# Doing well by doing good?

# On the scope and limits of Ethics in business – (2015)

Hartmut Kliemt

All rights reserved ©

Assistance by Katharina Brütt, Sarah-Lea Effert and Michael Kirmes is gratefully acknowledged

Comments welcome … new and expanded version pending

**To whom it concerns, i.e. the reader:**

1. Readers of the subsequent text should be aware of its somewhat preliminary character. Citations must still be checked and stylistic infelicities be removed, but otherwise it is basically done.
2. Preaching morals is easy justifying it is serious business in particular if complex issues of business and management decisions are involved. To make problems more transparent they are presented in elementary decision and game theoretic terms whenever appropriate. I will not hesitate to express my own classical liberal economic and philosophical normative views yet only for critical examination not with the purpose of indoctrination.
3. Without discussing concrete cases and problems this text is incomplete. Practical ethics must be applied to practical cases. Some exemplary cases are included below, yet others should be discussed. Otherwise the subsequent very short introduction to ethics in business is self-contained.

**Beyond the present text:**

**Shareholder vs. Stakeholder approaches, Friedman vs. Freeman**

Freeman, R.E. (1984/2014). Stakeholder Theory of the Modern Corporation. (pp. 160-168). &

Friedman, M. (1970/2014). The Social Responsibility of Business Is to Increase Its Profits. (180-183). Both texts reprinted in W.M. Hoffman, R.E. Frederick, & M.S. Schwarz (Eds.), *Business Ethics. Readings and Cases in Corporate Morality* 4th. Edition, Boston et al.: McGraw-Hill Higher Education.

**Reflective Equilibrium**

* 1. Hahn, Susanne (2004), ‘Reflective Equilibrium – Method or Metaphor of Justification’, in: Winfried Löffler and Paul Weingartner (eds.), *Knowledge and Belief, Proceedings of the 26th International Wittgenstein Symposium Kirchberg am Wechsel*: Wien, 237-243.
	2. Rawls, John (1951), ‘Outline of a Decision Procedure for Ethics’, *Philosophical Review*, 60(2), 177-197.
	3. Daniels, Norman (1979), ‘Wide Reflective Equilibrium and Theory Acceptance in Ethics’, *The Journal of Philosophy,* 76(5, May), 256-282.

**Markets and their moral limits**

1. Stodder, James (1998), ‘Experimental Moralities: Ethics in Classroom Experiments’, *Journal of Economic Education,* 29(2, Spring), 127-138.
2. Buchanan, James M. (1999 ff.), *Property as a Guarantor of Liberty*, in collected Works (ed. Brennan, Kliemt and Tollison) vol. 18
3. Kliemt, Hartmut (1997), ‘The Politics of the Market’, *Current Legal Theory*, 14.

**Firm Commitments**

a. Frank, Robert H. (1987), ‘If Homo Economicus Would Choose His Own Utility Function, Would He Want One with an Conscience?’, *American Economic Review*, 77(4, September), 593-604.

b. Kliemt, Hartmut (1993), ‘Constitutional Commitments’ in: Philipp Herder-Dorneich, Karl-Ernst Schenk and Dieter Schmidtchen, *Jahrbuch für Politische Ökonomie*, 12, Mohr Siebeck Verlag: Tübingen, 145–173.

c. Mayer, Colin. *Firm Commitment: Why the corporation is failing us and how to restore trust in it*. Oxford University Press, 2013.

**General texts relevant to the topic**

* + 1. As a brief introduction to theories of ethics it is useful to consult the modern classic: **Frankena, W.K.: Ethics. Prentice Hall. 1973 or later** eg. Frankena, William K. (1988), *Ethics 2nd edition*, Pearson: London.
		2. Brenkert, George G., und Tom L. Beauchamp, Hrsg. *The Oxford Handbook of Business Ethics*. Oxford University Press, USA, 2012 may be helpful, too.
		3. Of general interest and highly recommended because normative codes of business are very much comparable with tribal codes: Greene, Joshua David. *Moral Tribes: Emotion, Reason, and the Gap Between Us and Them*. New York: Penguin, 2013 or later.

Note for course work

Participants of courses based on this text are strongly advised to start reading the text in particular the exemplary cases and the questions following them as soon as possible. The material of the preceding reading list may be helpful in exploring issues somewhat more deeply but is not essential for the course.

Part I: What it is all about, examples and issues of business and corporate ethics

***1. Some examples and cases***

1.1 Are whistleblowers always heroes?

1.2 Environment, employment and loyalty

1.3 The Big Question: What is short selling, and is it a practice that should be stamped out?

*2. Setting the stage for practical ethics in business*

2.1 A plea for modesty in business and corporate ethics

2.2 Shareholder and moral value

2.3. The long- and the short-term

2.4 Obliquity

2.4.1 Obliquity and indirectness of success in ethics

2.4.2 Obliquity in an experiment

2.4.3 Obliquity in business

***3. The fundamental divide in business ethics: Friedman vs. Freeman***

3.1 The responsibility of management and the diversity of values

3.2 Transparency of operations as most fundamental requirement of “responsible” management

**Part II: The economics and ethics of** interaction

***1. The traditional argument vindicated by Vernon Smith‘s Double Auctions***

1.1 A stylized sketch of experiments

### 1.2 Normative aspects illustrated by experimental markets

***2. Markets, Firms and Guangxi***

2.1 Firms on markets

2.2 Guangxi and substitutes for markets

***3. Getting in and out of social traps***

3.1 Two person exchange game and the hidden side of the invisible hand

3.2 N-person interactions

3.2.1 N-person PD

3.2.2 N-person Volunteer’s Dilemma

3.2.3 Voluntary Contribution Mechanism, VCM

Part III: Towards “(Wide) Reflective Equilibrium”

***1. Disequilibrium and dissent***

***2. Deliberative practices in ethics akin to those in science***

2.1 Rawls’ outline of a decision procedure for ethics illustrated

2.2 Two examples

2.2.1 Trolleys and consequences

2.2.2 General principles of business

***3. A refined reflective equilibrium metaphor***

**Part IV: Substantive Ethical Approaches**

## *1. Basic concepts of consequentialism and non-consequentialism*

## *2. Moral institutions not moral judgments as the focus of ethics*

## *3. Ethics by exception*

## *4. Moral institutions as conventions*

## *5. A glimpse of ethical rigorism*

***6. Utilitarianism as classical teleological consequentialism***

## 6.1 Utilitarianism generally characterized

## 6.1.1 Utilitarianism and interpersonal comparisons

6.1.2 Utilitarianism descriptive and egoistic vs. prescriptive and altruistic

6.2 Act vs. rule utilitarianism

**6.2.1 Act utilitarianism**

### 6.2.2 Rule utilitarianisms

## *7. Should the numbers count and how?*

## *8. Rawls, the separateness of persons and consequentialist trade-offs*

**V. References**

Part I: What it is all about, examples and issues of business and corporate ethics

Adapting a famous statement of the German philosopher Immanuel Kant, we might say that business ethics without theory is blind and business ethics without cases is empty. In a course of business ethics participants clearly need to be exposed to cases in which they have to engage practical ethical deliberations. Such deliberations will typically start from a state of “**reflective disequilibrium**”.

A state of reflective *disequilibrium* is plagued by incoherencies and problems of several kinds. We do not know all the relevant facts or have contradictory signals of what the facts are. Moreover, there will be tensions between arguments representing our normative and factual convictions and our own several aims, ends or values may not cohere on the specific and/or the general level.

To the extent that striving for coherence is a requirement of our rationality as well as of our self-respect as moral actors we must want to get out of disequilibrium. We need to engage the search for a **“reflective equilibrium”, RE** (see Rawls, 1951, Hahn 2000, 2004). If we include in our search general ethical theories we may hope to reach a comprehensive, so-called “**wide reflective equilibrium**”, **WRE** (see Daniels 1970)eventually. A **WRE** brings together singular judgments concerning specific real cases, general normative principles, and general (ethical) theories in a coherent whole in which all elements support each other like the stones in a Roman arch hold each other.

Coherence typically requires that we weigh the pros and cons of several basic general principles like never “hurt an innocent person” or “promote the common weal” or, for that matter, “do not hurt the interests of customers of the company” and “promote the interests of the company’s employees and owners”.

To illustrate, assume that the owners of a company are pursuing a policy of “ecology first”. Now, an employee feels the moral obligation to help, say, poor producers in the supply chain of the company knowing that they produce in ecologically harmful ways. If such ecologically harmful production takes place within the bounds of law should a manager who thinks that in this case ecology is less important than other interests feel entitled to act against the owners’ intent of “ecology first” or should she decide otherwise according to her own moral principles?

The preceding illustration does *not* concern a “***conflict between ethics and economics”***. There is a conflict between two moral positions that both have an economic impact. In a world of scarcity one of the primary *moral* obligations is to act economically. Moreover, it is a *moral* obligation of employees to comply with the duties that they have accepted in their employment contracts.

If we take the moral obligation of keeping our (contractual) promises seriously it is ethically certainly not correct to act as if there were no ethical values other than our own. Though we are all partisan for our own ethical values we must be aware that others are as partisan for their competing views. It is natural to believe that our values are correct and the values others endorse are incorrect but it is as natural for them as for us to believe that they are right. We as they should take a stance on values, yet it is arrogant not to acknowledge the plurality and potentially conflicting nature of values.

When it comes to real world (ethical) decision-making in a context in which many individuals are involved, the claim to have found *the* ethically right answer merely indicates a misunderstanding of the powers of human reason and the extent of inter-individual agreement. There is always a **plurality of values**. Inter-individual agreement in practical affairs is mostly illusionary. Even though some of those who are partisan for their “parochial” ethical perspective may claim that every rational person would have to agree with them, this cannot be justified. Conflict between “us” and “them” rather than harmony prevails in the real world. We are social being adapted to tribal rather than universal ethics (see Greene 2013).

Even if all humans were to seek for wide reflective equilibria and coherent views on ethics the hope that universal agreement will emerge from their searches is certainly in vein. So, why bother with seeking for a WRE in a systematic ethical analysis at all? And why not leave such efforts as impractical to the philosophers at universities? Are ethical studies not simply bad business?

It clearly depends whether engaging practical ethical issues in business is good rather than bad business. Yet in a world in which ethical blame is rather swiftly allocated ethical risk analysis may in fact be good business. An analysis of ethical risks before finalizing business decisions will reveal what can and what cannot easily be defended in public. In particular if a scandal breaks it will certainly be helpful if the management of a company can demonstrate that in reaching decisions whose results are scandalized it had been weighing different arguments systematically. Those, who decided only *after* looking into all risks and consequences of their decisions and can document that they did, have a public convincing reason to have acted “responsibly” if unsuccessfully.

If external actors will hold an organization or its management responsible for substantive results of its decisions it will matter not only legally but also in its public relations if the company can convincingly show that it took practical ethical problems seriously and had analyzed them in a “systematic” way before taking action. If its internal processes at least broadly embody a systematic search for reflective equilibria on ethical issues – including an effort to avoid blind spots in ethical awareness (see Bazerman and Tenbrunsel 2011) – then the company will in all likelihood more easily get away with failure. Yet even if not, those who have to make decisions might want to know for their own personal moral sake what is ethically up when making the decisions.

In line with the quasi Kantian initial statement on the complementary role of theories and cases, it seems reasonable to look at specific cases, first. So let us start by addressing a few cases in more or less intuitive ways before turning to more general economic examples of interactions in a market economy and then on to other ethical issues.

# For more cases see the useful anthology of a leading practical ethicist:

Beauchamp, Tom L. *Case Studies in Business, Society, and Ethics*. Revised. Upper Sadle River, N.J: Prentice Hall, 2003.

## 1. Some examples and cases

### 1. 1 Are whistleblowers always heroes?

(By Julian Friedland, Ph.D.)[[1]](#footnote-1)

During his first year on the job, Jim Balast, a young engineer at Electromyte, had discovered a flaw in one of the company's newest products. He noticed that thousands of new fuel-injection computer chips already being shipped to several auto manufacturers were likely defective. He only discovered this by chance after taking a bagful of discarded chips home for his 6 year-old son Lyle, who loved to build imaginary inventions with various machine parts.

Once home, he took one of the chips apart to show Lyle how it looked inside. But when he examined it closely, he saw that the conductive metal gave off an uneven reflection in several places. He reached for a magnifying glass and discovered that the metal had hairline cracks close to the spots where the chip was to plug into the computerized fuel-injection system. It was obvious to Balast that if placed inside a car, the chip would eventually fail, cutting off the vehicle's engine and possibly causing a serious accident. Electromyte had just updated this chip design as part of a new fuel-injection system that provided increased fuel efficiency and durability. The new fuel-injection systems on the production line were created to only fit the new chip model.

Balast opened all the other chips he had in his bag. After all 25 had been examined, four other chips had similar defects. If this had been a representative sample, then 20 percent of all the chips placed inside new vehicles would prematurely fail, endangering the lives of countless people on roads. Jim had to tell his supervisor immediately, so he phoned Tim Walt, the head engineer. He told him that from what he could tell, it looked like a design flaw and/or manufacturing defect that had somehow passed unnoticed, since the location of the cracks seemed to be associated with the chip's new plug-in system. Strangely, Walt did not seem very appreciative that Balast had taken the trouble to alert him of the problem. Since it was already Friday evening, Walt wearily replied that he would inform management on Monday but that Balast should not tell anyone else about it. He added that little would likely be done about it this late in the game.

Balast returned to work on Monday and did not hear back from Walt, who was away on business for several days. So, he finally called him again on Thursday to see what had been decided. Walt apologized for not getting back to him and said he'd been assured that the chips had undergone a rigorous testing process and that defects were likely to be extremely rare.

But Balast protested that he had found 5 serious defects in 25 chips. Walt replied that at this point it was out of his hands and that they were both under strict contractual non-disclosure agreements. Balast said he was not sure he could keep quiet and suggested they meet to discuss the matter further. Walt ended the conversation by saying that if the company decided to go ahead with the chips and not delay production, then it was neither of the two men's faults should any accidents occur as a result. Each had promptly spoken up about the problem. There was nothing further they could do.

Balast was shocked at this attitude, which seemed reckless and inexcusable to him. He suspected that Electromyte was gambling that should the flaw be discovered by victims bringing lawsuits, the company could share the blame with both the fuel-injection manufacturer and the auto company. This scenario seemed less costly than having to take the entire blame alone immediately by recalling the chips and putting old fuel-injection systems back into production. Auto companies were counting on the new system to increase the fuel efficiency of their fleets by several miles per gallon. They would be quite angry if they had to abandon that goal and suffer production delays to boot. Electromyte might lose future contracts as a result.

Balast decided to contact the car companies himself to let them know of the flawed design. Within a few days he had reached all four companies and felt a measure of relief. But, he was fired shortly thereafter and later sued by Electromyte for stealing company information, violating his contractual non-disclosure agreement, and defamation.

What's worse, none of the four car companies halted production and presumably defective chips found their way into the market. Balast retained an attorney and countersued Elecromyte, citing recent whistleblower protection laws in his home state of California. A year later, Balast settled out of court for $150,000 in damages. In return, he had to write to each car company to withdraw his charges and apologize for making unwarranted conclusions based upon dubious evidence.

When a reporter asked him why he accepted the settlement, he said that he had to bear the full financial brunt of legal fees and being unemployed for more than a year. He was anxious to put the matter behind him and begin seeking a new job. No one it seemed wanted to hire him anymore as an engineer knowing he had blown the whistle on his past employer. It was looking more and more to Balast that he would have to find an entirely new profession.

**Discussion Questions**

Do you think Balast acted “ethically” in alerting the auto companies about the chip defect?

Why or why not?

Could Balast have acted differently to be more effective in getting the defective chips off the market? If so, how?

As a professional, does Balast have a greater duty to protect the public than himself and his family?

Should Electromyte have a mediation procedure via an ombudsman to protect whistleblowers?

Would this have been to the company's advantage in this case?

Why or why not?

Is there anything else the government should do to protect and prepare whistleblowers like Balast?

If so, what?

### 1.2 Environment, employment and loyalty

### Case 1

Chantale Leroux works as a clerk for Avco Environmental Services, a small toxic-waste disposal company. The company has a contract to dispose of medical waste from a local hospital. During the course of her work, Chantale comes across documents that suggest that Avco has actually been disposing of some of this medical waste in a local municipal landfill. Chantale is shocked. She knows this practice is illegal. And even though only a small portion of the medical waste that Avco handles is being disposed of this way, any amount at all seems a worrisome threat to public health.

Chantale gathers together the appropriate documents and takes them to her immediate superior, Dave Lamb. Dave says, "Look, I don't think that sort of thing is your concern, or mine. We're in charge of record-keeping, not making decisions about where this stuff gets dumped. I suggest you drop it."

The next day, Chantale decides to go one step further, and talk to Angela van Wilgenburg, the company's Operations Manager. Angela is clearly irritated. Angela says, "This isn't your concern. Look, these are the sorts of cost-cutting moves that let a little company like ours compete with our giant competitors. Besides, everyone knows that the regulations in this area are overly cautious. There's no real danger to anyone from the tiny amount of medical waste that 'slips' into the municipal dump. I consider this matter closed."

Chantale considers her situation. The message from her superiors was loud and clear. She strongly suspects that making further noises about this issue could jeopardize her job. Further, she generally has faith in the company's management. They've always seemed like honest, trustworthy people. But she was troubled by this apparent disregard for public safety. On the other hand, she asks herself whether, maybe, Angela was right in arguing that the danger was minimal. Chantale looks up the phone number of an old friend who worked for the local newspaper.

**Questions for Discussion**

What are the reasonable limits on loyalty to one's employer?

Would it make a difference if Chantale had a position of greater authority?

Would it make a difference if Chantale had scientific expertise?

**Case 2**

Jonica Gunson is the environmental compliance manager for a small plastics manufacturing company. She is currently faced with the decision whether or not to spend money on new technology that will reduce the level of a particular toxin in the wastewater that flows out the back of the factory and into a lake.

The factory's emission levels are already within legal limits. However, Jonica knows that environmental regulations for this particular toxin are lagging behind scientific evidence. In fact, a scientist from the university had been quoted in the newspaper recently, saying that if emission levels stayed at this level, the fish in the lakes and rivers in the area might soon have to be declared unsafe for human consumption.

Further, if companies in the region don't engage in some self-regulation on this issue, there is reason to fear that the government — backed by public opinion — may force companies to begin using the new technology, and may also begin requiring monthly emission level reports (which would be both expensive and time consuming).

But the company's environmental compliance budget is tight. Asking for this new technology to be installed would put Jonica's department over-budget, and could jeopardize the company's ability to show a profit this year.

**Questions for Discussion**

What motives would the company have to install the new technology?

What motives would the company have to delay installing the new technology?

Why might the companies in this region prefer for the government to impose new regulations rather than to self-regulate behavior?

What kinds of collective action problems are involved in the case?

For the public at large?

For the companies in the region?

**Case 3**

Robert Brink is head of a group of twenty employees in his bank. A crisis has hit and his superiors decide on reducing the staff of the bank by ten percent. Robert gets the order to suggest three individuals for layoff within the next two weeks. There is no way to appeal the decision. It has to be executed. It is also clear that the order is within the bounds of law.

Questions for Discussion:

How should Robert or anybody else go about such decisions?

From the point of view of Robert’s own interests and ethical convictions as head of a group and member of the staff of the bank?

From the point of view of the company and its future prospects?

From the point of view of universalistic ethics focusing on what everybody should do in a comparable situation?

Is there a good reason for Robert or the company to go by general ethical principles when deciding such issues?

Is there a good economic reason?

Are there any principles according to which we should in general decide on layoffs?

What is the relative weight of the following concerns that may apply when laying of employees:

 Marginal contribution of the worker to the collective product

 Age of employee

 Number of dependents

Is it possible to use expected profits directly as a criterion for layoff decisions?

### 1.3 The Big Question: What is short selling, and is it a practice that should be stamped out?

From *The Independent Business* (Wednesday 23, July 2008). The article below can be found on:

<http://www.independent.co.uk/news/business/analysis-and-features/the-big-question-what-is-short-selling-and-is-it-a-practice-that-should-be-stamped-out-874717.html>

**By David Prosser, Deputy Business Editor**

Wednesday, 23 July 2008

**Why are we asking this now?**

It has emerged that Morgan Stanley, one of two huge investment banks that has been helping HBOS to raise £4bn from shareholders, was at the same time selling the company's shares short – betting that the Halifax Bank of Scotland group would fall in value. Morgan Stanley has not broken the law, or breached any City rules, but its dealings have certainly raised eyebrows. And this is the latest in a series of controversies in recent months relating to short selling.

**What is short selling?**

It is a way of profiting from a fall in a company's share price. Most stock market investors buy shares in the hope and expectation that their value will increase – "going long" in the jargon of the City – but it is also possible to make money when the opposite happens. Shorting means selling a share that you don't own in order to buy it back once it has fallen in price, netting a profit in the process.

**How can you sell something you don't own?**

In a conventional short sale, the investor – usually a hedge fund or large investment bank – takes the view that shares in a particular company are set for a fall. The investor then borrows the shares from someone who does own them – most often a large pension fund or insurance company – and sells them in the market. Once the shares have fallen in value, the investor buys them back at the lower price and returns them to the lender.

If all goes according to plan, the investor is paying less to buy back the shares than it received for selling them. There are some costs involved, notably that the lender charges a fee for loaning out its shares, but in an ideal world the shorter still makes a tidy profit.

There's a variation on this theme, known as "naked short selling" – a form of shorting where the investor doesn't even bother to borrow the shares it is betting against. This is possible because share deals are often not settled immediately. The seller promises to deliver the stock after a short delay – say three days. If a short seller buys the stock back before it has to make good on the original delivery, no shares need actually change hands.

**So what's all the controversy?**

The biggest concern is that short selling has often been associated with market abuse. The clearest victim of such abuse was HBOS itself, earlier this year. Over the course of just an hour one day in March, its shares plunged when rumours swept the stock market that the bank had financial problems similar to those that caused the collapse of Northern Rock. The rumours were totally false and the share price recovered later in the day, but not before investors with short positions in HBOS shares had made a handsome gain. City regulators are still trying to discover who started the malicious gossip about HBOS, but there is widespread suspicion that the rumours were planted by a hedge fund keen to make a fast buck. Such behaviour is illegal, but also very difficult to investigate.

**What about Morgan Stanley?**

The ethics of Morgan Stanley's shorting of HBOS is much less clear-cut. The investment bank knew that if HBOS's fund-raising was not a success, it would be held to a promise to buy the bank's shares at a higher price than their prevailing market value, booking a nasty loss. Selling the shares short was one way of making some of these losses back.

Morgan Stanley argues that its shorting of HBOS was sensible risk management, conducted by a separate department to the part of the bank handling the fund-raising. But HBOS must wonder why a bank to which it is paying considerable sums for work on its fund-raising has been simultaneously making money from a fall in its value. It's also clear that by Friday, when Morgan Stanley began shorting HBOS shares in a major way, it knew that the fund-raising was going to be a serious flop.

**Any other problems with this practice?**

Critics also have little time for pension funds and insurance companies that facilitate the process by lending out their shares. These large investors buy shares on behalf of their clients – you and me – and presumably hope they will rise in value. So it seems bizarre that they're prepared to lend stock to people who want to make money from falling share prices.

The pension funds' argument is that they buy shares with a long-term view. The sort of short-term ups and downs caused by short sellers does not affect this, and besides, they say, they make money for clients from stock lending. Still, in the worst cases, short selling can totally destabilise a company, damaging its prospects for years to come.

**What sort of money is involved?**

The hedge funds and investment banks that dabble in short selling do so with astronomical amounts of money. Analysis conducted by The Independent shows that hedge funds have made more than £1bn betting on a fall in HBOS's share price in the past two and a half months. One single hedge fund is thought to have made £1bn betting on the collapse of Northern Rock last year.

Remember that stock-market investment is a zero-sum game. For every pound of profit made by a short seller, there's an equal loss for shareholders who have gone long.

**So what can be done?**

Regulators around the world have already attempted to police short selling more closely. The Securities and Exchange Commission in the US, for example, has prosecuted traders for spreading false rumours about companies they've sold short. The Financial Services Authority in the UK still hopes to bring similar charges against the HBOS raiders.

New rules have also been introduced. Naked short selling is illegal in the US in most circumstances. In the UK, the FSA announced earlier this year that anyone shorting the shares of a company holding a rights issue to raise new funds would have to disclose their trading positions once they exceeded quite low thresholds. It has promised additional regulation if these rules do not prove sufficient to prevent market abuse.

Pressure has also been brought to bear on stock lenders. Many pension funds now operate on the basis that they will not loan out any of their shares. But the fact remains that short selling, when done within in the law, is a legitimate investment practice and an outright ban would be a draconian intervention. Buying shares in a company is, in essence, a simple gamble that the price of the stock will rise. Why should investors not also be allowed to bet on a fall in share prices?

**Query of the journalist: Is short selling just another example of City excess?**

**Yes...**

\* Making money as companies lose their value is simply profiting from other people's misery

\* Short selling is very often associated with murky – and even illegal – market practices

\* Hedge funds, often owned by a small group of traders, make massive profits from driving down share prices

**No...**

\* Betting on a share-price fall is no less legitimate an investment than gambling that the market will rise

\* City regulators already police the financial markets to curb illegal practices and catch those who flout the rules

\* Some instances of short selling are just a way of reducing the risk of very large "long" investments

**Other queries**

1. What do you think of the behavior of Morgan Stanley in the affair if you look at it from an ethical point of view?
2. Is the ethical issue rather one of constitutional or institutional ethics of regulators or of within institutional choice:
	1. is it an issue of fixing the rules of the game,
	2. is it an issue on the level of the participants in the game
	3. is it both?
3. Is it convincing that Morgan Stanley argues that another department than the one helping HBOS to raise capital was betting on a decline in value of the company?
4. What is the relationship between ascribing responsibility to a company as a whole and separating decision-making in a company to prevent the use of insider information?
5. Would the actors in Morgan Stanley have violated the fiduciary relationship to their shareholders if they would have refrained from pursuing the profit opportunity?

## Additional topics for discussion: Extend the discussion to cover cases like HFT (high frequency trading). Is there anything wrong with it or could we say that participating in HFT or activities that are exploited by high frequency traders is simply up to those who trade?

## What about arbitrage trading in general? Is there something wrong if somebody exploits an opportunity not noticed by those who are in control of resources? (Hint, think of two people who stretch out their hands and want to connect with some other hand – yet unspecified – and somebody steps in and holds the hands receiving a slice of the cake that otherwise could have been split between the two original potential traders.)**2. Setting the stage for practical ethics in business**

“**Ethics**” is a word with many meanings. Sometimes the term is used to refer to the activity of ethical theorists who discuss ethical problems. Sometimes “ethics” refers to the actual rules and norms of behavior that are meant to guide practices. For instance medical doctors use the term “unethical” meaning that a behavior is in violation of the domain-specific norms of medical practice. „Ethical“ may cover strongly opposing views. For instance an ethical theorist may be of the opinion that what doctors call “unethical” is morally right even though the specific ethical code of doctors forbids it. Things may get even more complicated if we include the use of words like “moral" and “immoral”. Again the terms are sometimes used for referring to actual practices and sometimes for the theories that describe or pre-scribe behavioral practices.

Rather than presenting a long list of tedious definitions that of necessity will not cohere with the literature or common parlance the present text will accept the **ambiguity of terms like “ethics”, “morals”, “ethical”, “unethical”, “immoral”,** etc. It should be more or less clear from context what is meant. There will often be clear hints of what the discussion is aiming at. For instance, if practices of people are described rather than criticized this will be pointed out. It will also be (made) clear whether a statement is meant as suggesting that something should be done as opposed to predicting what would be done.

What as a matter of fact is done in business practice needs not be deemed right from a critical moral point of view. Likewise, what seems to be desirable from the point of view of some ethical theory or other will not necessarily guide actual human practices nor will every reasonable person agree that it should.

Quite generally **prescriptive** or normative ethical views must be distinguished from **descriptive** accounts of the ethical views that are as a matter of fact around (either in theory or in practice). Explaining why people do certain things is the task of empirical or descriptive ethics, criticizing or endorsing what they do is a prescriptive (at least evaluative) activity.

### **2.1 A plea for modesty in business and corporate ethics**

As a discipline modern philosophical practical ethics has contributed considerably to our understanding of the moral dimensions of practical business problems and the role of **ethical codes** in dealing with such problems.[[2]](#footnote-2) As a theoretical endeavor ethics does, however, not provide fixed rules or standard recipes for action that would be readily applicable in practical business affairs.

Those working in the discipline of business administration should not scoff at business ethics’ failure to come up with recipes or standardized general solutions for practical problems. More often than not, approaches discussed in management – as a discipline reflecting practical administrative issues – do not provide ready-made directly applicable standard solutions to problems either. Attending a course in management is often more like building up general capabilities in the gym than practicing specific routines of a specific sport. The muscles of the sportsman can be used for all sorts of activities even if they have been built up in the gym with no specific purpose.

Likewise, the purposes to which the know-how of a course in business ethics can be put remain unspecified. The trainer in the gym should not tell which sports his trainees should perform after training, the conductor of a course in business ethics should not tell attendants what to do with their know-how in practice either. The trainer is not in the role of an advisor to benevolent despots nor should his theories be designed to further specific purposes. Nevertheless the general advice on procedures that is created by a course in business ethics should be useful for the purposes of its attendants.

**In sum**: *Attending a course on practical business ethics should yield useful advice of how to construe a solution in pursuit of the attendant’s own aims, ends, or values rather than presenting pre-fabricated solutions furthering specific aims, ends or values.*

### 2.2 Shareholder and moral value

Most people presently working in the field of business and corporate ethics tend to regard an exclusive focus of managers on shareholder value as morally inadequate if not as frankly absurd. They intuitively reject Milton Friedman’s thesis that the business of business is not ethics but business (see Friedman 1970). This seems one of the reasons why they chose to focus on ethical issues in the first place. Most business(wo)men of today feel the same. They also reject the view that going for share value has priority of other business concerns.

In resolving conflicts between the pursuit of shareholder and other values, quite a lot depends on how we understand the term “**shareholder value”**. According to one prima facie appealing explication[[3]](#footnote-3) of the concept of “shareholder value”, it is measured at each point in time simply by the prices quoted at the stock exchange. According to the so-called “efficient capital market hypothesis” (see on this, in a critical and popular vein, Fox 2011) markets, at each point in time, present the best judgment available. Since managers should act to the best of their knowledge this seems to imply that managers in pursuit of shareholder value should focus on increasing market capitalization of the company.[[4]](#footnote-4)

Pushing the preceding line of argument to the extreme, managers would, however, have to “slavishly” follow market signals on a short-term basis. If present share value is all that matters they can immediately observe whether according to the best publicly available judgment, the stock price, their actions were conducive to the aim of increasing the value of the shares or not. Assuming that markets would perfectly aggregate all information available to market participants share prices tell all that can be known (for a popular presentation of the more general “wisdom of crowds” argument, see Surowiecki 2005).

Obviously this argument could – if at all – only hold under appropriate information conditions and after full adjustment of markets to new information. If there is private information and if it takes time for the market to digest the information, if there are cascades of information (e.g. in the sense of Bikhchandani et al. 1992) and “delays” in response behavior – while the information sinks in –prices, even if they were always heading towards equilibrium might never express the equilibrium.

In general, with some inertia of responses, immediate reactions in the short term will not tell the full story of what the “market knows” once it has fully digested the information. In a dynamically changing world initial instantaneous reactions of markets to management decisions may not even directionally indicate what will emerge after markets fully absorbed the information.

**In short**, the hypothesis implicit in the market price is simply one hypothesis among many possible ones. Like individuals “**markets” can err**.

On well working markets errors in individual judgment tend to cancel out. Diverse private information is taken into account on markets and better than with any competing mechanism. Nevertheless **“management by tracking market prices”** seems rather foolish. In any event, the argument that markets generally “know better” than each market participant does not imply that it would be a priori irrational to bet against the market in any instance. The prior is that the market knows better but there may be specific information that justifies the updated posterior that a specific individual knows better.

For instance, a firm that is in command of private information may rationally go against market bets. From the point of the common weal it is in fact desirable that this happens (for example, not all stock market participants can rely on passively managed ETFs). It is like in science. If there is a prevailing opinion among scientists that has been formed in an open process of criticism and counter-criticism then this is inter-subjectively the best evidence that is available at any point in time. We trust in science because we understand the process of criticism and elimination of error present in the system as a whole. Yet, as scientific competition invites challenging any hypothesis, for the market process to work there must be people who challenge the common wisdom embodied in market prices. And, this is part of the reason why owners in particular shareholders of firms hire managers.

### 2.3 The long- and the short-term

Homo (o)economicus is willing and able to instantaneously go for what seems better at the point in time when a decision is made. He suffers from no behavioral inertia. He does what seems the best opportunity among what is on offer at any time. But he does so only after taking into account all his knowledge concerning the causal effects of his choices. And, most importantly, at any point in time homo (o)econmicus takes into account all information concerning the future. However distant an event in time may be once he gets to know of it he will instantaneously choose any opportunity that he might be able to detect in view of that knowledge.

Fully rational economic man is farsighted and strong willed, i.e. fully non-myopic. In this sense those who complain about the alleged short-termism of economic *rationality* are barking up the wrong tree. Nevertheless the ideal of fully rational homo (o)economicus is entirely misleading when it comes to deriving good management advice.

In particular when a manager encounters a realistically complex problem of human decision-making the prescription to take into account all future consequences and to go for long-term wealth maximization does not tell him what to do. First of all, it is practically impossible for cognitively limited decision-makers, like real human beings, to tell how long-term maximization translates into practical action here and now. The abstract prescription to maximize the expected value of a monetary stream of income over an indefinite time-horizon does not amount to a concrete prescription of what to do. Neither will it help a decision-maker with limited cognitive capacities to insist that at any point in time the full expected causal consequences of all acts *should* be taken into account. The decision-maker still does not know how.

We live in a “large world” in the sense that we often do not have a clue what the more distant causal consequences of our actions might be (see for an account of the revival of the large world problem in present formal decision theory elegantly and instructively Binmore 2009, and as a lucid presentation of modern decision theoretic explications of rational choice, Gilboa 2009, 2010). We often suffer from so called Knightian uncertainty. We either face ambiguity in that we do not know which from a class of probability functions concerning the states of the world applies or we face structural uncertainty because we even do not know the possible states of the world. Lumping together all the consequences that we cannot foresee together in one rubric “else” is a trick that has been used time and again. But it does not help. For using it makes it impossible to meaningfully assign a value of desirability to the rubric. What is in it remains as unknown as before putting a name on it.

The consequence of the preceding behavioral prescription for homo (o)economicus is obvious: If we cannot meaningfully assign a value (monetary or other) to the outcomes, the prescription to increase value is useless.

To the extent that we are unable to foresee consequences or cannot assign value to them, we have good reasons to decide in a non-consequentialist manner according to some procedural rules rather than by seizing apparent opportunities case-by-case. That is, we will have to act in an **oblique** manner by not directly going for foreseen consequences but rather by following rules (see Kay 2011, and generally Hayek 1973 ff.). These rules may be of a very abstract and unspecific kind like “if it ain’t broke, don’t fix it”, “never change a winning team”. If the rules have worked in the past that may be taken as indicating that they might work in the future. The rules and heuristics are a kind of social memory. If they survived past experiences, then this selective process stores information.

For example it seems that the long-term success of the car manufacturing company Mercedes Benz – as long as it was sticking to its original ways – was based on pursuing the aim of “building the best cars that can be built” (given the state of technology) at “fair” prices. Engineering and fair value concerns rather than profit motives were dominating the actions of the company.

Without going over details it seems obvious that Mercedes was foregoing many short-term profit opportunities. At the same time it was very profitable over the long haul. But was it so profitable over the long haul because or in spite of its high quality at fair prices strategy?

One *possible* explanation for the former success of Mercedes and many other companies is: By not intending to maximize profits but rather pursuing different aims, ends or values a company can as a matter-of-fact be rendered more profitable than it would be if going for profits directly. Note carefully that in the case of Mercedes obliquity or indirectness did not only rule out choice making in a case-by-case manner. The practices of the company implied that maximizing share value and/or profit were not acknowledged as decisive long-term aims at all. Yet the strategy was for a long time economically very successful.

The example of Mercedes seems to indicate that the old philosophical insight that many good things in life can only be reached if we do not strive for them directly applies in the greater realm of business life. It seems often the case that profits are increased by following norms that do not in any way refer to profit.

Now, the prescriptions and aims, ends, or values described in the stylized Mercedes case are not ethical norms. But it is rather obvious that the argument concerning non-ethical prescriptions could in principle be extended to ethical ones. There seem to be good empirical reasons to believe that **doing well by doing good** is not an exotic possibility. Ethics may be economic “value generating”.

### 2.4 Obliquity

### 2.4.1 Obliquity in ethics

The Greeks of classical antiquity understood that it is often in the self-interest of an individual to act in intrinsically motivated ways that deviate from the opportunistic pursuit of extrinsic rewards. As far as instant gratification by material rewards was concerned they were “**deontologists**”, i.e. were advocating the view that what is ethically right or wrong is not exclusively determined by its non-ethical, material consequences. Nevertheless, other than many modern deontologists the Greeks were typically indirect “**consequentialists**”**.** To acknowledge that an alleged virtue was a virtue it had to be shown that the actor would *ultimately* do “well” by doing “good” in following *proximate* ethical rules. To cultivate intrinsic motives that do not directly aim at proximate external rewards was justified in terms of long-term rewards.

**In short**, in an indirect consequentialist ethics there is room for obligations that are directly (proximately) independent of consequences but to impose ethical constraints must indirectly (ultimately) be in the interest of those who are subjecting themselves to them.

What the Greek intellectuals believed to be true for their private lives may be true for the corporate actor’s modern business life, too. It seems plausible that more often than not imposing “internal ethical constraints on opportunism” and “short-termism” will serve a company well. To the extent that this is true, there is no clash between ethics and interest. Quite to the contrary **ethical codes** and the “virtuous” behavior brought about by their acceptance are **instrumental to the pursuit of long-term interests of the company**.

There is not only “experiental” (i.e. experience based) knowledge as in the Mercedes case. There are even experiments that corroborate that those who seek long-term success better do so by not going for short-term success.

### 2.4.2 Obliquity and indirectness of success in an experiment

There is experimental evidence that seems to support the preceding view. Reinhard Selten and his collaborators let individuals play a management game in the laboratory (see Hohnisch et al. 2011/2012). Subjects were paid according to their long-term success or wealth at the end of the game. The game went on over many periods. It posed a complicated interaction problem that could not be solved analytically by participants.[[5]](#footnote-5) Those who were most successful started with a kind of trial and error period and then adopted a more or less *fixed* set of variables. Choosing among several options like investing into advertising, increasing reputation for quality etc. they had to find a mix that would fulfill their aspirations.

The most interesting result of the experiment was that those who were not willing to commit to such a set of variables were not doing as well as those who pursued their **ultimate** ends more indirectly by not searching restlessly for opportunities to improve the situation. Subjects could not translate the long-term aim of maximizing wealth into a short run profit maximization prescription. They knew that they would be paid at the end of the game according to the wealth accumulated during the whole game but that did not tell them anything about how to find wealth maximizing strategies in each period of time. On the **proximate** level of motivation they had to pursue aims other than wealth maximization. They knew, of course, that wealth would accumulate by making in-period profits. However, that knowledge would not tell them what to do. In particular they were not told when to go for the intermediate opportunity an when not. Sometimes direct profit would have to be forgone as an investment to bring in future profits. There was too much interdependence among periods and the decisions made in them to rely on the short-term signal of each period taken separately.

### 2.4.3 Obliquity in business

Pursuing such an aim like building a reputation for quality is very to an individual’s effort to establish a reputation for sticking to rules of honest dealings. For the present context it is important to note that this may require that the corporate actor cultivates the virtues of its employees. The reputation for fairness in dealings with customers may partly be fixed in company rules dictating fair dealings. But it may and as a rule will also require ethically motivated employees who forgo direct advantages in favor of following moral precepts and long-term prospects (as in the case of doing another costly quality check) when they have to make individual judgments on behalf of the firm.

If the previous remarks on (efficient) markets are correct, taking into account the market evaluations will not provide a sufficient criterion to tell right from wrong and to single out management errors. This cuts both ways of course. Even if markets would evaluate a new **ethics strategy for the company** very positively that need not be compelling evidence that the management got it right concerning the long-term. Or vice versa, if the markets respond negatively to the announcement of some form of ethical orientation or other by the management of a company that need not be compelling evidence that the management got it wrong with respect to the long-term prospects of the company.

The management should be able to show that it has a plausible case “against the markets”. Like in all other cases of management decisions the possibility of error has to be taken into account when adopting certain strategies. Strategy choices that are motivated by so-called ethical concerns are no exception.

Good and “responsible” management will try to create a mechanism by which it can gather evidence. According to *ethical standards* adopted in this short text, management should engage so-called evidence-based whenever possible at acceptable costs (though possibilities may be rather restricted). It will be aware at the same time that there is still insufficient evidence concerning the contributions of ethical strategies to profitability.

Before the evidence is out (or in) we simply cannot know whether pursuing some kind of ethical management strategy is good business or not. We can, however, confidently say that giving up some amount of short-term profit need not necessarily be bad business in terms of long-term wealth maximization. It may be an investment like other investments. Costs are incurred since the company bets that the effect of the investment will be beneficial over the long haul.

A homo (o)economicus who could indeed invest in choosing his own utility function would want one with a conscience (see Frank 1981, Güth and Kliemt, 2000b). Under certain conditions the slogan “doing well by doing good” will apply. To the extent that this is true many basic conflicts identified by business ethics are less pressing. As shall be shown next business ethics’ “defining” controversy between shareholder and stakeholder approaches is a case in point. For instance, if virtue and ethical commitments serve the long-term interests of the company, an adherent of a shareholder first approach like Milton Friedman would endorse the corresponding strategies in the same way as those who justify them directly in terms of interests of stakeholders other than the shareholders.

## 3. The fundamental divide in business ethics: Friedman vs. Freeman*[[6]](#footnote-6)*

According to Friedman’s *moral* position, it is ethically illegitimate to use the funds provided by shareholders for purposes other than increasing shareholder value. The **fiduciary relation to shareholders** requires this. Why this relationship should have precedence over the relationship to other stakeholders of the firm and why it does imply an obligation to further the long-tem interests of shareholders has to be shown, however. It is contested whether there is a sound argument in favor of the “**shareholder interests first” conception of legitimate management**.

Contrary to folklore, Friedman was not endorsing “ruthless” capitalism as a value in itself. Friedman rather endorsed capitalism and a shareholder view because he believed in respecting individual autonomy and the moral duty to respect agreements among individuals. Respecting contracts with owners left no room for a position other than pre-ordering shareholder interests over those of other stakeholders in management decisions.

Other theorists like in particular Edward Freeman (1984) believed that values of other stakeholders are relevant not only for instrumental reasons (or justified “ultimately” as conducive to shareholder interests). The company is to be framed by its management as a contract between *all* stakeholders without special priority for contractual relations with and in the shareholder subgroup.

The controversy centering around shareholder and stakeholder interests is of such fundamental relevance for the practical ethics of business that it deserves to be addressed separately.

### 3.1 The responsibility of management and the diversity of values

If anyone utters the battle cry of the moralists that there should be “more ethics!” most people will concur. All are in favor of more ethics yet with respect to substantive ethical precepts people are typically of very different opinions. They all want “more of it” but the “it” is not the same for each of them. As emphasized above already, ethical views are not in interpersonal agreement but rather in disagreement. More emphasis on ethics may in fact exacerbate the disagreements.

More concretely, the manager will be asked to do x according to the values of stakeholder group A while according to the values endorsed by group B the action should be y. Typically y and x will be incompatible with each other. A coherent ethical view cannot contain both x and y, and a policy cannot realize them both. A manager cannot be loyal to all stakeholders by following one group. Which of the stakeholder groups should she let have its way, then? Even if she would restrict her loyalty to shareholders only, she would have to cope with the possibility that there might be a y-group and an x-group among shareholders.

To see a way out of the problem, assume you are a manager who is hired by a single entrepreneur who owns the firm. The contract with the entrepreneur creates a moral obligation to pursue the aims, ends or values of the principal within the limits of legitimate law. Therefore, if the principal’s aim is environmental protection, protecting the environment is what you should do. If it is protecting jobs then this should be what you have to aim for according to your promise to ratify in your own judgments the aims, ends or values of your principal. If the entrepreneur wants to make profit then you should help him in this but only then.

Nobody has forced you to enter an employment contract to serve as manager of the single owner of the firm. Since you engaged the **contractual obligations to act as an agent** by contracting into a fiduciary relationship voluntarily you are at least prima facie morally bound to them within the limits of legitimate law.

Now, assume that you are hired as an agent of a large company with a wide circulation of shares. As in the case of the company owned by a single entrepreneur a fiduciary obligation emerges. From your contract derives the moral obligation to pursue the aims, ends or values of the principals within the limits of legitimate law. However, now your principals will pursue a plurality of aims, ends or values. As a respectful manager who intends not to impose your own ethical values on others you have many masters to serve. Serving them equally is possible by one measure only: you have to increase the value of the exit option for shareholders who do not like your policies. Would you go for something other than what increases shareholder value then you would indirectly impose your own values on shareholders at the expense of a less valuable exit option for those shareholders who are not sharing your values.

If the manager would follow some ethical theory rather than his own views this would not change the basic ethical violation of the respect norm. The manager would still go for some values at the expense of the values of at least some of his shareholders. *An altruistic manager is imposing his own values as much as a selfish one*.

**In short**, altruism does not solve the problem of a manager’s inability to respect conflicting values simultaneously.

If the manager adopts an ethical theory that claims to express some kind of moral truth, this will not alter the situation either. That theory or the actions suggested by it may still clash with the aims, ends or values of other shareholders.

**In short**, taking recourse to ethical theories does not solve the problem of our inability to respect conflicting values simultaneously (since theories diverge).

For example, imagine a manager who would on behalf of, say, a conception of “social justice” that is shared by some stakeholders – possibly including some shareholders – keep employees on board at the cost of a decrease in shareholder value (and his own bonus). This manager acts altruistically and on behalf of ethical principles but at the same time does not respect competing ethical views of other shareholders (and, for that matter, other stakeholders) who do not accept the same principles of justice as he or accept no principles of social justice at all.

The manager can be acting only on behalf of those shareholders who share the ethical theory of social justice he applied. There may be even some who not only think that the move of keeping employees on board would be right but also would be willing to pay a premium in terms of reduced share value for this action. The action would still be illegitimate in the eyes of Friedman. For, the action furthering “social justice” would be akin to a tax levied on behalf of collective public welfare. Friedman insists that levying taxes is something only a government that can be held politically responsible for its actions should do. If a person who acts as a fiduciary of other individuals uses their funds for the ethical purposes he or others deem ethically acceptable, that person imposes her will on others like a state but without the legitimacy and authority of the state. That it is for an ethical purpose does not change the fact that the action is imposed.

Friedman’s emphasis on shareholder value would obviously be right within any political ethics that accepts some privilege of private property. To the extent that we regard it politically as legitimate that individuals may decide as seems fit to them in some sphere then there is a legal protection for a private sphere. The private sphere is politically set free of political intrusion (see on this most basic political decision of any society, Kukathas 2003). Yet to define the fiduciary relationship between manager and shareholder as private does not as such rule out to pursue aims other than increasing share value. Given the oblique nature of all human pursuits (see section 2.3 above) we will almost of necessity have to pursue the aim of increasing the value of shares indirectly by pursuing other ends. But then the shareholder first orientation of legitimate management loses much of its bite. It may well be that pursuing aims other than increasing share(holder) value on the proximate level ultimately serves the interests of shareholders by increasing the value of shares over the long haul.

### 3.2 Transparency of operations as most fundamental requirement of “responsible” management

On a liquid market for shares those who oppose an ethical move against increasing shareholder value could, of course, sell their shares immediately. In that sense an action of the management would in all likelihood hurt them only marginally. And, marginal reductions in value by the trial and error process of firm policy making will as a matter of fact always occur. If we do not believe that short-term shareholder value tells the whole story then shareholders will have to put up with actions that are triggered by motives other than increasing shareholder value (yet may indirectly further the interests of shareholders).

An ethically sensitive management that is aware of value pluralism must make sure, that it gives shareholders the chance to **exit from the contractual nexus** of the firm without (major) loss. Then, as long as the management is working on the basis of openly and transparently declared policies, shareholders can always sell. If they do not sell the shareholders thereby “reveal” a preference for the firm policies as compared to their alternatives on the market. As long as easy exit remains possible the management can do almost anything without violating its fiduciary obligations if it sees to it that information would reach the markets in ways that will give those who would like to a fair opportunity to exit the contract nexus.

The management could even **openly declare the strategic pursuit of certain kinds of aims** – seemingly or truly – **detrimental to share value**. There may be “people” on the market who are willing to pay a premium in terms of depressed share values for such policies (to the extent that this willingness will not lead to increases in share value). In the same vein, an openly declared policy of endorsing a Freeman type stakeholder approach that grants shareholders no special, privileged role among the stakeholders, cannot be criticized by Friedman and his followers to the extent that exiting is more or less costless.

**In short,** liquid capital and in particular stock markets assure that owners of shares can exit at any time. If they do not exit in view of transparent management policies then it is their own responsibility. As long as the principles that are guiding managerial decision-making are openly declared, managers can perhaps be criticized for choosing the wrong principles but not for violating their ethical obligations.

Actions performed in pursuit of aims, ends or values openly declared beforehand can hardly amount to a breach of a fiduciary relationship standing under the general proviso of the exit option. The pursuit of a stakeholder approach in policies of a company is almost completely compatible with a shareholder first approach. Contrary to common perception **shareholder and stakeholder views** in practical business ethics do **not** seem to be **separated by a great divide**. It rather seems that practically all ethical concerns expressed in a stakeholder approach can be pursued in ways compatible with the fundamental *moral* values of the shareholder approach.

To the extent that pursuing the aim of respecting the interests of stakeholders other than the shareholders is simply good business the Friedman and the Freeman positions coincide. But even in cases in which they do not coincide, because measures turn out to be detrimental to share value, as long as plans had transparently be declared well beforehand, the owners who stick to their shares are not wronged.

**A glimpse of more advanced literature and topics**

The question why one has a good reason to *be* rather than merely to *appear* as if morally motivated leads beyond the present text. It can be addressed against the background of the following literature: Ainslee, George. *Break Down of the Will*. Princeton: Princeton University Press, 2002. Frank, Robert H. (1987), ‘If Homo Economicus Would Choose His Own Utility Function, Would He Want One with an Conscience?’, *American Economic Review*, 77(4, September), 593-604. Frank, Robert (1988), *The Passions within Reason: Prisoner's Dilemmas and the Strategic Role of the Emotions*, W. W. Norton.: New York. Güth, Werner and Kliemt, Hartmut (2000), ‘Evolutionarily Stable Co-operative Commitments’, *Theory and Decision*, 49, 197–221.

**Part II: The economics and ethics of** interaction

As was indicated in the section before, under adequate institutional regulations the overlap between striving for long-term success and what is dictated by certain ethical codes of good conduct may be quite large. If the rules of the market game are appropriately fixed then those committed to act on the proximate level of action more often than not will ultimately be better off than the uncommitted. “Virtue” though not necessarily its own reward, may over the long haul give its possessors a competitive edge over those who are not in command of it.

Contrary to common folklore the chances for virtue to survive are in all likelihood better in a free contract society than in a more primitive small (see on this Henrich et al. 2004) or large command society, say of the former socialist type (see on large markets and virtue Baurmann, 2002). There is, of course the old criticism that on markets everything may be on sale including things that should not be for sale (see Sandel 2012). However, if what one group would like to trade another will not want to be up for sale, who is to say what may and what may not be traded. Once we agree that we practically never agree on these matters collectively, why not let those potentially involved in such trades decide on them freely? Why not respect the free contracting of consenting adults whether it be in finding mutually agreeable sexual relationships, exchange of goods or financial transactions?

**In short**, allowing for free contracting, **markets** are an expression of **inter-individual respect** for the **autonomy of persons**.

Moreover, by coordinating human action and allowing for the division of labor by facilitating the exchange of the fruits of specialized labor markets contribute to the common weal. They are wealth creating institutions that typically outperform their institutional competitors. To the extent that they allow for institutions like firms operating within the market context competition can be extended to all sorts of intra-organizational rules (see, above I.2). Still, there clearly may be cases of **conflict between** what **ethics** requires **and** what **market incentives** dictate as proximate motives. Moreover, as many philosophers and laypersons tend to argue, the modern **ubiquity of market relations** of monetary incentives and rewards may crowd out moral commitments and virtues. A stylized model of market interaction may be helpful in getting a better understanding of what is ethically and economically at stake here.

## 1. The traditional argument vindicated by Vernon Smith‘s Double Auctions

In seminal early studies of experimental institutional economics Nobel laureate Vernon Smith showed that “markets work” (see Smith 1962, Smith and Williams, 1992). In his approach Vernon Smith was inspired by older experiments that had been conducted in his presence some years earlier in a class at Harvard by a critic of markets, Edward H. Chamberlin (1948). Chamberlin’s experiment was rather a “classroom simulation” of market behavior than a market. Though initially Vernon Smith referred to his experiments as “simulations” of real world market contracting and exchange he created a market-like institution in the laboratory. The actors he invited to participate in a laboratory experiment were earning real money and he meticulously specified the “rules of the game” (in particular the information conditions) such that there was a *laboratory market*. The following is not a faithful account of the original experiments but one that emphasizes certain aspects.

### 1.1 A stylized sketch of experiments

A relevant class of experiments can be easily characterized:

30 participants are invited to the laboratory.

15 participants are randomly assigned to the role of seller,

15 participants randomly are assigned to the role of buyer.

Each of the **sellers** gets a yellow card with a dollar amount on it. That figure, the “reservation price”, indicates how much the seller must receive to avoid losses. Should the seller realize a higher sale price than his “reservation price” (the “costs” at which he would break even), then the difference between the sale price and the reservation price is his.

The 15 individuals who are assigned to the role of **buyers** have red cards in their hands. The dollar amount noted on the red cards indicates how much they can pay at most before incurring a loss. Again if they buy cheaper than this, the difference is theirs.

For the sake of specificity and simplicity assume finally that the **amounts noted on the cards** of the sellers are in ascending order 1, 2, …, 15, respectively while the buyers have on their cards in descending order 17, 16,…,3. (see figure 1 for illustration).

In figure 1 all “infra-marginal” traders to the left of the two marginal traders should trade at the uniform price of nine – *regardless with whom specifically* – while all “extra-marginal” traders to the right should be excluded from trading if collective gains or “rents” from trade are to be maximized. In Chamberlin’s experiment **“rent-maximization”** – or making the green area indicated in the graph below maximal – did not prevail. He let participants approach each other in the seminar room separately and negotiate with each other without asks and bids becoming known to the others. Under these conditions, in which the price information remained strictly private among those who negotiated, the market did not work. Vernon Smith felt that this amounted to loading the experiment against the efficient market outcome. Recognizing that markets are basically auctions he changed the rules towards more reasonable ones. He created market institutions that will in all likelihood tend towards the “rent-maximizing” results associated with uniform equilibrium prices.

Imagine our 30 participants with there red and green cards are brought together in one room. They can call out prices at which they are willing to buy and prices at which they are willing to sell loudly. All utterances become *common knowledge* as public events. There are several rounds of bidding. Say, somebody will announce to *buy* at a price of $3. Assume that nobody will accept the **bid**. Then another participant might call out and state that she would *sell* at $12. There may be again nobody who would accept what she **asks** for. But there may be somebody else who would, say, bid to *buy* at $5 etc.

If such a “**double oral auction**” process is repeated it will “converge” towards the market-clearing price at which the *community* of traders realizes the greatest possible – “collective welfare maximizing” – benefit from trade in mutual agreement. In fact, after an initial round of play – a series of asks and bids – in which selling might take place somewhere near the “market clearing” price, but not exactly at that price, exchanges will take place almost always at equilibrium and rent-maximizing prices. This laboratory result seems to indicate that markets do work even if – as in the laboratory – the information conditions are not perfect.

**In short**, provided that appropriate institutional rules are implemented the introduction of market institutions will add the maximum (monetary) value that can be realized by social coordination (for more on the robustness of market efficiency, see Gode and Sunder, 1997).



Figure II.1

To the extent that markets empirically work in the described way this is *ethically* most relevant. Market organization of social co-operation seems to satisfy both the ethical ideal of increasing collective welfare and of respecting individual autonomy. Within the institutional market framework it is left to their free inter-individual agreements what people coordinate to do. Which collective result emerges is not planned and directed by collectively binding commands. Yet the result that can be predicted to emerge from autonomous individual choices under appropriate rules and conditions will realize the maximum possible value that can be added to the wealth of the community without violating interpersonal respect of individual autonomy. (The same could be achieved by a market-making clearing-house which arranges trades according to the asks and bids.)

**In short,** the market seems to be an ethically ideal world making society richer without collectively forcing individuals into a specific behavioral mold.

However, at closer inspection and after factoring in distributional concerns, one may have some second thoughts (see recently Piketty 2014 with a popular data-backed assault on markets based on their alleged inequity). To see which kinds of doubts may arise look at the following figure 2 (see also Stodder 1998)



Figure II. 2

In this figure the seller who has to ask at least for one dollar will ***exogenously* be paired** with the buyer who is willing to bid three dollars at most. The seller who has to ask for at least two dollars is paired with the buyer who is willing to bid up to four dollars and so on until the buyer who is willing to bid up to $17 is paired with the seller who is asking for at least $15. Assuming that buyer and seller in each of the pairs will always evenly split the difference each of the individuals will gain exactly one dollar in each of the trades. In all, the gain will be $30 and it will be distributed perfectly equally.

This is not the collective-rent maximizing gain since in the preceding figure II.1 the gain from the two first trades amounts already to $ 30. According to figure II.1 the gain from implementing the double oral auction would be $62 while who trades with whom is *endogenously* determined in free agreements among participants.

**In sum,** the ***ethical* argument for “free markets”** is based on their efficiency and on the fact that markets let “consenting adults” figure out with whom to trade and with whom not at what terms. The ***ethical* argument against “free markets”** is the inequality of outcomes they generate.

So-called “**free-marketeers**” who like Milton Freedman and most neo-classical economists (and the present author) endorse the ethical ideal of respecting individual autonomy found it always ethically compelling that inter-individual agreement figures very highly. Yet, most philosophical ethicists and most members even of market societies are unconvinced. They are critical of the distribution of “gains from trade” that emerges from “free contracting”. The “**blind workings of market institutions**” can lead to results they deem unjust. In particular they raise the objection that in the rent maximizing trades *not all* individuals get a share and certainly not an equal share of the gains from trade and social co-operation.

**In short,** there is fundamental disagreement on the ethical standing of markets.

### 1.2 Normative aspects illustrated by experimental markets

Vernon Smith could have fixed the market rules in the laboratory differently. The auctions could have taken place in quite different ways. Bidding rules could have been fixed in ways guaranteeing equal gains according to bids (see Güth and Kliemt 2013). But all sorts of less extreme ways to fix auction or market rules in favor of more egalitarian outcomes are conceivable and could be studied in the laboratory.

More generally speaking, the outcomes of market interaction depend on market institutions that can be fixed by society in (regulatory) rule choices (see on the distinction of **rule vs. within rule choices** for instance Brennan and Buchanan 1985/1999). Once the state enters the “game” as an enforcer of last resort it can choose which kinds of rules to enforce and which not. Market rules do not drop from heaven. Laws governing market interaction are not “laws of nature”. Since government can act as a “market maker” why not let it fix rules according to ethical criteria that take into account the distributional consequences of free contracting on markets?

To give a concrete example, the **frequency auctions** in the telecom sector of recent years followed rules fixed by governments. Governments could auction off frequencies according to their own rules since they claimed a public monopoly on the use of frequencies. As monopolists, governments tried to capitalize on the resource they controlled. Some auctions (markets made by governments) were better at raising money for the government than others. Since the main concern in these contexts has been to maximize yields for the public purse this mattered most to governments. Whether the market rules were complying with certain other normative standards did not matter much to them.

Nevertheless, since there were alternative ways to fix the market rules, aims other than maximizing yields could have been pursued and the rules could have been chosen accordingly. Since in the frequency auctions the bidders were large privately owned companies, nobody complained about the neglect of distributional issues and any inequity of results. Governments behaved as “yield maximizers”. They tested alternative auction rules in the laboratory and then fixed the rules such as to extract maximum payments for the public purse. They behaved as rent-maximizers – and in the case of frequency auctions presumably rightly so. That is not to say, however, that governments must behave that way.

Governments can influence the rules that they enforce in assisting citizens to reach mutually advantageous trades. They can include certain distributional concerns as constraints when specifying laws and other regulations. In doing so they express a political view on the ethical “**equity-efficiency trade-off**”.

To illustrate again, if one compares figure II.1 and figure II.2 the trade-off between efficiency in the double oral auction market case, II.1, and equity or equality in the “equal split pairing” case, II.2, is obvious. Many people believe that regulators should influence the market rules in ways that make sure that all individuals share the fruits of social co-operation in an “equitable and just” manner without impairing allocative efficiency “too much”.

It is instructive to take a look at examples of phenomena that arise within the sphere protected by the legal monopoly of the state without being directly regulated by the state. Firms and other private institutions can be authorized to create rules within publicly enforced so-called civil or private law (see also Benson 1990). Even if it is centrally enforcing certain rules and selects contracts, the state is well advised to leave room for private (contr)actors to do so.

## 2. Markets, Firms and Guangxi

As is well-known, Adam Smith related progress in society to progress in the **division of labor**. The extent of the market is crucial for how far the division of labor can be pushed. If the markets get larger, in particular due to a reduction in the cost of performing transactions, then new forms of specialization and new firms can emerge. Firms are – at least in a way – not only operating on markets they are also substitutes for market organizations.

### 2.1 Firms on markets

The costs of performing transactions come about in many ways. The costs of transporting some merchandise physically from point A to point B are one factor that constrains the division of labor and the reach of markets. When transport was very expensive still, the goods that could be traded had to be very valuable per unit of weight. It took until the development of steam ships before market trade and the division of labor enabled by it could be extended to more bulky goods of lesser value per unit of weight. Then in the later 19th century the first wave of globalization was under full steam (under full sail it was not anymore).

Transactions are also accompanied by costs of communication and, perhaps most importantly, **costs of controlling and monitoring**. If contracts could be specified completely and perfectly at zero costs, organizing economic activity would be quite different from our real world in which the costs of specification, monitoring and enforcement of contracts can be rather high. Due to such costs it may not be worthwhile to perform certain transactions that otherwise might be rather attractive.

Quite generally speaking in a world with well **working “strong” legal institutions** we should expect forms of coordination other than in a world with “weak” legal institutions. How existing institutions work is decisive for answering the question whether division of labor and specialization should be organized across markets or within a firm or company. Also, whether or not ongoing forms of cooperation with the same trading/co-operation partners should prevail over on-spot transactions with changing partners depends on the quality and form of the legal institutions. If on-spot contracts are easily and well enforceable then traders should keep open to always go for the best ask or bid rather than commit to preferred trading partners and permanent relations.

Besides the preceding aspects that relate to extrinsic motivation primarily, there are aspects of transactions that are internal to the trading partners. In an environment in which people are more or less guided by **internalized moral or ethical codes of conduct**, other forms of organizing social transactions are viable than in a world of ethically unrestrained opportunistic behavior. It is possible that legal enforcement of norms can be substituted by informal enforcement of norms if actors can commit without a formal contract enforced by the legal system. And, this rather commonsensical insight – though not necessarily one appealing to economic common sense – can assign a role to ethical codes that is most relevant in a business ethics and economics context.

*In fact, one of the reasons why* ***the organization of the******division of labor takes place within firms*** *or companies derives from the comparative advantage that coordinating behavior by informal norms of a broadly speaking “ethical nature” can have over legally enforced “formal coordination”.*

Informal enforcement will typically work well only under conditions of repeated interaction with the same partners in small group settings. In such settings **the future of interactions with the same partner casts its shadow on the present**. It is in the interests of the actors not to become subject to retributive acts of their partners in later interactions. If they are tempted to exploit them on an earlier round of interaction they risk retribution on later rounds of interaction.

The past influences behavior since we tend to develop certain attachments to those with whom we frequently interact. Typically these feelings as well as internalized norms and personal sympathies will work in the same direction as mere interests. **Doing well by doing good** will operate on the individual and on the level of firms.

In line with the preceding discussion (see I.2.4) the economic logic of mutual advantage is lurking in the background as an ultimate source of cooperation even though our proximate conscious motives are of a different kind. It is in our long-run interest that we play by the norms that facilitate cooperation and coordination. However, on the proximate level we do not take into account ultimate interests explicitly but rather use the norms directly as guidance for our behavior. And even in cases in which we are aware that our interests are directly affected by our actions and how they are affected we nevertheless are typically *also* guided directly by norms as accepted standards.

***In short****, the firm provides a natural platform for in a wide sense* ***norm-guided or morally motivated behavior****. Other than a large formal market it is a way to organize crucial aspects of interaction in permanently interacting small groups. Such modern* ***moral tribes*** *can make the best use of natural human organizational faculties.*

It is one of the great advantages of repeated interaction in small groups that compliance with normative standards, that constrain short-term opportunism, is conducive to long-term interests of group members in a rather direct way. In the small permanently interacting group the effects of actions on others are significant to them. They will be able to show retributive behavior targeted at specific other behavior. The observation of norms of cooperation and the coordination of economic behavior can be sustained in such circumstances without specialized organizations like courts etc.

Yet firms are not the only way of bringing the human tribal heritage back into the world of civilized inter-course. Formal contract enforcement can also be substituted by informal networks and will be substituted this way to the extent that formal mechanisms are not working well or are too costly.

### 2.2 Guangxi[[7]](#footnote-7) and substitutes for markets

Legal institutions of so-called “**private law**” in China are weak. Formal enforcement of contracts through courts typically is either impossible or too expensive to be worthwhile. Therefore, many transactions that could be performed across markets in other legal environments have to be organized either internal to firms or by forming long-run relationships with specific partners, “**Guangxi**”.

Broadly sketched Guangxi means a long-term relationship that is started by producing the sunk cost of mutual gifts of considerable value. The process is different from bribery and corruption in that the gifts are used not directly for getting a specific favor but rather for the generalized favor of entering a long-term relationship. In particular, there are trading networks to which outsiders can find access only after spending considerable sums of money, time and/or other resources etc. The whole point of the networks is to provide the faculty to internalize external effects of action where neither specific contracts across firms nor generalized contracts internal to firms can accomplish this.

Not by accident, Aristotle has classified certain permanent inter-personal relations of a commercial nature as **commercial friendships**. Such kinds of “friendship” depend on virtues and are guided by ethical codes rather than legal ones. In Guangxi ethics – as a system of social codes and regularities – if embedded in supporting networks can serve the same functions as formal contracting. It can do so because separating a network of interaction partners from the larger social network can make coordination viable even when central enforcement of contracts is lacking.[[8]](#footnote-8)

Even in societies with well working centrally enforced formal contracting through a non-partisan legal staff, practices akin to Guangxi are still existing but to a lesser extent. These practices all rely on **reducing the number of potential trading partners** to such a small number that behavior remains individually significant. Only in sufficiently small groups the “shadow of the future” likelihood to meet any individual again becomes dark enough to endogenously enforce norm compliant behavior. What is ethically desirable coincides broadly with self-interest in an ongoing relation only if the partners – like employees of a firm – remain associated. *Doing well by doing good* can again apply under such conditions.

Before turning to examples like eBay where the future casts a shadow according to elaborate reputation mechanisms let us look at the bare bones of the challenge to which ethical codes as well as the legal order and the institutions related to them provide an answer. First in a familiar two by two Crusoe-Friday economy and then more generally.

## 3. Getting in and out of social traps

### 3.1 Two person exchange game and the hidden side of the invisible hand*[[9]](#footnote-9)*

Imagine person A has a spare quantity of apples “a”, person B has a spare quantity of bananas “b”. A is informed of the availability of b and B is informed of the availability of a. A would prefer to have b rather than a and B would prefer to have a rather than b. If A co-operates with the wish of B he sends his apples, acting CA, and if B co-operates with the wishes of A she sends, acting CB, bananas b to A. Likewise non-co-operation or sending nothing is indicated by no-apples⬄DA, no-bananas⬄DB, respectively.

Assume that A and B chat over the phone and agree to mail off their spare fruits on condition that the other does so as well. The following table describes the possible results of choosing their options. Bear in mind when looking at the table that it must be read to say that the options are chosen independently of each other. None can make her choice conditional on that of the other none can trigger that of the other.

|  |  |  |
| --- | --- | --- |
| BA | CB | DB |
| CA | b , a | 0, a&b |
| DA | a&b, 0 | a, b  |

Table II.3 Prisoner’s Dilemma or Exchange (payoff A, payoff B)

As can be seen immediately whatever the other actor does it is always better not to send anything, i.e. to choose DA, DB respectively. Yet if both do what is individually rational for them they end up in the status quo that is with a result combination (payoff A, payoff B)=(a, b). This is worse for each of them than the exchange payoffs (payoff A, payoff B)=(b, a). Ethics that would motivate them intrinsically to comply with the wishes and interests of the other would make them choose CA, CB respectively and make tem both better off in their material payoffs. Yet being fully rational they are trapped in (DA, DB). They are not guided as if led by an invisible hand to the superior (CA, CB) outcome that would make both of them strictly better off.

**In short**, in an exchange dilemma situation individuals are trapped by their individual rationality in a dominant strategy equilibrium to which another result of interaction exists that would make both participants better off. Both have a rational reason to wish that the other outcome would come about but no reason to act accordingly.

### 3.2 N-person interactions

Though many encounters of our social life are two-by-two – i.e. there are but two actors with two options each – like the preceding, others are not. There can be more than two actions from which to choose. Another essential generalization emerges if there are more than two actors. In that case, a simple matrix representation might not do, yet there are ways to get around that problem and to represent the interaction of many in matrix form as well. It is possible in particular if the game is “symmetric” in the sense that for every individual r from a set of N+1 individuals, the game played with the N *other* players looks exactly the same. We must then just let one arbitrary, though specific, individual play as “representative agent” for all others.

### 3.2.1 N-person PD

Consider the following two-by-N table for a game played among N+1, N+1>2, individuals in which the N+1-th individual plays as a row player and the others serve as column players. It is only shown what the payoff to the one representative individual will be since all others symmetrically have to decide as she does (the presentation follows Buchanan, 1965).

|  |  |  |  |
| --- | --- | --- | --- |
| Number of other players cooperatingPlayer N+1 | 0 other Co-operators | N/2 otherCooperators | N otherCooperators |
| CN+1 | a | m | y |
| DN+1 | b | n | z |

Table II.4: N+1-person prisoner’s dilemma game form, N+1-PD

Assume that the interaction is characterized by monetary payoffs that fulfill z>y>n>m>b>a for each and every individual. Moreover, let us also assume that generally for all numbers of other cooperators the C is always smaller than the D payoff and also fulfilling the other ordering conditions analogously (see more precisely “VCM” below). Assume also that the individuals are only interested in gaining a monetary payoff from the set that is as high as possible.

Observe that all would be better off in a state of universal co-operation – a state in which everybody would receive y – than in a state of universal non-co-operation in which everybody would receive b since b<y. Assume also, that for any number of other co-operators in-between, i.e. 1 to N-1 others cooperating, it is always better for each player not to co-operate herself rather than to co-operate. Therefore, non-co-operative choices form a dominant strategy for each and every individual. Since it is not assumed that actions of any individual can influence the choices of any of the others, it is obvious that all individuals should plan on choosing the defection alternative. The outcome of this is a so-called “dominant strategy equilibrium” – an equilibrium in dominant strategies – that is “Pareto dominated” in payoffs – i.e., for which there exists another alternative strategy combination making every individual better off.

**In sum**, in the game, all will rationally try to be “**free-riders**” on the efforts of the others; at the same time, all would be better off should nobody take a free-ride. Still free-riding is a dominant strategy.

The story to be told here cannot be told better than by David Hume and so, instead of telling the story myself, I let the greatest of all empiristic philosophers speak (Hume, 1739/1978, book 3, chap. 7):

Two neighbours may agree to drain a meadow, which they possess in common; because 'tis easy for them to know each others mind; and each must perceive, that the immediate consequence of his failing in his part, is, the abandoning the whole project. But 'tis very difficult, and indeed impossible, that a thousand persons shou'd agree in any such action; it being difficult for them to concert so complicated a design, and still more difficult for them to execute it; while each seeks a pretext to free himself of the trouble and expence, and wou'd lay the whole burden on others.

As we all understand, this is the paradigm situation of a so-called “**collective good problem**” in a large group.[[10]](#footnote-10) It is well known to us from our experience with such matters as environmental protection. Even if all intend to do good, they all would have the suspicion that they might become “suckers” or victims of the “lawlessness” of others. They may demand some kind of “assurance” here to get to a state of affairs in which almost all co-operate.

There may be saints and heroes (see Urmson 1958) in our world who, in situations like the N-PD, might want to co-operate regardless. There may be people who on rare occasions may be willing to co-operate even if large numbers of other potentially free-riding individuals are involved while their own acts are as **insignificant for the collectively perceived result** as a single grain of sand on a heap. However, in particular if we are participating in such interactions again and again, most of us will expect that co-operation will deteriorate. As shown in the field as well as in laboratory experiments co-operation will tend to wash out over the long haul (see below also VCM).[[11]](#footnote-11)

### 3.2.2 N-person Volunteer’s Dilemma

Consider a game in which each of N+1 individuals can provide a collective good all on her own at fixed cost c>0. One individual is sufficient. If none does what is required, then the result is bad for all, for it yields “0.” If one volunteers, this individual expects U-c, while all others, if they can let her do the job alone, would get U. Assume U>U-c>0. It would indeed be pure waste to have more than one individual spend c, yet if there is no way to determine who is going to be the volunteer, for example, by lot, which would amount to adding an additional option to the interaction, then each might speculate on somebody else providing the collective good.

|  |  |  |  |
| --- | --- | --- | --- |
| Number of other players cooperatingPlayer N+1 | 0 otherCooperators | About N/2 otherCooperators | N otherCooperators |
| CN+1 | U-c | U-c | U-c |
| DN+1 | 0 | U | U |

Table II.5: Volunteer’s dilemma

In real life such situations are not as rare as one might think. For instance, one person may be sufficient for getting the information out that the emperor has no clothes on, but one must indeed speak out. The **potential whistleblowers** in a case of company misconduct often seem to be in such a situation (see the introductory example “are whistleblowers always heroes?”). “Kitty Genovese” type cases are obvious examples as well. Many individuals witness a crime that they could prevent but knowing that there are others who could do so as well, each and every individual speculates that somebody else will incur the risk or cost etc.[[12]](#footnote-12)

A more amusing story may be just as well suited. For **penguins,** it is risky to jump into the water since a leopard seal may be waiting there for a juicy penguin “burger.” So, one of the penguins must literally test the waters in the morning. Penguins have to eat and, therefore, have to get into the water eventually. On their own, they would prefer to jump in, even at the risk of being eaten. However, it is better if somebody else volunteers. So, all the penguins become very polite to each other, indicating in true gentlemanly fashion: “after you.” Since penguin volunteers are in scarce supply, they shove politely along the edge of the ice until one of them accidentally falls in. All watch to see whether or not that “fallen comrade” will reappear. If the penguin has not “fallen into a trap,” all are happy to jump in after him; if the penguin fell for good, the others will wait a bit longer or try to get in at a somewhat different spot.

Now, we are not necessarily interested in the behavior of penguins. The film “Happy Feet” notwithstanding, it is not too easy for us to identify with them directly; however, we should note that the example may give rise to a few interesting speculations concerning our own species. The penguins may push each other physically, and humans may do so by some kind of group pressure. They can focus their applause on the volunteers, and the free-rider flipped over is indeed the zealot (see Coleman 1983, 1988). There is “the whisper in us” that makes us prone to volunteer and the applause of others supports this for better (mother Theresa) or worse (suicide bombers).

In this context, it may be worthwhile to note that a male penguin might have an incentive to jump in if it were to his advantage that individuals from the other sex observe whether or not he is volunteering. If survival of such volunteering actions is differentially related to good judgment – the wise male penguin jumps in only if he has a fair chance to get out living – and physical fitness – the better and faster swimmers among the penguins can afford some jumps that would be lethal for others – then females with a penchant for heroes may make a better choice of mating partners and, thereby, have better and more progeny. In turn, it may reward the male penguins to be chosen that way by females with the right capacities (see on this Zahavi and Zahavi 1997 and Zahavi 1975).

Of course, it would be foolish to generalize too quickly from the animal kingdom to the world of humans, but remarks like the preceding are in order to warn even at the present stage of the argument those who tend to think that only selfish behavior could have been selected for in evolution. An evaluation function that is much more complex may have survived as well. In nature, *it is not clear that only the subjectively selfish can survive,* and it may be no different when it comes to the human species. In fact the evidence seems to indicate quite the contrary.

It seems safe to assume that there is a niche for being guided by ethical codes. Evolutionary arguments seem to indicate that biological dispositions supporting this may have survived “biological competition”. Likewise, in economic life **norm guided behavior as well as virtues** should pay under certain circumstances.

**In sum,** in all **cases of split responsibility** volunteer’s dilemma type structures may play a role. Cases in which **whistleblowing** may seem appropriate are of this type and make their ethical life difficult for business people.

If we are seriously interested in practical ethics in business issues we are well advised to keep the volunteer’s dilemma as a paradigm in mind. As important as this is another paradigm of voluntary contributing to the common weal.

### 3.2.3 Voluntary Contribution Mechanism, VCM

Imagine a group of n individuals n≥2. Each has some initial endowment of which each individual can pass on a contribution with to a common project. The individual gets back a share of the sum of all contributions. More formally, we have as what the individual i receives after she contributed of her initial endowment the amount into the common pool, summing to $\sum\_{i=1}^{n}a\_{i}$, of which the actor gets a fraction $λ\sum\_{i=1}^{n}a\_{i},$, back.

If the fraction is large enough to make it worthwhile to cooperate, i.e. if

all have a reason *to wish* that all worktogether.

The condition implies that in case of all putting in all, i.e. each individual could improve her situation. As the reader can immediately see, all will, however, be individually better off if they put in nothing rather than something for any amount of contributions of others. Whatever the other individuals do it is better not to contribute.

To illustrate, the reader might want to use and go over the argument again with the specific parameter values. Of his personal contribution the individual, i, gets back only ½. He gets his fraction of what the others have put in independently of what he put in. Therefore it is always better for him not to contribute anything.

“Choice logic” is one thing, empirical human behavior another. In the laboratory it can be seen that people do volunteer initially. Yet, contributions to the common pool go down when the game is played repeatedly and anonymously. In such a **tragedy of the commons** situation (see on this Hardin 1968 and Ostrom 1992) even well willing people tend to learn not to contribute. But it is also obvious that an ethical system that would lead individuals to prefer to fully contribute out of an **intrinsic motivation** to do what is good and right would lead to an improvement in material payoff over the sum that could be earned under individually rational play with contributing nothing. But even then, it would still be better for the individual to contribute nothing as far as the **extrinsic monetary motivation** countervailing the intrinsic motivation is concerned.

An ethical code can conceivably get society out of the tragedy of the commons. Those social groups (tribes, states, firms) that manage to actually follow criteria of doing ethically good will more often than not also do economically well. It is “nice to have” an ethical system that induces cooperation in situations in which all “should” contribute to a common project to make each and everyone better off. At the same time we need to be aware that appeals to do good will not be sufficient in view of the logic of free-riding and the social psychological pressure to avoid being exploited as a volunteer or a well-behaved ethical person by those who do not do good. It is one if not the greatest follies of all not to include these insights into our search for an ethical reflective equilibrium.

Part III: Towards “(Wide) Reflective Equilibrium”

## 1. Disequilibrium and dissent

That something would happen in the future may be predicted abstractly by some theory or concretely by some person knowing the theory. Such a *prediction* is categorically different from the *prescription* that something should happen. “The demand for bread is not bread” (Bentham, 1843) and bread does not come to exist in our world simply because it seems extremely desirable that it would. No doubt that most people want that bread should be provided to all hungry people. Yet – since “want is not supply” (Bentham again) – the desire that there should be enough to eat for all does not bring about the desired state.[[13]](#footnote-13)

The abstract *impersonal* ethical statement that it is desirable that there be bread does not tell us who is (or feels) as a concrete person responsible to provide for it. Moreover, only if we know that there is a specific person who feels responsible for doing something and has the means to do it can we predict that the prescriptive demand will have an effect on the real world.

Explaining and predicting are activities other than reflecting and demanding action. Though it is rather clear when an empirical theory fails – e.g., a predicted phenomenon is not observed –, it is unclear what it means that a prescriptive theory is mistaken. Agreement is not a clear criterion either. To be rather specific, not all are agreed that everybody should be fed. Some have dissented and some perhaps still dissent (e.g. in view of the overpopulation problem). Examples of other kinds of fundamental disagreement abound. Some favor capital punishment or the killing of certain minorities or religious groups. All these phenomena occur in our world and there are normative systems and theories that prescribe actions that bring them about.

Though most of us would not think that even wide spread occurrence of the phenomena described before would justify the normative demands going along with them we have to know them. The facts of what others do or do not, what they as a matter of fact demand or do not demand as the ethical thing to do, are important. It matters if most people follow certain rules and enforce compliance with them. Not only in view of our personal interests in avoiding their resentment and sanctioning behavior but also as moral persons who always could act otherwise we need to know the ethical views of others.[[14]](#footnote-14) For, to expose ourselves to plural alternative ethical views is the first step towards examining our own.[[15]](#footnote-15)

Those to whom an examined life seems better than an unexamined one have a good personal reason to reflect on the principles that might guide and justify their actions. Managers are presumably no other than ordinary people in this regard. Yet being managers they may be inclined to wonder whether it is good management to engage activities like moral reflection on business issues.

**In short**, whether business ethical reflection is indeed good for business is an open question. Whether managers who ethically reflect their actions are serving their shareholders or other stakeholders better in terms of certain aims, ends or values pursued by stakeholders of the company is a contingent empirical question that allows for no general answer.

The following discussion assumes that for what reasons ever a business (wo)man is interested to reflect on the ethics of practices of business. Assuming that this interest exists the subsequent sketch focuses on the central methodological method (or metaphor) of modern practical ethical discourse: searching for a wide reflective equilibrium, WRE. As indicated already above a WRE incorporates in a coherent way specific ethical intuitions and general normative ethical principles such that the person who is in equilibrium can without cognitive dissonance defend her actions and plans to herself and others.

## 2. Deliberative practices in ethics akin to those in science

The “reflective equilibrium” method is the best candidate for accomplishing the aim of bringing substantive normative discourse as close as possible to what takes place in science proper. Like other influential ideas in twentieth century practical philosophy, this one, too, is linked with the work of the late **John Rawls**, who provided at least the name for and some of the details of the reflective equilibrium method. Whether it is rightly seen as a *method* or whether it is merely a metaphor must be left open here. In any event, I will refer to it henceforth as a method, if only in an “honorary” sense.[[16]](#footnote-16)

In a way, philosophers – and for that matter, educated laypersons – have used the reflective equilibrium method for ages in their practices of discourse. However, Rawls seems to have been one of the first philosophers to reflect explicitly on the relevant practices of rational discourse while searching for a reflective equilibrium. In 1951, Rawls published “An Outline of a Decision Procedure for Normative Ethics.” The title quite precisely reveals the topic of that seminal paper. With some heroic simplifications, the core of Rawls’ outline can be sketched quite easily.[[17]](#footnote-17)

### 2.1. Rawls’ outline of a decision procedure for ethics illustrated

The basic insight emphasized by Rawls in his outline is that in all intellectual enterprises we distinguish between singular and general judgments and use the one to assess the other. General rules are subject to the test of their specific implications. But, vice versa general rules can be used to revise specific judgments as well. Once the general and the specific cohere, a kind of equilibrium is reached. In this state, the general and the specific support each other.

To see less abstractly what is going on, consider the following example. Somebody utters the norm “Thou shalt not kill!” Assume, he initially thinks of this as a norm that is to be observed without exception. Now, somebody else states that the killing of plants is necessary for human survival and, therefore, should be deemed legitimate. May be the norm thou shalt not kill gets modified in ways allowing for the killing of plants. Perhaps the killing of animals may be legitimate as well? Again a meat eater might want to modify the original rule…

Being out of equilibrium in view of the unacceptable implications of maintaining the norm in its strictest form, the norm is modified. Let us assume that after a few steps it reads: “Thou shalt not kill fellow human beings!” Is this a sufficient modification?

Imagine that somebody draws attention to cases of self-defense. He might ask for instance – using a specific case – what he is to do if a man with an ax is trying to chop off his head while he himself has a gun in his hand that he could use in self-defense. The result of presenting this specific example presumably will be that the general principle will be modified.

The *overall* purpose of the norm suggests that we modify it. Going back to the aims, ends, or values that are pursued by implementing the norm as an institution presumably a restricted norm emerges: "One must not kill another person *unless* it is in cases of self-defense against an imminent threat to one’s own life."

Next let us assume that a person Z is threatened by an imminent danger to his life by another person X, yet has to kill a hostage Y held by X to defend his own, Z’s, life. In that case we might tend to say that the institutional rule that the killing of Y is forbidden should be upheld. At the same time we could argue that there would be a moral excuse for somebody acting in breach of the institutional rule.

Assume now that a man with an ax tries to chop off “only” a leg. Is the person who is being attacked still entitled to kill the attacker? Most of us would say that a person who is “innocent” and attacked should be entitled to defend herself by killing the attacker. So, the general norm would become something like “You must not kill fellow human beings *unless* it is in self-defense against an imminent threat to life or limb brought about by *them*!”

But what if the attacker is only trying to chop off the pony-tail of the person being attacked, and it is credible that he intends only this and is quite able to do it without any further risk to vital parts. Is it still justified to kill in self-defense, or is there a good reason to say that in such a case the taking of a life is not justified to prevent the deed?

Most of us would say that a threat to one’s hair is not a good reason to kill. It would seem that the involuntary haircut does not justify far-reaching acts of self-defense, but that may still be disputable. After all, how certain can one be of the intentions of a guy with an ax in real life?

**In sum**, in rational deliberations we seem to start with a general principle that seems to be justified *prima facie* and then rule out certain specific cases from the scope of application of the principle. We say something to the effect that it is generally justified to proceed in a specific way *unless*… By specifying the exceptions we reach an acceptable principle defended in a temporary reflective equilibrium.

### 2.2 Two examples

### 2.2.1 Trolleys and consequences

Questions concerning the killing of other persons are merely the most dramatic forms of sacrificing the interests of some to further the interests of others. Whole philosophy and psychology departments have discussed them in recent years in the somewhat fancy „disguise“ of so-called trolley problems (see Taurek 1977). Imagine you are standing at the side of a switch of the rails and can divert a trolley from its track where it is going to kill five workers (whom you cannot warn or rescue otherwise) to a track where one worker is working (whom you cannot warn either or rescue otherwise). Should you not send the trolley down the track where only one worker is working rather than let it kill five?

Clearly, an institutional rule that would demand that in such situations trolleys should be *generally* rescheduled such as to kill fewer rather than more persons would lead to better survival chances. As long as we would not know who of us will stand on the tracks we should all favor the rule according to which one rather than five get killed. Behind the **veil of uncertainty** of how any individual would be affected always to go for saving more than fewer lives is rational.

Individuals who would not know whether they will be in the position of the one or the five but with equal likelihood will be in each of the *six* positions seem to have a good reason to favor a rule of sacrificing one and rescuing always the larger number. Therefore should we not ethically desire to have such an institution in place?



Figure III.1 The Trolley problem

The standard response that there will be consequences that have to be brought about by action and consequences that can be brought about by omission does not help much here. We can easily construe situations in which either five or one individual can be rescued by one of two mutually exclusive acts. Imagine five swimmers drowning to the left, one swimmer drowning to the right, and only one boat to rescue either one or five. Or, imagine a drug with which you can either rescue one who needs five units of it to survive or five individuals who each need one unit of the five units that you can allocate. Should you allocate all to the one, to the three or should you throw a fair coin to give each a chance of ½ of survival? What about other lotteries that you could construe as institutional rules?

 Problems like Trolley may not seem realistic. They are not meant to be. “Trolley” and its variants are designed to put us out of equilibrium. Their discussion is for the ethicist like going to the gym for the sportsman. Nobody thinks of gymnastic weightlifting as a sports contest but almost all accept gymnastic weight lifting as a preparation for real contests in which other skills than mere force of muscle are relevant. Likewise thinking about Trolley and related problems will help us get prepared for real world problems in the sense of acquiring capacities to identify and to deal with hard problems. For more on these issues see the discussion of utilitarian consequentialism below in part IV section 6.

### 2.2.2 General principles of business

Recalling the controversy between shareholder and stakeholder approaches we can say something general about searching for reflective equilibria in issues of business ethics. A manager has an obligation to act as an agent of the owners. As manager she is under a moral obligation – though perhaps not extrinsically motivated by appropriate incentives – to pursue owners’ rather than her own aims, ends, or values when making management decisions. Her **general fiduciary relationship** towards her employers is clearly always a relevant consideration. Yet it is not necessarily outweighing all other considerations. Even if we accept that owners’ values should have normative priority over the interests of other stakeholders the manager is a moral subject herself. Even if she puts the values of shareholders first she cannot avoid taking responsibility for her personal behavior altogether. She could always exit the employment contract or voice her dissent. She may have several loyalties and her ethical ones not necessarily must have precedence etc.

For example, assume that a natural disaster occurred in a foreign “third” world country in which the company is operating. People are starving and lack access to decent water supply and health care. It may seem then to the manager morally mandatory to use the funds of the company for purposes of helping people “there” rather than stockholders “here”.

If the owners agree then the manager can go ahead. In fact if they wish the help to be administered the manager is under an obligation to provide it. But if the intervention negatively affects jobs at home, if it puts at risk the interests of some shareholders is the manager also under an obligation to help?

**In sum**, any realistically complex practical ethical business problem requires a rather complicated weighing of several moral aspects and conflicting interests. Weighing the tradeoffs cannot be avoided and there will **never** be **consent** on it.

Since practically everything can be justified in principle with some good reasons relative to some ethical system or other the claim that everybody must be rationally agreed is outrageous. As a matter of fact the claim amounts to arrogant partisanship for one’s own favored ethical position.

We may expect of decent moral persons only that they are willing to make a serious effort at reaching something akin to a personal WRE. In this they should include empirical knowledge and normative general principles along with detailed knowledge of the particulars of time and place.

The moral obligations that arise for those who like managers have to act as agents of others require to give matters serious second thought. They require a serious effort to find a coherent position based on best knowledge of the fact and practices relevant (see also above on shareholders and stakeholders again).

## 3. A refined reflective equilibrium metaphor

Like other basic ideas that Rawls had developed in the years before, the decision procedure outlined for normative ethics in his dissertation before 1951 becomes somewhat more elaborate in his 1971, “A Theory of Justice”, the most influential work of 20th century moral philosophy. In the later version there is more emphasis on the social aspects of the search process and its movement towards equilibrium than in the outline of the procedure itself. Moreover, it becomes a process of *mutual* adaptation and revision of the general and the specific normative statements. The equilibria sought for are, finally, of a more global kind in that they involve systems of principles rather than single principles.[[18]](#footnote-18)

Rawls' basic idea is the following one: There is a group of generally “**competent moral observers**.” The general competence of a moral observer to pass considered moral judgment is not determined by the nature and substantive content of the judgments found by that observer. The moral observer is competent to pass judgment if she is of about average intelligence, has some minimum experience in life, is a grown up and so on. Fulfilling these requirements, the competent moral observers may pass judgment on any moral issue that is brought before them.

They may in principle come to any conclusion. In doing so they must be non-partisan though. It is assumed that they have no special stake in the issue under consideration. They may be generally interested in the matter but not particularly so. For instance, every person may be regarded as generally interested in what a good society is like. Yet, that is different from having a special interest in a case that is to be decided. The moral judge is like a judge in court who will hopefully be partisan for rule of law and share an interest in upholding basic rights of citizens but should not have a particular interest in any problem brought before him.

It is most important to note also that the idea of the justificatory one-way street of the outline is given up in the reflective equilibrium search presented in “A Theory of Justice.” Rather than “inductively” adapting principles to specific cases, to evidence, and to other principles until they seem to cover the basic judgments of morally competent observers, it is now assumed that sometimes the specific intuitive judgments become revised in the light of well-entrenched general principles that typically form parts of a system of principles supporting one another or at least coherent with each other. Going from the general to the specific and from the specific to the general and back a ***cyclical* form of adaptation** emerges.

**In sum,** well-entrenched general principles that are very strongly supported because they have been adequate in solving a great number of specific problems and are supported by other well-supported general principles may have their way against intuitive judgments as well. In such cases we revise the specific intuition in view of the principle rather than vice versa.

For instance, somebody may have been of the opinion that it is legitimate to kill a bystander if that is the only way to defend oneself against a hostage-taker. However, after additional reflection it may seem better to revise the intuitive judgment and to stick to something like the moral precept that we must never hurt the vital interests of innocent bystanders.

This may be justified because, for instance, we foresee that an institutional rule that explicitly allows for exceptions might be abused. This is a kind of empirical psychology argument based on the likely workings of institutions but it clearly is part of a WRE search as well. We are talking about the real rather than an ideal world.

Revisions of views on behalf of general principles are more common in daily life than most of us seem to be aware. For instance, once many people believed that it was absolutely wrong that unmarried couples live together. Many of these people may, at the same time, have been of the general opinion that it is wrong to interfere with decisions of others that concern their purely personal affairs. In Europe, the latter principle gained rather strong support from the general experience that in particular religious conflict could not be avoided unless people were willing to endorse mutual tolerance. The principle that one should not interfere in personal matters supported the conclusion that the people who were demanding that non-married couples should not live together needed to revise their views. The original verdict that they might have spontaneously uttered against non-married co-residence in a specific case did not cohere with the general tolerance principle to which they otherwise subscribed. Thus, the specific intuitions, at least of some people, were revised and general practice followed suite eventually.[[19]](#footnote-19)

In many instances the problems we are dealing with in search of a reflective equilibrium do not arise from tensions between specific norms concerning singular cases. They emerge because of tensions between verdicts of different generality. There are more specific and less specific or more general and less general principles involved and some kind of hierarchical system among such norms may seem desirable. In that sense the principle of **shareholder first is a hierarchy building principle**. It is always relevant but it may not always strictly apply.

**Two main concerns** **in all weighing of principles** will always be relevant. *On the one hand we should try to go for the most general principles we can find since this gives us some assurance of coherence and non-arbitrariness. On the other hand, we should always try to make our actions and the reasoning behind them transparent to the stakeholders concerned. For then a large part of responsibility is theirs.*

For example, if a company considers how to deal with *pregnancy*and is interested to do so in a very general way it might think of treating pregnancy as a case of temporary disability. This avoids the risk of having to set up special rules for women. Women who get pregnant are treated the same way as, say, individuals – either female or male – who had a sports accident forcing them to stay away from work for four months. All who are unable to work for some foreseeable time are guaranteed their job. If this principle is made *transparent* then those who accept a contract with the company or do exit from the contract nexus thereby signal their implicit (dis-)agreement.

The more general the better are the rules that lead to adequate solutions in all known specific cases and are thus in a very global reflective equilibrium.[[20]](#footnote-20) The more clearly general guiding principles are communicated the less is the manager at risk of deceiving others into going along with his own ethical judgments without noticing it. She avoids imposing her own views by making them transparent and by making them as general as possible she assures coherence of her actions and formally equitable treatment of those who are affected by them.

Along the lines inspired by Rawls, we can envision a process of seeking an equilibrium including our specific and our general views as well as “pre-fabricated” external theories. Reduced to its bare essentials, the process seems to be of the following kind:

Assume that a system of ethical principles T (and singular verdicts) is established as a temporary equilibrium.

Derive implied solution *s* from T for some specific “new” normative problem or case *c*.

Form a considered judgment of whether or not *s* is “intuitively” adequate to solve *c*.

If *s* is adequate to solve *c*, this corroborates T.

If *s* is inadequate to solve *c*, then

modify T to get T’ (e.g. "T, unless cases of form *c* occur…")

or

revise your intuition as being in error, in particular if T is well corroborated and hard to modify in the relevant regard.

Go on with the process as described until it converges to a state in which no further adaptation seems necessary. A new temporary reflective equilibrium is reached then. (Dealing with the trolley problem and its many variants is exactly illustrating the process of going back and forth in search of an equilibrium.)

What has to be done to reach RE can also be described more loosely and succinctly in the following way:

Keep your eyes on systematizing judgments, integrating them by generalizations while weeding out incoherencies.

Integrate your perspective on different levels of generality.

Go back and forth between modifying general principles and the acceptance or rejection of specific judgments until no need of modification arises anymore.

 An intra-personal temporary reflective equilibrium is reached once general and specific individual judgments support each other coherently. In a system of general and specific judgments, the general and the specific support each other *in equilibrium*.

Typically external empirical and normative ethical theories are included into the reflection in a next step. This then hopefully may lead to a wide reflective equilibrium, **WRE**, which is exactly what we are aiming for when we discuss the ethical theories and fundamental normative issues of philosophy and economics. Whole classes of complicated general theories are brought together with mid-range general and specific “local” knowledge of prevailing practices in the search for WRE**.**

When confronted with a **complex business ethics issue** it is presumably a prudent strategy to search for a reflective equilibrium before taking action. Being backed by a systematic reflection will not render the actions impervious to criticism. However, an actor who can demonstrate that the actions he took were supported in some **reasonably comprehensive reflective equilibrium** can more credibly raise a claim to moral integrity and responsible behavior. That this will convince the public at large cannot be guaranteed. It depends on empirical factors. However, it seems that there is a kind of general proclivity in humans to search for coherence and to avoid suffering from cognitive dissonance.

With the characterization of reflective equilibrium and our search for it we can now also provide a general and formal characterization of **responsible behavior** *that is not biased in favor of what we personally prefer*. Fully responsible behavior requires at least:

to fully study the empirical action situation (with reasonable effort),

to listen to all who legitimately should be heard,

to go about decision-making in a principled and empirically informed way

to provide documentation on a search for a reflective equilibrium.

 **In short,** the postulate that mangers should act “responsibly” amounts to requiring that they search for reflective equilbria to the best of their knowledge.

If we do not intend to impose our own specific values on them by holding managers responsible for acting in the ways we want them to act then we must go for a characterization of responsible action that is as neutral as possible with respect to specific values. Assuming that only those who agree with our own specific substantive views act responsibly seems flagrantly incoherent if we accept value pluralism and intend to respect the views of others to the possible extent even if we disagree with them.

It is in general very hard to make up a reflective equilibrium after the fact or after action has been taken. If the actions have not been systematically planned and integrated into a **reflective equilibrium beforehand**, coherence will be hard to achieve after the fact. The appearance of arbitrariness is presumably the strongest indicator against both ascribing “moral integrity” and “responsible action”. In fact, being respected as a “responsible” actor of moral integrity may be a most valuable asset if others disagree with the substantive content of what has been done.

**In sum**, any organization that intends to manage “moral risks” and to prepare for the not unlikely event that it might come under “moral siege” in a public scandal seems well advised to prepare for that contingency by gathering as many reflective assets as possible (again this is a technological advice contingent on the validity of certain plausible empirical hypotheses concerning opinion formation).

In searching not only for a RE but rather for a WRE some knowledge of ethical theory may turn out to be an asset, too. So let us turn to a sketch of select elements of ethical theory.

# Part IV: Substantive Ethical Approaches

The verdicts that are reached by reflective equilibrium search depend on time, culture, ideologies etc. They persist only temporarily in a specific context. Because of their space-time dependence they are called **“local”.** For business ethics purposes reaching **local equilibria**, say, for a single company operating in a specific cultural environment at a specific time is all we can hope for. Still, ethics has always had the aspiration to transcend local restrictions and to become **global and universal** (at least “in theory”). This aspiration has been pursued in terms of what may coarsely be classified as **virtue, deontological and consequentialist ethics**.

## 1. Basic concepts of consequentialism and non-consequentialism

Approaches that focus on action types – e.g. whether or not an act is of the type of lying – are classified as **deontological**. The focus on action type (beyond action consequences) leads to classifying deontological positions as **non-consequentialist** whereas normative ethical approaches that justify their conclusions in terms of consequences are generally referred to as **consequentialist.** These are the two main types of ethical arguments and approaches.

Slightly less but almost as coarsely as in the preceding classification we can say that modern ethical consequentialism broadly amounts to **utilitarian positions**. Non-consequentialism of the deontological kind is akin to **Kantian views**. **Aristotelian virtue ethics** are, in a way that will become clearer below, mixed. The main focus of the subsequent discussion of ethical theories will be on deontological and consequentialist approaches though, other than in most conventional discussions, within an **institutionalist perspective**. The focus is not on what is right and wrong per se but what would be “right” or “wrong” institutions.

**Thomas Hobbes** was arguably the first modern thinker focusing on making moral institutions. “Right” and “wrong” are man made. As far as the process of crating institutions and the behavior under them were concerned Hobbes was, in a way, the first economic imperialist. He intended to construct a theory of what people would and what people should do solely in terms of opportunity seeking behavior directed at future expected consequences.

Economists are to the present day Hobbesians since for them opportunities and consequences are all that matters. They share Hobbes’ view that the Hobbesian model of Homo (o)economicus can be applied to all human behavior across the board. Many modern ethicist believe that the opposite holds true. That motives other than striving for better substantive results matter. In particular there is genuine rule following behavior. The Homo (o)economicus model is fundamentally flawed as a description and explanation of human behavior. The prescriptions of ethics cannot be, say the ethicists, understood in terms of guiding human opportunism.

**In short**, for most present day ethicists a Hobbesian consequentialist approach to moral behavior and institutions is descriptively and prescriptively inadequate.

As an illustration of basic concepts of consequentialism and non-consequentialism imagine an actor who considers a choice in Trolley. The actor may focus on the consequences of that act and more or less neglect the ways and means by which these consequences can be brought about. Such an actor will tend to justify actions solely in terms of the consequences and will be motivated to perform the acts by the envisioned consequences. In cases like “Trolley” the consequentialist will compare the consequence of one life lost with that of five lives lost. Going for the larger number he will neglect that the rescue or sacrifice of persons once involves active killing and once an omission to act.

Alternatively, an actor might consider chiefly the acts that she might perform while neglecting the consequences. In a case like Trolley the deontologist will compare active killing and the omission of killing. That once one life is lost and once five lives are lost is a secondary concern for her. Moreover, she typically takes her action “personally” or relative to her *own agency*. Such a deontologist deems it relevant that *her* acts are of some *type*. She will be primarily motivated by the desire to act in conformity with the ethical status of the *action type* (never actively kill, regardless of consequences) rather than the specific consequences of an action taken. She feels primarily responsible for what is *her* doing

An institutional focus will lead to an intermediate position between consequentialism and non-consequentialism. It renders the “means to ends”-justifications in ethics *oblique* and thereby mitigates the conflict between the two basic ethical outlooks.

## 2. Moral institutions not moral judgments as the focus of ethics

If we adopt an “institutional perspective,” the tensions between consequentialist and non-consequentialist approaches become much less severe. For instance, the social practice or the institution of “honesty” requires that people in general tell the truth without calculating the consequences of specific acts of truth-telling. In order for the act to be classified as ethically right, it suffices hat an act can be subsumed under the *institutional* description or the general action type of “honest act” as specified by the institutional rules.

The preceding seems to be fully on the deontological side. Yet this conclusion is premature. The justification for the institutional classification may be in terms of the ***good consequences*** *but of* ***the institution***.

For instance, institutionalized social rules require that people act in conformity with the rules of truth-telling. Those who violate the **primary rules** – the rules that specify what to do – are (to be) criticized according to **secondary rules** that specify how to respond to rule violations as a victim as well as an observer. Both types of rules (see Hart 1961) jointly define the practice and characterize right and wrong unless certain types of excuses arise.[[21]](#footnote-21)

If we look at it this way then the clash between consequentialist and deontological positions is not as clear-cut as it may seem initially. All seems to depend on which level of the institution we locate the argument. Consequences seem to reign supreme on the level of institutional choice, obligations and rights on the level of within institutional choice making.

**In sum,** deontological and consequentialist arguments are located on different institutional levels. **Institutional rules may be justified by their good consequences while acts are justified by the rules**.

Yet *not* everything can be pre-programmed in rules. There will be always exceptional cases when the institutional rules do not work satisfactorily. Typically the institutional rules are therefore assumed to apply only prima facie, i.e. unless exceptional circumstances prevail and suspend application of the rules. Regrettably as in good management in ethics we cannot exclusively rely on rules. We need to provide for exceptions and the necessity to manage exceptions.

## 3. Ethics by exception

Under “unforeseen” circumstances, typically an appeal to the judgments of other competent participants in the practice may be sought. These others who serve as a real or as an “imagined jury” in such cases are assumed to come to the same conclusion or at least to the judgment that some proposed conclusion can be justified in principle even if they themselves might not concur with the particular verdict.

At this juncture general ethical theories might enter the picture. They can be used to fill gaps in the institutionalized practice rules. “Moral judges” who set out to search for a local reflective equilibrium may turn to these theories to widen the perspective beyond the specific context. For instance, they may consider how an exceptional problem not covered by the institutional rules themselves would be solved under, say, utilitarian ethical principles according to which overall utility should be maximized. In that case, they might even go back to an ethical principle or theory – here, utilitarianism – to provide a reasoned account of the institution – i.e. its general merits.

For example, promise keeping is a beneficial institution in our society. Once it is in place it increases utility on the whole. Even if sometimes breaking a promise would be better than keeping it, the keeping of promises in such cases is an investment in the overall reliability of the institution. It is a fair contribution to sustaining it as useful instrument in pursuit of the common weal (see already Sidgwick 1874/2010).

**In short,** the familiar concept of “**management by exception**,” translates into that of an “**ethics by exception**.”

To illustrate further the manager provides general rules for her executive agents who have to play by the rules unless something special comes up. The burden of proof that something exceptional occurred rests with the agents. A good manager will applaud her agents if they can prove that a case is indeed an exception to the managerial rule. She will make it worthwhile for them to show that an exception has indeed occurred. She will, however, also insist that agents should not *lightly* offer the argument from exception. For, then, the main reason for having rules or routines in the first place would be undermined.

An analogous reasoning applies in the field of moral decision-making in business contexts and beyond. The established moral practices with their primary and secondary rules tell us basically what to do. To the extent that we believe that a practice in general is morally sound, we should feel under an obligation to respect the practice obligations as *prima facie* valid. Still, there will be exceptional cases in which moral management by exception will be needed. Then it might be useful to **go back to the primary ethical principles** of the ethical theory justifying the moral institution in the first place.

**In short,** primary ethical principles, as opposed to the primary rules of the practice, are *not* part of the practice but merely of the search for a wide reflective equilibrium on the constitutional level of fixing the rules.[[22]](#footnote-22)

There is a regularity aspect as well as a rule-following aspect when it comes to institutions. On the one hand, individuals behave regularly in ***conformity*** with the requirements of the rules. On the other hand, individual behavior is motivated by the individuals’ knowledge and ***acceptance*** of the rules. Acceptance is bringing about conformity but conformity can be had without acceptance. Ethical rules that motivate individuals *intrinsically* to behave in certain regular ways and announced sanctions that are regularly inflicted to motivate individuals *extrinsically* work in the same direction. Both induce individuals to behave in conformity with rules. They rely on different motivational mechanisms that in a well-ordered society support the same institutions that in turn support the motivations etc…

If some individuals *beyond* this also come to the conclusion that the rules stand up to critical scrutiny in a wide reflective equilibrium, then this forms an *additional* element in support of an institution. It is an “opinion” that intrinsically motivates acceptance of the primary and secondary rules as well as the moral dispositions (virtues) that are complementary to them and become operative when gaps are to be filled.

Whatever the auxiliary role of theories and reflections: The established rules and practices are the **primary and first source** of evaluative and normative standards. The institutional (practices of all sorts) precedes the judgmental level and not vice versa. The **burden of proof** is always on those who intend to deviate from institutional practices in view of some other reasoning. There is a “conservative” normative principle involved: In search for reflective equilibria individuals should start from accepted practices as prima facie valid.

In an institution in the full sense of the term, individuals must regularly comply with its standards, but that is not sufficient; at least some must also endorse those standards and follow them from an **internal point of view** sometimes.They must basically be intrinsically motivated by the rules.

**In sum**, an institution or institutionalized practice is always characterized by both a regular way to *behave* and a regular way to *judge, reason or evaluate* according to some standard that *as a matter of fact* prevails. It can exist only if at least some individuals *sometimes* act *because* they adopt an internal point of view to its primary and secondary rules. Moreover, the complex web of primary and secondary rules that, along with the proviso of case-by-case judgments under unforeseen circumstances, forms an institution is backed up by behavioral dispositions to act accordingly.

## 4. Moral institutions as conventions

As far as the issue of **consequentialism versus non-consequentialism** is concerned, the preceding has a clear implication: One can be a consequentialist with respect to a moral institution as a whole and at the same time require that the rules of the institution be observed in a non-consequentialist way (not extrinsically motivated by the consequences).[[23]](#footnote-23) This is true even if we think of the institution as something that is conventional and could be otherwise.

Somebody may justify his actions by pointing out their conformity with an established practice. He refers to the rules and an institution of which these rules form a part. He can say that, *according to the rules*, it was the case that he ought to have acted in a specific way and that he did indeed act accordingly. If these rules represented the prevailing moral practice, or “the morals of the land” as one might say, such a person has acted according to his **prima facie moral duties** or the duties as defined by the established practice. If it is claimed that the prima facie duty is not one that should be fulfilled because of exceptional circumstances he might go one step further and argue that the *institution* requiring the acts he performed could be justified in terms of its general consequences. If this argument were also challenged, the person might argue that the institution was still the established institution and that it was more desirable to have at least one established institution than to have none even if the established institution is perhaps not the best one in terms of its consequences.

To give a simple illustration of the preceding (“Humean”) argument, think of the side of the street one drives on. Imagine somebody driving on the right-hand side of the street in a country in which that is the established practice – let us forget the legal aspects for the time being. If there has been an accident and somebody argued – perhaps with good reason – that the accident could have been avoided had the driver not been driving on the right-hand side but instead on the left, the driver would nevertheless have a good reason “on his side” for driving on the right-hand side since the general practice is to drive on that side of the street.

For the sake of the argument, let us assume now, as any good British gentle(wo)man would, that driving on the left-hand side of the street is, for some deeper physiological/psychological reason or other, the “natural” way of driving. In view of this assumption about human nature, let us also assume again for the sake of the argument that it would be better for each and everybody (a Pareto improvement) if all human beings drove on the left. This institution and the complex practices supporting it would be better than the established one of driving on the right. However, when challenged on his failure to follow the not yet established rules and practices of the “naturally” better institution, the person who got involved in the accident by driving on the right could still insist that the established institution has been that of driving on the right-hand side of the street and, further, that even if driving on the left is the better **conventional institution**, this could not provide a good reason for an individual to act in line with the requirements of the potentially better as opposed to the established institution. The individual might have a good reason to act in ways that might convince others to make a coordinated change of the practice but certainly not have good reasons to deviate from the practice unilaterally.[[24]](#footnote-24)

Again more generally speaking, even if a potential alternative institution were better if all should conform to it, it is “as a rule” better to conform with the established institution as long as all or most others are acting in conformity with established institutional practices. Game theoretically speaking, we are talking about equilibrium selection problems here. If an equilibrium is supported by an established practice, then this may be a very good reason – from the point of view of self-interest as well as *from a moral point* of view – to act in accordance with the rules of the established practice.

This is true even if acting according to the established equilibirum leads to an outcome that is dominated by another *equilibrium* that makes everybody better off should everybody change behavior. For the sake of specificity and simplicity, consider the following most simple case of a so-called coordination problem with two equilibria:

|  |  |  |
| --- | --- | --- |
| Column=Role BRow=Role A | left-B | right-B |
| left-A | 2, 2 | 0, 0 |
| right-A | 0, 0 | 1, 1 |

Table IV.1 Coordination with two equilibria, one payoff dominated

If people are paired randomly to play such a game in a one-shot manner, it is advantageous if there is an established institution according to which each knows what to do. Assume that the established institution requires the players to choose “right.” Then, the outcome will be “1” for each individual on each round of play as long as all the individuals always stick to the rules of the institution. If there is a large number of individuals involved and if, as may also be assumed here, the individuals are randomly encountering each other, then deviating individuals will in all likelihood end up with “0” themselves and will cause their co-player to also receive “0” instead of “1.”

It seems clear that the *practice of criticizing those who deviate from the established practice* will derive support from the fact that a unilateral deviation leads to a worse outcome. As long as the institution is established, there is a good, prudential (and often also moral) reason to criticize those who deviate: **When in Rome do like the Romans do!** is a rather plausible maxim. This is true even if each individual would understand that she and everybody else would be better off if the institution were abolished and substituted by a (Pareto) superior one that makes everybody strictly better off; i.e. in the present example everybody being advised to take left rather than right.

**In sum**, institutional rules that require certain acts in a non-consequentialist way can coexist with the insight that in terms of the consequences of a general or co-ordinated switch it would be better to switch to another institution.

Still, there may be no way to *choose* the new equilibrium. Each individual in her choices may be stuck with the old equilibrium as long as she acts unilaterally. At the same time, **deliberations** may lead everybody to the insight that a co-ordinated change would be advantageous. “Indirect” **constitutional interests** in a change of institutions may coexist with “direct” **action interests** and obligations that both dictate adherence to the established institutional practices and rules and, thereby, prevent institutional change.[[25]](#footnote-25) Obliquity is a fact of life that is in many practical business ethics discussions unduly neglected.

Since institutions have an existence of their own, many people (including many philosophers) seem to be deceived by the prevalence of established institutional practices that pervade our lives into erroneously believing that the world or our conceptual apparatus have certain features independent of institutional facts.[[26]](#footnote-26) However, at least in principle all institutions could always be otherwise (i.e. another equilibrium could be selected).[[27]](#footnote-27) Institutions despite their rigorous demands are in a fundamental way conventional. [[28]](#footnote-28) They are relatively absolute absolutes (Buchanan, 1999, vol.1)

## 5. A glimpse of ethical rigorism

One cannot discuss ethical rigorism and the refusal to make exceptions to principles without referring to Immanuel Kant. **Kant** had that most fundamental faculty of the truly great philosopher, namely to identify the crucial alternatives (never mind that he almost always chose the wrong alternative then). He understood that the grand issues of theoretical and practical philosophy, respectively, were, (a) whether there are synthetic judgments a priori (whether we can know something about the world “before” experience of it) and (b) whether or not there are any categorical imperatives (or norms that may claim validity independently of the fact that somebody desires, wishes, values, or aims something).

In the theoretical as well as in the practical realm, Kant was searching for “laws” with *overriding* antecedent clauses (“ifs” in “if-then” conditions): whenever the if-clause is fulfilled, this is sufficient to allow for the conclusion. In such a scheme of things a moral rule, “whenever A, then B is obligatory” must not allow for an “unless-sub-clause” like “whenever A and C, then non-(B is obligatory)”. Nothing, i.e. no C, can suspend the overriding force of A. Promises are to be kept regardless (“whenever A then B”) and not “if a promise was given, then it is obligatory keep it, unless condition “C” also holds (whenever A then B unless, C. That is if “A and C then not-B” in all other cases if “A then B”).

**In short**, since a posteriori, or empirically, something can always happen to suspend the force of “A then B,” a law lending overriding force to A must for Kant be of an **a priori** nature. A moral law must hold independent of contingent experience or before anything happens.

In his own scheme of things Kant was right in drawing the preceding conclusion that ethics must be a priori. It seems that Kant pointed out correctly what it takes to defend his position but was incorrect in believing that it can be delivered.

If a priori valid normative insights would exist, the “moral laws” that bestow overriding force on their premise could be followed without any further ado or interpretation. This means in particular that any consequence C that may come about beyond the primary consequence B does not matter. For example, in the case of the obligation of truth-telling (or not lying)[[29]](#footnote-29), Kant intends to have an absolutely binding rule. Quite in line with this, Kant states that the person who is asked by a would-be murderer where her potential victim can be found is not allowed to lie about it, whatever the consequences. So the truth must be told even though telling it is conducive to let a murder take place.

A possible construction to avoid practical extremes like this might again be reminiscent of the precepts familiar from practical management discussions. “As a rule”, it is right to fulfill conventional obligations *unless* there is a good reason to assume that an exceptional case has arisen. Only then, prudent judgment is required. As in case of the frequently invoked “management by exception,” it becomes a matter of prudently weighing consequences if there is a good reason to assume that the rule does not apply or is in conflict with another rule. But such weighting has always some arbitrariness about it.

In the preceding light, the Kantian deontological precepts could be treated as merely **prima facie valid**. They are then more or less reduced to burden of proof rules. This is presumably good policy and brings the Kantian position very closely to the indirect consequentialism of the **ethics of institutions** discussed before (and below again under the rubric of “utilitarianism”). Yet, it is a far cry from the original Kantian rigorism. Moreover, it is unsatisfactory with Kant’s aim of providing clear guidance in an uncertain world. As long as we do not have some kind of standard by which we can weigh the relative force of the prima facie obligations we will not know what to do still.

As opposed to the implicit aim underlying Kantian rigorism, human abilities to evaluate situations are such that humans will always encounter cases in which sticking to rigid rules seems inadequate. In such cases the orientation by rules must be substituted by something weaker like general dispositions to go about problems in a prudent way. This is what followers of **Aristotle** suggest. Aristotle emphasized the role of general dispositions and attitudes, i.e. of virtues rather than rules. He believed that rules without judgment could never *as such* guarantee “good” results. Management of ethical affairs by rigid rules cannot work, only management by exception can.

Other than in a Kantian, in an Aristotelian framework of sound moral judgment there are no overriding conditions that lead to certain conclusions no matter what. Conflicts between principles, rules and intuitions will clash and states of reflective *dis*equilibrium will occur. For instance, in the case of not lying and protecting somebody whose life is threatened by murder each of the duties may seem prima facie valid, but only one can be fulfilled. Both duties can claim to be non-hypothetically justified in a deontological ethical framework.

Aristoteles believes that a virtuous person should do the weighing here. Even though no operational criteria for correct weighing exist there may be individuals who command the **virtue of fair judgment** and can find the middle ground that takes into account all interests fairly.

Neither Aristotelian judgmental (virtue) ethics, nor Kantian rigorous a priori rule ethics will yield anything specific in the weighing of personal interests. The first does not intend to reign in trade-offs and the second does not allow for them at all. Utilitarians try to tackle the problem of weighing the interests of individuals on the basis of some operational criterion head on.

## 6. Utilitarianism as classical teleological consequentialism

At least in the Anglo-Saxon philosophical discourse, which became increasingly dominant in Western Philosophy in the inter-war and early post-war period, utilitarianism has been the most prominent practical ethical theory until the 1960's.[[30]](#footnote-30) That there has been a closely parallel welfare economic discussion should not be forgotten here. The two traditions are more or less two sides of the same coin of a **consequentialist approach to normative ethical argument**. Because of this rather intimate association of ethics and economics in utilitarianism it is obvious that some special attention should be devoted to utilitarianism in a business ethical context.

### 6.1 Utilitarianism generally characterized

Utilitarians are consequentialist. Like their brethren, the classical and neo-classical welfare economists, they stick to the moral **principle of beneficence** or the aim of doing good. Beyond the principle of beneficence, **common characteristics of typical utilitarian positions** are:

*First*, what is morally right is characterized in terms of the ***non-*moral** consequences of choices or in terms of ***non*-moral** “goodness.”[[31]](#footnote-31) Non-moral goodness is determined according to some **value** **theory** or **axiology** which stipulates that “pleasure and pain” (hedonistic utilitarianism), “happiness” (eudaimonistic utilitarianism), “preference satisfaction” (preference utilitarianism) are non-morally good in themselves.

*Second*, what is non-morally “good” is determined independently or “before” what is “right.” Thus, **the “good” is used to characterize what it is “right”** to do in terms of causal consequences of action.

*Third*, **the distinction between actions and omissions is ultimately irrelevant**. Provided that choices x (-action), y (-omission) bring about the same (non-moral) consequences and the consequences are intended, the two choices x, y are morally equivalent.

*Fourth*, utilitarians are “universalist and individualist” in the following sense: a. in determining or ranking what is good for all **each individual should “count for one and none for more than one.”** b. In evaluating states of the world **the good is good independent of its location in a specific person**; i.e. provided that it is the same “amount,” it does not matter whether the good is located in individual A or B. c. Beneficence is universal in that all affected individuals who can **suffer** or **benefit** in consequence of some intervention must be taken into account.

Due to the fourth condition – with some minor technical additions – the criterion for ranking alternatives amounts to a sum total of measurement values. Though modern welfare economists share the preceding views with the classical utilitarians they became cautious regarding the fourth. At least they do not believe any longer that one can form a “sum total” of the several individual non-moral goods without further ado (see Robbins 1935). To put it slightly otherwise, the “new” welfare economists – as opposed to the old who were more classical utilitarian – do not anymore believe that there is an inter-subjectively accessible standard according to which **inter-personal comparisons** could be made and compensations could be calculated.[[32]](#footnote-32)

### 6.1.1 Utilitarianism and interpersonal comparisons

To illustrate where the difficulty lies and where not, imagine a **decathlon event**. Other than in ice-skating the performance in each of the 10 competitions can be *objectively* measured. However, it must still be somehow fixed in an act of value judgment whether or not the centimeter in the high jump event is of higher or lesser value than a tenth of a second in the hundred-meter dash.

Now transfer this model to a group of 10 people. Like the decathlon events each of the ten individuals ranks different competing policies for the group. Each of the ten individuals may evaluate each policy differently. Assume that as in the decathlon example which of two measures performs better depends completely on the separate evaluations by the 10 people (“welfarism”). Even if these values may be measured individually, an over all performance measure is possible only after telling what the evaluation by each of the ten people is worth in terms of the evaluations by the other individuals and vice versa. (What was before a conversion rate between centimeters and tenth of a second is now a conversion rate between individual rankings.)

In the ranking of states of the world an increase in the (non-moral) good of one person could conceivably compensate the loss of another person. However, we cannot *objectively calculate* whether the increase for one person can be a sufficient compensation for some specified decrease in another person’s good. Such interpersonal comparisons are deemed illegitimate as judgments of welfare economic *science*.

As a consequence, in welfare economics policy recommendations guided by the principle of beneficence have become restricted to recommending changes that make each and everybody better off, **Pareto improvements**, that is. If an alternative A is better for each individual than an alternative B then we need not say how much relatively speaking A is better. A recommendation of A compared to B does not require direct inter-individual utility comparisons. Classical welfare economics believed still in the possibility of making such comparisons and was therefore a much more powerful normative criterion in discriminating between alternatives. But the criticism of incommensurability was well taken.

The criticism of classical utilitarian welfare **economics** transformed it into what became known as Paretian welfare economics[[33]](#footnote-33) while utilitarian ethics still goes beyond this. It does not shy away from making inter-personal comparisons (though modern utilitarian ethicists may be more cautious in doing so).

The criticisms of naïve aggregation are obviously justified. Yet one should note that whenever somebody opts for a specific policy rather than another one making better or worth off different people he or she “reveals” an implicit interpersonal comparison. In a world of scarcity the decision-maker cannot doe otherwise. He must fix some rate of transformation at which the loss of individual A can be compensated by a gain for an individual B (the centimeter in the high jump and the 10th of a second in the hundred meter dash). If not both policies can be realized.

### 6.1.2 Utilitarianism descriptive and egoistic vs. prescriptive and altruistic

As opposed to descriptive utilitarianism that explains what *will* be done or will happen in terms of utility maximization normative utilitarianism is a substantive moral theory of what *should* be done. It is characterized by the view that a choice of an alternative is morally right if (if and only if) it brings about good in fact optimal consequences for *all* concerned (as measured by some pre- or extra-moral standard of goodness).

This concept of utilitarianism rules out **egotism** and requires **interpersonal comparisons**. What matters are always the consequences for *all* concerned, and the evaluating individual is simply one of the “all”; i.e. in evaluating the consequences, the specific interests of the individual who forms the evaluation and possibly has to act upon it do *not* have a special role or weight. But utilitarianism clearly cannot avoid interpersonal comparisons of utility lest it would become the same as Paretian welfare economics requiring that all must be made better off by measure A as compared to measure B before we can say that A should be preferred.

***In short sum****, the term “utilitarianism,” is reserved for a universalist impartial nomative theory characterizing morally right (or wrong) acts in terms of the good (or bad) consequences that they bring about for all concerned.*

Speaking of “all” concerned, we must, of course, specify who would count as one of all “individuals” who are to be taken into account in evaluating consequences. We could think, for instance, of all sentient beings as the relevant collectivity of individuals. Perhaps, however, we might want to only include human beings. We could restrict ourselves to persons, as has been indicated before, or perhaps to those who have the potential to become persons etc.

Again, there are many variants conceivable here, each of which raises further questions. It is, for instance, quite contested, whether human or non-human **persons** (like great apes) should be given special consideration as compared to other sentient beings in measuring the weight of the consequences. It is also open as to whether a utilitarian should make a distinction between those who already exist and those who will or merely might exist. What about those whose future existence depends on a present utilitarian decision concerning progeny? Should we discount the consequences for ourselves or other future selves according to some measure of distance (including time and uncertainty of consequences) or not?

It is in fact one of the merits of utilitarianism that the preceding problems become more clearly visible within its approach. Other ethical theories often suffer from them as well. That they do so more intransparently is hardly an achievement.

### 6.2 Act vs. rule utilitarianism

Act utilitarianism, on the one hand, and rule utilitarianism, on the other hand, must be distinguished from each other and from institutional variants of utilitarianism. In particular, for those acquainted with the model of rational economic man as characterized in decision and game theory[[34]](#footnote-34) it may be helpful to start with act utilitarianism and to characterize two types of rule utilitarianism only afterwards.

### 6.2.1. Act utilitarianism

**Normative act utilitarianism**: The utilitarian value theory[[35]](#footnote-35) informs the actor about the rankings of alternatives as they affect other individuals. In the light of this she *should* consider how each of her acts causally affects all other individuals concerned and choose opportunistically the best act. In that regard, she *should* behave as descriptive rational choice theory assumes she would. However, note that she *should* do so in pursuit of *moral* (impersonal) rather than her contingent non-moral (*personal*) aims. The normative act utilitarian uses a utilitarian value theory for ranking those states of the world whose emergence she can causally further or hamper.[[36]](#footnote-36) Except for seeking the utilitarian rather than her own good, she is an opportunistically rational actor and conventional decision and game theory are directly applicable to the normative act utilitarian exercise.

In act utilitarian rational choice theory the “given preferences” of the actor are no longer those that nature or nurture have bestowed on the actor before the actor became “guided” by the act utilitarian theory. However, after the actor adopts the utilitarian normative theory, her preferences are “given” by the value theory on which the specific brand of utilitarianism is based. She makes these standards her own and in that sense no longer evaluates the consequences according to her (“previous”) contingent desires: *To the extent that* an act utilitarian actor is morally motivated, she pursues the utilitarian non-moral good (pleasure, happiness and statisfaction of non-moral preferences) in opportunistically rational ways.

In its original form preference utilitarianism leaves open the substantive content of non-moral preferences. Other than in standard economics the non-moral preferences of *all* individuals concerned matter for the rational decision-maker as modeled by utilitarianism.[[37]](#footnote-37) It is “good” to satisfy individuals’ preferences, whatever they might be (a variant of the principle of beneficence which has “interesting” implications in case of, say, masochists or sadists) and consequently what is right depends on the contingent values around in a society.

**In sum**, according to act utilitarianism, the moral actor is supposed to be individually rational, i.e. *opportunistic, with respect to the moral aim of maximizing the good for all concerned*.

The actor is causally responsible for the causal consequences of her choices and only for them. *In particular, the question of what would happen if the actor would always or in general do something is as irrelevant as the question of what would happen if everybody would always act like the actor.*

The latter questions are not among the questions that an act utilitarian moral actor should pose. The criteria invoked by these questions refer to counterfactual rather than factual considerations. They are not in line with the rational taking of moral opportunities. The opportunistic actor focuses on the causal and, therefore, the factual consequences that she expects to bring about by her *choices*.[[38]](#footnote-38) These consequences must be optimal for all and according to the standards of the utilitarian value theory.

As a means of further illustration, imagine an environment with a modest amount of littering (see the VCM above for the abstract general background model). A little littering doesn't matter to anybody concerned. However, if there is a lot of littering, everybody regards it as a nuisance. In a situation where there is a single actor who can either drop a piece of paper where he stands or bring it to the next garbage can, nobody would notice the additional piece of paper on the ground as such. It would be *insignificant* for grading the degree of littering in individual perceptions. Going to the garbage can requires some effort. The costs for the individual are higher if he goes to the garbage can than if he stays put and just drops the paper. Since nobody will notice the individual piece of paper as making a difference, dropping it has no direct negative consequences for anybody else.

Obviously, behind **the veil of individual insignificance** the act utilitarian moral person is under a *moral obligation* to drop the paper where he stands. This is the choice that will maximize the *overall* good for society (as long as there are no further influences causally exerted by paper lying on the ground like triggering more paper being put there).

Likewise, since the individual vote is insignificant, a voter who can do so without anybody else noticing his action should choose to stay put and not to vote.[[39]](#footnote-39) Of the two possible “lotteries of election results” that he can choose, he should choose the lottery in which he does not incur the cost of voting. The cost of voting is certain while the incremental increase of the likelihood that a specific party will win due to the *individual* vote or that the *individual* vote will have some other positive effect is so small that it is negligible. As long as there is a sufficient number of other individuals who are likely to go to the polls, the act utilitarian *should* choose not to vote. This is the action that will maximize *overall* good – of which his is a part – and is, therefore, morally right.

As has been pointed out frequently, if everybody were to act as an act utilitarian, problems could arise. For instance, were all to follow the act utilitarian precept, there would not be a sufficient number of people who avoid littering. By observing act utilitarian recommendations, even Pareto inferior results may arise, i.e. everybody would be better off if the act utilitarian precept to seek the best for all would not be adopted. Likewise, in the example of attending the polls (or not), the result might be a “mixed equilibrium” of very low voter turnout (see Palfrey and Rosenthal 1984, 1985). According to equilibrium play, there is a very low likelihood that everybody would go to the polls. In the end, so few people would go with sufficient likelihood that the individual contribution would be sufficiently significant to compensate for the cost incurred.

**In short,** it can be objected that an act utilitarian theory might suffer from the perhaps most serious defect of a normative theory: it might be self-defeating when realized.

The act utilitarian might still object that *all* consequences must be taken into account. For instance, negative effects on the legitimacy of democracy are among the consequences.[[40]](#footnote-40) There are other counter-strategies for the act utilitarian, like drawing attention to institutions and higher order choices concerning such institutions. Yet, rendering so-called “free-riding” a morally obligatory act – even if it should only be prima facie so – certainly raises a problem of “moral psychology” for an act utilitarian theory. Strict case-by-case, act-by-act opportunism in the pursuit of the moral good[[41]](#footnote-41) is intrinsically motivated by ethical consequences. But there is no appropriate place for intrinsic motivation by *rules*. Act utilitarianism must go out of its ways to construe a basis for the typical ethical commitment to rules. Psychologically, things are much easier if the basic justificatory theory is already cast in terms of rules and thus in terms other than rational opportunism.[[42]](#footnote-42)

### 6.2.2 Rule utilitarianisms

I will touch on merely two main classes of rule utilitarian ethical theories. The first variant to which I refer as “rule utilitarianism-1” or, following Hoerster, as RU-1 for short, emphasizes as a criterion what as a *matter of fact* is *regularly* leading to good consequences. The second variant, to which I refer as “rule utilitarianism-2” or RU-2 henceforth, relies on the *contrary to fact* criterion of what would have good consequences if everybody acted in a specific way (see Hoerster 1971/1977).

 **RU-1** allows for rules of thumb and institutions as well as institutional obligations. If a certain institution leading to regular behavior of a specific kind has *in general* beneficial consequences then there is a good ethical reason to accept it. In a way, the rules fix our *aspiration levels* as moral subjects. Rather than acting as fully opportunistic ethical choice-makers, subjects now follow established practices, institutional rules, and rules of thumb. In the pursuit of their ethical standards they will not anymore instantaneously switch their behavioral gears whenever they perceive a chance of improvement.

Relying on the previous analogy, we may again state that the management of our ethical affairs becomes “management by exception”. There is no strong moral obligation not to deviate from fixed rules nor is there any good reason against changing the rules if that seems to be clearly desirable. But there is a role for rule-guided behavior as an “organizational” tool (see Sidgwick 1874/2010).

**In sum,** in RU-1 established practices become a source of moral obligation. Though exceptions from established rules can be made whenever there seems to be a strong reason for doing that, rules have a *prima facie legitimate* claim to be observed.

According to RU-1, rules and institutions that have proved generally advantageous can command prima facie moral legitimacy: the burden of proof is on those who want to deviate from the rules or to change them as a whole. Utilitarianism becomes indirect in the sense that the institutional rules that as a matter of fact prevail have obligatory force because they do prevail. Even if *ultimately* only utilitarian criteria matter, on the *proximate* level of choice making established institutions define the agenda for the utilitarian. The burden of proof is on those who intend to change the established institutions defined by institutionalized rules.

RU-1 as an extension of act utilitarianism makes utilitarianism more realistic or workable. Like in the case of full as opposed to bounded rationality in explanatory or descriptive rational choice models – “explanatory utilitarianism” – the distinction between accepting rules as ethical standards and full ethical opportunism is far-reaching in normative utilitarianism. The intrinsic motivation by rules that can come into play in RU-1 as a real factor, changes the game of life. But the focus is on “average” causal consequences that *as a matter of fact* are expected to arise for all concerned. RU-1 is very much in the spirit of act utilitarianism in that regard.[[43]](#footnote-43) This changes once we turn to what has been called rule utilitarianism of the second type or RU-2.

In **RU-2** the criterion of the rightness of an act is *counterfactual*. Central is still the goodness of consequences for all concerned but we are dealing with consequences that *would* emerge if everybody did the same. It is not, at least not primarily, considered what everybody does and what the consequences of this will be. It is rather an anticipation of what *would* happen if everybody *would* act the same way *quite independently* of what everybody *in fact will do*.

In a somewhat simple-minded variant of the approach the expectation of the good that *would* arise from *universal* compliance renders participation morally obligatory regardless of how many others will in fact participate. In this view that something would have desirable consequences for all concerned if everybody acted accordingly is a sufficient reason for making the corresponding individual acts morally obligatory. Likewise if the performance of some act by everybody would have bad consequences then in the simple minded variant of contrary to fact reasoning this is sufficient for making the performance of that act wrong even if not all are going to perform the act.

For instance, if it would have good consequences for all concerned if all would participate in building a dike then any single individual would be under the obligation to build his part of the dike. However, since this is a so-called weakest link technology setting,[[44]](#footnote-44) if only one does not do his part the whole dike will be useless. It seems somewhat ridiculous that the others are under the same moral obligation to contribute when there are some who do not do their part as when all do their part. Nevertheless counterfactual generalization seems to require this.

Assume that some 50% of those who are necessary to build the dike just stand aside and leave undone 50% of what is necessary to have any protective effect. Can we plausibly claim that all are still obliged to participate since building the dike is what would be good if everybody did the same? It is true that it would have beneficial consequences if everybody would participate. Yet as long as not all do, the consequences of participation are not beneficial. Quite to the contrary, a considerable amount of effort that could have been avoided or been put to better uses is wasted.

The weakest link technology exemplified by the building of dikes is a special example. It makes universal participation more plausible than other technologies since it rules out “free-riding”. Allowing for individual insignificance and free-riding it is obvious how the counterfactual generalization principle would operate more generally if other kinds of technology would prevail. If everybody would do the same and take a free ride the consequences would be harmful. However, as long as only a few – and possibly in a concealed way – would not contribute to the collective good the merely potentially harmful consequences may in fact remain unfelt entirely. No harm is done and some good is gained by the increase of good consequences for those taking the free ride.

Nevertheless free-riding is ruled out, since it would have bad consequences if all would be taking the free ride. From the point of view of intuitive norms of fairness this may seem a very plausible result. It seems doubtful though, that even then the good consequences of general obedience should be deemed sufficient for rendering it a moral obligation not to take the free ride.[[45]](#footnote-45) In any event, regarding a contrary to fact assumption as sufficient for imposing a moral obligation goes against the grain of the original utilitarian impulse of evaluating the moral value of acts in terms of their *causal* consequences. The action must be evaluated in terms other than the direct causal consequences. And, under certain conditions it must be chosen despite causal consequences suggesting otherwise.

**In short,** RU-2 is a consequentialist ethical theory only by name.

RU-2 is cast in terms of counterfactual consequences that as a matter of fact are not at all consequences in the real world. They are consequences in a possible world. “Unreal” consequences should be “of no consequence” within a theoretical approach focusing on real causal effects. It seems that RU-2 is at best a secondary type or a far relation of utilitarianism: RU-2 should not count as utilitarian in the proper sense of that term at all.

Though it may not be **Kantian** in the full sense **RU-2** comes very close to a Kantian moral position in some regards. It is not Kantian in that Kant’s categorical imperative is a rule of reason that claims to be morally justified independently of any considerations of good or bad consequences. At least for Kant the categorical imperative is derived on a priori grounds from reason itself. It would be self-contradictory (whatever that means in the sphere of action) not to follow reason.

**In short**, whenever Kantianism is rendered plausible by Kantians, arguments of the type offered in support of RU-2 are typically used.

Requiring the maximization of good consequences for all concerned does not do in RU-2. RU-2 becomes a **dualistic ethical approach** combining the two principles of **beneficence and fairness**. “Mixing” can avoid some gross types of unfairness that might result from focusing exclusively on maximizing good.[[46]](#footnote-46) The downside is that the two principles of fairness and the maximization of good consequences may clash and the dualistic ethical theory in itself may not have the means to resolve such contradictions as may arise. Then we will have to go back to the search for a reflective equilibrium.

## 7. Should the numbers count and how?

The standard argument against utilitarianism is based on what has come to be known as “the survival lottery” which is in fact merely a variant of the Trolley problem discussed above. In the thought experiment of the survival lottery one has to imagine a patient in a hospital. The patient is there for the sole purpose of getting some routine checkup. The outcome of the checkup of this, as we assume, 50 year old man is that he is healthy, all of his organs are well working and his statistically remaining life expectancy is, say, 30 years. At the same time there are three other patients in the hospital who all will die on short notice unless they receive an organ transplant. All of these patients are 20 year old and would each have an additional life expectancy of another 30 years of a quality comparable to that of the 50 year old should they receive a suitable organ.

The standard argument against the utilitarian ethical theory is that by its inner logic utilitarianism has to require that the one should be sacrificed for the benefit of the other three. Utilitarianism is accused to endorse the precept that the 50 year old person should be used as a means to the end of saving the three other individuals.

Indeed, it is not easy to see how that conclusion could be avoided in a utilitarian context. The fifty year old had already an additional thirty years of life beyond 20 while the three 20 year old persons could each prolong their lives by that span should the fifty year old be used as an organ donor. Three times 30 years are clearly superior to one-time 30 years. If each counts for one and none for more than one why not use the one to rescue the three and the more so since the one had a longer life before?

Most people find such consequences outrageous. They think that utilitarianism could not be upheld in reflective equilibrium.

As a standard defense adherents of utilitarian approaches have pointed out that a utilitarian would take into account all causal consequences of an act like the one envisioned. If it would become known that people who go to routine checkups into hospitals might be used in the way described nobody would anymore go to a hospital. Such effects must be factored in and in the end the consequences of rescuing the three may be much worse than letting them die.

Though consequences as described in the utilitarian defense are likely to emerge the argument does not really suffice as a vindication of utilitarianism. For we might construe the example such that it would not become known how the one person disappeared in favor of the three. Then the counter example would still stand.

Here turning to institutional utilitarianism (a variant of RU-1 in concrete institutional terms) might help. An *institution* that would provide for the killing of innocent people to rescue others would *as a matter of fact* scare everybody and therefore might not be maximizing the good consequences for all concerned. As a matter of fact we could not rely on the assumption that certain institutional rules of action would remain secret. In an institutionalist utilitarian approach that takes fully into account all factual consequences of *institutionalized* *rules* the conclusion that we should favor rules that require abusing one for the good of others might not follow and under plausible empirical circumstances *in fact will not follow*.

Still, in a society in which rules that suggest the sacrifice of one for the benefit of others would be in place, the survival prospects of all concerned could be better than in societies without such rules. For that to be actually true people would have to learn and even would have to be trained to appreciate the benefits. If that could be brought about and if it could be secured that all would play by the rules, the society in which in case of emergency one would be used to rescue three may simply be the better lottery for each and every individual who does not have a clue in which role he is going to play in the lottery of life.

Now, there may always be spill over effects if we allow for institutional exceptions in institutional rules like those that forbid killings. But if that is the crucial argument against admitting the killing of one in favor of three then it is possible for the utilitarian to use the same argument to defend his position. He could simply say that he is well aware of such consequences and would draw attention to them in showing that utilitarianism would be maintained in a reflective equilibrium factoring in all consequences of institutional rules including the spill over effect of exceptions etc.

From an institutional point of view the preceding argument seems to be prima facie plausible. A world in which general rules demand that in situations such as trolley we should rescue more rather than fewer individuals seems to be preferable *from an ex ante point of view*. However, if we are in favor of such an institution in the case of the trolley problem why not in case of the hospital and the patients? Should we not openly declare that these things would be done if ever the necessity would arise and then let the game of life unfold?

It has been often said that we should not take innocent lives and therefore the patient who is in for the check up should not be sacrificed to rescue the three patients who are terminally ill. But the one on the tracks is as innocent as the three others in the trolley problem. Taking innocent lives is clearly a bad thing. However, as John Harris has pointed out in his original discussion of the survival lottery, not only the one is innocent the three to be rescued by the sacrifice are – or at least can be assumed to be – innocent lives as well (see Harris 1975). If the taking of an innocent life is bad then the taking of three innocent lives must be worse or numbers should not count at all. If the latter, what if we could rescue a thousand individuals at the price of sacrificing one or one hundred thousand at the price of one etc.?

**In sum**, to say that numbers never count is obviously strange. Yet, if they count, how precisely? How many lives does it take to justify the taking of one?

Without going into the details of the extended discussion centering around the previously mentioned questions it should be obvious that the so-called numbers’ problem is a rather serious one for our ethical intuitions. It is serious not only in utilitarianism but in all ethical theories. Most of us are pushed into a state of reflective disequilibrium if sacrificing or not sacrificing one for many or very many is the matter. Quite generally speaking, once the numbers are accepted to count there seems to be a latent danger of not respecting the fact that each person has only one life to live.

The sacrifice of a person may be seen as the ultimate violation of inter-personal respect. Yet it may be argued also that taking persons seriously as separate entities suggests that three such entities are even more important than one. (And, certainly ten million should be.) Precisely because persons are so important it is better to respect more of them than fewer, or so the argument could run.

All this, one might say, merely reminds us of the old wisdom that “hard cases make bad law” and that that wisdom extends to ethics. Indeed, we tend to think of ethical problems as tough problems. However, the primary function of moral codes and institutions that embody ethical views lies somewhere else. We need moral institutions to co-ordinate our daily activities rather than to solve extra-ordinary and hard problems. To put it slightly otherwise, as opposed to the ordinary problems of life the preceding examples involve *extreme scarcity of resources*. Not everybody can survive. In such a situation so-called “life-boat ethics” – if there is such a thing – may be necessary and adequate but ordinary ethics – adapted to cases of intermediate scarcity – will in all likelihood fail (see on Promethean ethics Garrett Hardin’s book of the same title (1980)).

In focusing on extreme situations which perhaps no ethical theory can adequately solve in a general manner critics of utilitarianism have always been rather unfair. They blamed utilitarianism for a failure in situations where their own theories would in all likelihood fail, too – though perhaps at the other end of the spectrum. For instance, if taken literally the principle “Do justice though the heavens fall!” is not very convincing.[[47]](#footnote-47) The Kantian logic according to which we have to tell the truth if a murderer on the way to his prospective victim asks where the victim currently is, does seem even more extreme than the implications of utilitarian views.[[48]](#footnote-48)

**In sum,** though utilitarianism does not resolve extreme problems convincingly it may be that no ethical theory can do so since the problems are so hard that almost any choice seems to be (in-)defensible.

The only “solution” would be that we should not be confronted with the tragic choice in the first place. However, we cannot by issuing ethical demands that they should not exclude **tragic scarcity** from emerging. We better prepare ourselves for such situations. Tragic scarcity exists even in the rich Western World. It emerges on a daily basis in cases like organ allocation and with an increase in our technical powers situations of tragic scarcity may – ironically – become more than less frequently.

As far as cases like that of organ allocation are concerned it is significant, however, that there are *institutions* to deal with them. It is also significant that the rules of the institutions that allocate scarce organs are developed in the shadow of basically utilitarian considerations.[[49]](#footnote-49) A stronger focus on institutions rather than on theories of ethics – on moral institutions rather than moral theory – may in general be the best way to deal with real world ethical problems and in particular in a business context. It is clearly here that an integrated management philosophy and economics approach to ethics should have a comparative advantage over traditional ethics.

Quite in line with that an institutional utilitarian theorist might say that he intends to form an ethical theory that suggests *moral institutions* for situations of “intermediate scarcity”. The situations are not like paradise, but at the same time the scarcity is not catastrophic. The situations are neither like hell nor like heaven. Exactly for such situations of intermediate scarcity that ethical theorists from David Hume to John Rawls intended to formulate their ethical theories. A final remark on the Rawlsian as the presently most influential ethical theory may be useful.

## 8. Rawls, the separateness of persons and consequentialist trade-offs

John Rawls’ basic criticism of utilitarianism focuses on what may be called the implicit collectivism of that theory. He detects this defect even in variants of utilitarianism that start from quasi-contractarian or liberal norms of equal interpersonal respect. Rawls concedes that he and modern liberal utilitarians like the economics Nobel laureate John Harsanyi share many general background conceptions of their search for (wide) reflective equilibrium.[[50]](#footnote-50) Harsanyi treats individuals as the ultimate source of value. As each of the ten contests in the decathlon ranks an alternative in that contest, each individual ranks an alternative as it is affected by its implementation. Rawls accepts that Harsanyi is paying respect to individual values in that regard. But Rawls believes that Harsanyi goes astray when he introduces specific forms of interpersonal comparisons of utility (the conversion rates in the decathlon example).

**In short**, Rawls accepts that Harsanyi respects persons as the ultimate source of value and takes them seriously in that regard. But from a Rawlsian point of view he does not take their “**separateness**” as seriously as it should be taken since he searches for ways to make their “values” commensurable.

Unlike some economists, Rawls, in offering his criticism, does not say that inter-personal comparisons of utility are “impossible”. This would be a mistake anyway since even according to Lionel Robbins’ discussion intra-personal comparisons of the “utility” (pleasure, preference satisfaction, etc.) of other individuals are “possible” or can be made *subjectively* by *some individual* or other. Any person can form her own “subjective” moral value judgments concerning the well-being of society at large as based on her assessment of the well-being of individuals. However, there is no “scienitific”, there is not “the” interpersonal comparison of well-being. The evaluation is valid only from the point of view of a specific evaluator.

In short, within the limits of modern “Robbinsian” welfare economics the **comparison of utilities** is always **evaluator relative**.

Assume that there is a collectivity of individuals who as evaluators intend to pass ethical judgment concerning the welfare of a collectivity of r=1, 2, …, n individuals. Each individual r *can* form ethical preferences and represent them by for a collectivity of individuals r=1, 2, …, n as based on her own welfare-comparing judgments. Therefore personal welfare functions *can* be formed without going beyond Robbins’ meta-ethical constraints on normative economic argument. Every individual can on herself form a comparison between, say, the relative urgency of the needs of separate persons who compete for the satisfaction of their needs. In particular by putting herself in the shoes of each of those in need, each individual r, can form her own ***personal* ethical preference** about how interpersonal conflicts of interest should be resolved. So, it is clearly not “impossible” to form *personal* comparisons of the utility of different persons. Nor is it something alien to everyday life and experience. It is a very common activity we perform permanently as human beings.

The “interpersonal” comparisons cannot claim inter-personal validity for themselves. Moreover, there are no compelling reasons why evaluator-relative judgments of different individuals should always converge.[[51]](#footnote-51) And even if there were a process bringing about convergence there is no reason why the comparisons that are consented would be “right”.[[52]](#footnote-52) To put it slightly otherwise there is – at least at the present stage of the argument – no reason why of two diverging personal interpersonal comparisons of utility at least one must be wrong. There are simply two diverging views on the relative urgency of satisfying the preferences of the one rather than satisfying the preferences of the other individual.

According to the preceding, there could be as many utilitarian views of the world as there are utilitarian evaluators. Each of these utilitarian evaluators forms a personal utility function concerning the welfare of the collectivity (including interpersonal comparisons as made from his personal perspective).

Rawls main criticism concerns the form of the judgment. By forming a welfare judgment for the collectivity as a function of individual measures of well-being it does not take the “**separateness of persons**”seriously. Rawls reservation is based on the complaint that utilitarians in producing their (personal) welfare functions treat *several persons as if they were one*. This criticism applies completely independently of the issue of whether or not there are good reasons for inter-personally valid inter-personal comparisons of utility or not. The criticism concerns the inter-personal comparison that each person might make separately. Neither is the focus of Rawls’ criticism evaluator-relativity per se.

In sum, for Rawls the problem with utilitarianism is the *form of the judgment* on substantive ethical issues*.* A moral person would not base her judgment on indifference levels and rates of *substitution of value between* individuals.

According to Rawls, in our personal ethical judgment human persons are not substitutes for each other. They are perceived as so unique and separate that ethical judgments concerning the basic structure of society cannot be represented as functions of individual measures of well-being, of preference satisfaction or the like.

Though one can understand what Rawls is driving at, one basic argument against his position is entirely obvious: Like in the race between Achilles and the Tortoise for which a certain theory shows that Achilles will never be able to pass by the Tortoise, human individuals who act in ways that affect a collectivity cannot but benefit in practice some at the cost of not benefitting others. The dictum “solvitur ambulando” (it will be solved by walking) as applied to the Achilles paradox applies to collective action analogously. Even if costs and benefits are not explicitly calculated some will benefit at the cost of others once collective action under conditions of scarcity takes place. Unless a Pareto inferior state is changed by a Pareto improvement some will win and others will lose. The idea of a social consent/contract is nonsense upon stilts then.

# Closing the loop: For those who are interested in exercising what has been sketched in this very short introduction to business ethics it may be interesting to return to the simple initial cases now and to actually search for wide reflective equilibria on them. The proof of this pudding is as that of all others in the eating or the practice of ethical reflection on practical ethical business problems.V. References

Alchurron, Carlos and Bulygin, Eugen (1971), *Normative Systems*, Springer: Berlin et al.

Arendt, Hannah (1951), *The Origins of Totalitarianism*, Harcourt: New York.

Baurmann, Michael (2002), *The Market of Virtue: Morality and Commitment in a Liberal Society*, Kluwer Law International: Dordrecht.

Bazerman, Max H. and Tenbrunsel, Ann. E. (2011), *Blind Spots*, Princeton: Princeton University Press.

Benson, Bruce (1990), *The Enterprise of Law. Justice Without the State*. San Francisco: Pacific Research Institute for Public Policy.

Bentham, Jeremy (1843), 'Anarchical Fallacies; Being an Examination of the Declaration of Rights Issued during the French Revolution', in: J. Bowring (ed.) *The Works of Jeremy Bentham*, 2, William Tait: Edinburgh.

Bikhchandani, Sushil, Hirshleifer, Jack and Welch, Ivo (1992), 'A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades', *The Journal of Political Economy*, 100 (5, October), 992-1026.

Binmore, Kenneth G. (2009), *Rational Decisions (Gorman Lectures in Economics)*, Princeton: Princeton University Press.

Brennan, H. Geoffrey and Buchanan, James M. (1985/1999), 'The Reason of Rules', *The Collected Works of James M. Buchanan,* 20, Liberty Fund: Indianapolis.

Brennan, H. Geoffrey and Buchanan, James M. (1984), 'Voter Choice: Evaluating Political Alternatives', *American Behavioral Scientist*, 28(2, November/December), 185-201.

Brennan, H. Geoffrey and Lomasky, Loren (1984), 'Inefficient Unanimity', *Journal of Applied Philosophy*, 1(1), 151-163.

Brennan, H. Geoffrey and Lomasky, Loren E. (1993), *Democracy and Decision*, Cambridge University Press: Cambridge.

Buchanan, James M. (1965), 'Ethics, Expected Values, and Large Numbers', *Ethics*, 76, 1-13.

Buchanan, James M. (1999 ff.), *The Collected Works of James M. Buchanan*, Liberty Fund: Indianapolis.

Buchanan, James M. and Vanberg, Victor J. (1990), *Constitutional Choice, Rational Ignorance and the Limits of Reason,* Mimeo: Fairfax.

Chamberlin, Edward H. (1948), 'An Experimental Imperfect Market', *Journal of Political Economy*, 56, 95-108.

Coleman, James S. (1983), “Free Riders and Zealots.” In *Ökonomische Erklärungen Sozialen Verhaltens*, edited by W. Sodeur, 135 ff. Duisburg: Verlag der Kooperative.

Coleman, James S. (1988), 'Free Riders and Zealots: The Role of Social Networks', *Sociological Theory*, 6 (Spring), 52-57. (abridged version of Coleman 1983).

Davis, Douglas D. and Holt, Charles A. (1993), *Experimental Economics*, Princeton University Press: Princeton.

Diekmann, Andreas (1985), 'Volunteer`s Dilemma', *Journal of Conflict Resolution*, 29 (4, December), 605-610.

Fox, Justin (2011), *The Myth of the Rational Market: A History of Risk, Reward, and Delusion on Wall Street*, Harper Business: New York.

Frank, Robert (1988), *The Passions within Reason: Prisoner's Dilemmas and the Strategic Role of the Emotions*, W. W. Norton.: New York.

Frankena, William K. (1988), *Ethics 2nd edition*, Pearson: London.

Freeman, Edward R. (1984), *Strategic Management: A stakeholder approach,* Pitman: Boston.

Friedman, Milton (1970), ‘The Social Responsibility of Business is to Increase Its Profits’, *The New York Times Magazine*, (September).

Gode, Dhananjay K. and Sunder, Shyam (1997), 'What Makes Markets Allocationally Efficient?', *The Quarterly Journal of Economics*, 112 (2, May), 603-630.

Güth, Werner and Kliemt, Hartmut (2000b), *From full to bounded rationality. The limits of unlimited rationality*, Center for Interdisciplinary Research (ZiF): Bielefeld.

Güth, Werner and Kliemt, Hartmut (2013), 'Fairness That Money Can Buy: Procedural Egalitarianism in Practice', *Rationality, Markets and Morals*, 4, 20-30.

Greene, Joshua (2013), *Moral Tribes: Emotion, Reason, and the Gap Between Us and Them*, Penguin Press: New York.

Hahn, Susanne (1996), 'Überlegungsgleichgewicht und rationale Kohärenz', in: Karl-Otto Apel and Matthias Kettner (eds.), *Die eine Vernunft und die vielen Rationalitäten*, Suhrkamp: Frankfurt a.M., 404-423.

Hahn, Susanne (2000), *Überlegungsgleichgewicht(e). Prüfung einer Rechtfertigungsmetapher*, Karl Alber: Freiburg i.Br.

Hahn, Susanne (2004), “Reflective Equilibrium: Method or Metaphor of Justification?” In *Wissen und Glauben*, edited by W. Löffler and Paul Weingartner, Hölder-Pichler-Tempsky: Wien, 237 – 243..

Hardin, Garrett J. (1968), 'The Tragedy of the Commons', *Science*, 162, 1243-1248.

Hardin, Garrett J. (1980), *Promethean Ethics: Living With Death, Competition, and Triage*, University of Washington Press: Seattle.

Harris, John (1975), 'The Survival Lottery', *Philosophy*, 50, 81-87.

Harsanyi, John C. (1976), *Essays on ethics, social behavior, and scientific explanation*, D. Reidel Pub. Co.: Dordrecht, Holland; Boston.

Hart, Herbert L. A. (1961), *The Concept of Law*, Clarendon Press: Oxford.

Hayek, Friedrich A. (1973-79), *Law, Legislation and Liberty: A New Statement of the Liberal Principles of Justice and Political Economy*, Routledge & Kegan Paul: London

Hayek, Friedrich A. (2001), *Road to Serfdom*, Taylor & Francis.

Henrich, Joseph, Boyd, Richard, Bowles, Samuel, Camerer, Colin, Fehr, Ernst et al. (2004), *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*, Oxford University Press: New York.

Hobbes, Thomas (1651/1976), *Leviathan,* Ullstein: Frankfurt am Main.

Hoerster, Norbert (1971/1977), *Utilitaristische Ethik und Verallgemeinerung*, Karl Alber: Freiburg/München.

Hohnisch, Martin, Pittnauer, Sabine and Selten, Reinhard (2011/2012), 'Dealing with Dynamic Decision Problems when Knowledge of the Environment Is Limited: An Approach Based on Goal Systems', *Journal of Behavioral Decision Making*, 25(5), 443-457.

Hume, David (1739/1978), *A Treatise of Human Nature*, Clarendon: Oxford.

Jakob, Thomas (1996), *Die Ethik John C. Harsanyis,* Gerhard Mercator Universität Duisburg.

Janssen, Pottee and Ostrom 2010

Jasay, Anthony de (1995), *Social Contract – Free Ride*, Oxford University Press: Oxford.

Jouvenel, Bertrand de (1994), *On Power: The Natural History of its Growth*, Liberty Press: Indianapolis.

Kay, John (2011), *Obliquity. Why Our Goals Are Best Achieved Indirectly*, Profile Books: London.

Kliemt, Hartmut (1986a), 'The Veil of Insignificance', *European Journal of Political Economy*, 2 (3), 333-334.

Kukathas, Chandran (2003), *The liberal archipelago : a theory of diversity and freedom*, Oxford University Press: Oxford.

Lehrer, Keith and Wagner, Carl (1981), Rational Consensus in Science and Society, Kluwer Academic Publishers: Dordrecht.

Mackie, John L. (1980), *Hume's Moral Theory*, Routledge: London.

Mackie, John L. (1991), Ethics. Inventing Right and Wrong, Allen Lane: London.

Mayer, Colin (2013), Firm Commitments. Why the Corporation is Failing Us,

 Oxford University Press: Oxford.

Mill, John Stuart (1992), *On* Libertyand *Utilitarianism*. Everyman’s Library. New York: Knopf : Distributed by Random House.

Olson, Mancur (1965), *The Logic of Collective Action*: Cambridge, Mass.

Ostrom, Elinor (1990), *Governing the Commons. The Evolution of Institutions for Collective Action*, Cambridge University Press: New York.

Ostrom, Vincent A. (1971), *The Political Theory of a Compound Republic*, University of Nebraska Press: Lincoln and London.

Palfrey, Thomas R. and Rosenthal, Howard (1984), 'Participation and the Provision of Discrete Public Goods: A Strategic Analysis', *Journal of Public Economics*, 24, 171-193.

Palfrey, Thomas R. and Rosenthal, Howard (1985), *Private Incentives in Social Dilemmas*, Carnegie-Mellon University: Pittsburgh.

Piketty, Thomas. *Capital in the Twenty-First Century*. Cambridge Massachusetts: The Belknap Press, 2014.

Popper, Karl (2011), *The Open Society and Its Enemies*, Taylor & Francis.

Rawls, John (1951), 'Outline of a Decision Procedure for Ethics', *Philosophical Review*, 60, 177-190.

Rawls, John (1971), *A Theory of Justice*, Oxford University Press: Oxford.

Robbins, Lionel (1935), *An Essay on the Nature and Significance of Economic Science*, Macmillan: London.

Sandel, Michael J. (2012): *What money can’t buy. The moral limits of markets*, Allen Lane: London.

Selten, Reinhard (1986), 'Institutional Utilitarianism', in: Franz-Xaver Kaufmann (ed.), *Guidance, Control, and Evaluation in the Public Sector*, de Gruyter: Berlin and New York, 251 ff.

Sidgwick, Henry (2010), *The Methods of Ethics*, Nabu Press: Charleston.

Skyrms, Brian (1996). *Evolution of the Social Contract. Cambridge*, University Press: Cambridge.

Smith, Vernon L. (1962), 'An Experimental Study of Competitive Market Behavior', *Journal of Political Economy*, 70, 111-137.

Smith, Vernon L. and Williams, Arlington W. (1992). Experimental market economics, *Scientific American*, 267(December), 116–121.

Stodder, James (1998), 'Experimental Moralities. Ethics in Classroom Experiments' *Journal of Economic Education*, 29 (2, Spring), 127-138.

Sugden, Robert (1986), *The economics of rights, co-operation, and welfare*, Blackwell: Oxford/New York.

Surowiecki, James (2005), *The Wisdom of Crowds*, Trade Paperback: Anchor.

Taurek, John M. (1977), 'Should the numbers count?', *Philosophy & Public Affairs,* 6, (4, Summer).

Taylor, Michael (1976), *Anarchy and Cooperation,* John Wiley and Sons: London et al.

Taylor, Michael (1987), *The Possibility of Cooperation*, Cambridge University Press: Cambridge.

Taylor, Michael and Ward, Hugh (1982), 'Chickens, Whales, and Lumpy Goods: Alternative Models of Public-Goods Provisions', *Political Studies*, 30, 350-370.

Urmson, James O. (1958), 'Saints and Heroes', in: I Melden (ed.), *Essays in Moral Philosophy*, University of Washington Press: Seattle/London, 198 ff.

Zahavi, Amotz (1975), 'Mate Selection - A Selection for Handicap', *Journal of Theoretical Biology*, 53, 205-214.

Zahavi, Amotz and Zahavi, Avishag (1997), *The Handicap Principle. A Missing Piece of Darwin's Puzzle*, Oxford University Press: New York, Oxford.

1. Reference of the provider of the case: *The Cases in the Business* & *Society series provide fictional situations which illustrate ethical issues in business designed to prompt discussion in college classes. This series is provided without charge to educators for noncommercial use through financial support from the George and Judy Writer family. For additional cases, please visit our website:* [*http://leeds.colorado.edu/businessandsociety.*](http://leeds.colorado.edu/businessandsociety.)© 2006 Center for Business and Society, Leeds School of Business, University of Colorado, Boulder, Colorado, USA. All rights reserved. This article may be duplicated and disseminated for strictly non-commercial and educational purposes. [↑](#footnote-ref-1)
2. The reader might want to note that „ethics“, „ethical“ etc. are used here in the way medical professionals use the term. The term is not primarily referring to critical ethics as an academic discipline but rather to codes of conduct and institutionalized sets of norms. [↑](#footnote-ref-2)
3. By an „explication“ a pre-theoretical concept is substituted by a similar, more precise, theoretically fruitful and simple concept. For a fine presentation of this basic philosophical „technique“ introduced explicitly by Rudolf Carnap, see Alchurron and Bulygin 1971. [↑](#footnote-ref-3)
4. Setting aside the issue whether the value of the single share or the sum of value of all shares should be increased. Issuing new shares or buy-backs may exert an influence here. [↑](#footnote-ref-4)
5. There was an analytical solution based on the application of crunching power but that did not play a role for actual decision-making. [↑](#footnote-ref-5)
6. The reader might want to consult the two original articles of Freeman and Friedman here. I should like to point out that I fully acknowledge that I do not have a legitimate claim to present “the” right answer to the questions involved. I claim only that mine is one defensible answer. [↑](#footnote-ref-6)
7. Whether the subsequent sketch that owes quite a bit to the presentations of Chinese students in business and corporate ethics courses at the Frankfurt School of Finance & Management is really true to the facts is less relevant than the principles illustrated by what may be regarded as an exercise in economic story-telling. [↑](#footnote-ref-7)
8. In some ways this echoes the polycentric model of social organization propagated by Vincent Ostrom (1971) and, of course, Elinor Ostrom, see for instance Janssen, Pottee and Ostrom 2010). It “brings” elements of the original human organization of so called “primitive” societies into the modern world. [↑](#footnote-ref-8)
9. Similar ideas are developed independently under the somewhat misleading title of *Property as a guarantor of liberty,* by James M. Buchanan; see vol.18 of Buchanan 1999ff; my own work on this can be found e.g. in Kliemt 1985. [↑](#footnote-ref-9)
10. For the time being, it should suffice to point out some excellent literature, Olson 1965, Taylor 1976, Taylor and Ward 1982, Taylor 1987, de Jasay 1995 and in an applied way Ostrom 1990. [↑](#footnote-ref-10)
11. This is one of the most robust findings of experimental economics and already well documented in Davis and Holt 1993. [↑](#footnote-ref-11)
12. See on the crime story Frank 1988 and on volunteer’s dilemma more generally Brennan and Lomasky 1984, Diekmann 1985). [↑](#footnote-ref-12)
13. 1. Bentham reads more fully: “But reasons for wishing there were such things as rights, are not rights; -- a reason for wishing that a certain right were established, is not that right -- want is not supply -- hunger is not bread.“ Jeremy Bentham, *Anarchical Fallacies*, vol. 2 of Bowring (ed.), Works, 1843, Article 2 [↑](#footnote-ref-13)
14. Predictions of consequences of actions are the most important factor for our deliberations. Predictions become possible to the extent that we have an explanatory theory of human behavior that contains behavioral laws. Therefore we have a moral responsibility to research such relevant facts, too. [↑](#footnote-ref-14)
15. A point made in particular by John Stuart Mill in „On Liberty“,1992 [↑](#footnote-ref-15)
16. see Hahn 1996, 2004 [↑](#footnote-ref-16)
17. We are postponing the addition of some of the bells and whistles until a later stage of the present discussion. [↑](#footnote-ref-17)
18. In the preceding cases the reflective equilibrium, RE, was always strictly local in the sense of involving only one general principle and several specific instances. [↑](#footnote-ref-18)
19. Obviously, factors other than coherence played a role as well, but perhaps coherence was at least one of them. [↑](#footnote-ref-19)
20. It should, however, be noted that entering the course of ethical discourse has its own internal dynamics. There may be what Peter Singer aptly called an “escalator effect” here; once we step onto the escalator of moral discourse, it may take us somewhere against our own will or doing. [↑](#footnote-ref-20)
21. The use of the terms of “primary” and “secondary” rules is inspired by Herbert Hart’s corresponding use of the terms in relation to law. It may be useful to note that there is also John Stuart Mill’s distinction between primary and secondary principles. This distinction focuses on the fundamental ethical theory, on the one hand, and the practice guiding the rules, on the other. We will come back to that. [↑](#footnote-ref-21)
22. It is, to use the nice metaphor of Joshua Greene (2013), as if we switch from “preset motive” to the “manual mode” of a camera. Normally we will work with the preset routines but are able to switch to the manual mode. [↑](#footnote-ref-22)
23. This is again close to Mill’s distinction between the primary and secondary moral principles but even closer to Hayek’s self-conception as an indirect utilitarian. Yet, the position developed here will differ from both in that it will be particularistic or non-universalistic. [↑](#footnote-ref-23)
24. More specifically, at least under *normal circumstances* he would not have good reason to deviate from the established practiced unilaterally as assumed here; however, sometimes the best way to change a practice may be deviant behavior. [↑](#footnote-ref-24)
25. These terms are used by Buchanan and Vanberg (1990) when discussing these issues. The terms, however, reflect only much older insights already emphasized by Hobbes. [↑](#footnote-ref-25)
26. This is clearly Mackie’s view as presented in 1980 and 1991 [↑](#footnote-ref-26)
27. Incidentally the observation that as a rule there are always several possible reflective equilibria is of broadly the same nature; see above. [↑](#footnote-ref-27)
28. The game theory underlying this is presented in Sugden (1986) and also Binmore (2009) or Skyrms (1996) to name a few in the Humean tradition. [↑](#footnote-ref-28)
29. The transparency requirement in the account of Friedman above shows how essential honest dealings are in business ethical contexts. [↑](#footnote-ref-29)
30. All subsequent remarks are deliberately one-sided in leaving out the Marxist variants of normative ethical and normative political thought. From a polemical point of view, see Friedrich August von Hayek’s *Road to Serfdom* (2001), Karl Popper’s *The Open Society and Its Enemies* (2011), Bertrand de Jouvenel’s *On Power* (1994) and, last but not least, Hannah Arendt’s *Origins of Totalitarianism* (1951). [↑](#footnote-ref-30)
31. A non-teleological consequentialist evaluates actions in terms of consequences. However, the ranking of the consequences may itself be based on moral evaluations rather than restricted to non-moral considerations. [↑](#footnote-ref-31)
32. For the welfare economically interested: The strict adherents to modern welfare economic priciples would not accept the Kaldor Hicks criterion according to which an improvement can be diagnosed if all losers of a collective measure could conceivably be compensated. They would rely on compensation actually paid and agreement of all actually received. It is in this sense rather Buchanan-Wicksell unanimity and compensation. [↑](#footnote-ref-32)
33. Classical welfare economics is often called Pigovian welfare economics as well, for it is associated with Pigou’s works. [↑](#footnote-ref-33)
34. See for instance the relevant parts of my *Philosophy and Economics I,* available on the net [[http://www.frankfurt-school.de/content/de/education\_programmes/bachelor/mpe/ mpe\_content.html](http://www.frankfurt-school.de/content/de/education_programmes/bachelor/mpe/%20mpe_content.html)] [↑](#footnote-ref-34)
35. A theory of what is valuable is sometimes also called an “axiology”. [↑](#footnote-ref-35)
36. She must ask herself two questions, namely who is concerned (humans, sentient beings, persons etc) and what it is that is of concern (happiness, pleasure etc) before she then evaluates the causal consequences of her actions according to the answers to the two questions. [↑](#footnote-ref-36)
37. The principle that only Pareto improvements should be brought about is one way to respect the preferences of all. For then everybody has an implicit veto against any change. No ranking of collective alternatives can go against the preferences of any single individual. [↑](#footnote-ref-37)
38. Of course, the very concept of causality is intimately related to contrary to fact assumptions about what are necessary and sufficient conditions that can or cannot lead to alternative „effects“. [↑](#footnote-ref-38)
39. Of the rich literature on this, the work most relevant in the present context can be found in Brennan and Lomasky (1984). For a short account of the condition of individual insignificance, see Kliemt 1986a. [↑](#footnote-ref-39)
40. Then, however, the problem may re-emerge on the level of determining who is to contribute by going to the polls. [↑](#footnote-ref-40)
41. It may not be so for descriptive utilitarianism if the perspective of expressive voting or a factual rather than ethical “boo and hurrah” theory is adopted; see Brennan and Buchanan 1984, Brennan and Lomasky 1993). [↑](#footnote-ref-41)
42. Practices are always rule-guided and restrain the local maximization of choice makers whether they are ethically motivated or not. [↑](#footnote-ref-42)
43. But one should not be led astray here to the conclusion that rule following even if it consists in following rules of thumb that are not regarded as binding would not change the descriptive or explanatory account of what happens in the world. Even if individuals have the mental reservation that they could always go back to “square one” and calculate in an act utilitarian case by case manner what the morally most advantageous action in a situation might be as long as they do not go back there they reach the other squares of the playing field by means other than case by case calculation. This is already sufficient to bring into play a new type of behavior. [↑](#footnote-ref-43)
44. This is, in a sense, the “reverse” of a volunteer’s dilemma in which a single act suffices to bring about the collectively desirable and desired result but nobody volunteers because it is costly for each individual and all speculate that somebody else might be doing the job, see Dieckmann 1985. [↑](#footnote-ref-44)
45. Very interesting treatments of related problems can be found in de Jasay 1995 and in Taylor and Ward 1982 [↑](#footnote-ref-45)
46. Basically the ethics supported in the standard introductory text on ethics, Frankena 19882. [↑](#footnote-ref-46)
47. See on the latter emphasis on low cost in utilitarian thought in particular Harsanyi 1976 and Selten 1986. [↑](#footnote-ref-47)
48. If we try to rescue Kant by referring to moral institutions that should operate under such rules because only such rules can make commitments sufficiently strong in practice we can as well accept that on the institutional level utilitarianism and institutional Kantianism converge. [↑](#footnote-ref-48)
49. For example HLA matching as a criterion of kidney allocation is defended in terms of increasing the survival prospects of the organ which in turn will increase the utility that can be derived from an organ transplant. [↑](#footnote-ref-49)
50. It may even be argued that there is in fact a conception of the search for RE in Harsanyi, see Jakob, Thomas 1996). [↑](#footnote-ref-50)
51. Two functions  may both represent impartial judgments without being identical. It is a rather common error to assume that individuals who are all aiming at universalization in judgment thereby will reach unanimity. Unanimity and universalizability are independent. [↑](#footnote-ref-51)
52. As in case of the rational consensus models of Lehrer and Wagner 1981 [↑](#footnote-ref-52)