter, Boettke closes by providing a glimpse of what will feature in Austrian economics in the future. This chapter forms, in many ways, an eleventh proposition to augment the previous ten. By outlining two central methodological tenets of Austrian theorizing—praxeology and apriorism—Boettke overviews the epistemological issues that the Austrians of yesteryear have defended against, in order to direct the reader to a possible future. Methodological concerns are certainly central to most Austrian-type analyses, but this reviewer has mixed feelings about their isolation to the final pages of the book. On the one hand, these concerns receive so much attention elsewhere that it is not completely necessary to

rehash them afresh. Then again, since they are undeniably distinct and defining features of this school of economics, a more thorough inclusion is almost certainly warranted. At the end of the day, this is one topic that will likely please some readers while irking others.

Perhaps the largest misgiving I have of the book is the attempt to pigeonhole these ten Austrian propositions into categories more at home in its mainstream counterparts. By splitting the propositions into three separate parts of the book—the science of economics, microeconomics, and macroeconomics—Boettke is forced to almost redefine concepts to accommodate non-relevant categories. As an example, proposition 10 (concerned with spontaneously formed social orders) does not seem at home in a macroeconomics oriented section. At the same time, it does not seem overly at home in any of the other sections. It could be that Boettke's ten propositions do not easily fit into the more mainstream micro/macro divide. Trying to fit them into that mold makes the flow of the book a little clumsy, but does not meaningfully affect the knowledge being imparted.

In sum, the book is a concise overview of some of the distinctive features of Austrian economics. One could debate whether other facets should have been included (the aforementioned methodological concerns spring to mind). Likewise some will take issue that some topics are not sufficiently "Austrian", or uniquely so, to warrant inclusion. Whatever one's take on that, forming a central core of propositions that a group can identify with is a necessary condition for that group moving forward as a cohesive whole. In that regard, Boettke has moved his group in the right direction.

REFERENCES

Boettke, Peter J. (ed.). 1994. *The Elgar companion to Austrian economics*. Aldershot (UK): Edward Elgar.

Boettke, Peter J. 2008. Austrian school of economics. In *Concise Encyclopedia of Economics*, ed. David R. Henderson. Library of Economics and Liberty. Liberty Fund, Inc. http://www.econlib.org/library/Enc/AustrianSchoolofEconomics.html (accessed May 26, 2012).

Huerta de Soto, Jesús. 2010. *The Austrian school: market order and entrepreneurial creativity*. Cheltenham (UK): Edward Elgar.

David Howden is assistant professor of economics at St. Louis University at its campus in Madrid, Spain. His research interests include the Austrian theory of the business cycle and its implications for

banking policy. His most recent book is *Institutions in crisis: European perspectives on the recession* (editor, 2011, Edward Elgar).
Contact e-mail: <dhowden@slu.edu>

Erasmus Journal for Philosophy and Economics,
Volume 5, Issue 1,
Spring 2012, pp. 106-111.
http://ejpe.org/pdf/5-1-br-4.pdf

Review of Iara Vigo de Lima's *Foucault's archaeology of political economy*. New York (NY): Palgrave Macmillan, 2010, 274pp.

RYAN WALTER
Australian National University

Vigo de Lima has set herself the difficult task of selling Michel Foucault to historians and methodologists of economic thought. Foucault's name is closely associated with the 'theory boom' of the 1960s and 1970s, which is enough to arouse the suspicion of most economists. A second marketing challenge is to overcome the effrontery embodied in Foucault's decision to write histories of extremely varied terrains, including medicine, grammar, classical sexual ethics, and political thought. Worse, all of Foucault's histories offer powerfully revisionist narratives, but seemingly from the perspective of an outsider sticking his nose into another discipline's past. A final issue for the Foucault salesperson is that, except for the final two volumes of *The history of sexuality*, the narratives of all of his book length histories are driven by philosophical schemas that can both alienate readers and compromise the empirical credentials of his conclusions. What is Vigo de Lima's sales strategy? It is to pitch to a small minority who are prepared to buy high-grade Foucault, and this is both a strength and a weakness of the book.

To understand why the book's strategy should be seen in these terms, the first point to note is which part of Foucault's work Vigo de Lima has studied. As her title indicates, it is the archaeological phase, or the 'early Foucault'. Three monographs originally published in the sixties are typically assigned to this stage (*History of madness*, *birth of the clinic*, and *The order of things*), along with an unwieldy methodological treatise (*The archaeology of knowledge*). This instruction manual effaced most of the methodological diversity of the earlier studies to present a codified method complete with specialised terminology, neither of which Foucault would ever use, since he promptly shifted method and commenced his 'genealogies' or 'middle period'. Yet this method text—obsolete by the 1970s—is worked through patiently by Vigo de Lima, taking the reader through the ostensibly oxymoronic labels, such as 'historical *a priori*' and 'system

of dispersion'. The other text central to this study is *The order of things*, where Foucault's principal engagement with the history of economic thought is found.

From the marketing perspective, the trouble here is that *The order of things* is Foucault's most inaccessible text and the most historically unsound. Foucault asserted the existence of an *episteme*, a matrix existing beneath concrete forms of knowledge that structures the form these surface knowledges can take. On this basis, shifts in the underlying *episteme* will be reflected in shifts in its surface manifestations, and Foucault described two such shifts—from the Renaissance to the classical, and from the classical to the modern. This is sweeping enough, but Foucault's case studies were three distinct disciplines: grammar, biology, and economics. The resulting scale of *The order of things* is what makes it so breathtaking to read and yet so empirically flawed, as specialists from the relevant fields have noted. These concerns are acknowledged occasionally but they are not seriously confronted (see pp. 36, 57, 150, 209, 211, 214-215, 225, 243-244). Vigo de Lima's primary tactic is instead to repeat what Foucault said in *The order of things* and then suggest that his account is broadly supported by mainstream historians of economic thought, particularly Mark Blaug and Joseph Schumpeter (for example, pp. 195-196). This approach may not persuade the majority.

An alternative path for linking Foucault with the history of economic thought is to centre his studies of 'governmentalities'—the rationalities that inform particular practices of government. One might think here of the way a discourse of sovereignty is invoked to justify acts of war, while tax cuts tend to be discussed using economic language; war and taxation are particular practices of government and they correspond with distinct bodies of knowledge. There is now a sizeable literature examining this aspect of Foucault's thought, which derives primarily from his lectures at the Collège de France, circa 1977-1979 (Foucault 2007; 2008). The Anglophone reception of this research program was greatly spurred by the publication of *The Foucault effect* (Burchell, et al. 1991), and aided by a rich vein of work centred on the journal *Economy and Society*. Much of this literature is synthesised by two invaluable survey works—Mitchell Dean's *Governmentality* (1999), and Nikolas Rose's *Powers of freedom* (1999). This relevant literature is missing from Vigo de Lima's book, and this is a shame, for the distinguishing mark of

the governmentality material is that it builds on Foucault's insights and often subjects them to severe revision.

This last point leads to the second reason to question Vigo de Lima's marketing vision. The balance between, on the one hand, recapitulating what Foucault had to say about method and the history of economics and, on the other hand, picking up some of his pointers to develop a new historical narrative is overwhelmingly weighted in favour of recapitulation. Of the five substantive chapters, only chapters 5 and 6 relate to the history of economic thought, while the other three develop a reading of Foucault's archaeology. Chapter 5 takes the reader through Foucault's account of how mercantilism and physiocracy belonged to the classical *episteme*, analysing the representation of wealth in relation to the circulation of bullion and goods in the body politic. We are told that Foucault portrays Smith as partially escaping this epistemological epoch by examining the role of labour in creating wealth, but that ultimately Smith remains tethered to classical thought by his treatment of exchange. To understand this part of Foucault's argument it is necessary to know that he treats Ricardo and Marx as ushering in political economy, in which the finitude of man as a being that must labour for his survival is the central analytical figure, replacing the circulation of wealth (the key theme for earlier classical thought). Smith is therefore a transition figure in this aspect of Foucault's work, having one foot in each camp.

Yet in his later work on governmentality Foucault makes Smith emblematic of a shift to classical liberalism and the rejection of sovereign-centred conceptions of government, such as mercantilism. The point to note here is that Foucault's reading of Smith is largely determined by his unit of analysis—*epistemes* in *The order of things* and rationalities of government in the later work. This internal variation is one reason why a lighter reliance on the Foucault of the archaeology might have served Vigo de Lima better.

Nevertheless, the great merit of *The order of things* is that the reader is blocked from reading seventeenth-century writers as if they were early economists. This enduring anachronism in the history of economic thought is precluded by Foucault's assertion of a fundamental discontinuity, such that political economy is only possible once we enter the modern *episteme* around the turn of the nineteenth century. Foucault has certainly over-egged his pudding, and perpetuated the idea that mercantilism is a coherent unity, but the lesson regarding when we

can start to speak sensibly of political economy is a valuable one, most powerfully made by Keith Tribe (1978), who drew a little inspiration from *The order of things*.

Returning to the marketing perspective, the difficulty arising from an emphasis on exegesis over new history is obvious. The economist is asked to subscribe to the existence of *epistemes* as underlying epistemological unities that govern what can be thought in particular fields, for example: "[m]ercantilism was determined by this context of the classical episteme" (p. 210). This is a major metaphysical commitment for anyone, but it is even more daunting in view of the fact that Foucault backed away from it himself soon after publication, although Vigo de Lima finds the *episteme* enduring in Foucault's work (p. 237). The economist is therefore asked not to read the authors of the canon in the more familiar ways—for theoretical concepts, in view of the intellectual or social context, through biography, and so on—as this would be to stay at the surface level of knowledge and not penetrate to the archaeological depths. The economist who would embrace all this is Vigo de Lima's boutique market.

An alternative pitch would have involved exploring some of Foucault's insights in relation to contemporary economics, thereby demonstrating their certain capacity to stimulate new research. As an example, Vigo de Lima helpfully draws out the idea that analyses of money and language were highly analogical in the classical period (roughly, 1650-1750). In particular, in classical thought language/money was treated as representing objects/wealth transparently, while in modern thought language is no longer seen to name things without difficulty, but is instead recognized as having a history and autonomy of its own. Foucault's tantalizing claim is that since language is still the medium for scientific knowledge this non-transparency is a problem, and one that has been addressed in two different ways. One was formalisation, as in the positivist's effort to forge a value free language, or Bertrand Russell's attempts to create a symbolic logic. The other was the turn to interpretation, the terrain occupied by Freud, for example. Note that Foucault has little to say about the language/money analogy in the modern period, but a quick look at macroeconomic theory might join up the dots. For in general, money is treated as either a more or less accurate reflection of a *real* economy of goods and services (Friedman); or, in contrast, as a thing in its own right, that should not be reduced to a mere reflection of some other process (Keynes, Marx). The majority of

the mainstream analysis of inflation specifies money in this first sense, and hence much of economics is concerned with the problem of 'money illusion', while general equilibrium models introduce money into barter models with minimal changes in results. If, however, we look across at the new economic sociology, then we find that money is to be interpreted for its social meaning, using the suite of sociology's techniques (see, for example, Smelser and Swedberg 2005). Thus, one might claim that the analysis of money again reproduces the intellectual tactics brought to the study of language.

Broad schemas of this type are what Foucault furnishes in *The order of things*, but they need to be put to work, revised, and developed. Vigo de Lima hands this task on to others, hence she writes that Foucault "offered a novel perspective, together with a range of inspiring notions [...] which cannot be ignored by anyone intending to study the methodology and historiography of economics" (p. 243). This book is useful as a guide to Foucault's archaeology of political economy, which is what the author set out to provide. But my sense is that this conceals the simplicity of Foucault's overall method: first, suspend commitment to the self-evidence of economic thought and practice in order to narrate the emergence of economics in historical terms; second, do this by reading canonical texts for what they talk about and how, with an eye for shifts in the objects of discourse and the concepts that make these objects intelligible. The key idea is to substitute looking for the origins or immanence of economic thought and instead open up the possibility that there may have been something entirely different before political economy that we nevertheless persist in seeing as early economics. We may not need the theoretical machinery of archaeology to pursue this task, but only a humbler form of historiography.

REFERENCES

Burchell, Graham, Colin Gordon, and Peter Miller (eds.). 1991. *The Foucault effect: studies in governmentality*. London: Harvester Wheatsheaf.

Dean, Mitchell. 1999. *Governmentality: power and rule in modern society*. London: Sage.

Foucault, Michel. 2007. *Security, territory, population: lectures at the Collège de France, 1977-1978*, ed. Michel Senellart, trans. Graham Burchell. Basingstoke: Palgrave Macmillan.

Foucault, Michel. 2008. *The birth of biopolitics: lectures at the Collège de France, 1978-1979*, ed. Michel Senellart, trans. Graham Burchell. Basingstoke: Palgrave Macmillan.

Rose, Nikolas. 1999. *Powers of freedom: reframing political thought*. Cambridge (UK): Cambridge University Press.

Smelser, Neil, and Swedberg, Richard (eds.). 2005. *The handbook of economic sociology*. Princeton (NJ): Princeton University Press.
Tribe, Keith. 1978. *Land, labour, and economic discourse*. London: Routledge and Kegan Paul.

Ryan Walter teaches history of political and economic thought at the Australian National University. His monograph *A critical history of the economy* was published by Routledge in 2011.
Contact e-mail: <Ryan.Walter@anu.edu.au>

Erasmus Journal for Philosophy and Economics,
Volume 5, Issue 1,
Spring 2012, pp. 112-120.
http://ejpe.org/pdf/5-1-br-5.pdf

Review of Mark D. White's *Kantian ethics and economics: autonomy, dignity, and character*. Stanford University Press, 2011, 288pp.

NICOLAS GRAVEL
Aix-Marseille University

This interesting book presents itself as an introduction to economists of the moral philosophy of Immanuel Kant. I believe however that this self-presentation is not very reflective of the actual content of the book. Indeed, the book spends most of its 190 pages (excluding notes and references) in criticizing the ethical foundations of normative economics from a Kantian perspective. The presentation of Kant's moral philosophy *sensu stricto*, while present, serves mainly as a reference point from which the author develops his critique of the ethical foundations of economics. Hence, this book is more a Kantian critique of the ethical foundations of modern mainstream economics than an introduction to Kantian moral philosophy.

The ethical foundations of economics that are the object of the "Kantian" critique of the book are, essentially, those generally known under the heading of "welfarism"—using the term coined by Amartya Sen (1979)—and which White refers to as "utilitarianism". Welfarism is a rather coherent approach to social ethics. It asserts that the distribution of individual welfare levels is the *only* information that is relevant to evaluating the "ethical goodness" of social arrangements. Individual attributes such as rights, freedom, discrimination, and the like, are only important for welfarism insofar as they affect individuals' welfare. Similarly, welfarism attaches no intrinsic importance to the procedure by which a particular distribution of individual welfare levels is brought about. Welfarism is thus a *consequentialist* approach to social ethics, and this feature is at the very heart of the critique developed in the book:

> In this chapter—and, more broadly, this entire book—I propose to construct an economic model of decision-making based on nonconsequentialist ethics, specifically the moral theory of Immanuel Kant, in which the nature of actions themselves, rather than their consequences, determines their moral worth (p. 16).

Given the strong anti-welfarist stance taken by the author, it is somewhat surprising that the book makes no reference to the vast non-welfarist literature that has developed in the last forty years or so in economics and philosophy. I am referring here especially to the literature on equality of opportunity and responsibility (see, e.g., Rawls 1971; Arneson 1989; Arneson 1990; Dworkin 1981a; Dworkin 1981b; Roemer 1998; and the survey Fleurbaey 2008); on freedom of choice (see, e.g., Sen 1988; Sen 1990; Sen 1991; Pattanaik and Xu 1990; Arrow 1995; Sugden 1998; and the survey Barberà, et al. 2004); or on the modelling of individual rights (see, e.g., Sen 1970; Gibbard 1974; Gärdenfors 1981; Deb, et al. 1997; or Peleg 1998). The absence of reference to this literature is all the more surprising as several contributors to it—including Rawls himself (1971) but also Kolm (1997) and Van Hees (2003)—have made explicit reference to Kantian moral philosophy in developing their own non-welfarist approach to social ethics.

Be that as it may, the reference point from which the author develops his critique of the ethical foundations of mainstream economics is a quite sophisticated (if not abstract) view of Kant's conception of morality. This view, presented in the first chapter of the book, follows Kant's writings closely. It is founded on the notion of the autonomy of the human person, described as her "capacity [...] to make choices according to laws that she sets for herself, without undue influence from either external pressures or internal desires" (p. 19). According to White, autonomy—which distinguishes humans from animals or mere "biological machines" (p. 84)—imbues persons with 'dignity': an "unconditional and incomparable worth" that as such "demands respect from all persons" (p. 21). The autonomous agent imposes on herself a moral law or a maxim. In order to be moral, such a maxim must satisfy the "universalization" test provided by Kant's first formulation of the categorical imperative: "Act only according to that maxim whereby you can at the same time will that it should become a universal law". The maxim must also satisfy the second formulation of Kant's categorical imperative—which White seems to prefer: "act in such a way that you treat humanity, whether in your own person or in the person of another, always at the same time as an end and never simply as a means" (p. 26). Duties are maxims adopted by an autonomous agent which satisfy the two versions of the categorical imperative. A perfect

duty "permits no exception in the interest of inclination" while an imperfect duty is one which "the agent has some latitude in executing, both in degree and method" (p. 29). For instance the well-known—in Kantian writings—maxim "do not lie" is a perfect duty, while the maxim "do not be indifferent to others" is an imperfect one. White recognizes the possibility that various duties (perfect or imperfect) may contradict each other. For instance, the duty "do not be indifferent to your friends" may conflict with the duty "do not lie" for a person whose lie (say to a possible murderer of her friend) may save the life of this friend. White calls the ability of a person to solve the "problem of conflicting obligations or rules" 'judgement' (p. 34), but does not provide much insight as to how the exercise of this judgement may work in practical situations.

I mentioned that I find this view of individual morality rather abstract. Are there really such moral rules "which permit no exception in the interest of inclination"? I cannot think of any. Even the most basic—and widely discussed—Kantian duty "do not lie" seems to suffer several exceptions. After all, who can seriously maintain that it was the "duty" of partisans arrested by the Nazis in World War II to reveal truthfully to their enemies everything they knew about their comrades? If duties permit exceptions, perhaps in order to make them consistent with other duties, this means that moral maxims cannot be unconditional. The important question therefore becomes identifying the circumstances that could allow a moral individual to deviate from his/her maxim. I cannot see how this question can be answered once and for all, independently from the circumstances and the "inclinations" of the agent who is supposed to adopt the moral behaviour. For instance, in the specific context of prisoners' dilemma games, discussed in chapter 1, several authors, including myself (see, for instance, Bilodeau and Gravel 2004), have suggested that an interpretation of the first formulation of the categorical imperative would impose cooperative behaviour on moral agents. Indeed, any agent placed in an interactive context in which all agents have preferences like those that give rise to prisoners' dilemma situations would prefer that the maxim "cooperate", rather than the maxim "defect", becomes a universal law. But clearly this particular use of the universalization test depends upon the inclinations of the participants of the interaction, as would, I believe, most ethical rules of conduct.

The second chapter of the book discusses the difficulty that Kantian persons may encounter in practical situations when the maxim resulting from the exercise of their judgement, in the sense just given to this word, conflicts with their spontaneous inclination or preference. The author insists that the conflict between the aggregate duty resulting from the exercise of judgement and that person's inclination is of a different nature than the conflict between ethical duties solved by judgement. The latter conflict is managed by the autonomous person out of rational deliberation. The former conflict is solved every day by any person who must confront his or her spontaneous inclination with the general 'duty' that she imposes on herself.

White discusses at length the various forms of "weakness of will" that a person may experience in her day to day life, and sketches a "probabilistic" model of weakness of will. In this model, an individual is depicted as having two preferences: one resulting from duty, and one resulting from inclination. In any choice situation, the individual will choose her inclination-most preferred option with a certain probability p, and her duty-most preferred option with the probability 1-p. In White's view, the probability p can be seen as a measure of the weakness of the will of the individual. The author calls "character" the ability of a person to solve the conflict between the pursuit of spontaneous inclination and the call of duty. The author illustrates the conflict by discussing at length the phenomenon of procrastination.

While the issue discussed in this chapter—that of a conflict between a "superior" goal decided in an *ex ante* situation by the individual and the spontaneous inclination that may, at any single moment, distract the individual from the pursuit of her goal—is interesting and important, it is somewhat surprising that the author discusses it without any reference to the economic literature on the subject. After all, what the author is describing in this chapter is akin to a problem of time inconsistency. A plan of action—say to eat vegetarian food—is judged optimal once decided initially (perhaps out of a thoughtful Kantian judgement) but, when appetite preceding lunch time manifests itself, the initially recommended plan of action no longer seems appealing, especially in view (or in smell) of the hamburgers served in a neighbouring restaurant. Economic theory has provided several models of this situation that could have been, I believe, usefully connected to the discussion of this chapter. For example, Gul and Pesendorfer's

(2001) article on temptation and self-control seems particularly relevant to this issue.

The third chapter discusses some issues concerning the "social embeddedness" of the ideal Kantian person depicted in the two preceding chapters. It argues that the Kantian approach is founded on an atomistic view of the individual whose "character is self-created and maintained by her choices and actions" (p. 88). The chapter, whose main ideas are essentially those developed in the excellent book by Christine M. Korsgaard (2009), also discusses various views of individual identity, including those of John B. Davis (2003). The picture that one gets from reading this chapter is that the ideal Kantian person is indeed minimally socially embedded. The only "duty" that Kantian morality imposes on a socially embedded person is respect for the dignity of other persons. But it is unclear what this "respect" is supposed to mean in practical situations of real (social) life. White argues, repeatedly, that the working of the market—voluntary transactions between free private owners of resources—satisfies this requirement of mutual respect for individual dignity. White even writes that "the market represents a kingdom of ends which is limited but nonetheless complete within its scope—ensuring the maximal freedom *from interference* consistent with the same freedom for all" (p. 117). But in order to be convincing on this, White should have discussed examples where free exchange between owners may not appear at first glance to be spontaneously compatible with mutual respect. A good example of this is prostitution. Is the person who satisfies their sexual desire by using the body of a prostitute really treating the prostitute "as an end and not simply as a means to an end?" It would have been interesting to read the author's thoughts on this.

The last two chapters of the book present the author's rather violent critique of the economic approach to law and, more generally, normative economics. The starting point of the critique is an unashamed economic libertarianism according to which only minimal intervention by the state is compatible with the Kantian ideal of the "kingdom of the ends". I must confess that I do not find the author very convincing on this front.

Consider for instance his discussion of (negative) externalities, based on a distinction between "wrongful" externalities and those which are not "wrongful". I find the distinction rather blurred. The author for instance considers "crime" to be a "wrongful" activity, while the action

of driving a car is not "wrongful". It is hard to understand the basis of this distinction. A city like London, for instance, has imposed significant congestion pricing on private vehicles to protect the inhabitants of the city against the harms of air pollution (among other things). Hence, at peak hours, driving a car can indeed be "wrongful", perhaps even more "wrongful" than some other activities that are considered criminal in the U.S. (e.g., low-level pick pocketing, marijuana dealing, and so on). What is the criterion that would make one consider that selling a joint of marijuana is more "wrongful" than driving a car at peak hours in central London? The author does not provide any clue about what this criterion might be, or how a Kantian ethics can help one in identifying it.

White frequently criticizes the economic approach to crime and regulation for advocating the "management" or "optimization" of the level of criminal activity instead of its "elimination" (p. 127). But again, I do not see how one can avoid "management" of the human activities that are harmful to others, be they considered "criminal" or not. The amount of material and human resources that would be needed to "eliminate" homicides—assuming it were even possible—is just too great to make it a worthwhile objective. There seems therefore to be no alternative than to reduce the number of homicides to a level that is deemed appropriate, given the resources that we wish to devote to this endeavour. As the Nobel laureate James Mirrlees wrote more than twenty year ago: "A good way of governing is to agree upon objectives, discover what is possible, and optimize" (Mirrlees 1986). It is hard to see how one can oppose this basic wisdom.

The author also spends some time discussing the Kaldor-Hicks potential efficiency criteria, as well as the book by Kaplow and Shavell (2002). The discussion of Kaldor-Hicks is rather surprising, as most economists have been convinced that these criteria—proposed in the forties, and since shown to suffer from severe logical inconsistencies—are flawed (see, e.g., Bossert 1996; or Gravel 2001). It is therefore rather misplaced to base one's critique of normative economics on the Kaldor-Hicks criteria, or on the defence of welfarism provided by Kaplow and Shavell (2002).

The last chapter is concerned with challenging the ethical plausibility of the Pareto principle, according to which unanimous consent by the persons concerned is a sufficient condition for recommending a policy or a reform. Here again I found the argument very difficult to follow. The author emphasizes on many occasions that

the working of the market instantiates something close to the Kantian ideal of the kingdom of ends. Yet a distinctive feature of the working of the market is that it respects the Pareto principle: if there is a transaction between two parties, it is because both parties consent to it. In this sense, markets bring about what economists call "Pareto-improvements". Given this, it is very difficult to understand the author's arguments against Pareto improvements, some of which are very surprising. For instance:

> At the same time, persons who are not affected by the policy must be denied any right of consent, for policy decisions based on their input would again violate the dignity of those who are truly affected by the policy and deserve the right of consent (p. 179).

This sentence is strange. Of course people who are not affected by a policy are not affected by it. There is therefore no need to "deny them any right of consent". Sentences like this abound in the chapter, and give to the reader the impression that the author's artillery barrage against the Pareto principle largely stems from a basic misunderstanding.

Overall, the main criticism I would make of this book is of the gap and disconnection that seems to exist between the interesting and reasonably deep philosophical account of Kantian morality provided in the three first chapters and the rather confusing and dogmatic criticism of normative economics that is developed in the two last chapters. As I wrote above, a better starting point for criticizing the consequentialist and welfarist underpinnings of mainstream normative economics would have incorporated the extensive non-welfarist literature.

REFERENCES

Arneson, Richard J. 1989. Equality and equal opportunity for welfare. *Philosophical Studies*, 56 (1): 77-93.

Arneson, Richard J. 1990. Liberalism, distributive subjectivism, and equal opportunity for welfare. *Philosophy & Public Affairs*, 19 (2): 158-194.

Arrow, Kenneth J. 1995. A note on freedom and flexibility. In *Choice, welfare, and development: a festschrift in honour of Amartya K. Sen*, eds. Kaushik Basu, Prasanta K. Pattanaik, and Kōtarō Suzumura. Oxford: Oxford University Press, 7-16.

Barberà, Salvador, Walter Bossert, and Prasanta K. Pattanaik. 2004. Ranking sets of objects. In *Handbook of utility theory, vol. 2: extensions*, eds. Salvador Barberà, Peter J. Hammond, and Christian Seidl. Dordrecht: Kluwer, 893-977.

Bilodeau, Marc, and Nicolas Gravel. 2004. Voluntary provision of a public good and individual morality. *Journal of Public Economics*, 88 (3-4): 645-666.

Bossert, Walter. 1996. The Kaldor compensation test and rational choice. *Journal of Public Economics*, 59 (2): 265-276.

Davis, John B. 2003. *The theory of individual in economics: identity and values*. London: Routledge.

Deb, Rajat, Prasanta K. Pattanaik, and Laura Razzolini. 1997. Game forms, rights and the efficiency of social outcomes. *Journal of Economic Theory*, 72 (1): 74-95.

Dworkin, Ronald. 1981a. What is equality? Part 1: equality of welfare. *Philosophy & Public Affairs*, 10 (3): 185-246.

Dworkin, Ronald. 1981b. What is equality? Part 2: equality of resources. *Philosophy & Public Affairs*, 10 (4): 283-345.

Fleurbaey, Marc. 2008. *Fairness, responsibility, and welfare*. London: Oxford University Press.

Gärdenfors, Peter. 1981. Rights, games, and social choice. *Noûs*, 15 (3): 341-356.

Gibbard, Allan. 1974. A Pareto-consistent liberal claim. *Journal of Economic Theory*, 7 (4): 388-410.

Gravel, Nicolas. 2001. On the difficulty of combining actual and potential criteria for an increase in social welfare. *Economic Theory*, 17 (1): 163-180.

Gul, Faruk, and Wolfgang Pesendorfer. 2001. Temptation and self-control. *Econometrica*, 69 (6): 1403-1435.

Kaplow, Louis, and Steven Shavell. 2002. *Fairness versus welfare*. Cambridge (MA): Harvard University Press.

Kolm, Serge-Christophe. 1997. The values of liberty. In *Freedom in economics: new perspective in normative analysis*, eds. Jean-François Laslier, Marc Fleurbaey, Nicolas Gravel, and Alain Trannoy. London: Routledge, 25-46.

Korsgaard, Christine M. 2009. *Self-constitution: agency, identity and integrity*. Oxford: Oxford University Press.

Mirrlees, James A. 1986. The theory of optimal income taxation. In *Handbook of mathematical economics*, eds. Kenneth J. Arrow, and Michael D. Intriligator. Amsterdam: North Holland, 1197-1249.

Pattanaik, Prasanta K., and Yongsheng Xu. 1990. On ranking opportunity sets in terms of freedom of choice. *Recherches Économiques de Louvain/Louvain Economic Review*, 56 (3-4): 383-390.

Peleg, Bezalel. 1998. Effectivity functions, games and rights. In *Freedom in economics: new perspective in normative analysis*, eds. Jean-François Laslier, Marc Fleurbaey, Nicolas Gravel, and Alain Trannoy. London: Routledge, 116-132.

Rawls, John. 1971. *A theory of justice*. Cambridge (MA): Berknap Press of Harvard University Press.

Roemer, John E. 1998. *Equality of opportunity*. Cambridge (MA): Harvard University Press.

Sen, Amartya K. 1970. The impossibility of a Paretian liberal. *Journal of Political Economy*, 78 (1): 152-157.

Sen, Amartya K. 1979. Utilitarianism and welfarism. *Journal of Philosophy*, 76 (9): 463-489.

Sen, Amartya K. 1988. Freedom of choice: concept and content. *European Economic Review*, 32 (2-3): 269-294.

Sen, Amartya K. 1990. Welfare, freedom, and social choice: a reply. *Recherches Économiques de Louvain/Louvain Economic Review*, 56 (3-4): 451-485.

Sen, Amartya K. 1991. Welfare, preferences, and freedom. *Journal of Econometrics*, 50 (1-2): 15-29.

Sugden, Robert. 1998. The metric of opportunity. *Economics and Philosophy*, 14 (2): 307-337.

Van Hees, Martin. 2003. Acting autonomously versus not acting heteronomously. *Theory and Decision*, 54 (4): 337-355.

Nicolas Gravel is professor of economics at Aix-Marseille University and director of the research unit GREQAM. His research interests include normative and public economics and the intersection of moral theory and economics.
Contact e-mail: <nicolas.gravel@univ-amu.fr>

Erasmus Journal for Philosophy and Economics,
Volume 5, Issue 1,
Spring 2012, pp. 121-127.
http://ejpe.org/pdf/5-1-br-6.pdf

Review of Philippe Steiner's *Durkheim and the birth of economic sociology* (Trans. Keith Tribe). Princeton (NJ): Princeton University Press, 2011, 249pp.

JARED L. PEIFER
Rice University

Émile Durkheim believed that even the seemingly individual act of "thinking" is a social activity (Calhoun, et al. 2007, 142). Philippe Steiner extends this Durkheimian emphasis on societal level explanations to the burgeoning field of economic sociology. Just as thinking depends upon external categories and meanings provided by society, Durkheim instructs us that we likewise cannot truck, barter, or exchange without knowledge that is inscribed by society.

This is the first book on Durkheim's economic sociology, perhaps because Durkheim was less than explicit in his intent to create such a subfield. Steiner, a prominent Durkheimian scholar, shows that simply reading Durkheim's four most popular books that are translated into English—*The division of labor in society* (1893); *The rules of sociological method* (1895); *Suicide* (1897); and *The elementary forms of religious life* (1912)—will not yield a comprehensive view of Durkheim's economic sociology. Instead, Steiner expertly marshals period journals, including Durkheim's own journal *L'Année Sociologique*, and Durkheim's personal letters to fill in the holes. True to his communitarian sensibilities, Durkheim did not carry out his economic sociology project alone. Steiner shows how Durkheim's students, Marcel Mauss, Maurice Halbwachs, and François Simiand (among others), pushed his economic sociology research project forward during and after their tutor's lifetime. Since these scholars are more commonly studied by anthropologists, Steiner unearths fresh classical material for economic sociologists.

The original French version of Steiner's book is entitled, *L'école Durkheimienne et l'économie: sociologie, religion et connaissance* (2005). This more comprehensive title, roughly translated *Durkheim's school and the economy: sociology, religion, and knowledge*, clearly signals the important role religion plays in Steiner's interpretation of Durkheimian economic sociology. Present day sociologists have specialized in their

respective subfields (such as the sociology of religion and economic sociology) and have consequently mostly ignored the relationship between these seemingly disparate realms. Steiner presents the case that not only Weber, but Durkheim too, placed great importance on the relationship between religion and the economy.

Max Weber is widely viewed as the first economic sociologist (Swedberg 2011), so Steiner's argument that Durkheim also deserves such credit makes this book of special interest to scholars in the subfield. Despite the attention both Weber and Durkheim paid to economy and society, few early sociologists focused on the economy. In fact, Talcott Parsons, a towering sociologist of the mid twentieth century is reported to have made an influential pact with economists along the lines of, "You, economists, study value; we sociologists, will study values. You will have claim on the economy; we will stake our claim on the social relations in which economics are embedded" (Stark 2009, 7). After a long period of dormancy, the field of 'new economic sociology' sprang up in the 1980s. Accepting the risk of over simplification, this sociological subfield largely rallies around 1) articulating a critique of neoclassical theory, and 2) fashioning a more positive social scientific theory of how sociological factors explain economic phenomena.

While Durkheim is rarely associated with the subfield of economic sociology, he is often closely associated with the sociology of religion, stemming mostly from his tour de force, *The elementary form of religious life* (1912). There Durkheim erects a functional definition of religion, whereby we recognize religion whenever social actors come together in unified attention to something that is larger than themselves. The accompanying emotional states of this unified social activity denote the sacred (i.e., religion), which Durkheim pitted against the profane (e.g., mundane economic activity). Durkheim instructs sociologists of religion to abandon more substantive definitions of religion (such as Weber's) and recognize the "religious" or "sacred" in unexpected places.

Durkheim is also considered a seminal secularization theorist, asserting that "if there is one truth that history has incontrovertibly settled, it is that religion extends over an ever diminishing area of social life" (1966 [1893], 118). In this particular instance, Durkheim is referring to the organized institution of religion (e.g., Catholicism), not his functional notion of religion. Assessing the veracity of the secularization

thesis has kept many sociologists of religion busy for nearly half a century.

SUMMARY OF THE BOOK

Steiner argues that Durkheim's economic sociology consists of two research programs. The first is a critique of political economy, or the economists of his day. The second sees Durkheim turn to religion, instead of the economy, as the key to understanding society. The first two chapters outline each program in turn. Chapters 3 to 6 demonstrate how these programs were taken up by Durkheim's students. The final two chapters of the book focus on extending the material covered to an economic sociology of knowledge, with an emphasis on the schooling system. Here Steiner enters into dialogue with performativity scholarship, a growing area of interest among economic sociologists.

First research program: critique of political economy

Durkheim's first program of economic sociology articulated a critique of political economists for their penchant for reductionism, wherein complex social reality was replaced with simplistic assumptions of rational actors operating outside of society. However, Durkheim did not stop there. Similar to Durkheim's broader conception of a "social fact", which refers to "ways of acting, thinking, and feeling external to individual consciousness that are imposed upon individuals" (p. 22), Durkheim's economic sociology focuses on societal level "economic facts". For instance, Durkheim argues that a *contract* is more than the sum of each party's self-interest, "but is possible only thanks to the regulation of contracts, which is of social origin" (*Division of labor*; quoted in Steiner 2011, 28).

Durkheim also envisioned a moral component to his conception of economic fact. Namely, a healthy society's *collective conscience* constrains immoral economic behavior. For Durkheim though, modern industrial society was unhealthy. To fill this moral regulative void, Durkheim proposed a new social institution he called the "professional group". This was to be an association of workers within the workplace that would create cohesion, restrain individual passions and foster a principle of justice. This was Durkheim's early attempt to envision a remedy to the immoral aspects of the modern world.

Simiand and Halbwachs, both students of Durkheim, were principal contributors to the economic sociology section of *L'Année Sociologique*

for almost 50 years (1897-1942), extending Durkheim's critique against the orthodox economists of their day. Most importantly, they "put forward an approach that would draw upon empirical investigation together with contributions from history, statistics and sociology" (p. 107). Three of the book's eight chapters are comprised of long, detailed analysis of the work of Durkheim's students. Incidentally, this raises the question in the reader's mind about how much of the school of economic sociology Steiner outlines should really be credited directly to Durkheim.

Second research program: religion and the economy

In his doctoral dissertation, which was published as *Division of labor* in 1893, Durkheim looked to the economy to address his perennial concerns regarding societal cohesion, particularly focusing on 'organic solidarity', the interdependence stemming from vocational specialization. However, just a couple of years after finishing his dissertation, Durkheim dramatically announced an important change in his intellectual orientation in a letter to the editor of a periodical.

> It was in 1895 that I clearly understood the leading role played by religion in social life. It was in this year that, for the first time, I found a means to approach the study of religion sociologically. This was a revelation to me (quoted in Steiner 2011, 39).

Durkheim disagreed with the Marxian notion that the economy could be at the root of society. Instead, Durkheim set out to show that religion was the foundation from which the economic grew. In the closing pages of *Elementary forms* Durkheim writes, "the idea of economic value and that of religious value cannot be unrelated, but the nature of these relationships has not yet been studied" (quoted in Steiner 2011, 58). While this "revelation" was an important event, Steiner makes the case that it did not bring an end to Durkheim's interest in the economy.

While Durkheim never completed his project for linking religion to his economic sociology, Steiner argues that Durkheim's nephew, Marcel Mauss, did. In *The gift*, Mauss (1923) focuses on a unique kind of economic exchange (i.e., gift exchange) that comprises both interest and disinterest; egoism and altruism. This particular form of exchange can be aptly approached from the perspective of multiple societal spheres: religion (gifts to the gods), economy (exchanging goods), and family (dowries). This leads Mauss to the concept of a *total social fact*, which

tells us no sphere of social life is more fundamental than any other. Steiner argues that Mauss's conclusion parts ways with Durkheim's contention that religion is more essential than the economy (when it comes to explaining the origins of social life.)

While Mauss's stalemate answer (neither religion nor the economy trumps the other) may have been satisfying to its first readers, who were also interested in the origins of society, it strikes me as somewhat less than satisfying. Steiner does not end his story here, however. He emphasizes how Durkheim's "sociology of knowledge" provides a more intriguing link between the sociology of religion and economic sociology. Durkheim believed social action (or social reality) is necessarily preceded by knowledge. The prime examples being religion and myth, which, in Steiner's words, "make and express social reality in one move, and do so through the intermediary of symbols that unite different individual consciousnesses so that they might communicate and feel things in unison" (p. 177). Durkheim sees the schooling system replacing organized religion's role in producing knowledge. As such, the educational institution actively creates the knowledge that is planted in the next generation's heads and also takes over the reins of society's moral order from organized religion. Here we see Durkheim's secularization thesis taking shape.

In a clever thought experiment, Steiner considers how Durkheim's economic sociology might provide a more comprehensive explanation of how the modern profit-driven "economic man" came into being. Weber's (1930) *Protestant ethic and the spirit of capitalism* explains the religious spark that helped start modern rational capitalism, but acknowledges that those religious forces quickly dissipated. What sustained capitalism in the following centuries? Durkheim's emphasis on sociology of knowledge answers that "capitalism cannot survive without ideological support" (p. 213) and "that the schooling system plays a decisive role when it becomes a central institution through which the cognitive capital embodied in the individual is produced" (p. 213).

Performativity

It is with this emphasis on the education system that Steiner most directly engages with current day performativity scholarship, a growing area of interest within economic sociology. Performativity scholars seek to uncover the extent to which economic knowledge is helping create

(or 'perform') economic actors whose behavior often approximates that of the caricatured utility maximizer. As MacKenzie (2006) puts it, economic theory is better depicted as *An engine, not a camera* (the title of his book exploring how financial markets put economic theory into action).

For instance, Simiand shows how Fredrick Taylor's theory of scientific management in the 1920's produced the institutions that fostered its rational behavior. This peculiar workplace environment came about "not because of some miraculous coincidence of the ideal of an isolated theorist and the society's law of progress, but rather through social inscription of theory in institutions in whose terms individuals are led to act in the economic world" (p. 187).

Steiner's emphasis on economic *experts* as central performative actors in the economy closely follows the present day performativity research program. But this narrow implementation of performativity theory is tantamount to consideration of religious ideals only as they apply to clergy. The education system, broadly speaking, has the potential to inculcate economic knowledge to a much larger swath of social actors, not just those with formal economic training. Steiner's emphasis on the education system shows promise for a wider application of performativity theory.

Steiner's overtures to performativity scholarship will be of great interest to economic sociologists. Does Durkheim's work and that of his students provide a sufficient foundation for performativity scholars? Will it push the field in new directions? I would like to have seen Steiner engage with questions of this nature more systematically. Nonetheless, economic sociologists owe Steiner a debt for formally introducing Durkheimian thought into the subfield's institutional repertoire. In true Durkheimian fashion, and the mark of good scholarship, Steiner leaves ample room for creative research to grow.

REFERENCES

Calhoun, Craig, Joseph Gerteis, James Moodly, Steven Pfaff, and Indermohan Virk (eds.). 2007. *Classical sociological theory*. Malden (MA): Blackwell Publishing.

Durkheim, Émile. 1966 [1893]. *The division of labor in society*. New York: The Free Press.

MacKenzie, Donald A. 2006. *An engine, not a camera: how financial models shape markets*. Cambridge (MA): MIT Press.

Stark, David. 2009. *The sense of dissonance: accounts of worth in economic life*. Princeton: Princeton University Press.

Steiner, Philippe. 2005. *L'école Durkheimienne et l'économie: sociologie, religion et connaissance*. Geneva: Librairie Droz.

Swedberg, Richard. 2001. Max Weber's central text in economic sociology. In *The sociology of economic life*, eds. M. Granovetter, and R. Swedberg. Boulder (CO): Westview Press, 62-77.

Jared L. Peifer is a postdoctoral fellow in the sociology department at Rice University. He specializes in economic sociology and sociology of religion with a particular focus on how social actors bring these two realms together in their everyday lives. More specific research projects include analyzing the economic and sociological determinants of monetary contributions to local congregations, and investor loyalty to religiously affiliated mutual funds in the United States.
Contact e-mail: <jared.peifer@rice.edu>

Erasmus Journal for Philosophy and Economics,
Volume 5, Issue 1,
Spring 2012, pp. 128-130.
http://ejpe.org/pdf/5-1-br-7.pdf

Review of James E. Alvey's *A short history of ethics and economics: the Greeks*. Cheltenham (UK): Edward Elgar, 2011, 184pp.

SPENCER J. PACK
Connecticut College

This rather brief work on the relationship between classical ancient Greek thought and economics is apparently the first part of a much larger grand history of the relationship between ethics and economics (p. vii). The book basically follows the template of S. Todd Lowry's magisterial *The archaeology of economic ideas: the classical Greek tradition* (1987). Indeed, it is sort of an abridged version of Lowry's work, though much thinner—more like a sketch or an outline.

As with Lowry, there are sections dealing with the Sophists, Xenophon, Plato, and Aristotle. For the most part Alvey agrees with Lowry's interpretations. There are just a few differences. One is that Alvey stresses more the differences between the latter Plato, as found in the *Laws* and the *Statesman*, and Plato's early and middle dialogues, particularly the well known *Republic*. Alvey explains how the *Republic* is a utopian ideal of a perfectly virtuous city, with rule by philosophers; the abolition of private property for the warrior class; and gender equality, so some women will be philosophers and hence rulers. These are radical, revolutionary proposals on education and gender. Alvey paints Plato here as an early capability theorist in a long tradition which today manifests itself in the work of Amartya Sen and Martha Nussbaum. On the other hand, Plato's *Laws* and the *Statesman* are more practical: they detail what might be achieved in a real city, although the city's goal is still nobility. The later Plato realizes that, since the rule of an outstanding individual looking out for the benefit of others is rare and such people are not easily identified, there is a need for the rule of law. Hence, the rule of law is the best possible regime, replacing the utopian standard of rule by philosophers in the *Republic*. The *Laws* provides a model for law-givers, in which it is the written law and various important customs which shape the behavior and character of those who are compelled to obey them.

In another divergence from Lowry, Alvey distinguishes two major traditions relating ethics and economics, both stemming from Aristotle. One currently manifests itself in the Straussian school, with the work of Leo Strauss, and his followers Joseph Cropsey, Richard Stavely, and, more recently, such theorists as Peter Minowitz, Richard Temple-Smith, and Athol Fitzgibbons. The Straussians are basically elitists, stressing the aristocratic side of Aristotle, and the fundamental inequality of humans in terms of their potential for and achievements in developing the intellectual and moral virtues. The other tradition manifests itself in the work of Sen and Nussbaum. In their more egalitarian capabilities approach, Sen and Nussbaum stress that everyone has an entitlement or right to certain basic functionings or capabilities required for human flourishing. Hence, Sen and Nussbaum stress a more democratic side or interpretation of Aristotle. Nonetheless, both of these grand narratives have commonalities. For Alvey, both show the existence of a once vibrant ethical tradition in the history of economic thought; they agree that the ethical tradition in economics has largely faded out, and consider the current state of economics to be either ethically thin or amoral; and hence they both largely endorse the past over the present.

Alvey himself is a bit torn between these two grand narratives. Alvey was a student of the Straussian Stavely, and adopted his interpretation for a decade or so, before abandoning most of the doctrine (p. 8). Alvey is now more sympathetic to the Sen-Nussbaum approach. Nonetheless, Alvey concludes that "Nussbaum's social democratic interpretation of Aristotle is a distortion" (p. 154). Although in a footnote Alvey does accept that the approach of Sen and Nussbaum "is not a mere history of ideas" but contemporary political theory which is "Aristotelian in spirit", and that Aristotle is used by Sen and Nussbaum as a "tool for contemporary critique and progressive social reform" (p. 156, 8fn.).

I think the key difference between these two interpretations of Aristotle is who should be a citizen of the state; or, how broadly should citizenship be defined. For Aristotle the best regime allows for every citizen to live a flourishing life. Aristotle himself tended to take a narrow view of who should be a citizen. Yet if one takes a broad based conception of who should count as a citizen, then the Sen-Nussbaum capabilities approach is definitely Aristotelian. As Jill Frank persuasively argues in her *A democracy of distinction: Aristotle and the work of politics* (2005), Aristotle's activity-oriented philosophy does harbor democratic possibilities. She writes:

> some commentators conclude, on the basis of Aristotle's account of practical wisdom and virtue, that the political order he endorses must be exclusionary and antidemocratic. There is plenty of evidence in Aristotle's writings to support this conclusion: he excludes from citizenship women and slaves on the basis of their inferior practical wisdom. He excludes artisans and laborers on the ground that what they do interferes with the cultivation of the virtues necessary for citizenship. There is, however, nothing intrinsic to Aristotle's understanding of practical wisdom to support these exclusions (Frank 2005, 122).

Moreover, as with Hegel, the great systematizer of the 19th century, we might also expect that Aristotle, the great systematizer of the ancient world, would have his system split apart into what may be called left and right wing, or egalitarian and aristocratic, interpretations (Pack 2010). Readers interested in a brief account of the relation between classical Greek thought and economics will find Alvey's book worthwhile. However, they would probably be better served by dipping into Lowry's work; or, better yet, studying Lowry's (1987) entire book including his extensive endnotes. Meanwhile, we can look forward to Alvey's larger grand history of the relationship between ethics and economics; the material in the current book will no doubt have an important place at or near the beginning of that ambitious story.

REFERENCES

Frank, Jill. 2005. *A democracy of distinction: Aristotle and the work of politics*. Chicago: University of Chicago Press.

Lowry, S. Todd. 1987. *The archaeology of economic ideas: the classical Greek tradition*. Durham: Duke University Press.

Pack, Spencer J. 2010. *Aristotle, Adam Smith, and Karl Marx: on some fundamental issues in 21st century political economy*. Cheltenham (UK): Edward Elgar.

Spencer J. Pack is professor of economics at Connecticut College, New London, Connecticut, USA. His recent publications include *Aristotle, Adam Smith, and Karl Marx: on some fundamental issues in 21st century political economy* (Edward Elgar, 2010) and "Aristotle's difficult relationship with modern economic theory" (*Foundations of Science*, 13 (3-4), 2008).
Contact e-mail: <sjpac@conncoll.edu>

Erasmus Journal for Philosophy and Economics,
Volume 5, Issue 1,
Spring 2012, pp. 131-134.
http://ejpe.org/pdf/5-1-ts-1.pdf

PHD THESIS SUMMARY: Essays in the economics of knowledge.

SAMULI LEPPÄLÄ
PhD in economics, December 2011
University of Turku

"The economics of knowledge", as outlined by Kenneth Boulding (1966, 1), "studies the role of knowledge in social systems, both as a product of the past and as a determinant of the future". Within this extensive and challenging field, there have been numerous contributions since the early pioneers such as Kenneth Arrow, Friedrich Hayek, Fritz Machlup, and Joseph Schumpeter. The field experienced a rapid growth with the development of the knowledge industries, and Foray (2004) argues that the economics of knowledge has now been firmly established as a distinct discipline.

In conjunction with the expansion also came fragmentation. Different but interrelated terms—such as knowledge, information, belief, expectations, uncertainty, technology, innovation, invention, skills and human capital—appeared in the literature. Many papers, studying specific questions regarding the role of knowledge, seemed to have little in common. In particular, research on market knowledge (i.e., knowledge in the general decision-making sense) was largely separated from the research on technological knowledge (i.e., the causes and consequences of technological development).

This evident fragmentation is also the basis of Mirowski's (2009) thesis, contra Foray, that there is (as yet) no such thing as an economics of knowledge. While fragmentation along different research questions and schools of thought is regrettable, it is unsurprising considering the wide relevance that the economics of knowledge is claimed to have for every sub-field of economics (Lamberton 1971; Stiglitz 1985).

The introductory chapter of my thesis surveys the literature in this field, identifying the main themes and contributions. Besides the apparent fragmentation, it is perhaps also a sign of immaturity that we have nothing close to general economic theories of knowledge as yet. Nevertheless, it cannot be denied that there is a continuing and growing

tradition of research on how incentives and institutions affect the generation, dissemination, and use of knowledge in the economy.

While the importance of knowledge issues for economics was recognized early on, the assumption of perfect knowledge held the field. The reason for this, argues Stiglitz (2002), was that while knowledge can be perfect in only one way, it can be imperfect in infinitely many ways. As such, it was unclear how this assumption could or should be relaxed. Technological development finally prompted economists to analyze the incentives and institutions behind it. A prevalent feature in R&D, at least as much as for any other economic decision, is uncertainty, which implies lack of knowledge. This prompted economists to consider decision-making in situations where knowledge was imperfect or incomplete.

Uncertainty implies not only a lack of knowledge or available information, but also the issue of truth-worthiness regarding our beliefs and communicated information. While economists have noted that knowledge, being nonrival and (partially) nonexcludable, is a very special kind of an economic good, the issues related to the quality of beliefs and information require more work. After all, the concept of knowledge proper implies justified true beliefs, not merely any beliefs, be they true or false. If veracity is not trivially verifiable then it becomes an important issue, and we require a theory of justification explaining a tendency towards acquiring true beliefs and discarding false ones.

At first, this may seem something largely beyond the domain of economics. However, knowledge acquisition and communication depend on incentives as well, creating a niche for economists to study the role of incentives and institutions in truth-seeking and truth-telling activities. The demand for studies on epistemic efficiency, i.e., how closely beliefs and information approximate truth in various situations, seems extensive.

The subsequent essays in my thesis address more specific issues. Essay 1 takes an epistemological view on knowledge transfers, now in focus in many areas of economics and related sciences. While tacit knowledge is widely used to explain the main difficulties in knowledge transfers, besides the unarticulable nature of some knowledge, equally important are the incentives in knowledge acquisition and communication. Successful transfer of codifiable knowledge requires that the sender's belief is true (capability) and that she sincerely reports

that (reliability). The analysis of incentives can explain the success or failure of knowledge transfers in different contexts.

In Essay 2, I set out to find the level of excludability that creates the optimal allocation of research investments in nonrival knowledge. Taking into account the tradeoff involved in greater or less excludability the optimal level can be found; though, being a function of the research costs and benefits, this would typically vary between different knowledge industries and types of knowledge. While this result can explain the continuing controversy regarding the desirability intellectual property rights, it also illuminates a major challenge for IPR law.

Essay 3 (Leppälä 2010) re-examines Hayek's thesis on the informational role of the price system. Drawing an analogy from his psychological theory, I argue that the informative role of a price is not to communicate the same knowledge to everyone, as is generally thought in the subsequent literature, but to provide a reference point to which each can compare their local knowledge. This idea is further applied in addressing some common interpretations of the informational role of prices.

At a general level, Essay 4 (Leppälä and Desrochers 2010) argues that if one takes a methodologically individualistic perspective (seldom applied in regional economic analysis) the basis for publicly promoted regional specialization largely disappears. Furthermore, by ignoring the actual spillovers between individuals, the research on localized knowledge spillovers has been unable to prove or document the existence of these spillovers. An attempt to correct this shortcoming is made in Essay 5 (Desrochers and Leppälä 2011), where, based on a qualitative survey of individual Canadian inventors, three broad mechanisms conducive to inter-industrial knowledge spillovers and subsequent inventions are identified.

These essays are hardly the last word on the topic but hopefully illustrate the importance of these questions and provide examples of how they can be addressed from an economic point of view.

REFERENCES

Boulding, Kenneth E. 1966. The economics of knowledge and the knowledge of economics. *The American Economic Review*, 56 (1-2): 1-13.

Desrochers, Pierre, and Samuli Leppälä. 2011. Opening up the 'Jacobs spillovers' black box: local diversity, creativity and the processes underlying new combinations. *Journal of Economic Geography*, 11 (5): 843-863.

Foray, Dominique. 2004 [2000]. *The economics of knowledge*. Cambridge (MA): The MIT Press.

Lamberton, Donald M. (ed.). 1971. Introduction. In *Economics of information and knowledge*, ed. Donald M. Lamberton. Harmondsworth: Penguin Books, 7-17

Leppälä, Samuli. 2010. Hayek on prices and knowledge: supplementing "The use of knowledge in society" with The sensory order. In *The social science of Hayek's The sensory order* (Advances in Austrian Economics, Vol. 13), ed. William N. Butos. Bingley: Emerald Group Publishing Limited, 237-259.

Leppälä, Samuli, and Pierre Desrochers. 2010. The division of labor need not imply regional specialization. *Journal of Economic Behavior and Organization*, 74 (1-2): 137-147.

Mirowski, Philip. 2009. Why there is (as yet) no such thing as an economics of knowledge. In *The Oxford handbook of philosophy of economics*, eds. Harold Kincaid, and Don Ross. Oxford: Oxford University Press, 99-156.

Stiglitz, Joseph E. 1985. Information and economic analysis: a perspective. *The Economic Journal*, 95 (supplement): 21-41.

Stiglitz, Joseph E. 2002. Information and the change in the paradigm in economics. *The American Economic Review*, 92 (3): 460-501.

Samuli Leppälä obtained his PhD from the School of Economics at the University of Turku, Finland. He was supervised by Hannu Vartiainen (Helsinki Center of Economic Research), and the official opponent at his defense was Suzanne Scotchmer (University of California, Berkeley). Samuli is currently a postdoctoral researcher at the Department of Economics in the University of Turku. In August 2012, he will be joining Cardiff Business School as a lecturer in economics.
Contact e-mail: <samuli.leppala@utu.fi>
Website: <http://sites.google.com/site/samulileppala/>
PhD thesis available online at: <http://info.tse.fi/julkaisut/vk/Ae8_2011.pdf>