

# The Limits of Economic Self-interest

The Case of Open Source Software

Jeffrey P. Powell

# The Limits of Economic Self-interest

The Case of Open Source Software

Proefschrift ter verkrijging van de graad van doctor aan de Universiteit van Tilburg op gezag van de rector magnificus, prof.dr. Ph. Eijlander, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op maandag 20 december 2010 om 10:15 uur door Jeffrey Parker Powell, geboren op 24 juli 1962 te San Francisco, USA.

Promotor: Prof.dr. J.F. Hennart

Promotiecommissie:

Prof.dr. S.W. Douma

Prof.dr. J.J. Graafland

Prof.dr. N.G. Noorderhaven

Prof.dr. B. Nootboom

## Acknowledgments

The four years I spent writing this thesis at Tilburg were very enjoyable. I was essentially left to my own devices to formulate and implement a research project. My research question was originally inspired by an encounter in the dormitory cafeteria at the university where I was a freshman undergraduate student. A fellow economics student observed that he didn't think that it was right that students with mobility disabilities were given golf carts by the university to help them get around campus. He reasoning was that it was acceptable that student athletes get golf carts upon receiving injuries because they can further contribute to the coffers of the university, however, those with disabilities can't make such contributions and therefore shouldn't receive any special treatment. Such calculative reasoning is indicative of a type of self-interest that is common and acceptable within standard economic theory. Economic man is often portrayed as non-ethical and as pursuing aims which are narrowly focused on a very select set of interests. But why have economists saddled themselves with such a narrow and mean view of human nature? Answering that question involves establishing the limits of self-interest in economic reasoning, the subject I attempt to address in the following pages. The contention presented is that such a limited view of motivation causes economists to misidentify important economic phenomena. Economics is often identified as the study of self-interest and rationality. I concentrate on the role of self-interest in economic theory and determining precisely what it means to be economically self-interested. What self-interest entails is a fundamental question in many disciplines, particularly in ethics; therefore, my approach has been to analyze economic interpretations of self-interest through the theoretical lens of ethics. Given the breadth of the topic, my more limited aim has been to set the broad boundaries of economic self-interest.

I would like to take the opportunity to thank several people who have helped me during the process of completing this manuscript. Thanks to professors Barkema and Hennart for allowing me to enter the economics program at Tilburg. I thank professors Douma and Noorderhaven for offering support when it was most needed. Thanks to professors Graafland and Nooteboom for their close reading of the manuscript and helpful comments. In particular, the last two pushed for clarity and precision, together they have made this a much better manuscript.

Most importantly, I would also like to thank my extended family, particularly professor Roeser, whose unwavering support and encouragement gave me the energy to complete this manuscript. Finally, I would like to thank my parents, who endowed in me the joy of learning.



## Contents

Problem Statement .....	1
Does it Matter? .....	3
Practical and Theoretical Implications .....	3
Normative Significance .....	6
Proposed Motives of Open Source Contributors .....	7
Internal Motivation .....	8
External motivation .....	9
Contribution .....	11
Methodology .....	12
Introduction .....	17
Basics of Self-Interest .....	17
Dimensions of Self-interest .....	20
Structural Dimension .....	21
External Dimension .....	23
Characteristics of Self-interest .....	24
Instant Hedonism .....	25
Pure Egoism .....	26
Prisoner's Dilemma and Pure Egoism .....	27
Self-Interest Theory .....	30
Practical Egoism and Altruism .....	32
Instrumental Egoism .....	33
Sacrificing Interests .....	36
<b>CHAPTER TWO: ALTRUISM AND ALTERNATIVES TO SELF- INTEREST .....</b>	<b>39</b>
Introduction .....	39
Particular Affections .....	39
Duties .....	41
Special Relationships .....	42
Natural Altruism .....	43
Thomas Nagel's Possibility of Altruism .....	45
Altruism .....	46
Practical Examples .....	53
Identifying Motivation .....	56
Focus of an Action .....	57
Agent Uncertainty and Self-Deceit and Community Standards .....	58
When is an act altruistic? .....	60
Pseudo-altruism .....	61
The Duration of an Act .....	62

Discussion and Conclusions.....	64
<b>CHAPTER THREE: SELF-INTEREST IN ECONOMIC ANALYSIS.....</b>	<b>67</b>
Introduction .....	67
Section One: The First Principle of Economics? .....	68
Self-interest in Economics .....	68
Preference Theory .....	73
Instrumental and Scientific Realism .....	77
The Purview of Economic Analysis.....	81
Two Models That Assume Self-interest.....	82
Transaction Cost Economics.....	83
Agency Theory.....	89
Conclusions Chapter Three .....	92
<b>CHAPTER FOUR: DIMENSIONS OF SELF-INTEREST.....</b>	<b>95</b>
Introduction .....	95
Structural Dimension Of Self-interest.....	95
Psychological Egoism .....	97
Normative Approaches.....	105
Other Approaches.....	114
External Dimension of Self-interest.....	116
Parfit's Objectives of Self-interest.....	118
Some General Aims .....	123
Regard for Others.....	126
Conclusions .....	130
<b>CHAPTER FIVE: TRACING THE FOUNDATIONS OF MODERN UNDERSTANDINGS OF SELF-INTEREST .....</b>	<b>133</b>
Introduction .....	133
Bernard Mandeville.....	134
Thomas Hobbes.....	138
Bishop Joseph Butler .....	143
Butler's Hierarchy of Motives .....	144
Adam Smith .....	153
Ronald H. Coase's Analysis of <i>The Theory of Moral Sentiments</i> .....	161
Self-interest in TMS and WN .....	162
Self-interest and the Advantages of the Market.....	167



<b>CHAPTER SIX: SELF-INTEREST IN HELLENISTIC PHILOSOPHIES.</b>	<b>169</b>
Introduction .....	169
The Cyrenaics.....	170
Epicurus.....	174
Stoics .....	180
Conclusions .....	187
<b>CHAPTER SEVEN: FRAMEWORK FOR ECONOMIC SELF-INTEREST</b>	<b>189</b>
The Purview of Economic Analysis.....	190
Structural Dimensions .....	195
External Dimension.....	197
General Framework.....	199
Identifying motivation.....	199
Does it matter? .....	199
What is the context under consideration? .....	200
What is the time period over which an agent's interests are measured?	
.....	201
Why isn't preference theory enough?.....	202
Is the objective to describe how agents act or how they should act?.	202
Is a subjective or objective view taken?.....	203
To what degree are the interests of agents interconnected?.....	204
At what intensity are interests pursued?.....	204
Conclusions .....	205
<b>CHAPTER EIGHT: THE CASE OF OPEN SOURCE .....</b>	<b>207</b>
Introduction .....	207
Practical Relevance .....	207
Open Source and motives.....	209
Basics of Open Source .....	213
Economic Context: Networks .....	215
History of Open Source.....	218
Learning from the Fragmentation of UNIX.....	218
History of Linux .....	221
GNU-GPL .....	224
Licenses and the Motives of Open Source Programmers .....	229
Business Motives .....	231
Individual Motives .....	233
The Analysis of Lerner and Tirole.....	233
Other Motives .....	245
Self-interest .....	245

Internal motivation and self-interest .....	246
External motivation and self-interest .....	253
Applying the framework .....	254
Conclusions Chapter Eight.....	258

<b>BIBLIOGRAPHY .....</b>	<b>261</b>
---------------------------	------------

\

## INTRODUCTION

### Problem Statement

The case of open source software raises questions about the fundamental economic assumption of self-interest. In general, open source programmers combine their talents via the Internet to produce high quality, financially valuable software which they then make conditionally available to anyone who wishes to download it. There are no employment contracts among the programmers and there is no formal hierarchical structure within a given software project, programmers are free to join and quit projects as they choose. Perhaps even more surprising from a business perspective, anyone, whether participating in the original software project or not, can freely acquire the code and begin their own, perhaps rival, software project. In effect, the code that results from open source projects exists independently of any one person or group. The licenses commonly allow anyone to use the resulting software for whatever purposes they desire and to examine, change, and redistribute those changes. One important category of open source license even allows licensees to take open source code and reissue it under proprietary licenses. Such licenses have created a large pool of publicly available software which anyone can use for whatever purposes they deem fit provided they abide by the relatively liberal open source licensing conditions. However, it is important to realize that not all open source licenses offer the same conditions. Regardless of the type of license used, open source software is an important economic phenomenon to which proprietary software firms have been forced to respond whether as contributors, collaborators or competitors. In particular, open source is theoretically important because of what it implies for the basic motivational assumption of economic models.

Economics has a specific theory about motivation, it assumes agents are self-interested. Furthermore, it assumes that they are rational. Self-interest and rationality are closely related; self-interest is a theory about rationality as opposed to say, a theory about ethics. At its most general level, economics gives to agents the aim to maximize their economic interests by optimally applying the scarce resources they control. Open source software is a significant economic production method that raises challenging questions about the overall applicability of the motive of self-interest in economic theory as that term is generally understood.

The question that Lerner and Tirole (2000, p.3) raised in one of the first papers on open source still stands:

“Why should thousands of top-notch programmers contribute freely to the provision of a public good?” Any explanation based on altruism only goes so far. While users in less developed countries undoubtedly benefit from access to free software, many beneficiaries are well-to-do individuals or Fortune 500 companies. Furthermore, altruism has not played a major role in other industries, so it would have to be explained why individuals in the software industry are more altruistic than others.”

The answers provided by Lerner and Tirole to the issues they raise rest on the assumptions of standard economic theory. The conclusions which they draw from the case of open source depend, and indeed support, their understanding of what it means to be economically self-interested. Whether the motive of self-interest is sufficient to explain open source is an open question and will depend on one's understanding of what it means to be self-interested. The question is not whether economic theory or its application is wrong, that would be presumptuous, the question is whether the assumptions of the theory are sufficient to explain the success of open source. Ambiguities in the understanding of what it means to be self-interest have led to the belief that nearly *all* activities can be explained using the economic apparatus and its narrow definition of self-interest. Self-interest is a powerful assumption; why would or should anyone choose to work against their own interests? If a convincing argument can be made that open source contributors are motivated by motives other than narrow self-interest, then perhaps other important economic production methods are the result of similar motives.

The general methodology I will use in this manuscript is to compare the explanatory power of the standard economic model with competing models. According to Popper, almost every problem of explanation in the social sciences requires an analysis of the social situation. He writes, (1994, p. 166): "The fundamental problem of both the theoretical and the historical social sciences is to explain and understand events in terms of human actions and social situations." Popper further states that the social situation and aims of agents under investigation need to be clearly stated and understood, they represent the initial conditions in the social sciences. The general social situation in the case of open source is clearly the competitive software market widely defined to include its accompanying institutions and history. The market represents the parameters or constraints under

which economic agents operate. The aims that are assigned to open source contributors under the given constraints are what need to be explored. Accordingly, the research question is: Does the assumption of self-interested agents optimizing the use of their scarce resources adequately explain the success of open source in the software market or are there better explanations?

My aim is not simply to define a better definition of economic self-interest. Although Popper's (1985, Chapter 6) argument on the futility of defining terms in the social sciences goes too far (agreement about the meaning of terminology at a general level can help to avoid misunderstandings and advance arguments despite what Popper writes), he is correct to warn that a fixation on getting the terminology right can lead to an infinite regress as the terms of terms need to be defined. Refining and adding nuance to the term "self-interest" as it is understood in economics is only the starting point, the important step is to "test" how different interpretations of self-interest and other motives are able to explain and predict behavior. If a better understanding of agent aims leads to a better model, then it should be used. Self-interest is a heavily loaded term with both popular and technical meanings, the implications of these various definitions need to be worked-out and alternatives need to be explored.

Does it Matter?

#### Practical and Theoretical Implications

Does it matter whether an act that is motivated by something other than self-interest or self-interest broadly understood is taken to be motivated by a narrow interpretation of self-interest? From a personal standpoint, probably not. If I falsely believe that an act which I perform is done for altruistic reasons when in fact it is motivated by self-interest, then little damage is done other than the issue of living an unexamined life. However, if a researcher presumes that the behavior he observes is motivated by self-interest when in fact it is motivated by something else, then problems can arise. Self-interest is the basic behavioral assumption used in economics. If, in fact, it is not always the prime motivator and alternative motives lead to substantially different behaviors, then the assumption should be adjusted or thrown out. Motives are about why people do what they do. If the belief is that people are self-interested, then the means used to motivate people will appeal to those interests, but if different motives are present, different incentives may need to be applied. For instance, as I will show below, if open source

contributors are assumed to be motivated by altruism, then appeals to self-interest may be counter-productive. Economists, certainly the researchers I discuss in this manuscript, and researchers in many other disciplines, presumably use the assumption of self-interest because it allows them to explain and predict behavior, they rightly go to great measures to justify their use of that particular behavioral assumption because assumptions matter when it comes to interpreting behavior and making practical recommendations.

From the perspective of doing business, it is the practical implications of widening the scope of motives at work which, finally, matters. If employees are motivated by a wider set of interests other than narrow, economic self-interest, then employers should be aware of it. For instance, the question of whether open source production methods can be incorporated within a firm or exploited from outside the firm, will depend on what motivates people to develop and contribute open source software. If employees or potential employees are motivated by something other than the extrinsic interventions that standard economic theory prescribes, then it might be possible to change the nature of work by changing employee incentives. For instance, if intrinsic motives, such as altruism, are motivating behavior, then it might be counter productive to appeal to an employee's narrow self-interests (Frey, 1997, Ryan and Deci, 2000). Frey's idea, simply put, is that extrinsic, monetary rewards may crowd-out intrinsic motives. If true, this may have direct implications for the use of open source software and production methods. His starting point is that economics is concerned with external types of motivation, for example wages, as opposed to internal motivation such as the enjoyment of, for example, performing the activity itself. It is a well documented and robust finding that external rewards carry a hidden cost in that they undermine intrinsic rewards (Lepper and Greene, 1978, McGraw, 1978). Experiments conducted by Frey and his colleagues at the University of Zurich show that a subject's altruistic tendencies were undermined when they were forced to share with others (Frey, 1997, pp. 14-16). Therefore, if open source is motivated in part by altruism, firms which attempt to use open source should be aware of these crowding-out effects.

Firms are beginning to explore possible business links with open source production methods. West and O'Mahony (2008) show how open source projects sponsored by proprietary firms have led to new methods of productions, while Dalle and Jullien (2003) explore the potential of creating efficient economic institutions based on motives similar to those found in open source projects. In addition, there is a growing body of evidence which suggests that firms in many industries benefit

from contributions from community members, a concept closely linked to open source (Jeppesen and Frederiksen, 2006; von Hippel, 2005; Franke and Shah, 2003; Lakhani and von Hippel, 2003). Other firms are incorporating community developed open source software in their products and services (e.g. West and Dedrick, 2001), and there are many examples of open source communities attracting donations of code and development participation from firms as well as individuals (O'Mahony, 2005; Dahlander and Magnusson, 2005). For example, virtually all of the large software and hardware computer firms including, Microsoft, IBM, Hewlett-Packard, and Sun Microsystems have contributed software to open sources projects and some have issued software under their own open source licenses, although the degree to which such licenses fulfill the requirements of open source as defined by the Open Source Initiative is a point of debate. If the goose that lays the golden egg is driven by intrinsic, altruistic motives, then appeals to extrinsic, self-interested motives may be counter-productive.

In general, ambiguous or vague definitions of self-interest can cause researchers to misidentify the motives of agents which can then lead to poor theoretical understanding and misguided policy recommendations. Economists and other social scientists need to be informed by a proper understanding of human nature. At the practical level, a basic question that needs to be addressed is whether every economic act is necessarily motivated by self-interest. Empirical evidence would seem to suggest that some acts with economic implications are not motivated by self-interest, for instance, anonymous gifts to charities, economic provisions for future generations, and the work of some non-profit organizations, are difficult to explain entirely with reference to the common understanding of self-interest. Jon Elster (1989) has offered evidence against self-interest by observing that: "parents have a selfish interest in helping their children, assuming that children will care for parents in their old age - but it is not in the selfish interest of children to care for parents in their old age. And many still do." Issues of fairness and justice have been offered as a constraint to profit seeking ((see e.g., Kahneman, Knetsch and Thaler (1986a, b), Dawes and Thaler (1988)). However, there are theories of self-interest that claim that all acts are at core motivated by self-interest. An example of such a theory is psychological egoism; if that theory is accepted as a good description of motivation, then, for example, policies aimed at deflecting behavior away from self-interest are at best likely to lead to only temporary changes in behavior. In the case of open source, a psychological egoist would expect that a copyright holder who decides to issue his code under an open source license

expects to receive benefits greater than the costs of producing the software, where both benefit and costs should be widely interpreted. A narrow economic interpretation would understand those benefits and costs to be financial. A researcher who accepts psychological egoism would seek to explain anonymous gifts, provisions for future generations, non-profit activities and every other action in terms of how those acts serve the interests of those involved. Accordingly, if psychological egoism is accepted, then it may be good business policy to promote certain forms of self-interest in favor of other forms rather than trying to appeal to non-interested motives.

At the theoretical level, regardless of whether one is an instrumentalist or a realist, it is important to use unambiguous definitions of key assumptions. Self-interest, as the term is used in economic and business analyzes, is often poorly defined and in need of clarification if only to conduct meaningful theoretical debates. I will give numerous examples throughout this manuscript of important business models which use vague definitions of self-interest and altruism. For example, one particularly interesting example of an error in reasoning which is often heard in unthoughtful economic debates is called tautological egoism. This is the idea that because agents act to satisfy their preferences, then those acts must necessarily be motivated by self-interest. Such reasoning confuses the idea that an act is performed by an agent with the motives for performing the act. It is also quite common to hear economists say that a person who incorporates another person's well-being into his own utility function is acting in an altruistic manner. However, utility functions tell us nothing about a person's values and motives, a self-interested individual can also include the utility of others into his utility function. Theoretical debates about self-interest that are based on such reasoning are bound to be confused and largely unenlightening.

#### Normative Significance

Besides the practical and theoretical implications of the role of self-interest in human nature, there are also normative issues underlying much economic reasoning in regards to self-interest and selfishness. For instance, many economists appear to promote rational or ethical egoism as theories of motivation. These theories claim that people should or ought always to seek opportunities to benefit themselves without necessarily considering the needs of others. Normative positions need to be justified. It is possible to ask, for instance, whether an agent needs to be self-interested in order to be rational or ethical as opposed to having



some other motive. If, in contrast to rational or ethical egoism, a position such as predominant egoism or some other mixed theory of self-interest is accepted, then space is created for a variety of motives besides that of self-interest. Consequently, if a mixed theory is accepted, then it is important for researchers to clarify the contexts in which the various motives are active.

### Proposed Motives of Open Source Contributors

Many explanations have been offered as to why people contribute to open source projects (FLOSS, 2002), but even at this early phase in the manuscript, it is important to be clear about what the word contribution means in the open source context. Contributions can take many forms, from coding and managing a complex open source project such as a part of the Linux kernel or the Apache server, to contributing pieces of code to projects, to answering user questions in newsgroup forums. The vast majority of people who use open source software do not contribute to open source projects in any way. For example, there are millions of users of open source browsers, but no one argues that there are anywhere near that many open source developers. I won't be primarily interested in the motives of the users of open source.

The motives of those making relatively large contributions are likely to be very different than those making smaller contributions, although both are critical to the success of open source. I wish to avoid making distinctions among the motives of the various subsets of contributors by focusing primarily on open source licensors for reasons that I will soon elaborate. Briefly, my claim is that the choices licensors make about the conditions contained in their licenses can tell us something important about their motives and in particular, the assumption of self-interest as it is often used in economics. An important question that I need to address is: What motivates the first owners of software, those exercising dominion over the software in question, whether they are programmers or the people who hire them, to distribute their software under liberal, open source licenses? Or, inversely, what can open source licenses tell us about the motives of those who use such licenses?

Several types of motives have been identified in the literature, in this section I will attempt to organize those motives into broad categories. In general, there are many systems to organize motives, Maslow's (1943) hierarchy being the most famous. Motives identified as drivers of open source contributions are only slightly less

expansive than those proposed by Maslow. For instance, in the introduction to their edited book, Feller et al. (2005) list the following four distinct clusters of reasons for contributing: enjoyment and learning; feelings of obligation to the open source community and the belief that software should be free; satisfy non-work-related user needs; and, work-related needs and concerns. While their list is helpful, it doesn't emphasize the motives behind the activities of contributors. Hars and Ou's (2001) division of motives into external and internal lends itself better to the research question that I am addressing. In addition to the theorizing that has occurred in literature, empirical studies have offered support for both the internal and external postulated motives, but the evidence is heavily dependent on the meanings attached to, at times, ambiguous survey questions (Ghosh, 2005a; Lakhani and Wolf, 2005; Hertel and Herrmann 2003; Lakhani and Von Hippel, 2003); I will offer a critiques of these questions in the case study chapter, once I have analyzed what the postulated motives actually entail.

#### Internal Motivation

Ryan and Deci (2000, p. 70) offer the following explanation and definition of internal motivation: "Perhaps no single phenomenon reflects the positive potential of human nature as much as intrinsic motivation, the inherent tendency to seek out novelty and challenges, to extend and exercise one's capacities, to explore, and to learn." The conceptual framework developed by Deci (1975) and Ryan and Deci (2000) supposes three innate psychological needs--competence, autonomy, and relatedness. In the case of open source, Hars and Ou (2001, p. 2) interpret internal motivation to mean that open source participants are "selfless". They claim that advocates of this position mean that open source programmers are not motivated by monetary incentives but by their own hobbies and preferences or that they receive rewards from increasing the welfare of others. Hars and Ou (2001, p.2) claim: "Such motivations that are ultimately rooted within the individual himself are grouped under internal factors." They suggest (2001, p. 3) that intrinsic motivation might be a disadvantage for open source projects in comparison to proprietary software because open source participants may have not be driven to serve customer needs because of their focus on their internal well-being. Other motives mentioned that seem to fit under the category of internal motivation include Lakhani and Wolf's (2005), in which they refer to "the intrinsic satisfactions of doing the work" as a motive. Torvalds and Diamond (2001) book about the fun of open source programming fits neatly in this category. Finally, Feller adds learning as a source of pleasure as a motive (Feller, et al. 2005).

Hars and Ou (2001) categorize altruism as an internal motivator, which they define as a person seeking to increase the welfare of others. As is common, they regard it as the opposite of selfishness and quote Ozinga's (1999 p.5) definition of altruism as: "doing something for another at some cost to oneself." That is a good starting definition, but Ozinga later warns of the wastefulness of defining terms and settles on a definition of altruism as kindness. Ghosh (1998) claims that altruism is a motive for open source developers, but that it plays only a minor part in their overall motivational set. A more common belief is expressed by van Wendel de Joode (2005). Van Wendel de (2005, p. 6), writing about open source, claims: "the focus of research on individual motivation has now shifted and is currently much more centered on an explanation based on rational and individual profit-seeking actors." A theoretical argument against the motive of altruism in open source is made in Weber (2004), particularly chapter five.

Hars and Ou (2001) name community identification as a variant of altruism. This, they claim, is the need to be loved. As von Hippel and von Krogh (2003) note: "Contributors to open source software projects may also get private benefits from participating in the project "community" that are not available to free riders (Raymond 2001, Wayner 2000, Lerner and Tirole 2000, Moon and Sproull 2000)." Community identification is also apparent in "feelings of obligation to the open source software community" which Feller, et al. (2005) mentions and it is an idea emphasized in Barbrook (1998, 2005). Another internal motive mentioned is support for the open source movement's ideology that software should be free and open source.

#### External motivation

External factors are rewards and punishments induced from outside the individual (Ryan and Deci, 2000). For open source developers these are generally taken to mean the direct and indirect rewards gained by increasing their marketability and skill base for the purpose of making themselves more valuable to potential employers. In general, these motives are called reputation or signaling effects, the idea is that working in an open source project provides prestige and visibility that gives developers a chance to be noticed by software companies (Bonaccorsi and Rossi, 2003; Lerner and Tirole, 2002, 2002b, Lakhani and von Hippel, 2003). In addition, it is generally recognized that developers also seek improved reputations and honor from their peers. This point is outlined by the open source movement

ideology (Raymond, 2001). Raymond believes that prestige is central in the hacker culture's reward mechanism (2001, p. 89)

These sorts of explanations follow standard economic methodology, where the belief is that employers attempt to induce behavior by external intervention (Becker, 1976; Frey, 1992). Employees and potential employees attempt to increase their marginal productivity or the perception thereof in order to command a higher price in the labor market where, in its idealized form, the marginal value that an employee adds to the firm equals his wage. Developers attempt to raise their value through visible, valuable contributions to open source projects. In many cases, internal, "irrational" motives are used as a foil against the "rational" behavior offered by standard theory. For instance, Lakhani and Wolf (2005, p. 3) write: "Many are puzzled by what appears to be irrational and altruistic behavior by movement participants: giving code away, revealing proprietary information, and helping strangers to solve their technical problems." Their own position, as mentioned above, is that while external motivational factors such as better jobs, career advancement have been the main motivational explanations of contributions to open source, they suggest that intrinsic, enjoyment-based motivations, namely creativity, are the strongest motivators. Similarly, Ghosh (2005, p. 32), writes: "...hypotheses usually supposed largely rational, self-interested motives, among the most extreme being that open source is explained by the "simple economics" of signaling for better career prospects and hence monetary rewards." Van Wendel de Joode (2005) claims that contributors: "want to maximize their utility and act in their own self-interest" (Joode, 2005, p. 23), this, according to van Wendel de Joode, is the state-of-the-art research in regards to people's motives to participate in open source communities (van Wendel de Joode, 2005, p. 194).

The direct software needs of contributors have also frequently been named as a motive to participate in open source projects (Raymond, 2001, von Hippel 1988; 2001, 2005). This point is related to the fact that developers are often also users of the code they develop, while Feller et al. (2005) sensibly divides this motive based on whether the contribution is made for work or non-work related reasons. Other suggested motives are that open source is a gift economy (Barbrook, 1998, 2005), and what Ghosh (2005) calls "balanced value flow" in which rational self-interest is assumed, but self-interest can include a range of different types of rewards, not just monetary compensation (Ghosh, 2005b).

The number of motives attributed to open source contributors is diverse, they are rational and self-interested and/or irrational and altruistic. They are motivated by their own "internal" interests and "external" incentives as well as by the norms of a greater community. The meanings of these terms need to be clarified if motives are to be properly attributed to open source contributors in general; proper attribution, in turn, will help to us in making predictions about the future of open source and its economic role.

### Contribution

Confusion at the theoretical level is mirrored in popular and academic discussions on the motives of open source programmers. As I will show, neither those explanations that attribute motives to a narrow, economic, interpretation of self-interest or those that rely exclusively on non-self-interested (e.g. altruistic) explanations are entirely convincing. Incorrectly assuming that open source programmers are driven by a base form of self-interest can have practical as well as theoretical implications. My main contribution to the business literature will be a framework for understanding the range and types of motives behind economic behavior in general and open source in particular. The framework adds to and qualifies the intrinsic, extrinsic divide used by Hars and Ou (2001) and Lakhani and Wolf (2005). The framework is a starting point, it is meant to define the coarse boundaries of self-interest and other motives rather than be a complete description of all the facets of these motives. This limitation is due to the extent of the subject matter rather than a lack of desire on my part to address those details. A second contribution is to lend support to the empirical studies which show that motives other than self-interest influence important economic phenomena. This second contribution will, naturally, depend upon my ability to convince the reader that other motives play a significant role in the success of open source and that those lessons can be applied in other economic contexts.

Accordingly, I have two goals in this manuscript which will help me to answer the main research question which is, once again: Does the assumption of self-interested agents optimizing the use of their scarce resources adequately explain the success of open source in the software market or are there better explanations? The first goal will be to attempt to disentangle and clarify some of the confusing aspects of the motive of self-interest as it is used in the economic and business literatures. Once I have done that, I can categorize the motives offered in the literature on the basis of whether they are self-interested or non-self-interested. The second goal is

to provide a practical framework that will allow researchers to systematically consider the various elements of what it means to be self-interested in order to use that behavioral assumption in a more precise, meaningful, manner. The framework will provide a means to construct comparable, coherent, definitions of self-interest in place of the ambiguous and confusing terminology that allows economic self-interest to be simultaneously interpreted as both highly restrictive and all-inclusive. To construct the framework I will need to investigate the different roles that self-interest plays in important theoretical and applied economic and business models.

### Methodology

Ideally I would start my analysis by comparing the standard economic definition of self-interest to definitions used in other disciplines. I could then perhaps test the economic definition against the presumed motives behind the hypothesized behavior of open source programmers in order to identify any weaknesses in the economic definition. Unfortunately, as I have already hinted, there doesn't appear to be a detailed, standard economic definition of self-interest. The haphazard manner with which such an important assumption is made (self-interest and rationality are often taken to be the two characteristic which define economic analysis), is surprising given that for thousands of years the extent to which self-interest is presumed to play in human motivation, and the aims and goals at which it is directed, has been the topic of a continuous and sometimes intense debate within the field of ethics (see chapter four). There are, in other words, guidelines available to help explore the structural and external dimensions that I have identified.

There is no good reason to use a vague definition of self-interest. Historical and ongoing discussions within ethics have yielded a fairly precise understanding of the main issues involved in the motive of self-interest, if not always agreement on the answers to the issues raised (Griffin, 1988). Therefore, as an alternative to the ideal case sketched above, I will selectively choose from among the ethical theories of self-interest that address issues relevant to building a practical framework of economic self-interest. The amount of available material requires that I be highly selective, fortunately, two periods commend themselves to the case at hand. The first is a dialog which took place in Scotland and England in the Seventeenth and Eighteenth centuries. In particular, I will examine the positions of Thomas Hobbes (1588-1679) and Bernard Mandeville (1670-1733), both of whom advocated models of motivation that were based on the assumption that humans are at core

self-interested and are required by reason to seek their self-interest, in other words, they can be seen as defenders of the position known as rational egoism. Their conclusions about motivation as it relates to self-interests resemble, in large part, both modern popular and academic understandings of the subject. The next author I discuss is Joseph Butler (1692-1752), his critique of the positions represented by Hobbes and Mandeville provides the starting point for discussing the key issues that any modern theory of self-interest needs to address. In his response to the positions of Hobbes and Mandeville, Butler raised nearly every major critique of the popular position known as psychological egoism, a position that formed the basis of Hobbes's argument for rational egoism and still persists in many modern discussions of economic motivation. I will end my analysis of this period by discussing self-interest as it appears in the two major works of Adam Smith (1723-1790), namely, "An Inquiry into the Nature and Causes of the Wealth of Nations" and "The Theory of Moral Sentiments". Smith's work has had obvious implications for economics and business studies, but his views of human nature, and the part that self-interest plays in that nature, have been badly misrepresented. Smith had a very complex, surprisingly rich and generous view of human nature in contrast to what passes as common knowledge concerning his advocacy of self-interest. There is still a lot to learn from Smith's view of human nature and the boundaries of self-interest.

I should mention that although I will be using the writings of ethicists, Adam Smith was a moral philosopher as well as an economist, I will not be primarily concerned with the deeper ethical issue of the place of self-interest within a 'good' life, however good is defined if indeed it is possible to define it all (cf. Moore 1993). For instance, I won't be taking a stance on the issue of whether it is right or good for people to, on occasion, be willing to sacrifice their own interests for the interests of others or attempt to defend a Utilitarian analysis of self-interest. Although I do not intend to write a manuscript on the ethics of self-interest, the topics of self-interest and ethics are closely related because, on certain interpretations, they represent two sides of the same coin. For instance, Bernard Williams takes altruism as a necessary feature of morality (Williams, 1973). The fact that it is mainly ethicists that have developed the theory of self-interest should alone be enough to indicate the close link. A large part of the debate on self-interest in that field revolves around the issue of finding a convincing argument connecting self-interest to ethical behavior. Finding such a connection would then give an agent a good, self-interested, reason to act ethically. For instance, if a convincing argument could be made that living a good life means that people

should, under certain circumstances, be willing to sacrifice their own interests for the greater good, then this would imply that it might be in one's real, enlightened, interest, to on occasion sacrifice one's narrow interests. This sort of argument is an indication of the complexity involved and precision required to understand the issues relevant to the ethical debate on self-interest and why it is still an intensely debated issue in that field. Such issues are beyond the boundaries of the current manuscript. However, I will on occasion discuss the implications that certain forms of egoism popular in the economic literature have on ethical behavior. For instance, rational egoists claim, simply, that self-interest is the only rational motive. If the link described above between self-interest and ethics is accepted, then rational egoists are claiming that non-self-interested motives are irrational and thereby perhaps implicitly endorsing unethical behavior. I will argue that such a normative claim then needs to be defended on ethical grounds and should not be taken lightly.

Once I have dealt with what I take to be the historical foundations of the economic interpretation of self-interest, I will turn to the issue of building the dimensions of an economic definition of self-interest. Accepting that self-interest is a good description of human nature still leaves open the question of whether it is an adequate assumption to explain and predict economic behavior. This is a critical question given that economists, whether they are instrumentalists or realists, presumably make the assumptions of self-interest and rationality in order to arrive at a better understanding of economic events and to make better predictions. The authors whose work I will use to help me answer this question are the Hellenistic philosophers, in particular, the Cyrenaics, the Epicureans, and the Stoics. I realize that these philosophical schools may appear to be unusual choices given that this is a manuscript written for a degree in business studies, but I have chosen them for several reasons. First, these schools taught different forms of hedonism, meaning that they advocated the self-interested pursuit of happiness. Secondly, these authors were the first to see individuals as independent beings as opposed to elements of a greater society. They were, accordingly, some of the first to approach the topic of self-interest from a perspective that modern readers would understand and appreciate. Thirdly, they believed in living their lives according to their respective theories, in other words, they grounded their theories in practice. For them, living a good life was not an esoteric question, but a practical, open question with real implications for daily life. Finally, these authors believed in a rational approach to living one's life. By knowing themselves and their place in the natural scheme they believed that people, as opposed to outside forces such as the



gods or society, have the ability to control the level of happiness they experience. In summary, they argued that a self-interested, rational approach to life was the best way to make one's life go well; thereby subscribing to the two basic assumptions of modern economics. Because these authors attempted to live their lives according to their own theories, it is possible to ask whether the assumptions of self-interest and rationality alone are enough to describe and predict behavior, and if they are not, then we can ask why not, and adjust our theory accordingly. In short, these authors provide a small test of the different types of behavior that follow from different conceptions of self-interest.

The rest of my manuscript is divided into eight chapters; I will here briefly discuss the main themes of each chapter. In general, given the diversity of uses of the term of self-interest in the economic and business literatures, I will, wherever possible, discuss specific, important, economic models or critiques of economic models. I hope that this approach will allow me to avoid the charge that I am selectively choosing to discuss topics that no or very few economists defend. The first chapter discusses the basics of the theory of self-interest and the difficulty of assessing motives base on observed behaviors. The second chapter introduces economic models of self-interest and provides a brief discussion of the purpose of models within economic theory. Discussions of Oliver Williamson's version of Transaction Cost Economics and Meckling and Jensen's Agency Theory provide important clues to the actual use of the assumption of self-interest within two of the more prominent business economic models. The third chapter discusses the structural and external dimensions of self-interest in detail, while the fourth traces the foundations of the economic approaches to self-interest. Chapter six examines how the assumptions of self-interest and rationality were used by the Hellenistic philosophers to build models of the good life. The seventh chapter takes the essential features of the previous chapters and builds a practical framework of self-interest. The last chapter examines the case of open source software and demonstrates the need to clearly specifying what is meant by self-interest or face the risk of misidentifying motivation. In particular, I ask whether the economic conception of self-interest as it is used by Lerner and Tirole (2001) can explain the willingness of programmers to contribute to open source projects. A particularly interesting feature of the open source community is that there appears to be two principal motives at work, one part of the community is motivated by profit maximization, while the other is driven by a mixture of motives including a sense of duty to make and keep information freely available.



## CHAPTER ONE: THE BASICS OF SELF-INTEREST

### Introduction

Before proceeding further, I need to give a rough indication of what I mean when I use the term self-interest if only to provide a basis for later comparisons. I will discuss aspects of each of the following topics in greater detail in later chapters; the purpose of this chapter is to present a brief overview of the basic elements of self-interest. I put particular emphasis on setting the extreme boundaries of that motive before turning to some of the more subtle issues needed to develop a practical theory of self-interest. This approach will allow me to begin to remove some of the more fundamental misunderstandings about self-interest and facilitate later discussions. In keeping with my research question, I will emphasize those aspects of self-interest and altruism that appear in explanations about the motives of open source contributors, however, much of the discussion to follow has more general implications. The chapter is broken down into several sections. In the first I discuss the main characteristics of self-interest and begin to build a more detailed picture of that general motive. In the second section I examine some of the theoretical and practical alternatives to self-interest which may raise doubts about an exclusive reliance on that motive. In the third section I discuss the problems of identifying motives with special attention to the motive of altruism. The final section offers a general discussion and some conclusions.

### Basics of Self-Interest

I will presently introduce the basic issues involved in a proper understanding of self-interest and the relationship of self-interest to the domain of economic analysis. One simple message I wish to convey is that self-interest has never been a monolithic concept, it encompasses a large number of meanings which describe an equally large number of behaviors. The definition of what is in a person's self-interest has been the subject of debate for thousands of years, so it shouldn't be surprising that there are different ways of interpreting and understanding its meaning (Parfit, 1984, Dancy, 1997). The motive of self-interest, broadly understood, has always been an accepted explanation of a large part of human behavior, therefore it is also not surprising that economists and many other social scientists have adopted it as a key motivational assumption in their models. Having one standard assumption clearly has advantages in terms of parsimony. For example, Gary Becker's (1976, p. 112) view is that self-interest should be the

default position and when behavior seems to contradict that assumption, the researcher should, nevertheless, stick with it even if it cannot be proved. However, we need to also bear in mind Hirschman's (1992) warning that if self-interest explains everything, then it explains nothing.

Claiming that people are primarily self-interested is not a radical position that separates economic analysis from other social sciences or understandings of human nature. The issue at hand is that although self-interest is the central assumption in many important economic and business models, it is seldom clear from these models what sorts of actions and aims that term is meant to encompass. For instance, it is often ambiguous whether economists mean to imply that self-interested agents are taken to be psychological, rational or ethical egoists, and in some cases it appears that agents are modeled as a confusing and incompatible mixture of these different types of egoisms. As I will show, these three types of egoism, and others, represent very different assumptions and generally imply very different types of agent behavior. To take just one small example, psychological egoism is meant to be a description of actual behavior, while rational and ethical egoism are normative positions telling people how they should behave in order to be considered, respectively, rational or ethical. For instance, as previously mentioned, Ghosh (2005) and van Wendel de Joode (2005) respectively present and defend self-interest as a rational, utility, maximizing approach in contrast to other, by implication, irrational motives. The intentions and implications behind taking a descriptive versus a normative position are presumably very different and should be made with care--they clearly have different policy implications. Therefore, even if it is granted that a context is economic and that the agents concerned are motivated exclusively by self-interest, such an assumption would still leave room for a widely divergent set of interpretations of behavior and implications for policy development.

Another particularly important omission in many economic models is an explicit standpoint regarding the degree to which agent's interests are intertwined. In broad terms, the interests of agents can be correlated or not. For instance, it is widely accepted that certain relationships, for instance those among family and friends, often imply a dependence among interests--an agent's interests are affected by the interests of family and friends and vice versa. In a business setting, the interests of colleagues and competitors often influence perceptions of self-interest; we generally want to see colleagues and friends succeed, and may feel a corresponding obligation to promote their interests on the basis of the relationship itself. This

type of explanation can be seen in the arguments mentioned by Feller, et al. (2005) and Barbrook (1998, 2005). Many of the internal motivators previously mentioned, particularly those involving the beneficial aspects of being part of open source communities, attempt to make these relationships explicit.

However, interests clearly are not necessarily positively correlated in the sense that, for example, one person's interests are served through the promotion of the interests of others, often the situation is quite the opposite. Even within the open source community there are conflict resolution procedures (Weber, 2004). Interests may be negatively correlated as, for instance, described by a Pareto optimal situation in which an agent's utility can only be increased by harming the utility of one or more of the other agents. A more interesting example that applies in both business and non-business settings is the motive of revenge. A desire for revenge may cause agents to damage their own, informed, interests for some actual or perceived slight. Revengeful agents may be willing to damage their own interests in order to damage the interests of wrongdoers or perceived wrongdoers. Revenge may be too strong a word for motivation within the open source community, but the motive of being the underdog against established, proprietary, software firms has been mentioned as a possible driving force behind open source contributions (Raymond, 1999). Whatever the correlation, untangling the range of potential dependencies among interests will need to be addressed. In addition, the overall aims or goals of agents are often left open in economic models, for example, it is not always clear whether economic agents are taken to be aiming at maximizing or satisficing their happiness, preferences, desires, wealth, esteem or some other general objective; in fact, in some important economic theories no aims are specified. The aims of an agent who desires to maximize his wealth are very likely to be different than the agent who, for instance, wants to maximize his happiness or lead the best life possible.

Although self-interest might be assumed to be a necessary component to achieve these different aims, such widely divergent aims will likely require considering very different, irreducible types of interests. It is therefore necessary to investigate whether the economic interests of open source contributors differ from their non-economic interests; if interests are found to vary then the appropriate tools used to identify those interests should vary as well. A clearer demarcation of the different categories of interests may thereby help to clarify the applicable domain of economic analysis. For instance, is the fact that Linus Torvalds enjoyed programming an economically driven interest (Torvalds and Diamond (2001)?

Clarifying what it means to be economically self-interested will help to explain the actions of open source programmers, more generally, it will allow researchers to clearly state their intentions and properly use this most basic of motivational descriptions.

#### Dimensions of Self-interest

For expository purposes, self-interest can be thought of as varying along two principal dimensions, a structural dimension and an external dimension. I use the term dimension for two reasons, the first being that the word calls to mind a sliding scale which is appropriate for the components of self-interest. The interests of people can, for instance, be more or less self-directed and will vary according to the circumstances in which they find themselves. The second reason I use the word dimension is that many of the components of self-interest overlap, by which I mean that identifying someone's interests requires identifying their location along various independent dimensions. The two dimensions I am using emphasize the difference between assumptions about the role of self-interest in an agent's psyche versus assumptions about what objects are to be included in an agent's set of interests.

The structural dimension is, simply, the way self-interest in general is taken to work within person's psyche. On the other hand, the external dimension, (I use the word external for historical reasons), is the objects at which self-interest is directed. Using a mathematical analogy, the structural concept can be thought of as a function describing how a person is taken to combine his desires for external objects, while instances of objects are the variables within the function. For example, as previously mentioned, a psychological egoist (the structural component) claims that no matter what the goals or aims of an agent are, the agent uses the apparatus of self-interest to rank those objects (external concept), no other system of reasoning is allowed. The difference between these two concepts becomes apparent when I discuss the problem of identifying motivation. In order to facilitate later discussions I will briefly explain the basic elements of both dimensions. I will proceed by placing the elements of self-interest that I raised in previous paragraphs into their appropriate dimension.

### Structural Dimension

The structural dimension is a modeling assumption about the manner in which agents combine the objects of their desire. In regards to self-interest, it is used to define the different types of egoism that have been identified in the philosophical literature. Categorization of the different egoisms is a simplification, from everyday experience we know that the capacity for egoistic behavior varies per individual and the contexts in which they operate. For example, many economists seem to draw a distinction between agents acting in an economic context in which agents are motivated exclusively by a narrow, selfish form of self-interest, and other contexts in which other presumably non-self-interested or less interested motives are operating (Williamson, 1998). Other economists, for example Gary Becker (1976), claim that economic reasoning, including self-interest, is applicable across a very wide range of social settings including, famously, the marriage market and within family relationships. On these views, the context determines, to an extent, the type of self-interested reasoning that is used. However, the self-interest that is expressed in stock market dealings is presumably very different from the self-interest expressed in social relationships. People may be motivated exclusively by self-interest in both situations, however, not only the external aims and goals of an agent may vary, but the reasoning used to justify an agent's interests may vary as well. It is therefore important to clearly state the degree to which the context under consideration affects the egoistic tendencies of an agent.

I have previously mentioned a few of the structural dimensions of self-interest; including psychological egoism, rational egoism and ethical egoism. The type of egoism that is taken to describe an agent in a context will have repercussions on the sorts of behaviors that an agent believes are appropriate. Similarly, a researcher's beliefs about which structural dimension describe an agent's actions will influence his interpretation of the behaviors he observes. For instance, a researcher who accepts psychological egoism may believe that people are programmed by nature to act from self-interest regardless of the context; accordingly, even acts that appear to be altruistic should be understood to be, at their core, motivated by self-interest regardless of the context. As a consequence, such a researcher will interpret all behaviors as motivated by self-interest. In the specific case of open source software, a researcher who accepts psychological egoism as an accurate description of motivation will be inclined to reduce the motives of all programmers to self-interest. In contrast, advocates of rational egoism claim that it is necessary and sufficient for an action to be rational that it contributes to the self-interest of an

agent. On this normative view, acting for other than self-interested motives, for example altruistic motives, should be interpreted as irrational. Researchers who hold the view that agents are rational egoists may interpret the actions of open source programmers as non-self-interested but, as a consequence, the programmers are acting irrationally. Finally, proponents of ethical egoism take the normative position that people should or ought to pursue their self-interest, to do otherwise would be unethical. In effect, ethical egoism claims that you owe it to yourself to be self-interested and any other motivation is probably disingenuous and certainly unethical. In the case of ethical egoism, open source programmers, to the degree that they are not promoting their interests, are acting unethically. Elements of each of these positions, and more, are found in the economic literature and need to be sorted out.

Finally, the structural component of self-interest also relates to the time frame or intensity over which self-interest is expressed. For example, even an agent acting as a psychological egoist, who by assumption is always maximizing or satisficing his interests, will have to have a means of incorporating the time frame over which his interests span. The point of interest is that accepting the claim that all motivation is based on self-interest leaves open the question of the time frame under consideration. Each of us has many different and sometimes competing interests. For instance, a very intense form of self-interest, one in which an agent is exclusively directed at his own interests while ignoring the interests of others, comes close to what is commonly called selfish or egoistic behavior and leads to actions that are generally met with disapproval. Such motives may be economically, and otherwise, self-defeating in that they work against more important, perhaps longer-term, interests. A selfish person will pursue his aims without considering the effects his actions have on others and perhaps even the effects those actions have on his own longer-term interests. For instance, a selfish person may be indifferent between two acts each of which he finds equally worthy but each of which has different (positive or negative) effects on the interests of others. In contrast, more inclusive forms of self-interest are, to various degrees, seen as more socially acceptable and may help agents to achieve their greater long-term goals even at the cost of not achieving certain short-term interests. The point is that interests have various time frames over which their costs and benefits should be, to the extent possible, calculated or considered.



### External Dimension

The second dimension of self-interest, which for historical reasons I call the external dimension of self-interest, encompasses the objects at which agents aim to promote their interests and how agents incorporate the interests of others into their own interests. Whether there are many irreducible aims or one common aim, e.g., happiness, will have to be explored. Different basic aims and goals may lead to different sorts of behavior, for example, a person aiming at living a virtuous life or acquiring as much knowledge as possible may have an entirely different set of interests than a person who aims at maximizing his expected wealth. The other important aspect of the external dimension is the degree to which people include the interests of others in their own interests. I place this in the external dimension because it concerns how agents incorporate aims of others, aims external to those of the agent under consideration, into their own interests. Both extremes (not including the interests of others and equating one's interests entirely with the interest of others) are straightforward to understand, although equally unrealistic and probably self-defeating for their own reasons. Pure egoists are defined as those agents who only think of their own needs, without reference to others even if, for example, others can help them to realize their interests. Of course, this would be a self-defeating, irrational, strategy to the degree that others are able and willing to promote the interests of the pure-egoist. In other words, this sort of self-interest is frequently not really self-interested at all. The other extreme is a type of altruism in which agents completely sacrifice their own interests for the interests of others, a self-defeating and self-destructive approach for obvious reasons. Both of these extremes are often used as straw men by one side or the other to create a false dichotomy, but within these two extremes there are many possible combinations of overlapping motives which will need to be examined.

I claim that both the structural and external dimensions are necessary in order to arrive at a proper understanding of what it means to be self-interested. To its detriment, economics has largely failed to take a coherent stance on either dimension. The confusion over the meaning of self-interest is compounded by the increasing range of activities over which economic analysis is held to apply including sociology, political science and law (Mansbridge, 1990). In other words, the difference between economic and non-economic contexts is blurred. For instance, on occasion economists appear to have adopted an extremely narrow definition of self-interest in which agents are modeled as solitary, non-social, non-normative, calculators of net-interest where interests are defined in terms of wealth

or something similar. Such a description may give a good account of agent activity in the stock or commodity markets, but will be less applicable the more that social interaction or aims other than wealth become measures of well-being. Other economic scholars appear to take a very broad view of self-interest in which nearly anything that an agent prefers or desires is in his interest, such an understanding can easily lead to tautological reasoning or confusion in terminology. The wide range of behaviors attributed to self-interest has led to a confusing range of interpretations of 'economic' behavior. On the one extreme, self-interest is seen as sort of a mean, aggressive motive, while at the other it is a catchall for all types of behavior.

#### Characteristics of Self-interest

Ignoring, for the moment, the structural dimension of self-interest, there are two basic characteristics that need to be kept in mind when analyzing the general motive of self-interest. The first characteristic, the external motive mentioned above, is the aim or goal being analyzed. As a default I will assume the aim is happiness, an assumption that greatly simplifies the discussion by removing the need to consider a nearly unending and sometimes incompatible list of possible aims that agents can have. It also has well-known limits, namely, it rests on the shaky psychological foundation in which agents are assumed to pursue those aims that will make them 'feel' happiness. It has been argued that, at times, people will prefer, for instance, harsh reality to happiness based on false pretenses. For example, even if I expect the worse, I may prefer to know what a person really thinks of me rather than hearing the praise of a sycophant (Griffin, 1988). That said, happiness plays an important role in any practical understanding of the aims of agents. The second characteristic is what I call the focus of an action. At one theoretical extreme an agent is exclusively focused on his own happiness while at the other extreme he is exclusively focused on the happiness of others. The focus of an action includes the issue of whether other agents are considered ends in themselves or the means to achieve an aim. The majority of the discussion to follow is devoted to addressing these two characteristics with only a passing reference to the issue of measurement. I will begin by discussing three alternatives to typical theories of self-interest, hedonism, pure egoism, and altruism. The intent of discussing the alternatives is to provide a stark point of contrast from which to view self-interest.

### Instant Hedonism

Instant hedonism is at one motivational extreme. A simplistic form of hedonistic theory gives people the aim to do what would give them the most immediate happiness or pleasure or, most generally, do what leads to the most improved state of mind. These aims may include the happiness of others, by far the most commonly accepted assumption, or not, that distinction is not relevant to the case I am now considering. The important point is that an instant hedonist is focused on his instantaneous happiness, as opposed to say his net happiness over a lifetime. Immediate hedonism is interesting because of the contrast it emphasizes with generally accepted theories of self-interest.

The instant hedonist assesses the available alternative set of actions, orders those actions based on the happiness or pleasure that he expects to receive at that moment, and then pursues those acts accordingly. The happiness of the hedonist is confined to what he believes will make him most happy at the moment and is therefore a subjective measure. Hedonism, like all egoisms, can be seen as a type of rationality that aligns a person's set of motives and aims. Although the hedonist attempts to obtain the most happiness he can, he may fail to achieve his goal because he is misinformed or because he makes poor decisions or something along those lines, nevertheless, the failure to obtain his goal does not detract from the validity of the theory, although a theory can be too demanding. By valid I mean only that happiness is a readily achievable goal for most people even if circumstances, self-made or otherwise, occasionally prevent one from achieving it. The hedonist would not be satisfied with second best, he would not, for example, be content with being just happy when he believes that it is within his reach to obtain a state of greater net happiness, that is, after all, what it means to have the exclusive aim of being most happy. Similar reasoning holds for cases of less unhappy; the instant hedonist prefers those acts that he expects to bring him less unhappiness now.

Is hedonism a self-defeating theory? In other words, does the theory fail in its own terms? No, hedonism gives a person the readily achievable aim of doing what will give him the most happiness at the moment. In terms of acts, suppose a person is faced with the choice of receiving a thousand dollars (or units of happiness) immediately or foregoing the thousand dollars now and receiving a million dollars in a few years time, hedonistic theory says take the thousand dollars now. Note that the immediate choice facing the hedonist is between receiving nothing or a

thousand dollars at this moment, simplistic hedonism says take the money now because that will bring the most happiness at this moment. In a sense, the simplistic hedonist has an infinite discount rate rendering anything beyond the immediate a value of zero. While immediate hedonism isn't self-defeating or irrational, it is ludicrous. In contrast to immediate hedonism, and the point of discussing it, is that any theory of self-interest needs to include the future interests of the individual, further still, excluding discount rates, a self-interested person needs to see all interests as equally valuable regardless of when they are realized. In other words, the interests I will have in ten years are just as important as the interests I have now (Parfit, 1984)

### Pure Egoism

Pure egoists can be defined as those who only considers their own interests without reference to others; they are completely focused on their own interests to the exclusion of all else. Unlike the instant hedonist, those interests will include future interests. The important feature of this motive is that a pure egoist's interests do not include the interests of others. Clearly, pure egoism is generally a self-defeating strategy. Whenever interaction among people is required to achieve one's aims, the pure egoist is unlikely to be successful. To the extent to which his happiness is dependent on interaction with others, his is a bad strategy. The pure egoist may feel most at home as a lone trader in a stock or commodities market where the price is a sufficient statistic. By fixating only on his own interests the pure egoists may be damaging those very interests and defeating his own goal of maximizing his aims. Perhaps only a completely isolated character, such as Robinson Crusoe (before Friday's arrival of course), could be successful following such a strategy. However, even a person in total isolation who focuses exclusively on his own happiness may be following a self-defeating strategy. It has often been observed that trying to be exclusively happy (or any hedonistic aim) at all times and in all contexts may not be a good strategy to achieve happiness; at times distance from one's aims is required to achieve those aims (Sidgwick, 1981, p. 136.). It is not difficult to imagine economic circumstances in which it may be self-serving to understand and accommodate the happiness of others--the example of international trade comes immediately to mind. Economic market systems are successful precisely because agents intentionally interact with one another to achieve their individual goals, but it is not necessary to belabor this obvious point. Pure egoism is an extreme approach that probably no sensible person follows in reality; it represents a boundary case that emphasizes the importance of understanding the

focus of an agent. It also points out the importance of properly interpreting the motives of an agent in order to properly assess the value or goodness of an act.

### Prisoner's Dilemma and Pure Egoism

One of the important implications of self-interest theory for the case of open source can be seen in game theory, particularly the prisoner's dilemma. I take for granted that the reader knows the basic features of game theory and will immediately proceed to the point at hand (for an introduction see Gintis, 2000 or Osborne and Rubinstein 1994). The prisoners in the prisoner's dilemma need not be instant hedonists, the game can and has been extended over time, most famously in the tit-for-tat game by Robert Axelrod (1984). However, it is possible to ask the extent to which the prisoners are pure egoists. Given that the prisoners are aware of the choices they face and aware that the other is aware as well, they are not pure egoists. The point of game theory is that hypothesized players react to one another's expected decisions.

The prisoner's dilemma is typically represented in the following normal form, where smaller numbers represent better outcomes. I will assume that the game is played once and choices are made simultaneously.

		You	
		B1	B2
Me	A1	I 3,3	II 1,4
	A2	III 4,1	IV 2,2

Fig. 1: The Prisoner's dilemma.

The reasoning of each prisoner is that no matter what the other does it is better for me to choose A1. So, if I assume you will choose B1 I will choose A1 because 3 is a better outcome than 4 and if I assume that you will choose B2 I will choose A1 because 1 is a better outcome than 2. Accordingly, it makes no difference what I expect your choice to be or the expected impact that those choices will have on your happiness. In either case I make the same choice and you follow that same

strategy so we both land in zone 1, the third best alternative for us both. This game makes it possible to construct some rather unsettling examples. For instance, imagine the situation in which I am nearly indifferent between alternatives A1 and A2, but the choice I make has a tremendous impact on your happiness or well-being (Figure 2). I will choose A1 no matter what you do, but by my choosing A2 instead of A1 I would "save" you 200 units of happiness at a cost of only 0.01 units of my happiness. Considerations for the effects of one's actions on another are not a consideration for the self-interested person. Even knowing that he is causing the other person to suffer for a very small gain to himself does not offer the pure egoist a reason to suffer the smallest of losses. The point of this example is to emphasize the independence of this type of self interested person; he simply chooses what will make him happiest in a given situation and the effects of his behavior on others is irrelevant for his decision. The pure egoist is not malevolent, he does not intentionally set out to damage or use others to achieve his goals, however, his callousness is disturbing and his complete detachment from the wider implications of his behavior is unethical at best and probably dysfunctional in most social settings. He only sets out to do what is best for himself at that moment, while the consequences of his behavior for others, good or bad, are not his concern. But this doesn't imply that this type of self-interest is irrational, only that it may be morally challenged.

Prisoner's dilemma 2:

		You	
		B1	B2
Me	A1	I 3.00, 300	II 1.00, 400
	A2	III 3.01, 100	IV 1.01, 200

Fig. 2: The Prisoner's dilemma 2.

It might be argued that pure egoism is self-defeating because both prisoners would both do better if they cooperated. In terms of Fig. 1, we would both be better by choosing A2 and B2 respectively, and landing in quadrant IV instead of quadrant I. That argument fails, self-interest theory gives people the aim of achieving the

greatest expect benefits for himself, that is achieved by choosing A1 whatever it is you do. "We" could do better by cooperating, but the self-interested person is not concerned with "we". This issue is what makes common pool or open access resources interesting. For instance, it is common knowledge among fisheries economists that open access in the North Sea has led to an uneconomical and unsustainable level of fishing of important species such as Sole and Cod (Daan, 1997). The fishery would be better off, more profitable, if fishing was restricted and catch was significantly reduced. If, for the benefit of the group, a fisher chooses not to fish while others do fish, then the community minded fisher is sacrificing his own interests from which others gain. The reasoning in other prisoner's dilemmas is similar, from the perspective of the fisher or prisoner, the problem isn't that he's self-interested, the problem is that others are. It's interesting to note that several of the authors whom I have mentioned previously and I will again discuss in the last chapter view open source as a common pool resource, however, I believe that there are some problems with their approaches.

The point is that while this form of self-interest is not individually self-defeating, it can be collectively self-defeating in the sense that if everyone is a pure egoist then all are worse off than if they would have had another disposition. This is not the same as claiming that whenever interaction among people is required to achieve one's aims that pure egoism is a collectively self-defeating strategy. There are many examples in which parties pursuing their own interests produce results which provide better outcomes for everyone. For example, open source can be seen as a coordination game. Coordination games illustrate situations in which all parties can realize mutual gains, but only by making mutually consistent decisions. A common application is the choice of technological standards and I believe that this type of game is an accurate representation of some types of open source licenses in which open source software is made incompatible with proprietary software. For example, two players facing a choice between using open source as opposed to proprietary software will gain if they both use open source licenses and lose if their actions are uncoordinated.

Lower numbers are better outcomes.

		You	
		B1	B2
Me	A1	I 1,1	II 4,4
	A2	III 4,4	IV 1,1

Fig. 3: The coordination game.

A cooperative game is a game in which groups of players may enforce cooperative behaviour, hence the game is a competition between coalitions of players, rather than between individual players. In the case of open source and software in general, enforcement is maintained through copyright and patent laws.

#### Self-Interest Theory

A popular and robust theory of motivation is the theory of self-interest. The approach of self-interest has been the default position for most analysts of human behavior for thousands of years (Parfit, 1984, Rogers, 1997). Many philosophers and others have found it, for better or worse, to be a good guide to explaining and predicting all aspects of human and non-human behavior. Although it is seldom explicitly stated, I believe that this motive, rather say than pure-egoism, is what well-informed economists have in mind when they use the term egoistic or self-interested. Purely self-interested persons, in contrast to pure egoists, attempt to achieve their aims within a much broader context, using the terminology above, their focus is much broader. Whereas the pure egoist is focused on his own narrowly defined needs to the exclusion of all else and therefore generally fails to achieve his overall aims, the self-interested person orders and chooses his set of desires in such a way as to achieve those aims that will, for instance, make his life go best. According to self-interest theory, the disposition that is supremely rational is that of someone who is never self-denying. It gives to each person the aim of achieving those outcomes that will bring the greatest expected benefits for himself (Parfit, 1984). The question that a purely self-interested person should ask himself



is whether his set of desires (or preferences) promotes his given lifetime aims as well as an alternative set of available desires or preferences.<sup>1</sup>

Several questions immediately arise, for instance, is the purely self-interested person interested in his overall, long-term, happiness or in his immediate happiness? The theory of self-interest clearly states that a person should maximize his interests over his lifetime, so in contrast to the instant hedonist, self-interest theory is forward looking. Self-interest theory requires people to be able to evaluate the value of aims over a lifetime. It doesn't require super-rationality, only that people do the best they can. For instance, if a person could have made his life go better by being more reflective, then the person should become more reflective. And in contrast to the pure egoist, the theory of self-interest can include, and for most people, almost certainly does include the interests of others. The theory recognizes that the interests of people are often intertwined and a person is made better off when those close to him are made better off. I emphasize this point because in common language, and in some implicit definitions used by economists and business scholars, being self-interested often seems to exclude benevolent acts. Whether one is using the theory of self-interest or preference theory, this would be an arbitrary restriction that contradicts everyday experience. For instance, a person's interests can and often do coincide and depend on the interests of others; when a friend's happiness increases, my happiness increases as well. I then have a motive to make him happy if only because it will make me happy. In contrast, a non-self-interested act would include the case in which I want my friend to be happy regardless of how his happiness impacts my own.

The theory of self-interest is a very inclusive theory but not as inclusive as preference theory. Like preference theory, the subjective account of self-interest excludes little in terms of specific desires and aims that an agent can have. As long as it contributes to a person's subjectively determined aims then the theory allows it. Aims can include the maximization, minimization, satisfaction, etc., of wealth,

---

<sup>1</sup> Preferences and desires are not the same thing. A preference in contrast to a desire is a relative concept. However, having preference seems to presuppose having desires in that it is difficult to imagine preferring something that one doesn't desire. It is, however, possible to imagine cases in which a person may prefer something he doesn't desire. For instance, all else equal, a person may prefer a quick death to say a slow drawn-out illness followed by death, but he doesn't really desire either, at least in the immediate future. That said, I will generally use preferences and desires inter-changeably.

power, happiness, pain or any other imaginable aim. Aims such as maximizing pain are admittedly odd, but I see no reason to exclude them and other less conventional aims as a viable interest if they are assessed subjectively. Viewed objectively, it is reasonable to conclude that a person maximizing the pain he feels is working against his own interests. Similarly, a person interested in, for instances, hurting others, will face community sanctions, but such interests are not excluded by the theory. Despite their apparent similarities, the theory of self-interest and preference theory are very different. In the case of self-interest, the argument has to be made that the agent's interests, viewed either subjectively or objectively, are being served. An aim has to be specified before we can determine whether an agent's set of desires will help him to achieve that aim. In terms of happiness, the definition might be that the self-interested person adjusts, to the degree possible, his desires in such a manner that they will allow him to achieve the greatest amount of happiness or pleasure that he believes is possible over his lifetime. The definition excludes acts that will intentionally work against one's net interests. Preference theory, in contrast, only requires that an agent is able to compare two objects and choose the one he most prefers or declare that he is indifferent in regards to the objects. It doesn't necessarily say anything about a goal or aim. Preference theory is a much more general theory than the theory of self-interest. However, if we want to be able to conclude that a series of preference are rational, we will also have to know or make an additional assumption about a person's aims. In other words, we have to know if the person is ordering his preferences in such a way as to achieve the maximum, minimum of whatever his aim is (if he has an aim at all). To conclude, preference theory is even less restrictive than the theory of self-interest because it does not specifically state an aim or goal, while self-interest theory does.

### Practical Egoism and Altruism

Taken together, these two features, the ability to incorporate benevolent acts and the broad reach of the theory, are what help to make the theory of self-interest such an enduring theory. The theory also corresponds well with everyday experience. It is difficult, although possible I hope, to make a convincing argument that a person may intentionally performs acts which damage his own interests. However, given a broad enough definition of self-interest, it is possible to argue that every act is motivated by self-interest--a point to which I will need to return to in later chapters. The generality of the theory has one major drawback, namely, it faces the charge of emptiness as a guide to behavior in that it allows too much. In order to be of

practical use as a modeling tool its scope needs to be restricted. I will now discuss some subtleties to the definition of self-interest in order to emphasize specific features of the theory, particularly what the theory says about how others should be treated.

### Instrumental Egoism

An instrumental egoist differs from the pure egoist in that he is capable of considering and incorporating the expected effects of his actions on the happiness of other people. He is no longer so exclusively focused on his own aims that he ignores the desires of others. Although the instrumental egoist is no longer psychologically isolated, his consideration for the happiness of others extends only to the degree to which it impacts his own aims, say happiness. Just as with the pure egoist, the focus of the instrumentalist is still on his own happiness, but he differs from the pure egoist in that he is able to use the happiness of others as a vehicle to achieve his own aims. Within instrumentalism egoism several types of behaviors can be contemplated including callousness, destructiveness and no-harm instrumentalism. I will briefly discuss the first two types of instrumental egoism before turning to the more confusing case of the no-harm instrumental egoist.

A callous instrumentalist is indifferent as to the impact of his actions on the happiness of others; he doesn't, for instance, care whether his actions have a harmful, beneficial or neutral influence on others, as long as his own aim of happiness is promoted. For instance, taking some small act improves his own happiness he will do so, but not because the other person experiences happiness. If the other person's unhappiness would make the callous instrumentalist happy and giving conducting some act would achieve that aim, then the callous instrumentalist would follow that path. In a sense, he is the most general of the instrumentalists, anything goes as long as his happiness is increased. The callous instrumental egoist's decision concerning an action is not directly dependent on the happiness or unhappiness of the other person as an end state; for instance, he would not perform an act for the other person (and improve that person's happiness) if doing so would reduce his own happiness by even a very small amount. The difference between the callous instrumentalist egoist and the pure egoist is that the instrumentalist is now able to incorporate the happiness of the other person into his own decision making process, but that knowledge only makes a difference in his behavior via its impact on his own happiness. For instance, he will be indifferent between a choice that improves or reduces the happiness of another person but

leaves his own happiness unchanged. In contrast, given the same choice, the destructive instrumentalist will intentionally choose the option that reduces the happiness of other person. He is no longer indifferent in regards the impact of his actions on the happiness of the other person, in fact, the unhappiness of the other person or his being made less happy increases the instrumentalist's happiness. He looks for those opportunities that damage or reduce the happiness of others as a means to increase his own happiness. The destructive instrumentalist will perform and act because he knows that the other person will be made unhappy or less happy which promotes his own happiness. The important point for both the callous and destructive instrumentalist, and for that matter all instrumentalists, is that they see others only as a means to achieve their own ends, the callous instrumentalist doesn't care about the impact of his actions on the happiness of others beyond their impact on this own happiness and the destructive instrumentalist will seek out opportunities to damage or reduce the happiness of others if it serves his needs.

It is possible to imagine some disturbing cases, for instance, both of the instrumentalists mentioned above would be willing to cause immense damage to others as a means to achieve even a very small gain in happiness for themselves. They are both even more ethically abhorrent than the pure egoist because they are keenly aware of the effects of their actions on others, while the pure egoist is incapable of empathy given that his focus is assumed to be completely inwardly directed. In effect, the pure egoist can't be held accountable for actions he can't help committing, while calculating instrumentalists can. The difference is that the instrumentalist considers, for better or worse, the happiness of the other person while the pure egoist is by definition incapable of doing so. Economists often appear to model agents as callous or even destructive instrumental egoists. Agents attempt to maximize their happiness even if it comes at the expense of the happiness of others and sometimes their happiness is promoted precisely because of a reduction in the happiness of others. Economists, see for example the above prisoner's dilemma case, generally tend to shy away from making normative claims, by which I mean that they are not willing to take a stance on whether such instrumental behavior is unethical and/or repugnant. As a consequence, they do not take a stance on whether it is ethically good, for instance, that a very small increase in the happiness for an agent is purchased at great cost in terms of the unhappiness of others. To the extent that they purportedly (many economist, as I will show, often do in fact take a rather strong ethical stance by condoning self-interest) ignore ethical issues, they are using an abstraction from real, observed, behavior in that people generally do use ethical reasons to condemn such callous

and destructive behaviors. There may be good reasons for economists to model agent motives as instrumentalists, but those reasons need to be argued for.

The case of the no-harm instrumental egoist is more subtle than the two cases discussed above, but understanding it is important to forming a proper conception of what it means to be an instrumentalist egoist. The no-harm instrumental egoist specifically wants others to be happy, but only to the extent it serves to increase his own happiness. He is no longer indifferent nor does he seek to reduce another person's happiness in order to achieve his aims. In other words, the happiness of the other person is an aim for the no-harm instrumental egoist. However, it is not an absolute aim as in the case of the altruist in that the no-harm instrumental egoist will not promote the happiness of the other person if it means that his own happiness is diminished. The focus of the action remains squarely on the happiness of the no-harm egoist, but the assumption in this case is that the no-harm instrumentalist has a reason to promote the happiness of the other person as a means to improve his own happiness and we can add the condition that he won't act if it means that the happiness of the other person is damaged. This case differs from the case of the callous instrumentalist in that the no-harm instrumentalist is now motivated to promote the happiness of the other person rather than in promoting any state of the other person that will increase his own happiness. In the previous two cases the instrumentalist could be happy because the other person was unhappy, while in the present case I am restricting the state under consideration to that in which it is the other person's happiness that adds to the instrumentalist's happiness. Is such an act an act of self-interest on the part of the instrumentalist or something else? I believe that this case also falls under the heading of self-interest because it is the instrumentalist's happiness which remains the focus of concern rather than the happiness of the other person. The no-harm instrumentalist makes his decision to forgo a benefit because he is concerned primarily for his own happiness; the point of reference is the instrumentalist's own happiness. The happiness of the other person is important, but only as a vehicle to increase his own happiness. This situation is in sharp contrast with the case in which an agent wants the other person to be happy not as an instrument to fuel his own happiness, but because he simply wants the other person to be happy--a case to which I will soon turn.

### Sacrificing Interests

The situation becomes even more complicated in cases involving the sacrifice of interests. In the case of the no-harm instrumentalist, the agent was not asked to sacrifice his own happiness for the benefit of the other person. By definition, instrumentalists, no matter what sort they are, act only if they expect to improve their net happiness. However, we know from experience that economic agents are accustomed to sacrificing short-term financial gains for long-term financial gains. Many other interests can be sacrificed for both personal reasons and for the benefit of others. An obvious example of the first case is the sacrifice of immediate interests, generally defined, in return for greater expected interests in the future, assuming away the complication of finding an appropriate discount rate for future interests. In the terms of happiness, we know that agents are willing to sacrifice short-term happiness for greater long-term expected happiness, although finding a precise measure of the resulting net sum of happiness is problematic. Similarly, a no-harm instrumentalist may be willing to act to promote the happiness of someone else if it means that he expects to receive a greater amount of net happiness in the future. For purposes of the present discussion, the more interesting case involves the trade-off of interests that occurs among different people, a point I address in the next few paragraphs.

Summarizing some of the cases discussed above might make the important differences easier to identify. The callous instrumentalist is not concerned with the specific state of the other person, only that the state which is expected to improve his own happiness. The destructive instrumentalist wants the other person to be unhappy because it will increase his own happiness. The no-harm instrumentalist wants the other person to be happy only because his own happiness is thereby improved. Complications discussed include cases in which trade-offs of interests occur and cases in which the happiness of others is seen as an end in itself and not a means to increase one's own happiness. Taken together, these cases begin to demonstrate how broadly the term interest can be construed and why it is important to draw such distinctions between self-interest and non-self-interested motives and among different forms of self-interest. Simply stated, most forms of self-interest do not exclude the possibility of promoting the happiness of others either for instrumental reasons or as ends in themselves. Determining the focus of the act, whose interests are being served by an act, will be important to assessing whether an act is motivated by self-interest or something else.

Self-interest is as a type of rationality. In previous discussions I have set the foundations upon which I can be more explicit about what I mean when I use the term self-interest. In standard, everyday language, the term self-interest is used to describe a motive to act, a type of behavior, and an emotion, all of which are partially true. However, I will follow Joseph Butler's lead and take self-interest to be a general motive to act. I will be more explicit about what I mean by general motive when I discuss Butler's: "Sermons on Human Nature", but for now I will say that self-interest should be understood as having an organizing function within human psychology. By which I mean that it serves to align more basic motives, such as passions, appetites, desires, preferences and dispositions, with aims such as, for example, the aim to make one's life as happy or pleasurable or as excellent as possible. In this sense, the definition I use corresponds well with Parfit's (1984) claim that self-interest is a type of rationality. I should stress that self-interest is a way to align basic motives with aims; it asks if the set of motives a person has are the best set of motives he can have in order to achieve his specific aims.

What does the previous discussion have to do with the case of open source? In the introduction I provided an outline of some of the motives researchers have attributed to open source contributors. For instance, Hars and Ou (2001) interpret internal motivation to mean that open source participants are "selfless". They claim that advocates of this position mean that open source programmers are not motivated by monetary incentives but by their own hobbies and preferences or that they receive rewards from increasing the welfare of others. However, now armed with the distinctions I have drawn above, it should be clear that the fact that open source programmers may receive non-monetary rewards or rewards from increasing the welfare of others in no way shows that they are non-self-interested. Claiming that someone is motivated by internal motivation says, by itself, little about whether that motivation is self-interested or not. The brief outline of the types of self-interest and the range of behaviors that that motive can entail should serve as a warning to those attempting to categorize the behavior of open source programmers and other economic actors. I will argue in the final chapter that Hars and Ou's (2001) attempt to categorize the motives of open source programmers along external and internal lines and attempts by the other authors mentioned to distinguish between self-interest and non-self-interested behavior are largely oversimplifications based on misunderstandings about what self-interest implies for behavior.





## CHAPTER TWO: ALTRUISM AND ALTERNATIVES TO SELF-INTEREST

---

### Introduction

In the previous chapter I showed the range of behaviors that can fall under the heading of self-interest. However, if self-interest is assumed to be the exclusive motivational driving force, as many economists and others assume, then it becomes difficult to account for many commonly observed phenomena. In order to find the limits of self-interest it might therefore be helpful to examine which actions an exclusive dependence on that motive would exclude. In other words, if self-interest is taken to be the only motive, which other motives are denied and what sorts of behavior become unintelligible? I will first discuss some relevant alternatives to self-interest before presenting a few practical examples that may call into question an exclusive reliance on self-interest as an explanation of many everyday behaviors. The alternatives are practically relevant in that they have been mentioned as motives for open source contributors and theoretically relevant in that they will help me to set the parameters of my proposed framework.

There are many reasons why people may not act in their own interest, but given my research question I will limit my discussion mainly to what are known as particular affections, duties, and special obligations as examined by Samuel Scheffler (1997). I will then discuss whether the motive of altruism is as 'natural' a motive as self-interest as discussed by Thomas Nagel (1986). I will have much more to say about particular affections when I discuss Joseph Butler; for the present, I will mention that particular affections can be thought of as certain types of passions and appetites that can be directed towards objects that favor or, more importantly for purposes of the current discussion, harm private or social interests. It is important to keep in mind that certain motives can damage as well as aid our interests. Butler's insight was to argue that such motives at times drive or push people to work against their own interests and to provide a framework in which to place self-interest.

### Particular Affections

I find everyday examples the most compelling, and the instance of overeating is something many people in richer countries face or will face at some point in their lives, but overeating is only one of many possible cases of weakness of will; feel

free to choose an example that you find most convincing. People often feel compelled to eat even when they know that it is not in their interest to do so. For example, if someone is overweight, it is not a good idea to eat a big bag of chips or cookies in one sitting, people know it is unhealthy to do so, yet they do it anyway. In terms of desires, it can be said that they desire, at that instant, to eat the chips more than they desire to maintain a healthy lifestyle or in terms of preferences, they prefer to eat the chips, at that instant, more than they prefer the aim of a healthy lifestyle. The assumption that people *know* that the act is unhealthy is important and critical to the argument. There is an internal conflict that takes place at such moments, upon reflection people know (perhaps even at the moment that they are eating the chips) that it is not in their best interest to eat chips and damage their health, but they do it anyway. Perhaps it is only a matter of separating the impacts of short-term versus long-term desires and preferences, but in any case, such examples then make it difficult to conclude that people always act in their best interest when they knowingly act against it. Again, I will have much more to say about these types of cases in later chapters particularly as it relates to the decision to use an open source license, the important point is that a theory of motivation should account for such behavior.

The motive of revenge raises similar issues, but the period over which the motive extends can be much longer than in the case of overeating or other urges and so may therefore be more convincing. Leaving aside ethical issues, interests can be negatively or positively served by the desire for revenge (Diamond, 2008). Although revenge can act as a sort of energy pushing people to do greater things to overcome or to 'get-back' at the object at which their revenge is directed, the focus here is on those cases in which revenge leads someone to a destructive fixation on the object inciting their revenge. The motive of revenge can last an instant or a lifetime; those that last an instant can be very destructive for all parties concerned. However, revenge can also last a very long time; the lives that revengeful people lead may in great part be formed by the real or presumed slights they have suffered. The sometimes long duration of revenge means that reflective people have the time to consider the source of their motives. They may *know*, just like the overeater, that such motives are destructive and counter-productive, but they cannot help themselves. Once again, although revenge can lead to good outcomes, it can also lead to destructive behavior for both the person experiencing such desires and the persons to whom those desires are directed; desires that to an objective observer may appear out of all proportion to the original damage suffered. Similarly, pity or sympathy can be of social benefit by providing comfort, but excessive or

misdirected pity can be personally or socially damaging. For instance, excessive pity can lead to self-neglect, while misdirected pity can mean that other equally or more deserving causes suffer.

To conclude, particular affections, whether for food or revenge or pity or something else, can damage what upon reflection the person himself and outside observers know to be in a person's real, long-term interests. They are not necessarily urges or instincts to which people immediately, subconsciously, react, their duration can last a lifetime. An advocate of the purely self-interested approach (e.g., an advocate of psychological egoism) would have the difficult task of bringing such motives under the umbrella of self-interest. It has been suggested, although not thoroughly argued, that people contribute to open source projects as a mild type of revenge against proprietary software (Raymond, 1999). We know, for instance, that a motivating factor for Linus Torvalds' decision to initially issue Linux free of charge was his disappointment at not being able to receive a UNIX type operating system for his desktop machine (Moody, 2001). Whether revenge is a good reason to contribute to open source projects will be addressed in the final chapter.

## Duties

Thomas Nagel mentions deontological reasons (duties and obligations) and special relationships as two categories that may give agents reasons to pursue goals other than their self-interest; I will very briefly discuss each of these in turn. Duties have been mentioned as a motive to contribute to open source projects explicitly by Feller, et al. (2005) and Barbrook (1998, 2005) and implicitly by von Hippel and von Krogh (2003). Many people take duties or obligations--whether they believe the duties are derived from a greater being, intuitions or rationality (or some combination thereof)--as fundamental, by which I mean that they are requirements that precede considerations of interest. These people consider duties to be intrinsically right or good and, by themselves, provide enough of a reason to act. For example, many people believe that telling lies is wrong and should be avoided; which is not to say that a person who believes that it is a duty to tell the truth will necessarily sacrifice everything or nearly everything to avoid telling a lie, it only says that the obligation gives a good reason by itself not to lie. This point is perhaps easiest to see by comparing obligations to the case of an ethical egoist. An ethical egoist might claim that it is wrong to tell a lie, but only in cases in which lying damages one's interests. If lying damages a person's interests, then the strict

ethical egoist will claim that it is not wrong to lie, while a person who believes that he has an obligation not to lie considers lying to be intrinsically wrong, whether it benefits or damages his interests. Many other duties, including considerations of justice, good faith, and the necessity of treating people as ends rather than means, are considered to provide good reasons to restrain or even disregard self-interest. For instance, the topic of justice or fairness has received much attention from experimenters in economic contexts (see e.g., Kahneman, Knetsch and Thaler (1986a, b), Dawes and Thaler (1988)). These experimenters have shown that a large minority of people appear to be willing to sacrifice their interests in support of fair play.

### Special Relationships

Some relationships, for instance those involving friends, family, citizenship or trusteeship can be a source of duties that further restrain interests. The issue at hand is more subtle than an economic obligation that results from some type of transaction. Samuel Scheffler (1997) has captured the idea that I would like to convey by drawing a distinction between what are essentially transactions and those obligations that derive from the nature of the relationship itself. In the first case we often feel that we have a special responsibility for certain groups of people because of shared interaction that derive, for instance, from something like receiving a gift or making a promise. These are cases in which something was done by one of the two parties and that act is the source of a responsibility. The other obligation that Scheffler describes does not derive from a transaction, but from the relationship. For instance, people may feel that they have special responsibilities to relatives, friends, neighbors, or even co-workers and clients because of their *relation* to those people or groups. It may be the case that members of a collective never meet or interact with one another and yet still feel a mutual obligation. For instance, people from the same country may feel a certain responsibility for the well-being of fellow citizens despite never having interacted or expecting to interact with them and perhaps sharing nothing else in common other than nationality. I need to stress that these feelings of responsibility do not come about because anything was done to incur them; nothing was given or received with the implicit or explicit intention of getting anything in return. Rather, such responsibilities often arise because one stands in a certain relationship to another person. These duties may be very strong despite their involuntary nature as, for instance, in the case of the responsibility that many family members feel for the well-being of relatives even in those cases where they didn't choose to be

related. As noted, in regards to open source, several authors have argued that membership in open source communities may create a duty or a special relationship among members of the community. If so, then Scheffler may argue that these relationships create obligations which may transcend interests.

### Natural Altruism

Human beings are by nature self-interested, but whether all human activity is motivated solely by self-interest has been a matter of debate for thousands of years and acceptance of the positions I outlined above would seem to raise some doubts about whether self-interest is an exclusive motive of behavior. I now turn to the specific question of whether altruism is as natural an explanation of human activity as self-interest. My point is to show that altruism or other non-self-interested motives are not necessarily any less rational than self-interest as has been suggested by some researchers working on the topic of open source (van Wendel de Joode, 2005). Accepting the theory of natural selection would seem to imply an acceptance, to some extent, that humans are hardwired to be self-interested in the sense that, at a minimum, they are in normal circumstances compelled to avoid injury, disease and death, and they actively seek to promote their own well-being through such activities as seeking food and shelter. There are extreme cases that provide exceptions to these rules, but they are rare and can be thought of the exceptions that 'prove' the general rule of self-interest. However, does natural self-interest necessarily exclude a natural form of altruism? The question is whether altruism, like self-interest, can be shown to be an inherent part of our reasons for acting. Where by altruism, following Nagel (1970, p. 79), I do not mean abject self-sacrifice, but a willingness to act in consideration of the interests of others without the need of ulterior motives and the possibility that one's own interests may be severely damaged.

I will briefly discuss two approaches that ask whether altruism is as natural a motive as self-interest. These two approaches, one empirical and the other philosophical, capture the gist of their respective general categories (see Searle 2001 or Kavka 1986 for other approaches). I do not claim to solve the complex issues within these respective debates, my modest goal is to show that there may be a place for real altruistic behavior.

The first approach discusses the possibility of altruism from the perspective of socio-biology (e.g. Dawkins, 1989). Defenders of the position known as

behavioral egoism claim that truly altruistic acts do not exist, although instrumental, pseudo-altruistic acts may (Kavka 1986). The theory predicts the elimination over time of genes that give rise to actions that sacrifice an organism's ability to pass on its own genes (Kavka 1986, p. 56). Therefore, according to this theory, altruistic acts that expend energy and resources for the benefit of other organisms will gradually be eliminated to the degree that they reduce the altruistic organism's viability. The validity of behavioral egoism has been called into question on at least three accounts, the first is through the observation of animals and the second is its founding on a narrow form of genetic determinism. Contrary to what the theory of behavioral egoism claims, many animals appear to engage in altruistic acts, e.g., they share hard-won food and defend co-species by, for example, uttering warning calls that focus a predator's attention on themselves (Kavka 1986, p. 57). If true, then even creatures whose behavior is more tightly linked to instincts derived from their genetic make-up than that of humans would seem to have the capacity to perform altruistic acts. More troubling for this theory is its narrowly defined genetic determinism. Behavioral egoism rests on the claim that actions are determined by genes; a claim that at some level is obviously true. Absent superhuman intervention, all of our actions are genetically determined in that we operate within the limits of our genetically determined physical capacities--our brains etc., are created from genetic blueprints. However, our genes have given us more than the ability to react in a purely instinctual manner, many believe that we have the ability to reflect and change our behavior within limits. As a consequence, our genetic code may allow us the ability to choose to act altruistically. The conclusion from Kavka's (1986) literature review indicates that altruism is not necessarily eliminated from the gene pool. He provides evidence that seems to indicate that there is evolutionary space for at least some members of a group to be altruistic some of the time. A third issue with this theory concerns its applicability to modern human life. Although an exclusive reliance on self-interest may accurately describe existence on the Malthusian edge between life and death, modern people living in wealthy nations can be altruistic without necessarily endangering their ability to pass on their genetic material. The degree to which one needs to sacrifice in order to be considered altruistic is a topic to which I will return.

### Thomas Nagel's Possibility of Altruism

The second approach that I discuss is based on Thomas Nagel's (1970) argument that altruism can provide a reason to act. Nagel's argument is complex and I will only be able to sketch it here. Nagel draws an analogy between prudential (self-interested) reasons and altruistic reasons. He summarizes his approach in the following, "I will argue that our own future interests provide us, by themselves, with reasons for present action to secure them, and that motivation of this sort cannot and need not be explained by intermediate present desires, or any other intermediate motive" (Nagel, 1970: 16). As I understand it, Nagel is arguing that we are able to place ourselves into our future selves because we are the same person through time; we are in a sense timeless. The argument follows, "... that prudence is not fundamental, but derives from the requirement that reasons be timelessly formulable, so it will turn out that altruism is not fundamental, but derives from something more general: a formal principle which can be specified without mentioning the interests of others at all" (Nagel, 1970, 87-88). His argument is that just as we have no difficulties acting prudentially because what makes us what we are is timeless, we can act altruistically because of the realization that we are one among many. It is this recognition that provides a structural reason to act in the interests of others. Searle (2001), I believe accurately, expresses doubts about the applicability of the analogy between prudence and altruism. There is clearly a significant difference between the prudential case, in which a person sees himself as timeless, and the altruistic case, in which he must see himself as one among many. However, rejecting the analogy does not imply rejecting the rest of Nagel's argument. The ability to, in a sense, disengage ourselves from time, may as well imply an ability to imagine the desires of others as providing sufficient reasons to act. Nagel's attempt is not perfect, there is the problem that his argument, for instance, rests on a somewhat awkward differentiation between motivated and unmotivated desires. However, I will have to leave the argument here, and only suggest that although Nagel's approach is not the final word on this subject, it raises the possibility that altruism is as natural a motive as self-interest (Searle 2001; Dancy 2000).

To conclude the previous theoretical sections, there appear to be reasons that people act in contrast to what they know to be in their self-interest, whether it is via particular affections or beliefs in certain duties or obligations that they have towards other people. In addition, there are also ongoing debates in genetics and

philosophy that leave open the possibility that people may be as naturally inclined to act altruistically as they are to act from self-interest.

## Altruism

The previous review of the motives of open source contributors raised two relevant issues in regards to altruism, namely, what is altruism and is it rational. Keeping in mind Popper's (1985) and Ozinga's (1999) respective warnings as to the limited value of defining terms in general and definitions of altruism in particular, I will develop a working definition of altruism that can be used as a guide to whether altruistic acts can be distinguished from other sorts of acts.<sup>2</sup> This will allow me to ask whether open source contributors are altruistic. If it turns out they are, then perhaps altruism can be stimulated or facilitated or at least not hindered in playing a role in other economic processes. If what is called altruism is in fact self-interest or pseudo-altruism, then the antagonism existing between motivating self-interest and altruism behaviors can be ignored (Frey, 1997). Altruism is sometimes represented as an irrational alternative to self-interest, (van Wendel de Joode, 2005; discussion in Weber 2004). I will therefore also need to ask whether altruism is irrational in the sense of being a self-defeating theory.

What is altruism? I suggest that a good working definition of altruism is that it gives to each person the aim of achieving those outcomes that will bring the greatest expected benefits to others over time. A distinction has to be made between living an altruistic life, and performing an altruistic act. A person attempting to live an altruistic life will be charged with finding a balance in his intentions, a sort of weighing of intentions that together, over his lifetime, leads to the intention to perform a set of actions the provide the greatest expected benefit to others. The intention of an altruist is to benefit others, whereas the intention of a self-interested person is to benefit themselves. In contrast, an altruistic act is simply an isolated action performed with the intention of helping others. I will be primarily concerned with understanding acts of altruism, but I will also discuss the

---

<sup>2</sup> David Collard (1978) defines self-interest as the case in which the marginal weight that a person attaches to the welfare of others is less than that which he attaches to his own welfare. Collard's approach to building economic models of altruism is quite interesting, but it is in a sense one step beyond mine in that he is interested mainly in preferences and utility functions, while I am interest in what stands behind preferences. Many of the articles in Mansbridge (1990) begin from a similar position.



more general case of living an altruistic life because it is often used as a scapegoat in the economic literature. Altruism is by definition irrational if the self-interest theory is accepted, however it might also be argued that altruism is irrational because it is self-defeating by which I mean that it fails by its own terms. This argument goes that by believing and acting to provide the most benefit to others the altruist deprives himself to such an extent as to be unable to perform altruistic acts in the future. For instance, an altruist living in the West sees the relative poverty in Africa and concludes that he should immediately give away his resources to the point at which the marginal value of a dollar he faces is equal to the marginal value of a dollar for the recipients of his charity. In this case, the actions of the Westerner might be self-defeating in that he damages his own interests to such an extent as to be unable to perform beneficial acts in the future. The actions of a self-defeating altruist, like the instant hedonist, are untenable over time. A person who is always willing in all situations to sacrifice his own preferences and desires for the benefit of others will quickly exhaust his ability to help himself and, as a direct consequence, others. Instant hedonism and self-destructive altruism both fail because of their short-sightedness. It is interesting to note that the actions of a self-defeating altruist, in contrast to the instant hedonist, are sometimes valued by society.<sup>3</sup> For example, a person willing to significantly sacrifice his own interests for a worthwhile cause, even at severe cost to himself, is often celebrated by society.

Although the argument that altruism is self-defeating works against the sort of self-destructive altruism outlined above, it fails to work against altruism as I have initially defined it. The theory of altruism which I propose proposes that an altruist needs to look at the benefits he can provide over time. If giving away all of his resources in a rash act of supreme sacrifice leads to the altruist's ruin then perhaps that act isn't the way to provide the most benefits over time and, as a consequence, not as heroic as it might first appear. An altruistic act, like a self-interested act, has to be seen as one act within a set of actions over time. This doesn't exclude the possibility that there might be circumstances when a great sacrifice is recommended (altruism must be voluntary), but those are the exceptional cases. In

---

<sup>3</sup> Although some individuals want to be known as altruists, that isn't generally an acceptable role for publicly traded firms. The role of for-profit firms is to maximize profits, a role that is in some cases legally enforceable. The question of why we don't expect or want firms to behave altruistically is interesting, but one that I won't address. My interest is in how the interests of individuals, not firms, should be modeled.

terms of open source and altruism, whether the act of contributing to open source projects is altruistic will depend on whether the altruist intends and expects his contributions to benefit others over time.

Can an altruistic act be useless or unwanted? For instance, is making a piece of code freely available an altruistic act if the code has no expected value? Casson and Giusta (2006) define altruism in terms of the obligations that actors undertake. While defining an altruistic obligation they write: "An altruist wishes to help another person whether or not that person's expectations have been disappointed, and whether or not they have a specific need for care." This understanding of an altruistic obligation raises a general issue that will appear again when I discuss the work of Jensen and Meckling (1994) and Jensen (1994). The definition doesn't mesh with altruistic theory, to be an altruistic act the intention of the action has to be of benefit others. Helping those that do not need help should not count as an altruistic act. However, if the code might be of use to someone in the future, and the contributor makes the code available for that reason, then the act is motivated by altruism. Not everyone has to benefit from the act, but the intention is that someone will. Although attempting to live an altruistic life implies that the lifetime set of outcomes should be directed at achieving the greatest benefit for others, an altruistic act doesn't require that the act is the most altruistic act available at a moment. Given the choice between two action, an altruist isn't required to perform the act yielding the greater benefit, just a person acting for self-interested reasons isn't required to perform the actions that is the most self-interested in order to be considered self-interested. As with any approach, one's expectations and calculations concerning an action might be wrong, but if an altruist intends that a certain act would provide benefit to another, all else equal, then the act should be performed.

Does an altruistic act need to involve sacrifice on the part of the altruist? Is an action altruistic only if it damages the actor in some way? This is a common stipulation in many definitions of altruism, however, an altruistic act doesn't necessarily require a sacrifice on the part of the altruist. The theory of altruism concerns the intention behind an action, whether the action involves a sacrifice is, I believe, largely irrelevant, given, as already discussed, that the altruist understands that he is charged with providing an amount of benefit to another. If a person performs an act that is "costless" for the reason that he wants to provide benefits to others, then that act is altruistic. Furthermore, it can be argued that all acts at some stage in their development or execution "cost" something, but what I have in mind

are significant costs that damage the altruist's interests. All activities, even thinking or writing a check, cost something, that alone doesn't make the action self-interested. A wealthy benefactor in the luxurious position of having the marginal value of an additional dollar close to zero can still perform altruistic acts by giving money away to those who need it. For instance, an altruist with the general intention of benefiting others makes valued software freely available to others, the fact that there are no additional costs whether ten people use the code or thousands makes no difference. The marginal cost is zero, but the benefits may be substantial.

Likewise, altruistic acts do not exclude the possibility that costs can be incurred by the altruist as Weber (2004) argues. I will discuss Weber's argument in greater detail in the case study chapter, but he seems to argue that because there are conflicts among open source programmers, they cannot be contributing to open source projects for altruistic motives. Because programmers are incurring psychological costs, the argument goes, they must be receiving compensation which Weber equates with improved reputations. The theory of altruism as I understand it, does not exclude the possibility that altruistic acts cost something, in practice they frequently do. Nor does an altruistic act need to be anonymous as Weber implies, it only depends on the intent of the actor. It is too easy, but common, to conclude that because a person receives some form of payment, an action cannot be altruistically motivated. Clearly, it makes it more difficult to convince a skeptic that a compensated act was conducted for altruistic reasons, but if the act was conducted for altruistic reasons, then compensation is a by-product, not a reason. Therefore, altruistic acts can even be profitable for the altruist. I see no theoretical reason to assume that acts that yield net benefits to an actor are necessarily self-interested; although I concede that such acts make it more difficult to practically, convincingly, invoke altruism as the key motivator.

Although the theory of altruism neither requires nor excludes costs on the part of the altruist, it is common that definitions of altruism contain some provision for net losses on the part of the altruist. This is because these definitions are assuming that a decision maker is facing a choice between an altruistic and self-interested act. The choice is made explicit in Gintis (2000, p.34): "By definition an altruist is an agent who takes actions that improve the fitness or material well-being of other agents when more self-interested actions are available." Given that choice, a person who chooses the altruistic act is, by definition, sacrificing personal interests for the benefit others. The situation Gintis draws is relevant in the case of open

source and most practical decisions. However, there are various important distinctions which need to be drawn between types of practical altruism, in other words, those situations in which a person is facing a choice between performing an altruistic act and a self-interested act.

Many practical cases of altruism involve an intentional sacrifice of one's own interest for the expected benefit of others. The practical altruist doesn't act as a self-defeating altruist because there are limits to his willingness to sacrifice his own, for instance, happiness for the happiness of others. I will leave open the types of trade-offs that the practical altruist is willing to make and only mention that there is probably some sort of nonlinear relationship at work. For instance, an altruist is probably more willing to make small sacrifices to his own happiness if it means very large increases in the happiness of others and much less willing to make large sacrifices to his own happiness no matter what the gains are to others. The issue of trade-offs aside, there are degrees of practical altruism that are important to distinguish. The first involves a reduction in the altruist's level of happiness, while the second concerns the case in which the altruist becomes unhappy. I call the first qualified altruism and the second unqualified altruism.

It is not difficult to imagine cases in which an actor cares for the happiness of another person, we do it all of the time, the question is why an agent cares about the other's interests, in other words, what is the focus of the action. Imagine two players trying to divide the contents of a fruit bowl.<sup>4</sup> The focus of the decision is on finding a method of sharing the fruit involving the influence of the altruist's actions on the happiness of the other person. Say, for instance, that the altruist marginally prefers an apple in the bowl to the banana, orange and pear that are there, but he also knows that the other person strongly prefers the apple; ignore for the present the complication of effects of the happiness of the other person on the agent's own happiness. In such circumstances, the agent might be willing to sacrifice his own preference for the apple for the less desirable banana because he knows that the act will improve the happiness of the other person. In fact, as long as his own expected net happiness remains positive he might be willing to take the pear which he least desires as long as it will still give him some happiness. The focus of his actions is split between his own happiness and the happiness of the

---

<sup>4</sup> See Amartya K. Sen's article, "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory" for the fruit bowl example that inspired my own.

other person. While he is still focused in part on his own interests (he will, after all, not act unless he expects his happiness to remain positive), he now considers the happiness of the other as well as an end in itself. He is willing to accept a lower amount of happiness if it means the other person experiences an increase in happiness. I think that this act can best be described as qualified altruism because the agent is willing to suffer a net-loss in happiness as long as his overall happiness remains positive and happiness of the other person improves. However, at the point at which a person is willing to reduce his net happiness for the benefit of another, I propose that we move to a different sort of altruism. For instance, if an agent is willing to promote the expected happiness of someone else while suffering unhappiness, then he is, by many definitions, an unqualified altruist, if not yet the sacrificial altruist described above. An unqualified altruistic agent gives the other person the apple even if it means that he has to eat the pear which he detests. I would add the stipulation that the actor must intend, his reason for performing the act, is to benefit the other person.

The next complication that I will examine is the case in which the happiness of an altruistic agent is influenced by the happiness of the other person. This is a slight variation of the case of the no-harm instrumentalist which I will soon discuss and emphasizes the difficulty of identifying motives. It might plausibly be argued that the altruist in the previous paragraph sacrifices the apple for the happiness of the other person and eats the pear because it makes him happy to know that the other person's happiness has increased. It can be argued that the altruistic agent then experiences an increase in happiness despite having eaten a pear which he detests. If that is an accurate assessment, then even this apparently altruistic act might arguably be motivated by self-interest. At this point the argument can become confused; the definition of interests is in danger of becoming over inclusive. The above analysis fails to address an important aspect of the matter, which is why a person is made happier when someone else's happiness improves. The agent might be pleased to see the other person's happiness improved because he cares for the other person, period. It is not his own happiness that motivates the action, but the happiness of the other person. I would call this an altruistic act; the focus is clearly on the other person. However, an egoist might also want to improve the happiness of another person because of the feelings of happiness that he himself experiences, in which case the action is motivated by self-interest. The focus in the first case is on the other person's happiness and is therefore altruistic, while the focus in the second case is on happiness that the egoist experiences which I would consequently label an act of self-interest. The idea that people carefully calculate their expected

net interest may seem far-fetched; however, people daily perform rough estimates of the implications of their actions on their expected interests. The ability to correctly make important decisions depends on the ability to reflect upon the given options and arrive at a measure of their relative worth. Indeed, the relevance of the science of economics depends in large part on people exercising such abilities when making decisions. People can be wrong about such calculations for a variety of reasons, but that is another matter.

How do altruism and self-interest differ? It is often difficult in practice, and sometimes in theory, to distinguish between self-interested and altruistic acts. The practical difficulties will be clear when I discuss the case study, here I will address some of the common theoretical understandings of altruism. Elias L. Khalil (2004) provides a helpful overview of how altruism has commonly been used in the economic literature. He divides the literature on economic altruism into egoistic, egocentric and alter-centric categories.

Khalil writes: "The egoistic perspective, best expressed in the work of Robert Axelrod (1984), maintains that altruistic assistance would be offered if one expects future benefit." As Khalil (2004:100) explains, this view presupposes repeated games where beneficence is modeled as a non-myopic self-interested strategy to ensure future cooperation (e.g., Bergstrom and Stark, 1993, Kurz, 1978). This type of altruism doesn't appear to differ from non-altruistic transactions. The conditions are strong, it doesn't, for example, explain many types of altruism including anonymous gifts, and aid given to strangers one never expects to see again. It doesn't explain open source either--where users are under no obligation to reciprocate if they only want to use the code for private consumption. The "egocentric" view, according to Khalil, is associated with Gary Becker (1976), who argues that the donor's utility function includes the utility of potential recipients. That is, "the donor would give if the enjoyment of watching the pleasure of others exceeds his satisfaction of consuming said community." The action, as its name implies, is not motivated by altruism, the focus of the action is on the actor's happiness; the reason the action is performed is to increase the actor's happiness. Finally, Khalil writes: "The "alter-centric" approach can be surmised from the works of Robert Frank and Herbert Simon. It views the benefactor's actions as stemming from a personality trait that arises from artificial selection. Actions performed for alter-centric reason may not be altruistic acts because the intentions of the agent are unclear. Is the actor free to make the altruistic choice or are his

action genetically determine like a bee's relationship to the hive? Altruism requires free will while alter-centrism appears to take that trait away.

Khalil's nomenclature, particularly his term alter-centrism, points out the importance of distinguishing between real and pseudo-altruism, those actions which appear to be altruistic but are not. There are generally thought to be two types of pseudo-altruism. The first, called hard-core or hard-wired altruism, is defined as altruism on behalf of kin (Wilson, 1975). For instance, an individual may be genetically programmed to behave 'altruistically' toward his relatives in order to promote the continuation and dissemination of his own genes--bees, ants and termites being obvious examples. This is a Darwinian approach that has been shown to predict certain human behavior (Trivers, 1974), and has had some success in explaining the extent to which certain insect cooperation is correlated to the degree of their biological relatedness (Wilson, 1975). These cases appear to provide evidence for a sort of programmed, determined, drive to pass on one's genes. They are pseudo-altruistic acts because although they appear to be altruistic they are, in fact, conducted to promote one's own genetic interests. This approach is also controversial; many of us would like to feel that we have the free will to overcome natural programming, while the degree to which we are free to do so is precisely the debate. A second form of pseudo-altruism is called reciprocal egoism, it is exemplified by organisms that only cooperate with one another, for example in hunting together, sharing food, and grooming, because these acts serve their own self-interested needs. Such acts can be seen as beneficial acts that generally 'pay' for previously received benefits or 'prepay' for future expected benefits. This is clearly a type of instrumental altruism rather than a form of real altruism. It is important to realize that not all behaviors that appear to be altruistic are indeed altruistic, but the fact that some behaviors are pseudo-altruistic does not detract from the observation that real altruism is possible even if it is not hard-wired into our being.

### Practical Examples

At this point, in terms of the overall purpose of my manuscript, it may appear as though I have painted myself into a corner. I have expended quite some effort in the last paragraphs arguing that it is difficult, at times even for the agent himself, to identify the motivation or intentions behind actions. This might be interpreted by an advocate of self-interest, say a psychological egoist, as confirming their position, and an indication that all acts are indeed at their core motivated by self-

interested. However, this would be a misinterpretation, all motives, whether self-interested or otherwise, face the same difficulties in regards to interpretation. The general historical and modern acceptance of self-interest as the default motive means that a lack of conclusive evidence that other motives are at work is often interpreted as support for the belief that self-interest is the only or predominant motive. As previously mentioned, Becker's (1976) recommends precisely such an arbitrary approach. However, there are both extraordinary and mundane examples that might cause even a psychological egoist to pause and perhaps reconsider his position.

The task at hand is not easy; even an extreme example, such as a soldier sacrificing himself for his comrades by diving on a hand-grenade can be construed as a self-interested act. The self-interest in the grenade example is sometimes explained by speculating that the soldier decides to sacrifice himself because if he chooses not to the guilt of knowing that he could have saved his buddies and did not will haunt him for the rest of his life and cause him only misery. The soldier is presumed to perform something like a net present value of pleasure calculation and decide that it is better to sacrifice his life. His decision to sacrifice himself is then presumed to be one of self-interest because he wants to save himself a lifetime of guilt. In addition, the type of calculations involved depends on the options being considered, large sacrifices, say for instance those that involve risking one's life, probably involve a different sort of calculus than day-to-day calculations of net self-interest. These points of calculation aside, I find that reducing all explanations of heroic events to cases of self-interest unconvincing. It is possible to explain such events using self-interest, but in such cases Occam's razor might be called upon to do its business of paring down the explanation to the more parsimonious explanation that the soldier or Mother Theresa or some other hero is acting from a concern for others and simply disregards his own well-being. There is equally no doubt that such extreme forms of altruism are the exception to the more common mix of motives found within the average person.

A more mundane example that is difficult to explain under the rubric of self-interest is the case of life insurance. Why would a strictly self-interested person buy life insurance? For example, it could be argued that a parent buys life insurance for his family because it gives him a good feeling while he is alive to know that his family will be provided for upon his death. However, a purely self-interested person should not care whether the policy pays out or not upon his death. Say for instance that the parent is only presumed to be dead and is in fact safe on a



deserted island. Would he be disappointed upon returning to civilization to find out that he had bought a bogus policy and that his family has been financially suffering? If he is purely self-interested he should not, because while on the island he had been under the false impression that the policy was valid and received the psychological benefits of so believing. In other words, he should be indifferent to whether the policy was valid or not, as long as he believed it was valid when he was alive. The thought experiment is admittedly a stretch, but I believe that most would agree that policyholders really do want their families to benefit from the life-insurance policies they buy for their families even when they are not alive and therefore unable to receive pleasure when the funds are dispersed. They do not buy insurance for their own, self-interested pleasure, but with the thought that their families will benefit from the policy after their death. In a similar vein, we may care about the interests of others we only briefly meet and would be disappointed to later learn that their interests had not been met. In general, we do not want to experience pleasure or happiness which is based upon false beliefs and, more generally, many do care and are moved to act by events that do not have a direct impact upon their interests.

Many people clearly do care about their legacies and such interests cannot be derived from purely self-interested motives. For instance, it is difficult to imagine how self-interest can explain anonymous gift giving to charities through provisions in wills. Anonymity is an important qualification to control for those cases when people give to an organization for the psychological benefits that they receive from knowing that others are aware of and appreciate their activities, e.g., receiving esteem because of their donations. However, once again, it could be argued that anonymous donators give because giving makes them feel pleasure today, although the recipients will not receive the benefits of the donation until after their deaths. At this point in the argument we return to the original question posed which is, why do acts of generosity give actors pleasure? Do people give because it makes them feel good and it is in their interest to feel happiness or because it makes the beneficiaries feel pleasure regardless of the feelings of the benefactor? Henry Sidgwick (1981: 51) wrote that "... men have sacrificed all the enjoyments of life, and even life itself, to obtain posthumous fame: not from the illusory belief that they would be somehow capable of deriving pleasure from it, but from a direct desire of the future admiration of others, and a preference of it to their own pleasure." The key to Sidgwick's claim is in the desire for future admiration, when the donator is no longer present to experience admiration. The case of the anonymous donator presents an even stronger case. Why, if a donator is concerned

with his own interests, does he make his donations anonymously? The advocate of self-interest would have to argue that the anonymity somehow increases the well-being of the donator. The more parsimonious answer is the one suggested by Sidgwick, they donate because they want to improve the lives of future generations and not because they expect to derive pleasure for themselves.

Besides making a case for real altruistic behavior, another point of these examples is to show how difficult it can be for outsiders to interpret the underlying motives of an individual, and in some cases even the individual concerned may not be able to completely identify his own motives. The crux of the problem is that observed behavior does not always reflect motivation and there may be cases in which people intentionally hide their true motives. So, while we observe behavior, we cannot directly observe motivation and the same action can be motivated by many different motives. One person may give to charity primarily because it makes him feel good regardless of the effect that the gift has on beneficiaries, while another gives because it makes the beneficiaries feel good. Identifying the motives that prompt action is difficult, but that is true of all motives whether self-interested or otherwise.

Although there are no guaranteed methods to identify motivation, several indicators of motivation are available. Good indicators of underlying motivation might include repeated actions by an agent that have predictable consequences, and actions that require obvious, substantial commitments on the part of an agent. A proper understanding of self-interest requires a thorough investigation of the role it plays in human nature. Ethicists have spent thousands of years investigating what it means to be self-interested and so provide the appropriate questions that need to be asked if a more comprehensive and coherent theory of economic self-interest is to be developed (see chapter 5).

#### Identifying Motivation

Although the question of whether people are motivated by self-interest alone would appear to be one that could be solved empirically, I have demonstrated that that method always leaves room for skeptics to interpret what would appear to be non-self-interested motives as self-interested. The brief discussion in previous paragraphs should provide some evidence of the difficulty of assessing what I am calling the focus of an act--whose interests are being served by an act. In short, there is no way that I know of to prove that a person is acting from some motive

other than self-interest, just as there is no way to prove that a person is only motivated by self-interest, the important point is that the same reservations hold for any motive. The question that I address in the following paragraphs is why motives are so difficult to identify. I mention several reasons including: the problem of identifying the true focus of an act; uncertainty on the part of the agent over the source of his own motivation due to self-deceit; and, social pressures to be self-interested. I also examine the related question of when an act should be considered altruistic as opposed to self-interested. This is an important issue because a wide enough definition of self-interest can, incorrectly, be made to incorporate altruistic acts. Therefore, I will attempt to define when an act should be considered altruistic and the problem of confusing altruism with pseudo-altruism. Finally, I mention several potential challenges to a narrow interpretation of self-interest.

#### Focus of an Action

As shown in previous paragraphs, it can be difficult to identify the motives that stand behind observed behaviors because doing so requires identifying the 'true' focus of an act. Suppose that we observe or read about an agent who makes anonymous donations to a charity. The first inclination might be to call this behavior altruistic, in fact, it might be an altruistic act according to all of the agents involved. However, if the donator experiences a feeling of pleasure by the act of donating it might be argued that the act is motivated, at least partially, by self-interest. The argument is that because the donator expected to receive the psychological 'payoff' of happiness then the act is motivated by self-interest. However, such reasoning does not get to the important question of why the donator expects to receive pleasure from his act. For instance, the donator might argue that he experiences pleasure from the belief that he has improved the well-being of the beneficiaries and not from the pleasure that he receives from giving the gift. His argument is not that he does not experience pleasure, but that the source of that pleasure is the belief that he has helped others. Perhaps another donator derives pleasure from making a donation because it gives him a feeling of superiority over those who benefit from his donation. The results of the donations may be the same, but the focuses of the actions are very different. In the first case the focus is primarily on the happiness of the recipients and is therefore an altruistic act, while the focus in the second act is on the happiness of the donator and is therefore motivated by self-interest. Becker (1976), as we have seen, clearly supports the second view. When researchers report that people experience feelings of

happiness when they give to charities we have to know why they feel the happiness they are experiencing. Do they experience pleasure because they are helping others or are their pleasurable feelings an end in themselves? Recent evidence seems to indicate that three areas of the brain are involved in altruistic decisions: the part of the brain associated with rewards associated with sex, money, food and drugs; the part of brain associated with the bonding of mother and child; and, most importantly for present purposes, the part involving conflicts between self-interest and ethical beliefs (Moll et al., 2006). This research would seem to suggest that there are a mixture of motives involved in the decision to perform an altruistic act, only some of which can be traced to self-interest.

#### Agent Uncertainty and Self-Deceit and Community Standards

Another difficulty with assigning motivation to a behavior is that the underlying reasons that induce feelings of happiness may be ambiguous, even to the person performing such acts. For instance, suppose that a donator experiences a feeling of happiness when he makes his contribution because his parents used to reward similar types of behaviors when he was young or perhaps he believes he will receive a better after-life if he does good deeds in this one. In the first case the donator is motivated to make a donation because of the earlier rewards he received, while in the second case the donator expects to receive a positive payoff for eternity for the good deeds he performs in this world. The second case is clearly one of self-interest, but I include it in order to emphasize that self-interested motives can have beneficial outcomes. The first case is the more interesting one because a proper assessment of the case requires an understanding of the intentions of the agent. Suppose that the person is unaware that he is responding to previous incentives, in that case his action may be described as altruistic because he does not know, and perhaps cannot be expected to know, that he is responding to deeply ingrained experiences. However, if he is aware that he makes donations only because of his past training, rather than a present concerns for the well-being of recipients, then I think his gift giving is a case of self-interest because his actions are motivated by the previous pleasure he received from making his parents happy, so his motives are internally focused, rather than focused on the well-being of the recipients. The argument revolves around the motivation of the donator, and the question is whether the focus is on his interests or on the interests of the recipient.

Another type of ambiguity is self-induced. Attempts to be truthful to others about our motives may be spoiled by self-deceit or living an unreflective life. We are all

victims (and beneficiaries) of self-deceit, Adam Smith regarded it as an important source of human misery: "This self-deceit, this fatal weakness of mankind, is the source of half the disorders of human life. If we saw ourselves in the light in which others see us, or in which they would see us if they knew all, a reformation would generally be unavoidable. We could not otherwise endure the sight" (Smith 1976: III.4.6). Presumably Smith believed that if people could see their behavior as others do, they would change for the better. On balance, I believe Smith's assessment is correct, although there are cases in which people underestimate the social benefit of their behavior. People may unintentionally deceive themselves into believing that they are behaving altruistically when in fact they are being self-interested--making the proper assessment of such acts by outside observers all the more difficult.

People may also be influenced by how altruism and self-interest are perceived in their communities. For instance, if altruists are considered naive or 'soft-touches' to be taken advantage of, then there may be good reason for an altruist to under-report his altruistic tendencies. If, on the other hand, self-interest or selfishness is portrayed as being "realistic" and a sign of a shrewd business mind, then there may be a tendency to over-emphasize and glorify the role of that motive. Self-interest is often portrayed as a source of economic good. It is not an exaggeration to claim that the most powerful message that many people take from popular economics is that individual self-interest leads to economic benefits for the group as a whole and is therefore good. As I will show, Bernard Mandeville and Adam Smith are often said, wrongly in the case of Smith, to have argued that individual self-interest always leads to benefits for the society as a whole. Clearly many self-interested acts improve the well-being of others, however, many do not. Actions by individuals motivated by self-interest can lead to socially beneficial outcomes, but self-interest is neither a necessary nor a sufficient condition to achieve beneficial economic outcomes. That it is not a necessary condition should be obvious, there is nothing to prevent altruistic individuals from performing altruistic acts that benefit society. That self-interest is not a sufficient condition for socially beneficial acts can be seen in practice by the number of regulations controlling such behavior. For example, even well functioning stock markets are closely regulated to prevent abuses of power and information, thereby limiting the ability of individuals with such power and information to maximize their interests. The point is that there may be reasons to distort, intentionally or not, one's motives of which community norms is an important example.

When is an act altruistic?

I begin this sub-section with quotes by Adam Smith and Thomas Reid. First Smith, from his book, *The Theory of Moral Sentiments*: "The only consequences for which he can be answerable, or by which he can deserve either approbation or disapprobation of any kind, are those which were someway or other intended, or those which, at least, show some agreeable or disagreeable quality in the intention of the heart, from which he acted. To the intention or affection of the heart, therefore, to the propriety or impropriety, to the beneficence or hurtfulness of the design, all praise or blame, all approbation or disapprobation, of any kind, which can justly be bestowed upon any action, must ultimately belong" (Smith 1976: 93). I believe that Smith has correctly identified the main issue at hand when assessing the motives behind an act, namely, the need to examine the intentions of the agent rather than the outcomes. Similarly, Reid writes: "Some figures of speech are so natural and so common in all languages, that we are led to think them literal and proper expressions. Thus an action is called brave, virtuous, generous; but it is evident, that valour, virtue, generosity, are the attributes of persons only, and not of actions. In the action considered abstractly, there is neither valour, nor virtue, nor generosity. The same action done from a different motive may deserve none of these epithets. The change in this case is not in the action, but in the agent; yet, in all languages, generosity and other moral properties are ascribed to actions. By a figure, we assign to the effect the quality which is inherent only in the cause (Reid, 1969, p. 773)."

As a consequence, an act should not be considered altruistic if it is motivated by self-interest. This is not to say that actions motivated by self-interest cannot have beneficial effects for others, clearly in the case of trade they can, but in those cases the act should not be called altruistic. On the other hand, acts motivated by altruism can have negative effects in at least two situations, namely, altruism can lead to a reduction in the agent's interests (a condition not, I have argued, required for an act to be altruistic), but it can also have negative implications for the well-being of others. For instance, should altruistic acts towards an individual who, perhaps foreseeable by the altruist, then goes on to perform heinous crimes still be called altruistic despite their consequences? The quote above by Smith clearly needs some subtle adjustment. Whether actions motivated by altruism that lead to negative outcomes should be labeled altruistic depends in part on whether the agent could have reasonably foreseen the implication of his actions. If the resulting negative effects could reasonably have been foreseen, then it is difficult to see how

these actions can be properly classified as altruistic. To take another example, an altruist should have a reasonable understanding of the impact of his actions on expected recipients. For instance, an altruistic act directed *solely* at persons who are in no need of altruism or inappropriate assistance to those in need is, at least, a less altruistic act, and perhaps not altruistic at all. The qualification of reasonable is critical and what is reasonable will depend on such factors as the amount of information available to make a decision and the importance of the decision to be made. If negative or positive effects could have reasonably been foreseen, then although the agent's intentions may have been altruistic, the execution of the intentions was misguided--the agent should and could have known better so his act is not altruistic. We should not label reasonably avoidable ignorance as altruism. Space considerations prevent me from presenting a thorough discussion of reasonableness, but I use the term as it used in common language.

The answer to the question, "Can an act be altruistic even if it is not motivated by altruism?", is no. An altruistic act requires an altruistic intention and the agent must be informed to a reasonable degree. At a minimum, the psychological state of an agent must be right, we would not, for instance, generally want to call acts that accidentally have altruistic repercussions altruistic. The psychological approach outlined above is not without problems; a person may be completely committed to performing altruistic acts but incompetent in executing those commitments. However, that is not a problem with the definition of altruism, but with the competence of someone in carrying out their intention to perform an altruistic act. In short, I believe that altruistic acts are dependent on the psychology of the altruist rather than the outcome achieved. Altruism should be evaluated subjectively, by the motives of the agent, rather than by the objective results of his actions.

#### Pseudo-altruism

To remind the reader, my intention in this manuscript is to provide a better understanding of self-interest, not to pass judgment on whether self-interest is a good or bad motive, clearly it can have both good and bad economic implications. However, at a practical level, it is interesting to note how pseudo-altruism can lead to beneficial results, where by pseudo-altruism I mean self-interested acts that are understood to be altruistic acts. Although we would not want to call everyone or every firm that gives to a charity an altruist, perhaps the psychological state of their benefactor does not matter to the beneficiaries. It might be enough that actions motivated by self-interest have beneficial implications. Imagine the case of a self-

interested person who gives money to a charity for his own benefit, say in terms of the good publicity he expects to receive from the donation, yet claims to be an altruist. From the point of view of those who know or presume to know the real, self-interested motives behind the contributor's behavior, such acts will appear crass. Recall that a dedicated psychological egoist will always be able to find the "real" self-interested motive behind every act. From the point of view of those who benefit from the donator's gifts, if they properly perceive the donator's real motives then they may feel slighted and used by the pseudo-altruist. They may accept the gift, but the resentment caused by the feeling of being used may not soon dissipate. There may also be cases in which the beneficiary does not care about the motives of his benefactor, the gifts are needed and gladly accepted. Pseudo-altruism can clearly provide many benefits and although someone who gives to a charity for the benefits that he receives might be less worthy of approbation than someone who gives for the sake of the beneficiary, the net benefit from such acts can be positive. In addition, pseudo-altruism, may, as a byproduct, lead to more pseudo-altruism. Assuming that firms value their reputation and altruism is seen to enhance that reputation, making a public promise to perform altruistic acts may give them an added incentive to carry out those deeds or otherwise risk damage to their reputation if they fail to do so. Furthermore, if consumers reward such acts, then if one firm performs altruistic or pseudo-altruistic acts, it may force its competitors to do likewise. Finally, Robert Frank (1988) reports that repeated pseudo-altruistic acts may produce beneficial side-effects for some. Although he goes on to say that if it is possible to distinguish between genuine altruists and pseudo-altruists, the genuine altruist will have an advantage over the pseudo-altruists because true altruism is generally more highly valued than pseudo-altruism. Frank is not the first to observe this phenomenon; it has long been suspected that acting in the interests of others, for whatever reason, may eventually lead to real altruistic behavior (e.g. J.S. Mill and Thomas Hobbes).

#### The Duration of an Act

The period over which the effects of an action are measured can be important in determining its status as being motivated by self-interested or another motive. For instance, it is possible that acts that are motivated by altruism can lead to longer-term, self-interested, benefits for an individual performing such acts. This would appear to contradict the definition of altruism which I gave earlier, but I do not believe it does. Once again, the intentions of the agent are the important factors for determining the status of an act. For instance, a friend of mine decided to donate a



photo to a charity that was raising funds for people with heart problems. I believe that her intentions were purely altruistic, however, before she could auction her photo she needed to get permission from her gallery to do so. The gallery agreed on the condition that they receive a quarter rather than their normal half of the proceeds from the auction. An unanticipated result of the auction, at least to my friend, was that her work became better-known and her sales increased. Perhaps she should have been able to foresee that her act might raise sales, in which case the act would be classified as self-interested with beneficial implications, but she did not. Altruism, in her case, led to greater long-term benefits for both her and her gallery. While I would classify my friend's actions as altruistic because her intentions were altruistic, the intentions of the gallery are more difficult to assess. I presume that the gallery's overall policy is to promote its own interests, however, the specific decision to accept a reduced percentage might have been an altruistic act, but without knowing the true intentions or perhaps precedents of similar acts of the individual or individuals making the decision to accept a lower commission it is nearly impossible to determine their motives. Determining or estimating whether an act is motivated by altruism or another motive requires determining the intentions behind an individual act, a part of which is a determination of the time-frame over which the agent measured the expected net results of an action.

One final note of clarification before I end this section, it is important to keep in mind that even if an act motivated by pseudo-altruism results in an equal amount of net benefit, there is still an important difference between self-interested and non-self-interested actions. Self-interest and altruism are two different motives that are best kept separate. It does matter if we call an act altruistic, pseudo-altruistic or self-interested. This can be observed in the case of the pseudo-altruist. A pseudo-altruist will only act altruistically if doing so is in his self-interest. If the situation changes and acting altruistically is no longer in his interest, then he will simply revert to self-interested acts based on his self-interested motives. This shifting of actions may be of importance to policy makers. One final remark related to this discussion is that if everybody "knows" that people are motivated only by self-interest, then there would be no benefit in pretending to be altruistic--people would see right through the deception. A pseudo-altruist must believe that there are advantages to appearing to be altruistic, however, what is the source of those advantages? I believe it is simply that people respect and admire altruists--it is something many people admire and aspire to be, it is more than just shadow-play. On the other hand, there is little to recommend egoistic, selfish, self-serving or greedy behavior, as those terms are commonly understood. On occasion, there

have been popular and academic advocates celebrating something similar to what I call pure self-interest, but in general, it is an attribute with which people would rather not be labeled. Even the famous business catch-phrase of the 1980s, 'greed-is-good', achieves its limited appeal because it contradicts the common understanding that greed is not good. Again, self-interest can be good for an individual and society, but it is not necessarily good and may be unethical or non-ethical and damage both the agent and society in general.

### Discussion and Conclusions

The argument is sometimes made that self-interest must be the basic motive because altruism leads to self-destruction, but, as I have shown, this is a false dichotomy. An extreme form of altruism would have an agent sacrificing his own interests for the benefit of others at all times and in all circumstances. This would clearly be a self-destructive and an ultimately self-defeating sort of behavior because an agent who is constantly sacrificing his own interests for the interests of others, without attending to at least a minimum level of his own needs, would quickly exhaust his ability to help anyone including himself. This position is a straw man; no philosopher that I know of defends this view. A more defensible form of altruism would allow an agent to make small self-sacrifices that yield large benefits to others. For instance, most people would find throwing a life-preserver to a stranger who has fallen into a cold stream a small sacrifice for a potentially large benefit to the stranger. To make a more convincing argument that the example is one of altruism, we would also have to assume that the rescuer has no expectations of receiving a reward (psychological or otherwise) greater than the costs of throwing in a life-preserver. In other words, the potential rescuer cannot expect to receive a net gain in well-being from his act; he would have to expect a net decrease in his well-being for an expected net increase in the well-being of the victim. Even if the victim experiences a net loss in happiness, say the life-preserver unintentionally lands on the poor man's head and causes him to drown, I believe that we should call this an altruistic act because the would-be altruist's well-informed intentions were altruistic. We would expect that the same potential rescuer might be less willing to jump in the same stream to rescue the stranger if it meant a major self-sacrifice, for instance, maybe the rescuer is an only marginally good swimmer or the water is dangerously cold. In other words, as the net costs of an altruistic act increase we would expect to see less altruism. In contrast, a purely self-interested person would not be able to make even a small sacrifice for a very large benefit to a recipient if it meant reducing his own estimated net pleasure by

even a trivial amount. The narrowly self-interested person believes that he is under no additional obligation to help friends or relatives, let alone strangers. Everyone is an instrument of equal standing to be used as a vehicle to benefit his chosen projects. In fact, all egoists, not just purely self-interested persons, may believe that they have no obligations to others. Egoists other than those that are purely self-interested can consider the interests of others for other than instrumental purposes, but they do not feel that they are under any obligation to do so.

In the preceding sections I have tried to raise theoretical and practical objections to an exclusive reliance on the motive of self-interest. Many of these objections can be overcome by simply expanding the scope of self-interest to include an ever greater set of reasons to act. I will emphasize in later chapters that such a strategy obfuscates an important difference that should be maintained between the motives of self-interest and altruism. I have shown that self-interest is a very general motive, it can incorporate many behaviors including selfishness and benevolence. However, that generality has the drawback that self-interest can become an empty concept; it needs to be properly specified if it is to be of practical use. In addition, I have argued that altruism is a rational concept and there is no good reason why contributors to open source projects couldn't be motivated by that motive. Other motives that may play a contributing role are special relationships and duties, both of which have been suggested by researchers. The problem of identifying the motives of open source contributors is always going to be an issue as it is whenever the aim is to identify intentions. Fortunately, open source licenses provide a good indication of intentions.



## CHAPTER THREE: SELF-INTEREST IN ECONOMIC ANALYSIS

---

### Introduction

The previous chapter was concerned with identifying the theoretical boundaries of self-interest and altruism and some of the problems associated with identifying those boundaries in practice. In a similar vein, this chapter examines two extreme theoretical views of self-interest found in the economic literature and then discusses two practical examples of how that general motive is used in important business-economic models. The purpose of this chapter is to identify significant inconsistencies and ambiguities in the use of the assumption of self-interest in both economic theory and practice.

The chapter is divided into five sections. In the first section I examine the general role that self-interest is assumed to have in economic analysis. I then attempt to locate the economic definition of self-interest within the more general understanding of that term as discussed in the previous chapter; it turns out that the economic self-interest as typically defined by economists, is more accurately labeled as selfishness. Surprisingly, self-interest, despite its popular and sometimes professional popularity, is not the basic behavioral assumption of every important economic model. In the second section I turn to the related issue of the purview of economics. Self-interest and rationality are often considered to be the distinguishing characteristics of “the” economic approach. The range of topics that are assumed to be the subject of economic analysis is ever-expanding and includes sociology, law and political science. It is therefore of increasing importance to get the assumption of self-interest right, and ask whether economic self-interest has any particular characteristics that distinguish it from non-economic motives. In the third section I briefly examine two views of the purpose of behavioral assumptions in economic models, namely, are they supposed to reflect the real motives and behaviors of agents or do they serve a merely instrumental purpose? This is clearly an important issue, if the assumption of self-interest is supposed to represent real behavior, as I believe it does, then it is possible to ask under which conditions it is an accurate assumption and when other assumptions might be more appropriate. If, however, self-interest is meant to serve an instrumental purpose, then the issue of accuracy is less important and under some interpretations of instrumentalism the degree to which an assumption reflects reality is largely irrelevant. Having discussed these important theoretical concerns, in the fourth section I address the issue of how the assumption of self-interest is defined and used in transaction cost

economics and agency theory. The advantages of using these models is twofold; first, they are two of the more influential models in current business-economic theory and practice; and, second, the principal developers of these models have taken explicit positions concerning their understanding of self-interest and its part in the decision making process of the agents whom they wish to model. In the fifth and last chapter I draw some general conclusions.

#### Section One: The First Principle of Economics?

The purpose of this section is to examine the role of self-interest in economic analysis. I briefly examine some historical and recent understandings of that motive before comparing the economic definitions to the more general understandings of self-interest. The main question that I address in this section is the degree to which self-interest is an integral part of economic theory. Perhaps surprisingly, preference theory, the core theory of economics, makes no claims about the motives of individuals.

#### Self-interest in Economics

A reader of economic history can be forgiven for concluding that economics and self-interest are essentially interlinked. No less than F.Y. Edgeworth claimed that the first principle of economics is that every agent is actuated only by self-interest (Edgeworth, 1881). Milton L. Myers (1983: 26), in his book *The Soul of Modern Economic Man*, nicely summarizes the place that self-interest is often awarded in the understanding of economic man.<sup>5</sup>

"What is the universal or fundamental element in the personality of economic man? What is it in his makeup that cannot be removed if he is to remain what he

---

<sup>5</sup> Mr. Myers and I agree on the importance of the subject of self-interest and even the approach; he also uses some of the same philosophers to ground his position. Our approaches differ in terms of their emphasis; his is to show how individual self-interest leads to the common good while I emphasize the importance of deriving a clear understanding of what self-interest is. Albert O. Hirschman's (1997) book, "The Passions and the Interests", is probably one of the more famous books on the subject of economic self-interest. He develops the idea how economic self-interest came to still the more violent passions. Both books offer interesting perspectives on how the perception of self-interest has changed through time.

is? The answer is self-interest. This is the one drive that is common to all our writer's depictions of economic man. Whether economic man is the introspective Puritan, the extroverted man of the world, the cold calculator of gain, the conspirator, the manipulator, or even the primitive, his underlying motive is self-interest. Our writers have either placed this characteristic of economic man in the foreground of their comments or they have assumed it to be so obvious that they have handled it in an implicit manner. In any event it is that feature of economic man which cannot be removed if he is to remain economic man."

I believe that Meyer is correct in his description of the historical role that self-interest plays in economic theory and popular interpretations of economics; however, the very generality with which the attributes of self-interest can be applied should give pause. Myers identifies many types of economic man including Puritans, extroverts, cold calculators of gain, conspirators, manipulators and primitives, all of whom, according to Meyers, share the attribute of self-interest. More generally, self-interest is an important characteristic of human behavior of which economic behavior is, presumably, a subset. The importance of self-interest as a reason to act has long been recognized; Derek Parfit claims that: "The Self-interest Theory has been believed by most people for more than two millennia" (Parfit, 1984: 194). Given its prevalence, it would be odd then if self-interest was not taken to be an important motive of economic activity.

The interesting issue is not whether self-interest plays an important role in economic motivation, but rather how interests in economic contexts differ from interests in non-economic contexts. This is not directly Myers' concern, his thesis concerns the interesting balance that many early economists tried to strike between the interests of individuals and the interests of the group as a whole. Yet that balance, I believe, brings to the fore one of the dimensions of self-interest that I am addressing, namely, the various and sometimes incompatible interests that a person can have, only some of which are economic. Utilitarians, for instance, despite the various forms that that theory takes, all agree that the interests of the individual must sometimes be sacrificed for the greater good. In contrast, the modern economic interpretation of self-interest is now clearly set on the sovereign interests of the individual and the implicit acceptance of the assumption that what is good for the individual is good for the group. Modern economic man is therefore never asked to sacrifice his interests for the interests of others.

What I take to be the modern, popular, economic understanding of the nature of economic agents is offered by Geoffrey Brennan and Philip Pettit (2004: 18). They claim that economics uses: "...an austere picture of human beings as centers of self-interest that operate out of society, without normative expectations or evaluations of themselves or one another." The point that I would like to make about Brennan and Pettit's description is that although they don't define self-interest, it appears from the above sentence, and the general theme of their book from which it was taken, to be used in a negative sense--self-interest is clearly not meant to be a flattering description of agent motives. The agents they describe are isolated from the rest of society and operate without reference to ethical guidelines suggesting how they *should* get along with others. There is no place for objective evaluations of the actions of others because agents are driven exclusively by their subjective preferences. In such a model, there is no reason to consider the interests of others except perhaps as instruments to be used to achieve one's own aims. Brennan and Pettit's definition fits well with the common understanding of self-interest; when we call someone self-interested in everyday language we are saying something to the effect that the person only considers his own desires or preferences--the person is purely self-regarding. If the interests of others are considered, it may only be to the limited extent that their actions affect one's own interests. The interests of other people, in other words, are not ends in themselves, but mere instruments to be used to achieve one's own interests. In short, Brennan and Pettit's description, I believe, represents the standard understanding of self-interest in both economics and common speech. However, self-interest has not always been given such a limited scope.

In sharp contrast to many modern understandings of self-interest as a part of an immoral or non-moral character, ethics and the early history of economics assigns self-interest a much broader meaning. Adam Smith, for instance, uses the term selfishness in a pejorative sense to mean harm or neglect of other people, and self-interest to mean that which is a proper 'regard to our own private happiness and interest' (Smith, VII.ii.3.16, pp. 20-25). Using Smith's definition, the sort of behavior that Brennan and Pettit describe as typical economic motivation is better thought of as selfish, rather than self-interested. Describing someone as selfish (as opposed to self-interested) implies more than just an unflattering attribute, it can also be interpreted as indicating unethical behavior (Schlick 1939: 85). Casson and Giusta (2006, p. 339), propose the following description of the unethical nature of neoclassical economic actors: "In standard neoclassical theory people are assumed to be selfish and to maximize their material gains. They will therefore have an



incentive to cheat, so sanctions are incorporated into material gains." The characterization of economic man provided by Casson and Giusta is extreme, and one that reflects the motive of selfishness rather than self-interest. They claim that agents in neoclassical models are assumed to maximize material gains, however, this strikes me as an overly restrictive view of economic analysis which ignores much recent work. At a more basic level, I believe their characterization of neoclassical economics is false. It is far more common to assume that agents maximizing their utility, where utility can include non-material gains. In any case, Casson and Giusta do not substantiate their claim.

This entire discussion may appear trivial, after all, I am writing about the meaning of just a couple of words, but the attention is justified given the implicit and sometimes explicit importance that the assumption of self-interest plays in economic reasoning. The point is that self-interest and selfishness are two very different descriptions of motivation with important implications for models of behavior. In addition, Casson and Giusta's claim also reemphasizes the need to be clear about the objects at which self-interest aims.

Self-interest, as opposed to selfishness, is used in the ethical literature to describe a very general motive; interests can lead to good, bad or neutral behaviors; in any case, not necessarily socially destructive behaviors. As I pointed out in the first chapter, a reliance on self-interest as an exclusive description of behavior would exclude some observed behaviors, but as a starting point it would seem to be a good description of most behaviors in most circumstances. The point is that self-interest can be used to describe nearly any sort of behavior. It has traditionally been used to describe, literally, the interests of people, which can be directed at all sorts of aims. As shown in the first chapter, a key to understanding motivation is to identify the focus of an action; in other words, does an agent have his interests in mind or the interests of someone else when he acts? Under the general understanding of the definition of self-interest, a person can be self-interested while incorporating the interests of others into their own interests. For instance, it is expected that people have an interest in the well-being of loved ones and that they attempt to promote those interests. An important stipulation, one made to distinguish a real difference in motives, is altruism--altruistic acts should not be classified as self-interested. Altruistic acts are defined as those that are intentionally performed for the benefit of others while the agent may suffer an expected net reduction in their own interests. In the case of altruistic acts, the focus is on the interests of other agents perhaps, but not necessarily, to the detriment of the interests of the acting agent. For instance, it is probable in normal

circumstances that we want to see loved ones do well and thrive even in cases in which our own interests suffer as a result. Further still, people may promote the interests of others without reference to their own interests. To take a rather mundane example, when a couple is deciding on which restaurant to go to in the evening, it might not be a question of balancing interests at all; rather, the partner deciding will choose the restaurant that they think their partner will enjoy--without regard to their own interests. I have mentioned other categories of actions that work against interests in the previous chapter, including particular affections, duties and special relationships.

If this were the extent of the misunderstanding--that economists and business scholars incorrectly use the general term self-interest to describe selfishness motives and that there is confusion about the important real differences between benevolent, self-interested, acts and altruism--then the confusion could be easily rectified by simply replacing the assumption of self-interest with the term of selfishness and clearly demarcating self-interested from altruistic acts. However, that change would only serve to realign the economic usage with its traditional usage, the underlying problems would remain. In fact, the question that I am addressing would remain open and another would arise, namely, why do economists make an assumption about human behavior (selfishness) that conflicts with every other previous attempt to describe self-interest? Derek Parfit (1984: p. 5) claims that no theory of self-interest has ever excluded the concept that the interests of friends and families are also a part of normal interests. The degree to which the interests of others affect our own is a subject of intense debate and I will return to it in later chapters, but for current purposes it is important to remember that in the thousands of years in which self-interest has been studied, no major scholar has ever proposed that our interests excludes the interests of others (see chapter 5). Selfishness is therefore an extreme position in that it restricts our interests to a very narrow sub-set of interests and conflicts with common experience and years of scholarly investigation. The fact that selfishness is an extreme assumption, by itself, is not necessarily a strike against it, but such a position needs to be justified if it is to be used in economic models. The more general point I wish to make in discussing the differences between self-interest and selfishness is to begin to show that there are a wide range of interests that a person can have, all of which fall under the general heading of self-interest; therefore, even if we accept that self-interest is an accurate description of behavior, a comprehensive theory of self-interest will need to specify the boundaries of the interests that it hopes to explain. Selfishness may be an appropriate assumption in

some economic contexts (however that is defined), but it is an extreme assumption that needs to be justified by a modeler.

I need to emphasize that my argument is not that people are not self-interested. I believe, as do all the scholars that I examine in this manuscript, that people are, in large part, motivated by self-interest. People are, in general, naturally self-interested, to argue otherwise strikes me as naive and indefensible after even a cursory glance at human behavior and the presumed intentions that stand behind those behaviors. As a practical matter, the burden of proof lies with those who claim that at least a part of human behavior is motivated by something other than self-interest. My purpose, simply stated, is to suggest that self-interest admits a wide range of motives and aims that need to be specified in order for self-interest to be a meaningful and effective assumption. Which dimensions of self-interest are appropriate to explain economic behavior will depend on what one takes to be the purview of economics; in other words, what interests are economic interests?

#### Preference Theory

I turn my attention from the very general theory of self-interest to another very general theory, namely, preference theory. Self-interest has been called the first principle of economics, however, it is not obvious how to reconcile it with preference theory, which, in turn, has been called the core of economic theory (Broome, 1999:8). In particular, revealed preference theory explicitly makes no assumptions about a person's motives. Consequently, the first principle of economics appears to be incompatible with a basic premise of revealed preference theory. In the next few paragraphs I will outline the basics of preference theory and offer a possibility to reconcile the assumption of self-interest and preference theory.

The developers of revealed preference, the core of empirically based consumption theory, intentionally made no assumptions about the motives of agents. Originally proposed by Paul Samuelson in 1938, the approach estimates consumer preferences by observing the consumption of goods, it thereby bypasses the issue of what economic agents believe and desire and purportedly avoids the problem of researcher subjectivity. The basic idea is that by carefully observing and drawing conclusions from the behavior of consumers, researchers can avoid their own biases about what consumers prefer yet still draw conclusions about consumer demand given certain minimum assumptions. The approach was met with

enthusiasm by many of the best economists of the time, for example, John R. Hicks was initially convinced of the advantages of studying economic agents: "only as entities having certain patterns of market behavior, it [revealed preference] makes no claim, no pretence, to be able to see inside their heads" (Hicks 1956). The approach is designed to circumvent the problems of linking motives with behavior and is valued precisely because it avoids the issue of what motivates economic agents to make the choices they do. An advocate of the method of revealed preference doesn't take a stance on whether an agent is an altruist or an egoist, either, for example, can prefer to give to charities for very different reasons, only behavior is observed. Lionel Robbins put it best when he stated: "So far as we are concerned our economic subjects can be pure egoists, pure altruists, pure ascetics, pure sensualists or - what is much more likely - bundles of all these impulses" (Robbins, 1935). Preference theory is a position firmly based on the sovereignty of the individual and the belief that an agent is (or should be) completely free to decide what he prefers and not required to explain to anyone why he prefers it.

The approach of revealed preference is valued precisely because it is a method of avoiding the issue of the motives and aims of economic agents. It is important to recognize that the assumption of self-interest is just that, an assumption added to the core of consumer theory, it is not a part of the preference framework. The term preference as used by economists simply indicates an ability to indicate whether one bundle of goods is preferred to another (e.g., Varian, 1999, 5ed.). Because the approach says nothing about the inner workings of economic agents, it is often assumed that people are self-interested must come from somewhere else. One possible explanation is that it creeps into the model of consumer preferences via the assumption of strong monotonicity which says that at least as much of every good, and strictly more of some good, is strictly better. However, strong monotonicity is not a requirement of economic rationality, which only requires preferences to be transitive and complete (Hausman and McPherson, 1996, 27). Preferences are transitive if when  $x$  is preferred to  $y$ , and  $y$  is preferred to  $z$ , then  $x$  is preferred to  $z$ . If someone prefers bananas to apples, and apples to oranges, then they must prefer bananas to oranges, according to the transitivity condition. Similarly, for the case of indifference, if  $x$  is indifferent to  $y$ , and  $y$  is indifferent to  $z$ , then  $x$  is indifferent to  $z$ . Preferences are complete if for all  $x$  and  $y$ , either  $x$  is preferred to  $y$ ,  $y$  is preferred to  $x$ , or the chooser is indifferent between the two options. For didactic reasons preferences are usually represented in two dimensions with, for example, the amount of one bundle of goods is represented along the  $x$ -axis and another along the  $y$ -axis. Different combinations of bundles  $x$

and  $y$  are represented by points in the positive quadrant of the  $x$ - $y$  graph. Different combinations and amounts of  $x$  and  $y$  allow indifference curves to be plotted and the typical well-behaved indifference curve is represented by a convex curve. However, in keeping with the generality of the approach, indifference curves can have many different shapes representing, for instance, perfect substitutes or perfect complements. 'Goods', as opposed to 'bads', are defined by those things which are consumed. Well-behaved indifference curves are also assumed to exhibit strong monotonicity which means that larger values of  $x$  (or  $y$ ), all else equal, and given that the point of saturation for either good hasn't been reached, are preferred to smaller values of  $x$  (or  $y$ ). That more is preferred to less may be interpreted as supporting the belief that people are self-interested, but that would be a restrictive and arbitrary interpretation of the general model and of the model of self-interest. In the previous chapter I mentioned several reasons that people may have to damage their own interests. For instance, I mentioned that they often prefer things that are not in their best interest. Drug addicts, for instance, prefer more drugs to fewer, but few would argue that by doing so they are acting in their best interest. Drugs are goods in that addicts like to consume them, but they serve their interests in only the most general of interpretations. The conclusion is that there are many things that people prefer that are not in their interest so that preference theory does not necessarily support the assumption of self-interest.

Economics generally deals with first order preferences, that is, those preferences that are revealed by agents when they buy goods and services (Hirschman 1985; Sen 1978). Higher order preferences, on the other hand, address the issue of why individuals prefer what they prefer and might include norms, values and beliefs. Higher order preferences are identified through reflection; they are recognized when agents ask themselves why they prefer A to B whereas first order preferences can arise without reflection. For instance, a person can prefer one thing to another because of instinct or habit; it is thereby possible to prefer things that are not in one's best interest. First order preferences are not necessarily immoral and selfish. Although Hausman and McPherson (1996, p. 52) are correct to conclude that: "Moral and altruistic preferences are ruled out of many economic models not by axioms of rationality, but by assumptions that the objects of preferences are bundles of commodities to be privately consumed and that there are no interdependencies among the preferences of different individuals." However, the assumptions that bundles are privately consumed and that preferences are independent are not assumptions of preference theory in its most general form. Again, modern axiomatic utility theory makes no assumption that people are self-

interested, all that is assumed is that agent preferences conform to a number of axioms, roughly, they simply need to be consistent. They can conform to the core axioms without being self-interested.

Clearly then, there is a potential contradiction between the narrow understanding of self-interest as it is commonly understood and represented in the Brennan and Pettit quote above, and the undefined and presumably irrelevant model of behavior found in preference theory. Confusion can arise when a narrow definition of self-interest is combined with a very broad understanding of preference theory. I contend that modelers of economic and business agents often implicitly assume both a narrow, selfish sort of self-interest, while simultaneously assuming that preference theory allows an agent to exhibit any sort of behavior--these two views may conflict. As a result of holding and confusing both of these positions at the same time, I believe that such modelers often conclude that all acts can be taken to be motivated by selfishness. The net of preference theory is cast very wide, while the term self-interest is used in a very narrow, opportunistic, manner. The result is that it is possible to conclude that the position of economics is that all motives are described by a narrow, aggressive form of self-interest and it is possible to conclude that nearly all behavior can be understood as economic.

I was careful to qualify many of the above remarks with words like 'potential' and 'may' because, perhaps surprisingly, there is not necessarily a contradiction between self-interest and preference theory. For example, psychological egoism, one of the more enduring theories of self-interest, claims that all motives are self-interested. It is a descriptive position that says that people are naturally self-interested. Accepting psychological egoism would mean accepting that everyone, whether acting economically or not, is always motivated by self-interest, thereby creating room for a slightly refined version of Robbins' understanding of economic theory as stated above, and the narrow, self-interest assumption to coexist. Robbins claimed that an economic agent can be motivated by anything including altruism, but a proper understanding of those terms means that altruism and self-interest are incompatible. That qualification aside, Robbins' position, along with psychological egoism, can be combined. To see this, remember that under psychological egoism all actions are presumed to be motivated by self-interest and, thereby, all of an agent's preferences would be motivated by self-interest. There could still be a difference between acting economically and acting in some other manner, but the difference would not detract from the fact that agents are always self-interested. In short, the position is that all acts are motivated by self-interest

and therefore all preferences are motivated by self-interest and so the theories of self-interest and revealed preference are potentially compatible. Similarly, models that assume that agents are rational only if they are self-interested can also be defended by advocates of a position known as rational egoism. I have already hinted at the potential problems of this compatibility thesis, the first is that people often do not act from self-interest and, second, there are many different, sometimes incompatible aims associated with the general motive self-interest--I will explore these issues in greater detail in later chapters.

I draw three general conclusions from the discussion thus far, the first is that there is a need for clarification in the economic and business usage of the assumption of self-interest in terms of both its structural, psychological aspects, and the external aims at which self-interest is directed. Second, the boundary between economic and non-economic behavior is vague, but seems to be expanding to include many social relations beyond its traditional subjects of firms, consumers and markets. Self-interest is increasingly being taken as the default position for many disciplines and perhaps by society at large. Rather than an assumption made in order to make parsimonious models, self-interest is taken a description of real behavior in an ever increasing number of social contexts. This is not a surprising development given that both the theory of self-interest and preference theory can be very broadly interpreted. Third, there are many sorts of self-interest expressed in economic models as represented by the contradiction between a narrow interpretation of self-interest and revealed preference theory. The default position in many economic and business models and popular opinion is that economic agents are narrowly self-interested, while the core of economic theory upon which many of those models rest, intentionally leaves the motives of agents undefined. Self-interest is generally not a formal assumption, it has been largely implicitly accepted and where it has been explicitly defended, as I will show below, it has often been mis-defined or reduced to an unhelpful tautology. Such a fundamental assumption needs to be clearly defined in order to properly understand its implications and limits.

#### Instrumental and Scientific Realism

Although the economic definition of self-interest or selfishness is seldom directly defended, another line of defense could be offered that would negate the importance of much of the discussion thus far. It could be argued that this entire issue of what sort of self-interest should be used to model the behavior of economic agents is irrelevant because self-interest, or any other assumption concerning

motives for that matter, is just a simplification made in order to build more tractable models; under this reasoning whether assumptions reflect reality is irrelevant, the only measure of importance is whether the resulting model allows good predictions. There are many viewpoints about the role of assumptions; I will limit my discussion to two extreme views that discuss the relevance of making realistic model assumptions, namely, instrumentalism and realism.

### Instrumentalism

At one extreme is the position that model assumptions are instrumental and do not need to reflect reality at all, the only measure of importance for models is the accuracy of their predictions.<sup>6</sup> Most economic and business students at some point in their training have heard the following sort of reasoning: "Assumptions are simplifications, they are not meant to reflect reality and should not reflect reality; they are simplifications made to build models that are more tractable." One of my instructors used the analogy of a road map to emphasize the point, he asked, "What would be the use of a life-sized road map which perfectly represents reality?". No one argues that assumptions should mimic reality; the issue is finding a good balance between parsimony and realistic descriptions of behavior. An instrumentalist, such as Bishop Berkeley, argues that the function of scientific theorizing is not to explain but to organize our experiences into packages. "On this view, theoretical terms are not abbreviations for observed ones, they are more like mnemonic devices, acronyms, uninterpreted symbols without empirical or literal meaning. And the aim of science is constantly to improve the reliability of its instruments, without worrying about whether reality corresponds to these instruments when interpreted literally" (Rosenberg, 2000: p. 94). Accepting this view would mean that the economic assumption of self-interest would not have to

---

<sup>6</sup> Milton Friedman's 1953 article, "The Methodology of Positive Economics", contains the controversial claim that the validity of positive (as opposed to normative economics) economic theory should be based solely on the accuracy of its predictions rather than the realism of its assumptions. However, even if we accept Friedman's advice that the test of whether one set of assumptions is better than another is in its ability to improve the power of a model, then, at a minimum, the assumptions have to be well-defined. See the last chapter of Mark Blaug's (1996) book and Daniel M. Hausman (1992), for just two of the many discussions of Friedman's claim. If Friedman is correct, then there is no need to discuss assumptions at all. However if assumptions are going to be made then, at a minimum, they need to be well defined. Saying that someone is self-interested is too broad and ambiguous a claim.



reflect “real” motives, assuming one could determine what the real motive is, rather, self-interest would only serve to symbolize the unobservable behaviors of economic agents. However, the overall aim, according to the instrumentalist position just stated, is to improve the reliability of instruments over time. Under this reasoning, it is still permissible to criticize the definition of economic self-interest, or what self-interest is taken to mean, on the simple grounds that it is too vague to be of much use; it allows too broad a range of behaviors. I have already given the example of the confusion that exists in economics between self-interest and selfishness, many more examples are to come. In short, I do not believe that even a die-hard instrumentalist would condone using ambiguous definitions when better ones are available. In the case of self-interest, better assumptions that can be used to develop more reliable instruments are available from ethics and should therefore be used.

#### Scientific Realism

The other extreme position is known as scientific realism, it argues that model assumptions should be based on empirical observations that describe reality. By building models based on realistic assumptions modelers hope to gain new insights into truths about the world in addition to making accurate predictions.<sup>7</sup> Economists and business scholars following this methodology are required to accurately define what it means to be self-interested and rational. Again, I will argue that the economic assumption of self-interest is ambiguous and may not reflect many important motives driving economic behavior. Therefore, I believe that both instrumentalist and scientific realist perspectives justify the search for better (predictive or realistic) assumptions. For the purposes of this manuscript, I am content to conclude that, for all but the most extreme views of scientific methodology, unambiguous definitions are required and the no view of science would condone making vague assumptions when better definitions are available. In addition, I will show below that many economists and business scholars appear

---

<sup>7</sup> Adam Smith's view of science should be placed in the scientific realist camp. Smith minimized the role of prediction in science, his view was that the main job of scientists was to resolve contradictions and tensions in past explanatory systems, rather than to learn new facts (Fleischacker, 2004, 34). This view seems overly conservative, why should, for instance, new explanatory models be excluded simply because they don't address contradictions in existing models? With that qualification in mind, I believe Adam Smith would argue that assumptions need to reflect reality.

to believe that the assumption of self-interest is meant, at some level, to describe real economic behavior and behavior in general. In other words, the assumption of self-interest is meant to be a description of reality, placing those scholars squarely in the realist camp.

The desire to build parsimonious models is a good reason to make simplifying assumptions. However, there are some obvious trade-offs between the simplicity of a model's assumptions and the value of its resulting descriptions and predictions. I wish first to emphasize that the assumption of self-interest, if properly understood, does not go far in reducing the sorts of actions that a modeler is allowed an agent to perform. Again, self-interest is a very general motive that is often taken as the default explanation of human motivation across many activities. Selfishness, on the other hand, if properly understood, is a very parsimonious description of motivation. Whether it is a good description, in terms of the purposes that a model is meant to serve, would need to be argued for by the modeler. In general, I find the modeling of economic agents as isolated, non-interacting units in all but the most extreme contexts to be of dubious value, but only if there is common understanding of the terms used, then the advantages and disadvantages of the assumed motives can be discussed.

An analogy might help to clarify the point I wish to make. The assumption that planets move in perfect circles might in many cases lead to reasonably good astronomical models, e.g., models that lead to good predictions. In fact, in some circumstances models based on such ideal assumptions might initially yield better predictions than models with more accurate, but more complex assumptions. It might be argued that self-interest is not the perfect assumption, but it allows modelers to build parsimonious models that yield reasonably good predictions. The issue that I am addressing is that self-interest, as it is used in economics, is not an ideal assumption because, in part, it is not accurately defined. In other words, we know the properties of an ideal circle, but economics and business models do not specify the properties of an ideal, self-interested, person. The current situation in economics would be equivalent to the case in which astronomers used many, and in some cases conflicting, definitions of a circle. A loose definition of a circle would mean that different astronomical models could not be compared and many shapes that are not circles would be called circles. Finally, even if we could arrive at an ideal definition of self-interest, there still should be room for building models based on less ideal and more realistic assumptions. At a minimum, we need to be

able to occasionally reassess the degree to which our assumptions deviate from reality and examine how people actually behave.

#### The Purview of Economic Analysis

The range over which economic analysis is taken to apply is ever increasing. This is not a surprising development given the very general assumptions of preference theory; again, an agent only has to have the ability to compare different bundles of goods where goods are widely defined, and to be considered rational, he essentially only has to be consistent in his choices. It follows that many economists take a very broad view of the scope of economic analysis. For instance, in a typical textbook, Douma and Schreuder (2002: 1) claim the following: "The point is that an economist would identify an economic problem in any situation where needs would not be met as a result of scarcity of resources: and resources are quite broadly conceived as all means that may contribute to the satisfaction of human needs." Given such a broad definition, it is indeed difficult to imagine situations in which economic analysis would not apply. The authors unnecessarily constrain economics to cases in which *needs* are in question. Economics concerns the aim of fulfilling preferences, not all of which are necessarily needs and it is probably better to think of them as desires. Defining resources as all means that contribute to the satisfaction of human desires is a particularly ambitious goal for any social science to try to address. There are significant differences among the many human needs and it seems unlikely that any one set of tools, particularly the standard economic tools of rationality and selfishness, would be equally applicable across the diversity of human desires. If that diversity is granted, then such a set of tools would at best be necessary rather than sufficient to understand human behavior.

Douma and Schreuder's definition of the scope of economics echoes that of Gary S. Becker, who is perhaps the best known advocate of an extensive definition of the scope of economic analysis. Becker (1976) writes: "Indeed I have come to the position that the economic approach is a comprehensive one that is applicable to all human behavior; be it behavior involving money prices or imputed shadow prices, repeated or infrequent decisions, emotional or mechanical ends, rich or poor persons, men or women, businessmen or politicians, teachers or students." Accordingly, Becker has applied economic analysis to discrimination, crime, marriage and the family (see, respectively, G. Becker, 1957, 1968, 1974 and 1981; see Becker, 1996 for a summary of some of these findings). In a discussion of why people give to beggars, Becker (1996), presents a good indication of his general

approach. He writes: "A person makes contributions [to beggars] to add to his own utility--otherwise he would not part with any wealth" (p. 232). The interesting question for the purposes of this manuscript, as always, is why does giving to a beggar increase the contributor's utility? In other words, what is the aim of the contributor? Two other notes in regards to Becker's approach already mentioned in chapter 1 are that not all reasons to act are based on "feelings" and people often prefer things that are not in their interest--both of which are significant problems for preference theory and therefore utility theory. More generally, Becker may be right, to the extent that it is possible to rank objects within a preference function, it may be possible to use utility theory to address many important questions, but that theory does not help us to understand what motivates people to make the choices they do. Incidentally, issues of comparability and commensurability call into question the ability to meaningfully compare relevant alternatives in many contexts thereby limiting the applicability of preference theory, but that issue is beyond the scope of the present discussion. People who give to beggars have many reasons for doing so, some of which may be to increase their own utility, but many of which are certainly not, and equally certainly not all reasons to give are readily identifiable as economic.

The critical point which I wish to emphasize is that both the theory of self-interest and preference theory are very general descriptions of motives and behaviors. Many, but by no means all, economic activities can be understood using these two theories, but so then can most human activities. In most contexts people are self-interested and they can prefer one object to another. However, interests vary widely and I would contend that economic interests, properly understood, are but a small part of the many interests a person can have. For instance, maximizing profits, a commonly accepted economic aim, is a very different aim from that of living a good life as the term good has been commonly understood. I argue, therefore, that the aims of an agent need to be specified, once those aims are specified the argument over whether the analysis should be labeled as economic or something else can follow.

#### Two Models That Assume Self-interest.

In order to ground the previous discussion of economic self-interest and avoid the charge that I have created a straw-man, I will examine how self-interest is actually defined and used in business economic models. Fortunately, there are two very influential, modern, models that I can use, namely, Oliver Williamson's (1975,

1985) version of transaction cost economics and Jensen and Meckling's (1976) agency theory. These authors are some of the few in the business and economic literature who explicitly address and defend their behavioral assumptions.

The ambiguity with which the assumption of self-interest is used in economic models means that almost any critique or defense of what self-interest is supposed to mean is supportable. Motives in economic and business models are a moving target. Revealed preference theory, as I have pointed out, intentionally says nothing about the motives of agents, while popular and scholarly interpretations often take a very narrow, almost mean definition of self-interest. As a consequence, economic and business models can be criticized for having both a very narrow, unrealistic, description of motives, and for having no stance at all on motives. Likewise, models can be defended from the charge that their assumptions are too narrow by appeals to the fact that (some) economic models say nothing about motivation, while any support provided by models for the assumption of self-interest can then be used as evidence that a narrow interpretation of self-interest is the proper interpretation for all economic models. It does neither defenders or supporters of the economic approach any good to use ambiguous, undefended, assumptions. To get around the problem of ambiguity, I will provide a detailed example of how the assumption of self-interest is used in transaction cost economics and agency theory, two of the more influential business economic models of the last decades.

#### Transaction Cost Economics

The behavioral assumptions of business economic models vary considerably from model to model so it is important to specify which model one is criticizing. In fact, many economic and business models make no explicit behavioral assumptions beyond vague references to self-interest or selfishness. One of the few modern business economists who explicitly states his assumptions is Oliver Williamson. Williamson (1975, 1985) is widely accepted as a leading scholar in the theory of transaction cost economics and is often credited with single-handedly reviving that theory which was originally developed by Ronald H. Coase (1937). Williamson makes two behavioral assumptions that he claims are the most critical elements of his model. Quoting Herbert Simon, he writes: "Nothing is more fundamental in setting our research agenda and informing our research methods than our view of the nature of human beings whose behaviour we are studying" (Simon 1985: 303). I take from this quote that Williamson believes that a study of economic

organization must begin with realistic (not purely instrumental) assumptions about the nature of human behavior. Williamson (1993) makes two such assumptions, the first follows from Simon's (1961, xxiv) claim that human agents are assumed to be "intendedly rational, but only limitedly so" and, second, that human agents are given to opportunism. According to Williamson, "But for these two behavioral assumptions - both individually but, especially, in combination - the study of economic organization would be vastly simplified" (Williamson, 1993, Chapter 7 p. 93). There is little disagreement about the first assumption, empirical studies by, among others, Kahneman, Slovic and Tversky (1982) appear to confirm inconsistencies in human behavior under conditions of uncertainty and, generally, in the comparative assessment of alternative decisions. I will therefore concentrate my efforts on explicating the second assumption.

It is along with Williamson's second assumption, opportunism, that bounded rationality contributes to his explanation of the existence of firms. Williamson defines opportunistic behavior as a: "deep condition of self-interest seeking that contemplates guile" (Williamson, 1993, p. 458). He identifies opportunism as an unflattering attribute but contends that it is: "basic to the logic of organization - in that, absent opportunism, there is no contractual reason to supplant market by hierarchy" (Williamson, 1985, pp. 30-32, 64-67). Williamson's insight was to recognize that bounded rationality provides 'space' for opportunistic behavior. His reasoning, if I have properly understood his argument, is that bounded rationality for economic organizations implies that even complex contracts are unavoidably incomplete (Williamson, 1998 p: 31). Therefore, according to Williamson, contracts will contain gaps, errors, and omissions that are difficult to enforce in a court of law. Mere promises, unsupported by credible commitments, will not be enough to ensure that parties to a contract behave nobly given Williamson's assumption that people may behave opportunistically. In other words, there is a non-zero probability that one or both of the contracting parties will behave in an opportunistic manner, i.e., they will in some cases stick to the letter of the contract rather than the spirit of the contract when it is in their interest to do so.<sup>8</sup> When it is difficult to write contracts that closely approximate the spirit of the contract, parties will expect their opposites to act opportunistically and will opt for other ways to enforce agreements. In short, bounded rationality ensures that contracts can never

---

<sup>8</sup> Thanks to Professor, Dr., Hennart for clarifying Williamson's assumptions and model.

be complete, in addition, behavior is assumed to be opportunistic. The result is that non-market hierarchies (e.g., firms) are created in order to forestall such behavior

I believe that there are four important concepts missing from Williamson's description of self-interest, all of which correspond with the ideas that I am addressing in this manuscript. First, as I will show, Williamson's definition of what he calls 'simple self-interest' is weak, arbitrary and appears to be a straw-man, furthermore, the definition he uses clearly ignores the literature on self-interest. Second, Williamson never adequately addresses the question of which objects are included in the interests of an agent (what I am calling the external component of interest). Third, Williamson doesn't clearly specify his assumption about the role of self-interest in the human psyche (the structural component of self-interest). Finally, Williamson is ambiguous about the scope of economic analysis. While he writes at the beginning of his most complete description of agent behavior, "Calculativeness, Trust, and Economic Organization" (1993), that he is interested in economic calculativeness, he subsequently extends the reach of calculativeness to include situations far removed from what I take to be economic, going so far as to include in his analysis on calculativeness the topic of love.

A particularly unsatisfying feature of Williamson's approach, given his emphasis on the importance of good behavioral assumptions, is the lack of evidence that he provides to support his assumption that agents are self-interested. In fact, his evidence amounts to too little more than references to other authors who share his assumptions. For instance, Williamson quotes H.L.A. Hart (1961, p. 193) who asserts that, "Neither understanding of long-term interest, nor strength of goodness of will ... are shared by all men alike. All are tempted at times to prefer their own immediate interests... ." Williamson offers another quote from Robert Michels (1966, p. 370): "[only] a serene and frank examination of the hazards of opportunism will enable us to mitigate these hazards." Such quotes do not amount to a coherent argument in favor of self-interest. However, there is little argument that decision makers usually act in their own self-interest, no one is questioning that, but without knowing what the intentions of the agents are it is not possible to claim that agents, even in economic contexts, are motivated exclusively by economic self-interest. Simply put, agents, even in economic contexts, as I will show in the final chapter, can have non-economic intentions. However, if a definition of economic self-interest is not offered it is impossible to provide a counter argument showing that people are not always economically self-interested. I believe that the mistake that Williamson and others make is to assume that self-

interest is a specifically economic motive. By collapsing complex human motives into the one unfaltering dimension, opportunism, which he calls self-interest with guile, Williamson has oversimplified human nature.

Simple "Self-interest seeking", for Williamson, means economic agents will, "continuously consult their own preferences but will candidly disclose all pertinent information on inquiry and will reliably discharge all covenants..." (1993, 158). This is a rather unusual interpretation of self-interest because while there are situations in which candidly disclosing pertinent information to others will damage an agent's ability to achieve his preferences, there are also situations in which revealing one's preferences will help an agent to achieve those aims. Therefore, under the definition of simple self-interest offered by Williamson, self-interest can either help or hinder an agent to achieve his aims and, in cases in which one's self-interests are damaged by candid disclosure, Williamson's definition would clearly lead to self-defeating behavior. I believe that Williamson's point is that self-interest seeking, without guile, is a naive approach. Guile allows agents to use deceit and cunning to achieve their interests at the expense of others. Guile is an admittedly unflattering qualification that Williamson makes to the general definition of self-interest. For Williamson, self-interest along with guile, means that agents may realize their aims at the expense of other agents. But again, it will often be in the interest of agents to reveal their preferences to one another in order to achieve their common goals, so a more general definition would be to leave the question of disclosure open rather than include it in a general definition.

I presume that what Williamson means when he says the agents will "continuously consult their own preferences" is that agents will consult their preferences with the aim of realizing their most preferred choice. Williamson may be implying that agents prefer those things that are in their self-interest, which, as I have shown, is not necessarily true; people often prefer those things that are not in their interest. I have also shown that it is difficult to identify motives by observing behavior. The task of deciding what one prefers is often influenced by one's higher order preferences including norms and values. Observed behavior can therefore be motivated by quite different motives, only some of which may be reduced to self-interest. There is a potentially more challenging problem for Williamson and others using the definition of preference. It is not unusual to hear the claim that because people choose what they prefer they are, accordingly, acting in their own self-interest. The idea is that because a person has chosen an object, it must, on the whole, improve their well-being and, therefore, must be in their self-interest. The



claim is partially, trivially, true, people choose what they prefer, but this does not tell us why they prefer what they prefer, we have to dig deeper into the motives that define preferences. For instance, people may prefer to do their duty even if it works against their interests, e.g., telling the truth when a lie would be more advantageous or helping someone in need if the costs to ourselves are not too high. An outsider viewing such benevolent behavior might reasonably conclude that the agent prefers to tell the truth and help others because it improves the recipient's well-being and not because it is in the agent's self-interest, it may in fact detract from the agent's interests. Agents do regularly consult their preferences, but they do not always choose what is in their self-interest. As a final minor point, I use "regularly" where Williamson uses continuously because in many situations people make their choices without actively appraising their preferences. I suggest that the act of consciously ranking preferences only takes place in extreme situations and that many preferences are the result of habits and the like.

Although Williamson's failure to provide an adequate definition of what he means when he claims agents are self-interested is odd given his strong emphasis on the importance of making accurate behavioral assumptions, in his favor he does give some hint of a difference between economic and non-economic motivation. Williamson distinguishes between benign "frailty of motive" which describes day-to-day activities, and opportunism (1998, p. 31). As I understand it, frailty of motive for Williamson, describes what people do most of the time, namely, they usually do what they say they are going to do without self-consciously asking whether the effort is justified by expected discounted gains. However, if the relationship involves not day-to-day affairs, but long-term contractual relationships, then, according to Williamson, people switch to making opportunistic calculations. It is clear that Williamson is right that not all activities involve calculation of expected discounted gains, but, importantly, Williamson doesn't specify what role self-interest plays in day-to-day, self-conscious, decisions. He seems to imply self-interest is present only during important economic calculations, while its role in day-to-day decisions is left undefined. The distinction between day-to-day and important economic activities strikes me as arbitrary, there are undoubtedly many non-economic decisions in which self-interest plays an important role; an advocate of psychological egoism would claim it plays a leading role in all decisions. More importantly, by limiting self-interest to important economic calculations, Williamson's view would contradict the common belief that self-interest is a very common motive in all activities. If that is Williamson's position, then he has greatly underestimated the role of self-interest in

daily life. In fairness, this is not Williamson's concern, perhaps he has something like selfishness in mind when discussing important economic events and self-interest in other situations, but his position has implications for the issues in this manuscript. It confirms the ambiguity with which self-interest is used and raises the corresponding issue of how economic behavior differs from non-economic behavior.

In contrast to Williamson, Amartya Sen contends that the economic conception of self-interest encompasses non-economic activity as well. Sen believes that a good deal of economic theory, to its detriment, assumes that agents identify only with themselves: "Indeed, this assumption has often been seen as adequate both in explaining human behaviour and in explaining the efficient operation of market-based economies" (Sen, 1999: p. 2). The importance of Sen's comment is that he believes that the assumption of self-interest is more than just an assumption made for the sake of building tractable economic models, rather, it is taken as a true representation of all human behavior. Again, the economic position of self-interest is nothing new, the assumption that people are primarily self-interested has been the default position since people have raised the issue of motivation. Sen's concern for the misuse of self-interest and rationality is exacerbated by the extent to which the "economic" assumptions of self-interest and rationality have been exported and adopted by a number of other social sciences including, law, political science and sociology, thereby, incidentally, compounding the problem of using an ambiguous definition for such a key assumption. The extension of these economic assumptions to other disciplines further blurs the understanding of what it means to act economically and further confuses the issue of the boundaries between economic and non-economic motives. Whether Sen is referring to self-interest or selfishness in the above quote is unclear. For instance, benevolent persons may identify only with themselves (their interests), but those interests include the well-being of others. He is much more explicit about his position in the following quote: "Universal selfishness as actuality may well be false, but universal selfishness as a requirement of rationality is patently absurd" (Sen, 1987, p. 16). I suppose that by "universal selfishness" Sen refers to the position of psychological egoism and thereafter he is clearly referring to rational egoism. However, the relationship between psychological egoism and rational egoism is probably more complex than Sen suggests. For example, if people are motivated as psychological egoists believe they are, then they cannot help but be self-interested, rationality and self-interest would therefore always go hand-in-hand. On this view, rationality might influence the choice of acts but not the motivation behind those acts which

would always be one of self-interest. If there is a choice about whether to act from self-interest or not, then psychological egoism is no longer an option as a description of all acts on all occasions. The important point is that Sen suggests, and I wholly agree, that self-interest for economists is more than just a simplifying assumption, it is taken to be a description of real behavior. Such a claim is, despite all of the qualifications I raised in the previous chapter, empirically testable.

### Agency Theory

In many respects Jensen and Meckling's agency theory provides the clearest statement in the field of business economics on the economic interpretation of self-interest. The authors explicitly address both the structural (they are rational egoists) and external ("almost everything" is an aim) dimensions of self-interest and they take a clear position on the scope to which their model is meant to apply (Jensen and Meckling 1994, Jensen, 1994). The purpose of agency theory is to describe how conflicts of interest can arise between the owners and managers of a firm and to suggest methods to control such conflicts (Jensen and Meckling, 1976). Although agency theory was originally directed at understanding and overcoming conflicts within corporations, the authors have since attempted to generalize their findings to all situations involving conflicts, including conflicts between members of families and other social organizations. In their general model, which they call the resourceful, evaluative, maximizing model of human behavior, they make a number of claims that are relevant to the issues at hand.

Their model consists of four propositions each of which I will briefly discuss (Jensen and Meckling, 1994). The first proposition consists of three sub-components: a. Agents care about "almost everything" including knowledge, independence, the plight of others, the environment, honor, interpersonal relationships, status, peer approval, group norms, culture, wealth, rules of conduct, the weather, music, art and so on; b. An agent is always willing to make trade-offs and substitutions among goods; c. Preferences are transitive. The second proposition describes the self-interest assumption of the model, and amounts to the assumption that the desires of agents are unlimited. Those things that agents want are defined as goods, so that agents always want more goods. The third postulate says that although agents want to enjoy the highest level of value possible, they are constrained by limits on such things as wealth, time and laws. The last postulate says that agents are resourceful and able to foresee changes to their environments and respond by creating new possibilities.

Each of the propositions raises a number of questions. For instance, the first proposition, sub-component “a”, states that agents care about almost everything, which, given the diverse list of goods that the authors mention, leaves one wondering what the "almost" refers to. The model appears to be a somewhat restricted version of preference theory. However, while preference theory makes no constraints on what is preferred, Jensen and Meckling specify that their model refers to “goods” which are defined as those things that people want. However, this is an odd restriction given that agents also care about avoiding “bads”. For instance, given two bad outcomes, an agent will presumably prefer the outcome that causes the least harm. A theory that only looks at what agents want will “miss” much of what drives human behavior.

A more substantial problem, especially given that almost everything is a good that is wanted and given the authors' transitivity assumption, is that there are bound to be comparability and commensurability problems. For instance, how much freedom is someone willing to give-up for what amount of wealth? What is the unit of measure of a norm or honor? For instance, duties are obligations that agents generally are not readily willing to negotiate away. This is not to say that a person who believes, for instance, that telling lies is wrong, will not lie under all conceivable conditions. At some point, for instance, on pain of death, people can be made to do almost anything. However, the authors specifically state that an agent is always willing, as opposed to forced, to trade goods. That assumption raises the issue of what the authors mean by willing, but even if we ignore that complication, we can certainly ask whether all goods should be tradable. For example, should we accept someone selling their freedom for a loaf of bread even if they willingly chose to do so? By increasing the scope of goods to include “almost everything” and by their claim that agents are always willing to trade and substitute all goods, the authors have raised fundamental normative issues. In that respect the model appears to be fundamentally flawed.

In terms of the dimensions that I defined at the beginning of this manuscript, the authors defend the structural assumption that agents are rational egoists while the aims of agents are "almost anything". The authors thereby reject what they call the economic notion that agents are only interested in money income or wealth maximization, while accepting what they claim is the economic assumption of self-interest maximization. I do not believe that most economists claim that people are only interested in money and wealth, certainly those who follow preference theory

do not make such a claim. Furthermore, preference theory says nothing about the motive of self-interest; however, I have already extensively discussed both of those objections and won't repeat them here. The claim that I would like to address is the one that people maximize their self-interest. It is not obvious to me that people maximize "anything" as the authors claim. It could (incorrectly) be claimed that no matter what people do, they are maximizing something. For instance, while deciding whether to choose cola A or cola B, at some point a person might decide that he is wasting his time and it really doesn't matter which cola he chooses given their similarity. So, the argument might go, the person has included the value of his time by deciding not to decide and simply taking the cola closest at hand. This would be an unusual definition of maximization. Maximization, on my understanding, involves a certain level of psychological commitment or involvement in arriving at a decision, while many daily decisions involve only a very low level of intention or are arrived at by habit. The point is that maximization, to be somewhat meaningful, requires a certain amount of vigor--vigor that is frequently lacking in daily decisions. I thereby reject as unrealistic the assumption that people are self-interested maximizers.

That agents in Jensen and Meckling's model are rational egoists is not surprising given the definitions they use. The authors claim that anything that a person wants is a good and that they attempt to maximize their good. Putting aside the problem of maximization, it would be difficult for anyone accepting the author's assumption to argue that it would be rational for agents to prefer those things that do not give them the maximum good. The main problem with such an argument is that by including almost everything as a part of self-interest, and by not specifying what "almost" entails, the terminology becomes confused. As I have stated previously, not everything that an individual prefers is preferred because it is in their interest, for instance, people may prefer things because it improves the well-being of others, despite the impact of that choice on their own interests. Take for instance the case of a preference for altruistic acts, a person cannot be an altruistic, self-interested, maximizer without hopelessly muddling the language. Altruism and self-interest are not the same things and their different meanings should be preserved. Furthermore, the authors do not offer any evidence that altruistic motives or actions motivated by duties or any other motive that works against one's interests are irrational--why should we assume they are?

In their model, Jensen and Meckling allow that people have the capacity for altruism which they confusingly define as: "They (people) care about others and

take their interests into account while maximizing their own welfare" (Jensen and Meckling, 1994, p. 18). This is a very loose definition that places the focus of an act clearing on the agent's own interests rather than the interests of those who are the beneficiaries of their altruistic actions. I suspect that their definition is too inclusive; for instance, even using a very generous interpretation of Jensen and Meckling's definition, instrumental self-interest is not ruled out. For instance, an agent might express an interest in the well-being of others only to the extent that it maximizes their own interests, allowing the possibility of actions that are closer to enlightened self-interest than altruism. Their definition, in short, allows many acts that are motivated by self-interest to be called altruistic, thereby, again, confusing two concepts that should be kept separate. Finally, the authors criticize an altruistic position that they call the perfect agent. Perfect agents are defined by the authors as people who maximize the public good as opposed to their own welfare. What appears to make agents perfect is their indifference to the type of altruistic acts they perform. According to the authors, since agents have their own preferences they cannot be perfect agents. "Perfect agents would be equally satisfied working to save the whales, feed the poor, make computers, or care for the musical interests of the rich through the local symphony orchestra at the bidding of their employees. Altruist that she is, Mother Teresa's devotion to caring for the poor of Calcutta does not make her a perfect agent." I find this a convincing argument. Although it is possible to imagine individuals who are motivated to act altruistically whatever the cause, it is more likely that people have some preference ranking among the sorts of projects that they support. In fact, a "perfect" altruist would, in my opinion, prefer to perform those acts that provide the greatest net benefit rather than haphazardly choosing altruistic projects. However, I do not know anyone who defends the perfect altruist position, being an altruist does not exclude the possibility of having preferences as the authors propose.

### Conclusions Chapter Three

For the most part, economists and those scholars using the "economic" assumptions of self-interest and rationality appear to take both of these assumptions for granted. While the understanding of what economists mean when they use the term rationality has been thoroughly scrutinized, few economic or business models that I know of offer more than superficial argumentation in support of the assumption of self-interest. The evidence that agents are self-interested often boils down to the claim that others have made the same assumption. However, the assumption of self-interest, at least in purely theoretical

contexts, is not a requirement of economic theory. That said, self-interest is generally accepted, and defended, by economists and business scholars. They are by no means alone in making either the assumption of self-interest or rationality, some form of these assumptions have been, more or less, the default positions since philosophers in Athens first began discussing how to live a good life. Although the dominance of self-interest as a description of motivation has occasionally been challenged, it has always been, and will always be, an accurate description of a large part of what motivates human action. The trend in economics is now towards a very narrow understanding of self-interest. This narrow definition is taken by modern scholars, and a large part of the general public for that matter, as such an obviously and naturally true description of motivation that it does not appear to even warrant comment among scholars. In fact, it so widely adopted that it appears to be accepted as a description of actual behavior rather than an assumption made in order to build more tractable models. The current situation makes the opening remarks of Joseph Butler's "Eleventh Sermon" of his book, *Sermons on Human Nature*, as applicable today as they were when they were written in the early eighteenth century: "... that vice and folly takes different turns, and some particular kinds of it are more open and avowed in some ages than in others; and I suppose it may be spoken of as very much the distinction of the present to profess a contracted spirit and greater regards to self-interest than appears to have been done formerly." (Butler, 1986, p 46). In short, self-interest as a default position is nothing new, forms of self-interest have been taken as prima facie explanations of human nature for thousands of years. What I believe is unique is the very narrow form that has been attributed to the term and the degree to which it has been casually adopted and applied, largely implicitly, by economists and other scholars. These scholars have neglected the many subtle differences in theories of self-interest, a topic that I address in the next chapter.





## CHAPTER FOUR: DIMENSIONS OF SELF-INTEREST

---

### Introduction

In previous chapters I have outlined some of the sources of confusion surrounding the economic and business assumption of self-interest, much of which is caused by an overly simplified and ambiguous understanding of self-interest. I have also claimed that the general motive of self-interest allows a very wide range of behaviors, in this chapter I will provide specific evidence supporting that claim. I believe that a finer demarcation of the different sorts of self-interest will allow modelers to build less ambiguous models which, in turn, will yield results that are easier to interpret and more illuminating. In the first two sections of this chapter I will provide the fundamental elements of a system for organizing the motive of self-interest along what I have called its structural and external dimensions. The separation that I draw between structural and external is made partially for expository purposes; the two dimensions are often linked in the sense that, for instance, a person's overall aims, those objects that a person uses as guideposts from which to subjectively measure his performance, may often affect how he orders his preferences or desires. However, in order to clarify the main issues at hand, I will discuss the two dimensions as if they were independent.

### Structural Dimension Of Self-interest

To remind the reader, the structural dimension is an assumption about the manner in which an agent combines aims and goals. I have chosen to limit my analysis to those types of reasoning that appear to represent positions expressed in the economic and business literatures. Although the terminology related to this topic is subtle and dense, two categorizations are helpful in emphasizing important differences in the various approaches. The first category divides what I will call pure theories of self-interest from mixed theories of self-interest. The theories of psychological egoism, rational egoism and ethical egoism are pure and therefore parsimonious in the sense that they are generally taken to describe what people are or should be in all contexts, namely, self-interested. Mixed theories are those approaches that allow motives other than self-interest to be expressed in certain contexts. The trade-off between pure and mixed theories may be described in terms of the greater accuracy that many believe mixed theories provide, but they provide that accuracy at the cost of greater complexity. I have mentioned several examples of both pure and mixed strategies that the reader may want to keep in

mind when reading this chapter. For instance, Gary S. Becker and many other economists have a very expansive understanding of the range of actions over which the motive of self-interest applies. Meckling and Jensen, developers of Agency Theory, believe that the pure position of rational egoism describes behavior in almost all contexts. On the other hand, Oliver Williamson, a primary developer of Transaction Cost Economics, at times claims that a narrow type of self-interest called opportunism operates when significant economic decisions are under consideration and another type of motivation, which is left undefined but apparently is non-interested, operates in everyday contexts.

The second category concerns positive versus normative theories of self-interest. This categorization is important because economists and others who assume self-interest may not realize that certain forms of that motive have normative and specifically ethical implications. Rather than being a norm free assumption, as I believe it is often taken to be, the assumption of self-interest, at a minimum, has ethical implications. Consequently, if self-interest is understood to be a position with ethical implications then it needs to be defended against other ethical positions. To take a highly relevant example, if an economic researcher's message to his audience is that they should be narrowly self-interested, then that researcher has opened himself to the basic question of why. For example, Meckling and Jensen believe that agents should be modeled as rational egoists, which is a normative position that implies that agents should act from self-interest if they want to be rational. Their position raises the obvious question of why an agent should act from self-interest as opposed to acting, for instance, as duty demands. For instance, a person can ask whether acting from self-interest necessarily makes his life go best or in what dimension of his life will being self-interested make his life go best. The answer to those questions will depend on one's aims, the topic of the section to follow; in this section I will confine my analysis to the ethical implications of the different types of self-interest under consideration. Economic and business scholars should be aware of whether they are taking a stance that has normative implications and, if so, they should be prepared to defend their position against other norms of behavior.

I should stress that the fact economists and business scholars make normative statements is not the issue, that, after all, is a large part of the entire economic exercise. The ultimate goal of normative economics (as opposed to positive economics) and business research, is presumably to derive a better understanding of economic and business processes and to promote better economic and business

practices. Economics, and more importantly business studies, is concerned with the positive study of how markets and organizations work, but also with how firms should operate in order to raise profits or share price or some other objective. The point at issue, however, is not the best procedures to reach an objective, but the normative assumption that economics makes about the aims of agents. Economics, rightly or wrongly, is often taken to advocate a very narrow set of norms centered on egoistic behavior. The message is that agents should be and ought to be selfish or at least should not feel inhibited from so acting. I am not suggesting that economic audiences are merely passive recipients to economic propaganda, however, there is evidence that by repeating a message enough times it becomes true or believable to some. According to Jane Mansbridge: "Interpreting all non-egoistic behavior as egoistic reduces the degree to which the readers of those interpretations are exposed to unselfish behavior. Therefore, interpreting non-egoistic behavior as egoistic diminishes a powerful incentive to unselfish behavior. This in turn reduces the likelihood of human cooperation" (Mansbridge, 1990; see also, Dawes et. al, 1990, and Marwell et al. 1981, Frank, et al., 1988). Although I do not agree with Mansbridge's claim that egoism and cooperation are necessarily opposed to one another (instrumental egoists for instance may work together to achieve their individual goals), her more modest claim that misinterpreting non-egoistic behavior as egoistic can lead to more egoistic behavior seems reasonable. More importantly, as I will show in this chapter, there is a strong link between self-interest and ethics, so by advocating the position that people are or should be self-interested, economists are advocating a position that at a minimum has ethical implications.

### Psychological Egoism

Given that it appears to be the default position for many economists and people in general, I will begin by discussing the position of psychological egoism before turning to some of the other theories of self-interest. Psychological egoism has some interesting and unique features, namely, it is a descriptive position that says that people are motivated purely by self-interest so that, in turn, all actions or behaviors are at their base motivated by self-interested. Advocates of psychological egoism argue that even those acts that appear to be motivated by some other motive are, if one digs deep enough, really motivated by self-interest. This position is meant to describe the motives behind all actions in all contexts, as a result, it is open to refutation by counter-example, although I have shown in

previous chapters that discovering the motive or motives behind a behavior is easier said than done.

There are several reasons that psychological egoism is a particularly interesting topic for discussion in this manuscript. An obvious reason is that the standard behavioral assumption made by many economists and business scholars seems to reflect a belief that all “economic” reasoning is based on self-interest. And, since it is becoming increasingly unclear how economic reasoning differs from non-economic reasoning, and the trend is to extend the range of subjects to which economic analysis applies, psychological egoism then appears to be an accepted description of all facets of life. It follows from this line of reasoning that to the extent that activities are based on the motive of self-interest and rationality, then they are potential subjects of economic analysis.<sup>9</sup> Second, as previously discussed, an acceptance of psychological egoism would have the benefit of aligning the assumption of self-interest with preference theory. Although preference theory says nothing about motives, the belief that all actions are motivated by self-interest would imply that all preferences are motivated by self-interest as well. An acceptance of psychological egoism would therefore mean that the assumption of self-interest and preference theory would no longer potentially be at odds. Third, an acceptance of psychological egoism would open the way for an acceptance of rational egoism, a position that has many adherents in economics; I will discuss the relationship between psychological egoism and rational egoism in a separate subsection below. Finally, if the arguments for psychological egoism are found convincing, then motives other than those based on self-interest are implicitly rejected, if not, then the assumption of self-interest in particular contexts needs to be defended or face the charge of arbitrariness. In other words, if psychological egoism is accepted, then there would be no reason to differentiate between different contexts in terms of their motivation because all actions in all contexts would be motivated by self-interest. If it is rejected, then those who assume that self-interest is the motive behind behavior will need to defend that assumption.

---

<sup>9</sup> See Brennan (1994) for the negative implications that some economists make by equating self-interest and rationality. Within economics it is often taken for granted that self-interested (or prudent) behavior is rational, while the argument that ethical behavior can be rational is often met with skepticism. That's my starting point as well, it should be emphasized that self-interested behavior can be irrational, and altruistic motives, which I believe to be a necessary condition for an act to be ethical, can be rational.

I need to be more careful about the basic claim of psychological egoism. I do not believe that a defender of psychological egoism has to believe that people are somehow bound or required to act from self-interest alone; there is no law which forces an agent to always be motivated by self-interest, by which I mean that an agent always has the choice to act from a non-self-interested motive. The argument, I believe, is that it is “natural” to be motivated by self-interest and acting from a motive other than self-interest is therefore “unnatural”. The term “natural” (like the term rational) is one of the more abused terms in philosophy and social sciences. I use it in a biological sense, that people are genetically inclined to be self-interested. Therefore, people are able to act from a non-self-interested motive, but doing so goes against their true inclinations according to this position. Perhaps, along similar lines, a psychological egoist might argue that there is no good reason to act from other than self-interest, thereby, a person can act altruistically, but there is no good motivating reason to do so. The second line of reasoning begins to express the view of rational egoism to be discussed below.

Psychological egoism, just as with all of the theories of self-interest that I discuss in this section, doesn't tell people which external objects to aim at, rather, it says something about how self-interest is presumed to function in human psychology. Psychological egoism differs from rational and ethical egoism in that it is a descriptive, non-normative, approach that holds that people can naturally pursue only what they take to be in their self-interest. In addition, because all motivation to act is taken to be self-interested, it also implies that all behavior is at base self-interested. However, as shown in chapter two, accepting psychological egoism would mean rejecting what many would find to be essential human characteristics, for example, there would be no place for truly altruistic acts, sympathy, duties and particular affections that damage one's interests, and special relationships that provide good reasons for a person to make self-sacrifices for others based on the relationship alone. In keeping with my previous attempts to define the boundaries of self-interest, psychological egoism can perhaps best be understood by examining the sorts of behavior an acceptance of that position excludes. Shaver (1999) has compiled just such a list which includes the following items: a. harm to the agent's own interests; b. doing what is morally right; c. and, promoting the well-being of others. These three acts are excluded because, respectively, they are self-destructive, morally motivated, or altruistic. As I will show, I do not entirely agree with Shaver's analysis, but his list brings to fore many important issues that need to be addressed if psychological egoism is to be understood. The procedure I will follow in the next few subsections is to examine and challenge each of Shaver's

points in turn, emphasizing their respective shortcomings; by this method I hope to give an impression of the facets and potential limits of psychological egoism.

According to Shaver, a theory of self-interest needs to exclude the possibility of intentionally harming one's interests. On the surface this seems like a reasonable requirement; it would be odd to argue that someone who intentionally harms himself is acting in his interest. However, people frequently, knowingly, harm their own interests. The psychological phenomenon known as weakness-of-will is a common occurrence, it describes those situations in which a person performs an act even though the person believes, even at the point of acting, that it is not in their interest to do so. An example is the case in which a person succumbs to a momentary urge to eat a bag of cookies even when they simultaneously believe that they need to lose weight to remain healthy. Many consumers feel a strong impulse to buy things that deep down they know they really don't need or even want. Supermarkets, for example, know that it is hard to resist the impulse to buy small trinkets and snacks and intentionally place those items near the checkout counter hoping that shoppers will succumb to a momentary urge to buy those items. In fact, many people go to great lengths to avoid temptation--to put it into familiar terms--they intentionally restrict their preference set. But that explanation is not quite right, because what seems to be a better description of what is happening at such moments is that a person has an immediate preference to eat a handful of cookies, although they simultaneously have a longer-term preference not to. I propose, in keeping with precedent, to call such immediate preferences urges, impulses or passions. The difference between these drivers of action and preferences is critical. Preferences, I believe, are informed by careful reasoning, while urges and passions may cause a person to temporarily bypass reasoning; they thereby represent a potential source of internal conflict between immediate and longer-term interests.<sup>10</sup> Succumbing to an urge means doing what you believe is not in your long-term interest. You have no good reason to eat that bag of cookies, and you are aware of that fact, but you do it anyway. People appear to simultaneously have at least two conflicting sets of interests, one set based on good reasons and another not.

---

<sup>10</sup> See Amartya Sen (1978) for a discussion of meta-preferences.

What makes weakness-of-will cases particularly interesting is that there is no confusion or lack of information about what is in one's interest; again, momentary urges overcome what people believe is in their best interest. It is not the case that a person exhibiting weakness-of-will isn't aware or has forgotten that their actions are harmful to their interests. Weakness-of-will is also not the same thing as taking an interest in activities that, perhaps objectively, may not be in one's best interest. I believe that the difference in these cases relates to the reasons that an agent uses to justify their actions. For instance, a person may gamble even though they believe that gambling is not in their long-term financial interest; however, they expect the short-term pleasure received from gambling to outweigh the expected future displeasure from being a little bit poorer. The gambler in this case believes that it is probably not in their best financial interest to gamble, but the pleasure derived from gambling overcomes the reduction in pleasure they expect to suffer from a reduction in wealth. They are aware of the calculation, proceeds to gamble, and have good reason to do so. This case differs from weakness-of-will because the expected pleasure of gambling overrides, for the person in question, the expected displeasure from losing some money. In contrast, cases of weakness-of-will involve internal controversies; for the gambler described above there is no controversy because they expect to receive an increase in net pleasure. Contrast the foregoing case with the case of a person addicted to gambling. Addicted gamblers believe that gambling will leave them worse-off. Put in terms of the aim of pleasure, they believe that their overall expected pleasure will be diminished by gambling but they do it anyway. They are aware that the momentary pleasure that they expect to derive from gambling will not be greater than the displeasure that they will experience from lower wealth, yet they continue to gamble. They don't have a good reason to gamble, in fact they have a good reason not to, but they do so anyway.

The relevance of this issue should be apparent: immediate interests, whether they are called urges, impulses or passions, don't necessarily reflect a person's wider interests. People have immediate interests that they correctly believe will damage their long-term interests and, although their reasons for performing such acts might not be good, that doesn't stop them from acting upon them. Weakness-of-will is an empirical strike against any theory of self-interest, including psychological egoism, which claims that people always do what they believe to be in their interest. As a descriptive position, psychological egoism is therefore flawed. It might be possible to modify psychological egoism to claim that, for instance, rational or reasonable people should always do what they believe is in their interest, but that maneuver

would mean that psychological egoism is no longer a descriptive position representing real motives. Rather, it would be a normative position that claims to tell people what they should do if they want to live, for instance, a rational life. To conclude, a theory of self-interest should perhaps exclude acts that harm one's interests, but such a theory will thereby exclude some observed behaviors.

The second exclusion mentioned by Shaver, that psychological egoism bars doing what is morally right, is more problematic. Many ethicists propose that ethical behavior can't be motivated by self-interest, thereby making this claim true by definition (Sidgwick, 1981, Williams, 1973). I agree, ethical behavior has to be predominately motivated by altruism and cannot be motivated by self-interest alone. However, contrary to Shaver's claim, it is possible that someone can perform acts that meet with moral approval while being self-interested. This point turns on the issue of what it means to be ethical. If a person is required to be in an ethical psychological state of mind in order to perform ethical acts, then behaviors that meet with ethical approval but are not based on moral reasoning, for example altruism, are not ethical. However, this doesn't necessarily exclude the possibility that behavior that is motivated by self-interest can result in acts which are ethically approved. For example, self-interested people and corporations often donate to charities because it is good for their respective images. If we take the aim of improving one's image to be exclusively motivated by self-interest, then these individuals and corporations are not acting ethically. However, the actions of these donors may meet with ethical approval. I have previously discussed the difficulty of determining the degree to which an act is altruistic or pseudo-altruistic, but the point remains that it is certainly possible to perform acts that have beneficial effects, without being an altruistic person or company. Psychological egoism excludes the possibility of being altruistic and therefore, by some definitions, morally right, but it doesn't prevent performing acts that meet with moral approval. In short, theories of self-interest bar a person from being ethical, but do not bar them from doing what is morally right.

An interesting twist on this issue which is sometimes heard is the argument that someone can't reasonably be held accountable for what he should do if he is incapable of doing otherwise. It is not, for instance, reasonable to tell someone that they should be willing to sacrifice their own interests for the interests of others if it is believed that people can only naturally be self-interested. It is always possible that some people might be able to suppress their natural inclinations and act according to some other motive than self-interest, but such behavior would be



beyond the reasonable limits of most people. Telling people that they should act altruistically would be asking them to perform acts that are known to go against their true nature. If this argumentation is accepted, then an acceptance of psychological egoism thwarts any discussion of ethical norms. Not all of which is necessarily a strike against psychological egoism; rather, it has often been used as an argument against the unrealistic requirements of many ethical systems.

In addition, 'original' self-interested motives can be manipulated to conform to standards of ethical behavior. For instance, incentives can be offered to bring self-interested motives into line with what appears, but of course isn't, ethically motivated behavior. In such cases, even granting that people are psychological egoists doesn't imply that self-interest can't be used to promote ethically approved actions. For example, many governments offer taxpayers incentives for contributing to charities. The idea being that if the reward is big enough, self-interested people will contribute to benevolent causes. In effect, a tax deduction transforms self-interest behavior into ethically approved activities. Even though psychological egoism may describe the motives behind original behavior, it doesn't necessarily describe the resulting actual behavior because self-interest can be manipulated to conform to ethically commendable behavior. One last comment on this subject, the case of tax incentives implies that at some point someone may have decided that benevolent, pseudo-altruistic, acts were worth supporting, which raises the issue of whether the psychological state of the person (or persons) making such a decision was altruistic. The issue then becomes whether an act with ethical implications is ethical if the state of mind of the person who induced such an act was ethical, but I'll leave the answer to that discussion to braver souls.

I disagree with Shaver's third exclusion that states that psychological egoism forbids promoting the well-being of others. I believe that it is often quite the opposite, a psychological egoist who refuses to promote the well-being of others would in many cases be damaging his own interests. I will give two examples that demonstrate this idea; the first is the qualification that a person's interests often include the interests of others if only for instrumental reasons, the second is an empirical exception involving special relationships.

Self-interested motivation, for most people, almost certainly includes considerations for the interests of selected others. To take a simple and obvious example, people's interests often include the well-being of their family, friends and other relations. This is not the same as arguing that people are altruistic in that

they are willing to make a net sacrifice of their own interests to promote the well-being of others, rather, this is the much weaker claim that people are at times willing to include the well-being of others in their own interests. A person can most certainly include the well-being of others in his own interests if only for his own, instrumental, reasons. To use the terminology of the previous chapter, while the focus of the action is on the interests of the agent and not on the interests of other person, the interests of the other person promotes the agent's interests. In short, a person can be self-interested and promote the well-being of others if only for self-interested reasons. There are many economic examples of cases in which promoting the well-being of others is a necessary component in promoting one's own interests. For instance, many economic trades, excluding those made under duress, are expected to promote the well-being of both parties. In such situations we expect that the focus of each of the trading partners is centered on promoting his own interests. However, an enlightened agent will realize that his own interests and the interests of others are often intertwined. One example should be more than enough to convey this simple idea; an employer attracts employees to his firm by offering incentives that promote an employee's well-being, the employee then, presumably, helps the employer to achieve his aims. In well functioning labor markets, the employer must account for and promote the interests of potential employees if he wants to realize his own interests.

I will make two final comments on psychological egoism before discussing the relationship of self-interest to ethics, one involving its execution and the other its extent. Even psychological egoists would have to admit that at times people don't achieve their interests due to such things as accidents or misinformation or ignorance. Although clearly true, this admission doesn't damage their position because a psychological egoist only claims that people are motivated to act by self-interest, not that they always achieve those interests. There are many factors, both within and beyond a person's control that can cause him to fail to achieve his aims and goals. For instance, people often make decisions that conflict with their best interests due to poor information or an improper use of information, but this is not a point against the theory of psychological egoism, only against its execution. Of course, a theory must take into account whether the goals and aims it sets are achievable, but under most circumstances this charge cannot be leveled against theories of psychological egoism. In other words, it is usually possible to act from self-interest alone.

Finally, it is unclear whether a psychological egoist would have to argue that people are required to attempt to maximize their self-interest or if a marginal or minimum increase in well-being is enough to motivate an action. One could argue that as long as an agent's interests are promoted, then that is enough to motivate a psychological egoist to act, but this would seem to go against the spirit of the approach. Faced with two alternatives, one of which does more to further his interests than the other, it seems reasonable to require that the psychological egoist choose, within reason, the one which yields the greatest net benefit or causes the least net harm. A complete definition of psychological egoism would have to distinguish between these different sorts of approaches to self-interest.

To summarize this long sub-section, there is empirical evidence in the form of weakness-of-will and special relations that call into question whether psychological egoism is a good description of all behavior. What psychological egoism probably does exclude is the reasonable or rational harming of one's interests, ethical motivation and the sacrifice of one's own interests for the well-being of others. The major benefit of assuming that people are psychological egoists is parsimony; all motivation is taken to originate from self-interest and therefore, to the degree that it represents actual motivation, it can be an efficient assumption.

#### Normative Approaches

In this subsection I purpose that to the extent that economists and business scholars claim that agents should or ought to be self-interested they are taking an ethical stance or a stance that at a minimum has ethical implications. Mine is not a new observation, for instance, Henry Sidgwick (1981) rhetorically questioned whether he should include the topic of egoism in his famous book, *The Methods of Ethics*. While initially offering arguments against including it because, "... there are strong ground(s) for holding that a system of morality, satisfactory to the moral consciousness of mankind in general, cannot be constructed on the basis of simple Egoism" (Sidgwick, 1981, p. 119). Sidgwick decided to include a discussion of egoism in his book because, there is "wide acceptance of the principle that it is reasonable for a man to act in the manner most conducive to his own happiness." Indeed, he notes that self-interest is expressly admitted into the systems of leading representatives of both Intuitionism and Universalistic Hedonism. More forcefully, Sidgwick notes that: "[it is] hardly going too far to say that common sense assumes that 'interested' actions, tending to promote the agent's happiness, are prima facie reasonable: and that the onus probandi lies with those who maintain that

disinterested conduct, as such, is reasonable" (Henry Sidgwick, edition 1981, p. 120). There are a couple of points to note in Sidgwick's statements. First, Sidgwick is not claiming that egoism is an ethically defensible position, only that many people take what is in their self-interest as offering a good reason to act. I believe that Derek Parfit (1984) echoes Sidgwick's statement when he makes the claim that egoism has ethical implications. Simply stated, as I have alluded to on numerous occasions, acting exclusively from the motive of self-interest potentially crowds out the possibility of exercising ethical motives. The second point to note is that Sidgwick is drawing a relationship between self-interest and rationality. He claims that the burden of proof lies on those who claim that actions that are disinterested are rational as well. Therefore, following Sidgwick, economists and business scholars who assume rational egoism are in line with "common sense" understanding when they link rationality and self-interest.

#### Self-interest and Ethics

The relationship between ethics and self-interest is as old as ethics itself. In this paragraph I will briefly discuss two general approaches that typify this relationship; the point being to demonstrate the flexibility with which the term self-interest has been used, a flexibility that allows confusing and ambiguous descriptions of behavior. On the one hand, a great deal of effort has been expended by ethicists to attempt to show that self-interest, if properly understood, includes acting ethically. In other words, one line of reasoning used by ethicists is to show that a well-informed, knowledgeable, reflective person will act ethically because he properly understands that doing so is in his enlightened self-interest. For instance, the reasoning may go as follows, what may objectively appear to be altruistic behavior is, in fact, self-interested behavior because it is in a person's interest to behave altruistically. A person who acts unethically has an improper understanding of where his interests lie, but once the misunderstanding has been addressed, the miscreant is expected to amend his behavior and act as ethics would have him behave. This strategy involves the difficult task of finding convincing links between what are considered ethical actions and self-interest. The second approach follows the opposite course by disassociating self-interest from ethics. Many ethicists contend that an act motivated by self-interest can't be ethical. The argument is, in order to be ethical, an action must be motivated by altruism, although altruism alone is not sufficient to make an action ethical. I have followed a path in this manuscript analogous to the second approach (that of separating altruism and ethics from self-interest), in order to maintain an important distinction

between the motives of self-interest and altruism. I define an altruist as someone who intends to help others. Admittedly, an act is easier to identify as altruistic if an actor knowingly sacrifices his own net interests for the interests of others, but it is not a necessary condition of an altruistic act. The danger that I am attempting to avoid is that of collapsing all behaviors, including those motivated by altruism, into the motive of self-interest, as is the case in the first approach described above. That approach incorrectly muddles the important, real, distinction between these two very different motives; in short, a person can't be, as I understand the terms, a self-interested altruist.

If the distinction I have outlined is accepted, then economists and business scholars who tell people that they should be motivated by self-interest are taking an ethical stance. By telling people that they should act from self-interest, they are, in effect, telling people that they shouldn't act ethically, while in some cases they are telling people that it is irrational to act ethically. Such an important normative step shouldn't be taken lightly by scholars and requires justification. If the argument is that people can only be self-interested, then we are once again back to the less than completely satisfying case of psychological egoism. However, if psychological egoism is rejected, then the argument for normative views of self-interest seems to be that people have the ability to choose whether to be self-interested or not, and given that ability, they should choose for the motive of self-interest. This raises the issue of what makes self-interest a superior choice over other motives. Several important questions naturally follow. Why, for instance, should a person believe and act as if their interests are more important than the interests of others? Are all of one's interests, no matter how trivial, more important than any of another person's interests? Are there circumstances in which the interests of others have priority? These are questions with important practical implications and the list of potential questions that arise is nearly endless, but a proper discussion would take me too far into the field of ethics. My point, however, is that there is a relationship between self-interest and ethics that needs to be considered by those who promote the motive of self-interest.

Another idea that often accompanies economic discussions of self-interest is the argument that self-interested actions by individuals are justified because they lead to greater good for the group as a whole; an ends-justifies-the-means type of argument. Note that it wouldn't be possible to argue that a person who is aware that his self-interested acts are leading to benefits for the group as a whole is behaving ethically because the intent of the actions are exclusively focused on his

own well-being. However, the “invisible hand” argument has the effect of rationalizing self-interested behavior by which I mean that it offers people a justification to act from self-interest. The argument might go as follows, even if acts motivated by self-interest are unethical, they lead to beneficial results for the group as a whole, and should therefore be considered ethical. The argument is not convincing from an ethical standpoint for what should now be obvious reasons; although a self-interested action by an individual group can have beneficial effects for a larger group, the important thing to consider are the intentions of the agent. A self-interested act can clearly have beneficial effects, but that's not the reason why the act was executed. The benefits to the group are not the motivating factors, they are a happy coincidence, and a truly selfish person will be indifferent between those acts that benefit the group and those that don't. Actions motivated by self-interest can undoubtedly have beneficial effects for the group, but they do not necessarily have such effects. Adam Smith, in fact, gives several examples in, *An Inquiry into the Nature and Causes of the Wealth of Nations* in which self-interested acts have led to negative consequences for the economy in general. In fact, it is so obviously true that self-interested acts can lead to a reduction in overall economic well-being that it shouldn't need to be said; the papers are full of examples in which such motives have damaged the lives of others and the overall economy. Altruistic acts, on the other hand, are intended to benefit others. This is not to say that altruistic motives always result in benefits for the wider group; it is also possible to be entirely other-regarding and still perform acts that damage the interests of the group. For instance, a man who sacrifices his own interests to free a dangerous criminal from prison would not be acting for the benefit of the community at large--he might be motivated by altruism, but his actions have negative consequences for the greater group. However, in practice, such cases are probably the exception, altruistic motives are intended to benefit others, as opposed to self-interested motives that are intended to benefit the individual alone.

To conclude, although acts motivated by self-interest can be beneficial for others, that is not the intention of the agent; his intention is clearly to promote his own interests. Acts motivated by altruism, on the other hand, are intended to benefit others. To the degree that people are successful in achieving their intentions, and supposing that people are as likely to achieve their altruistic intentions as they are their self-interested intentions, altruistic acts promote the well-being of others. Whether there is a net benefit for the entire group (including the altruist) will depend on such issues as the marginal returns of actions and multiplier effects. For instance, if an act yields great benefits for the rest of the group at a small cost to the

altruist, then there will be a net benefit to the group as a whole; I will return to the topic of measurement in a later chapter.

### Rational Egoism

Although rationality is not the main subject of this manuscript, the relationship between self-interest and rationality in economics is a close one, so close that the two taken together are often identified as the foundations of economic and business analysis. Some economists argue, but usually just claim, that rationality and self-interest go hand-in-hand. For instance, advocates of the position of rational egoism claim that it is necessary and sufficient for an action to be rational that it contributes to the well-being of the agent. In other words, agents should or ought always to seek their interests; in fact, the claim is that it would be non-rational for someone to pursue any other course of action other than that which promotes his self-interest. In contrast to psychological egoism, which is supposed to describe a real psychological phenomenon and is not meant to advocate a normative position, rational egoism describes how people should act. However, as shown above, it is a small, although not necessary, step from the acceptance of psychological egoism to the acceptance of rational egoism. If people are only naturally motivated by self-interest then they pursue only those actions that they expect will yield a net-increase in their well-being. It is then possible to conclude that those activities that intentionally damage one's interests are not natural and therefore are not reasonable. The drawback of linking rational egoism so closely to psychological egoism is that rational egoism is then susceptible to the same empirical deficiencies that plague psychological egoism. It is possible of course to base rational egoism on something other than psychological egoism, Thomas Hobbes, to whom I shall return, is often taken to have advocated rational egoism on the basis of the fundamental right to preserve and promote one's life.

Following the reasoning outlined several paragraphs above, the position of rational egoism clearly raises the issue of the relationship between economics and ethical norms. I believe that, in general, economists, particularly neoclassical economists, are wary of making normative statements. Advocates of modern axiomatic utility theory largely disqualify themselves from making normative claims other than their unwavering support for the objective of personal autonomy to the exclusion of all other values. Autonomy is the only value that the theory implicitly addresses and although autonomy is certainly an important value, for instance, few would want a benevolent dictator defining our lives even if he made our lives go better.

Economic autonomy, by which I mean having the means and ability to make meaningful choices in regards to one's economic interests, is an important sub-set of autonomy, but it is most certainly not the only thing that people value. For instance, in many countries people are quite happy to give-up a degree of their economic autonomy to the government in exchange for national health insurance and a smoothly working pension system. By relying on utility theory, economists largely forfeit the right to make value judgments. I believe that this is the price paid for the benefits of Hicks' remark, quoted above, of not wanting to see inside people's heads. Should economists be willing to forfeit the right to say, for instance, these acts will maximize profits, but are nevertheless wrong and should be avoided? Utilitarians were willing to make the claim that, on occasion, the rights of individuals need to be sacrificed for the greater good. Modern Utilitarians (the ethical school) and their opponents continue to argue over issues such as what is a right, the conditions under which a person should be willing to sacrifice his rights, and whether it is possible to compare and exchange rights. Should, for instance, a person ever be made to sacrifice his life or is that a basic, irrevocable right? The approach of modern economic utility theory avoids these issues and places autonomy as the central value.

An often heard analogy is that economists are interested in growing the size of the cake and not in the issue of its distribution. I take this to mean that economics, like baking, is taken to be a technical skill, while deciding who gets what share is a normative issue which positive economics doesn't claim to address. However, of course, economists and especially business scholars do make normative statements. For instance, they build models that tell people what they should do in order to maximize their profits or share value or some other objective. However, I believe that most economists would claim that they do not make ethical normative statements, in other words, they are in principle opposed to telling people how they should live their lives. However, given the close link between ethics and self-interest which I have outlined above, it is easily possible to stumble into making ethical statements. For instance, Meckling and Jensen's agency theory clearly advocates rational egoism, a position that has ethical implications.

To conclude this subsection, to the degree that economists subscribe to rational egoism, they are taking a stance that, at a minimum, has ethical implications. Psychological egoism, on the other hand, removes the ethical stance by all claiming that all behavior emanates from self-interest. An advocate of psychological egoism removes the normative element by claiming that people are self-interested, leaving



no place for the claim that they should act in one way or another. In short, you can't tell people how they should act when you believe that they can't help but act from self-interest. Rational egoism can't take this approach because its advocates are telling people that if they want to behave rationally then they should be motivated by self-interest. Therefore, to the degree that self-interest and ethics are correlated, rational egoism has ethical implications. The claim that people are rational egoists shouldn't be taken lightly because it places economics squarely within a thousand year old ethical debate. Not to be flippant, but economists who make such claims should be able to defend their "should" against other "shoulds"; they need to be able to argue why and in which contexts people should act from self-interest as opposed to some other motive (Hausman et al. 1994).

#### Uninformed Instrumental Egoism

An alternative theory to both psychological egoism and rational egoism is instrumental egoism. It claims that it is necessary and sufficient for an action to be rational that it best contributes to the satisfaction of the agent's desires (Shaver, 1999, p. 39). By stipulating that rationality is based on "an agent's desires", as opposed to "the well-being of the agent" as in the case of rational egoism, instrumental egoism avoids some of the pitfalls of rational egoism. In particular, it avoids the empirical problems of that theory because an agent has only to pursue his desires to be rational. This theory would seem to fit well with what I presume most economists mean when they use the term self-interest; however, even a glance at this theory reveals some major practical problems. Note that Shaver's definition allows for other-regarding actions. For instance, suppose a person desires that others do well and is willing to sacrifice his own net well-being for their well-being. The first thing to note is that accepting the instrumental theory would mean explicitly rejecting psychological egoism because it allows for the case that the well-being of others can alone motivate one to action. An instrumental egoist can desire to be an altruist, in other words, he can be a self-interested altruist, a conclusion that hopelessly confuses the terms self-interest and altruism. Secondly, an instrumental egoist doesn't need to benefit from his actions, he may in fact be harmed by them, as long as the action is desired it can fit within this theory. Instrumental egoism thereby rejects each of the conditions of self-interested listed by Shaver. An instrumental egoist can harm his own interests, do what is morally right, and promote the well-being of others. One final note, by not ruling out malevolent acts this theory has severe ethical implications. According to instrumental egoism, a person who desires to enslave or destroy the world is

rational. I suspect that even the most liberal economist might have troubles allowing the expression of such desires within his model.

### Ethical Egoism

Another normative position that some influential authors have argued for is ethical egoism. Advocates of this position claim that the necessary and sufficient condition for an action to be ethically right is that it maximizes one's self-interest. A nice example of this view is offered by Ayn Rand (2007) in her book, *Atlas Shrugged*: "Accept the fact that the achievement of your happiness is the only moral purpose of your life, and that happiness--not pain or mindless self-indulgence--is the proof of your moral integrity, since it is the proof and the result of your loyalty to the achievement of your values." Although happiness has often been equated with self-interest, Rand goes one step further; it is her opinion that pursuing happiness is a moral (or ethical) duty. Although she sets boundaries on self-interest, arguing that it shouldn't include "mindless self-indulgence" nor should it include pain, she offers no argument for these constraints. There are many problems with Rand's approach, her terminology is vague and she often attacks positions that nobody defends. However, Rand makes two major mistakes. The first is the by now the familiar refrain that she conflates self-interest and ethics. To see this, take the example of a person who experiences the feeling of happiness while conducting acts based on altruistic motives. That person, according to Rand, performs acts which are motivated by altruism for self-interested reasons; thereby mixing two motives that should remain separate. Her second mistake is to equate feelings of happiness with ethical behavior. Although it is reasonable to argue that people do those things that will bring them happiness (who would argue otherwise?), it is absurd to claim that all acts that bring happiness are ethical. If I have correctly interpreted and represented Rand's claims, then it is perhaps best to move on.

A subtle but important adjustment might be to add to the definition of ethical egoism the condition that people should follow their enlightened self-interest. Accepting this position would mean that a person should or ought do a specific action, all things considered, if and only if that action is in that person's overall, enlightened, self-interest (Kalin, 1968: 66). This approach follows the general strategy of broadening the definition of self-interest to include those things that people should know are in their best interest where the word enlightened now carries a heavy burden. The advantage of this and other similar approaches is that

it is possible to then claim that a person should have known, for instance, that drinking too much alcohol or eating too many french-fries is not in his best interest, while giving to charities promotes interests because part of being happy is showing a proper concern for the well-being of others. The problem in broadening the definition to include what people should know is in their best interest is that the theory of self-interest may lose the advantage of subjectivity, one of its more appealing points. Many theories of self-interest contend that a person should pursue their desires or preferences as they see fit. By adding the stipulation that people should know what's in their best interest, objective evaluations of actions may be implied. Who, for instance, decides what a person should know? Does the theory only apply to those with the proper knowledge and ability to reason? While the maneuver of adding the stipulation of enlightenment of appeals to common sense, there may be a basic set of norms to which most of us could agree, the potential loss of subjectivity is a heavy price that runs against the grain of much liberal economic reasoning.

#### Strategic Egoism

Strategic egoism is similar to what I have previously called instrumental egoism but with the twist that it is a position which specifically tries to reconcile the demands of self-interest with ethical attitudes. It is an enlightened, instrumental approach in which people come to realize that it is in their advantage to develop, maintain and act on other regarding attitudes. It is described as an enlightened position because people are taken to realize that while their interests are conceptually distinct they are causally interdependent. Once people realize their mutual dependence, the argument goes, they reason that it is in their advantage to cooperate with one another and restrain from particular activities; the result is the development of ethical systems of behavior. This is clearly an instrumental (non-realist) understanding of ethics in that ethical behavior is used as a vehicle to realize one's own interests. People may behave in an ethical manner, but by the reasoning of strategic egoism they are not motivated by ethical concerns and so such behavior is better labeled as pseudo-ethical rather than ethical. There are some other problems with this approach. The model stipulates that agents must come to realize that the compliance of others is conditional on one's own compliance. As an ethical system this position is thereby troubled by the possibility that it is dependent on an agent's belief that others are helpful to achieving his own projects, beliefs that might be missing. For instance, if agents do not believe that others can help their cause, then they have no reason to act

ethically. For example, there will be no ethical reason for strategic egoists to consider the effects of their actions on future generations. In addition, people who are powerful enough can simply opt-out of this pseudo-ethical system. These shortcomings call into question whether strategic egoism is an ethical system at all. Strategic egoism may be a good description of how pseudo-ethical systems develop, but as an ethical system it is a non-starter.

## Other Approaches

### Predominant Egoism

With the discussion of predominant egoism I turn to mixed as opposed to pure theories of self-interest. The appeal of mixed strategies is that they provide a potentially better reflection of reality in that they can accommodate both self-interested and non-self-interested motives--those who base their understanding of human nature on psychological egoism would undoubtedly disagree with this position. The corresponding down side of such approaches is that they are less parsimonious. I have previously shown examples of mixed theories, Oliver Williamson, for instance, differentiates between opportunistic behavior, which dominates in periods in which important economic consequences are contemplated, and what he calls the benign day-to-day periods when "frailty of motives" dominate. An important requirement of mixed theories, one that Williamson fails to adequately address, is a clear specification of the conditions under which a given motive is relevant. The theory proposed by Meckling and Jensen doesn't suffer from this problem to the same degree as Williamson's theory because they argue that rational egoism is appropriate in almost any context. Perhaps surprisingly, Gary Becker (1976), despite his advocacy of the general applicability of preference theory and self-interest, distinguishes between family contexts, when disinterested concerns dominate, and non-family contexts, in which self-interested motives dominate. In doing so, Becker appears to recognize the existence of special relationships, although limiting such relationships to just families strikes me as arbitrary. Specifying the contexts in which the motive of self-interest is expected to apply is only the first step, the next step is to specify which type of self-interest is applicable. Given the many understandings of the meaning of self-interest it is not enough to label motives as merely self-interested, more subtle descriptions are necessary in order to build an economic theory of self-interest.

Gregory S. Kavka (1986) describes predominant egoism as being a Hobbesian position. As the name implies, predominant egoism is the view that human action is predominately motivated by self-interest, although there is room for other-regarding motives. Kavka confusingly distinguishes between self-interest and other-regarding behavior, but as I have previously shown, there is nothing to prevent a self-interested person from incorporating, to an extent, the interests of other people. That issue aside, Kavka describes the position by referring to the following four propositions (Kavka, 1986: 64-65) which I paraphrase:

1. For most people in most situations, the motivation to act altruistically is only triggered if the gains to the benefactor are much larger than the costs to the altruist.
2. There are only a small number of people who are normally motivated by altruism as opposed to self-interest.
3. There are only a small number of situations in which the average person is motivated by non-self-interested motives.
4. The number of relations that motivate people to act altruistically is small and may include friends, relatives and close associates.

Although there is room for debate about the number of people to which these propositions apply, they appeal to common sense ideas about the rarity of altruism and emphasize the mix or distribution that self-interest has both within a person and across people. It is important to recognize that this position leaves open the question about the conditions under which a person is more or less willing to act altruistically. So while it is perhaps a more realistic position than say psychological egoism, it forces modelers to be more explicit about their behavioral assumptions. In particular, an economist adopting this position would have to take into account and specify when egoism is expected to predominate as opposed to some other motive and whether disinterested motives can also be labeled economic.

#### Tautological Egoism

The last type of egoism that I will discuss is tautological egoism, which is the argument that because people always act to satisfy their own desires or preferences they are thereby always motivated exclusively by self-interest. Tautological

egoism is a nagging problem that has a tendency to creep into discussions of self-interest. The problem is that if desires or preferences are interpreted in a broad enough sense, so that they refer to whatever motivational feature within an agent that produced an action, then it is expressing a truism without empirical content. The reasoning incorrectly assumes that whatever it is that people desire or prefer they do so because it is in their self-interest. It assumes, in other words, that the only motive behind an action is the motive of self-interest. An altruist, for instance, desires or prefers placing the needs of others before his own, a motive that is not intelligibly labeled as self-interested. Clearly, an advocate of psychological egoism wouldn't label this argument as empirically empty, but he would be equally unable to intelligibly use it as evidence to support his position. Tautological egoism, as I have shown in the case of both the theories of Williamson and Meckling and Jensen, is still heard today in defense of self-interest. Their arguments run something along the following lines: agents prefer whatever it is that they choose (trivially true) and they chose whatever it is that is in their best interest (not necessarily true), therefore they are always motivated by self-interest (not necessarily true).

#### External Dimension of Self-interest

A source of much of the confusion concerning self-interest is the failure to distinguish between the structural aspects of self-interest, and the aims of self-interest. Previous sections of this chapter have dealt with the different ways that can be used to understand and explain the rationality of self-interest. For instance, an ethical egoist orders his desires or preferences based on his belief that it is ethical for him to pursue his interests in a manner that promotes his designs. The structural component doesn't address the different types of aims or goals that an agent can have. The confusion rests with the knowledge that agents with different structural orderings can have the same external aims or vice versa. For example, an altruist and an egoist can both have the goal of maximizing wealth, while the first intends to promote the happiness of others, the second intends to promote his own interests.

The external component of self-interest concerns the aims or goals of agents. In previous chapters, for simplicity, I mainly assumed that the pursued goal was happiness or pleasure, but the number of aims that an agent can have is large. In addition, aims can have different time frames which may conflict with one another. For instance, a long-term aim might be to live a cerebral life, while a short-term

aim might be to read the latest Nicci French novel or listen to the newest Madonna CD; two sets of goals that clearly conflict. The general motive of self-interest, once again, is one method of reasoning or rationality that a person uses to link his motives with his overall aims. While the structural dimension represents the rationality aspect of self-interest, how preferences and desires are ordered, the external component represents the general aims that are preferred or desired. Separating the reasons or rationality that a person uses from the aims he pursues is a simplification, but I hope a helpful one. It is a simplification because what a person aims at is likely to influence how he tries to achieve it. For instance, as I have discussed previously, a person whose primary aim is to improve the well-being of others will probably use a different sort of reasoning or rationality than a person whose aims to maximize his own well-being. Vice versa, the rationality that a person uses may partially determine the aims of the person. In all likelihood, neither component of the relationship is fixed and one influences the other. That said, the message I wish to convey in this section is simple: the aims or goals that agents have affect the type of reasoning or rationality that they use to achieve those aims. It is important to remember that self-interest is a means to achieve an aim, it is not itself an aim. People do not attempt to maximize (satisfice, etc.) their self-interest, they attempt to maximize certain objects that they take to be in their interest. Popular use of the term self-interest may cause people to overlook the idea that self-interest doesn't specify a particular aim. I suggest that different irreducible aims are likely to result in different types of reasoning and rationality, implying that modelers of human behavior need to specify the aims of agents if they wish to draw conclusions about the rationality of their subjects. Claiming that an agent is self-interested, but not specifying the aims of that agent, may lead to ambiguous interpretations of the agent's rationality. Even if one accepts self-interest as the only type of rationality organizing motives, the breadth of that assumption is so extensive that its use as the sole assumption of a model is of limited value.

For reasons of simplification, in the following discussion I take for granted that self-interest, as opposed to some other motive, is the sole aim of the agents under consideration. Therefore, the external dimension of the general motive of self-interest concerns the various aims or goals of self-interest. Whatever the type of self-interest one assumes, the aims of agents are undefined. For instance, an agent can be self-interested and aim at leading a happy or pleasurable life, an academic life or a life full of experiences. Being exclusively self-interested probably prevents pursuing some aims, for instance living an ethical life or a godly life, but

the list of general aims that self-interest allows is nearly endless and so a method will have to be found in order to keep the discussion tractable. The procedure I will follow in the subsections to follow is to first discuss Derek Parfit's (1984) very general categorization of the objectives of self-interest before turning to a more specific list of objectives compiled by C.D. Broad (1930, 1952) and others. Finally, I will examine some of the default aims that economics is commonly taken to specify.

#### Parfit's Objectives of Self-interest

Derek Parfit provides an outline of the different objectives of self-interest in his book, *Reasons and Persons*, (1984, appendix I). Although he doesn't use the same terminology that I do, his review of the different sorts of self-interests coincides well with what I call the external dimension of that motive. Parfit identifies three kinds of theories of self-interest: hedonistic theories, desire-fulfillment theories and, objective list theories. In the terms I have been using, claiming that someone is self-interested leaves open the question of whether that person is attempting to aim at what will make their lives happy or pleasurable, or to fulfill their desires, or an objectively compiled list of objects that will make their lives go best. I will briefly discuss each of these theories and show how they demarcate different external aims.

Hedonistic theories say that what is best for someone is what will make their lives happiest or most pleasant. Henry Sidgwick has called these approaches *prima facie* reasonable, accordingly, in previous chapters I have been using happiness as the single goal pursued by agents, however, there are some well-known problems with hedonistic theories. The three I discuss are: immediate versus overall happiness, aberrant behavior and superficial happiness.

An obvious flaw with the hedonistic definition above is that what makes someone happiest at a moment in time may not bring them long-term happiness. A solution to this problem would be to amend the definition with the proviso that the aim in question is total, life-time, happiness. This solution, however, raises some problems of its own in that it is possible to imagine some troublesome cases. For instance, a person experiencing just a minimum of happiness, but living a very long life would, under the hedonistic definition, be living a better life than someone who lives a life of average length, but one which is intensely happy. Although the first person, given his very long life, experiences a higher total amount of happiness



over the course of his life, the second person has, by most accounts, led a better, although shorter, life. In other words, the definition doesn't account for the intensity of happiness during a lifetime, in a similar vein, this position doesn't address how to balance short-term versus long-term happiness or how to balance happiness and unhappiness. What, for example, are the respective discount rates of happiness and unhappiness?

Hedonistic theories also do not account for aberrant behaviors. For instance, a person who is made happy by causing misery to others or by restraining the possibility of others to achieve happiness may be leading the best possible life under this definition. A hedonist isn't required to take into account the happiness or well-being of others unless it affects his own happiness. He only has reason to consider the well-being of others to the degree that their well-being affects his own--in short, in this respect he is an instrumental egoist. For example, a dictator may be very happy and living the best life he can while causing his subjects great harm. Although he may believe that he is living the best life possible, most observers would find his life repugnant and would believe that he too should see that his aim is flawed. Incidentally, the dictator might disagree, he might, for example, argue that special circumstances entitle him to his position and he therefore has no cause to show remorse for his behavior. In that case, there is nothing more that can be said to such a person under this theory. One could try to find the source of the dictator's belief that his life entitles him to particular privileges and attempt to demonstrate to him that his belief is (presumably) incorrect, but there is no guarantee of success given his ilk. Convincing someone that a particular act is right or good can be difficult, and can't be the standard for judging the rightness or wrongness of an act; even knowing that an act is right or good may not give some people a reason to do it. It is impossible to prove that all lives are of equal worth or that some interests outweigh others unless some basic measure of comparison is adopted. Whether there are right and wrong, and good and bad ways to act, and how and if we are able to perceive them is, of course, at the core of ethics. I will leave this discussion to others, my simple claim is that a theory of self-interest that doesn't take into account the basic implications of the effects of one person's actions on another is seriously flawed. A practical theory of self-interest needs to set certain limits on the types of behaviors within its purview if it hopes to reflect observed behavior. This doesn't have to be a complicated task, it may mean specifying, for instance, that anything that is legally permitted or is a part of the generally accepted behavioral norms in a community is an allowed objective of self-interest. Such solutions won't solve the ethical problems (e.g.,

relativism versus moral realism), they merely shift the point of scrutiny to the legal system and norms of a community, but they do help to specify the limits of self-interest.

Hedonistic theories do not account for qualitative differences in happiness. A famous example that emphasizes this shortcoming is the case of the happiness machine. The idea is to imagine a situation in which someone is connected to a happiness machine his entire life. The machine keeps the agent alive and provides every conceivable need and, additionally, provides a drug from which they experience an intense state of happiness. The agent is never allowed to be unplugged from the machine, and may not even know that they are attached to the machine. Under the definition of hedonism this person would be living a best life, in fact, under the condition set by hedonism, such a life could be considered the best of all imaginable lives. However, many would consider such a life a very shallow sort of existence, not worthy of being called a life at all, and certainly not the best life possible. The point is that many believe that there is, or should be, some purpose or objective to life other than happiness. Happiness should be based on something worthwhile, but this admission allows for the possibility that there is an objective element to self-interest. In other words, being happy is not enough, a person should be interested in leading, for example, a fulfilling life full of experiences some of which may bring temporary unhappiness. An economic theory which claims that people should only be interested in their own happiness leaves open the possibility that people may choose to live lives that are shallow and unworthy. There should be some good method of comparing the quality of economic lives.

Incidentally, the same three critiques discussed above, immediate versus overall happiness, aberrant behavior and superficial happiness, can also be applied to preference theory with one major stipulation. Preference theory, to remind the reader, in its most general form only requires an agent to be able to make pair wise comparisons of objects and choose those objects which they prefer. The theory intentionally leaves open questions such as what it is that agents prefer and why they prefer them. In other words, no objective stance is taken on what a person prefers, the model is entirely subjective. The major difference with hedonistic theories is that preference theory doesn't require that an agent choose what will make their life go best, in fact, they are free to choose what will make their lives go worst or to first choose what will make their lives go best and then go worst or any other conceivable aim or set of aims. This difference in the theories is important

and a strike against preference theory.<sup>11</sup> Hedonistic theories may be criticized as too liberal in that they allow behavior that is inconsistent, ambiguous and unethical, but at least they contain an internal measure of success, i.e., doing what will bring about the most happiness. Preference theory lacks any such measure with the result that although the model is of nearly unlimited applicability, its generality is its downfall. Such a general theory may be necessary to explain observed behavior, but it cannot be a sufficient explanation of the behavior of agents; other restrictions need to be added to the theory if it is to provide a helpful guide to human behavior. Economists choose to add the stipulation that agents are self-interested, but, as I have shown, that too is a very liberal assumption which allows nearly any sort of behavior. To conclude, the combination of preference theory and self-interest is too general to be of much use.

The second category discussed by Parfit is desire-fulfillment theories. The idea behind these theories is that what would be best for someone is what, throughout this life, would best fulfill his desires. This is a more general theory than hedonistic theories because happiness is only one of many possible desires an agent can have. However, the theory suffers from some of the same flaws as hedonistic theories, for instance, it also has the disadvantage that it makes no distinction as to the quality of the desires; a desire to rule the world is on equal footing as a desire to eliminate hunger. An argument against these theories, analogous to that of weakness-of-will, is the claim that they may be self-defeating because they make the assumption that getting everything one wants will make one's life go best. Many, including Adam Smith, the Cyrenaics, the Epicureans and the Stoics, have argued against the benefits of pursuing all but the most basic desires. The argument is that once basic human needs are satisfied, anything over that level can be a source of unhappiness and should be avoided. Desires beyond what is necessary to lead a modest, comfortable life, the authors above argue, can never really be fulfilled and can lead to entanglements and frustrations. In a related concern, the Stoics argued that as knowledge grows desires change. The desires of

---

<sup>11</sup> Hausman and McPherson (1996, p. 5) make a different claim about the role of self-interest in preference theory: "Moral and altruistic preferences are ruled out of many economic models not by axioms of rationality, but by assumptions that the objects of preferences are bundles of commodities to be privately consumed and that there are no interdependencies among the preferences of different individuals." I do not believe that there is any rule in utility theory which restricts preferences to private consumables. If there is, then Hausman and McPherson have identified a serious limit to the applicability of preference theory.

a child are different from those of an adult, while a sage's desires are more rational and virtuous and therefore lead to greater happiness than those of a "normal" adult. Similarly, the desires of a conscientious, reflective adult are different from those of an adult that follows his base instincts. Therefore, desire-fulfillment theories need to account for the "value" of a desire. These differences will be important when I discuss Joseph Butler's idea of the role of reflection. Briefly, Butler argued that there is a hierarchy among the things that motivate a person to act. For instance, impulses and certain passions may lead a person to desire objects that are not in his best interests. The desire fulfillment theory makes no distinction between the sometimes self-destructive desires a person has as a child and those he has as a mature, reflective adult. We can also ask of this theory what happens if desires are not fulfilled. For instance, aiming for too much and missing doesn't necessarily mean that an outcome is not in one's interests. For example, trying to be the best person one can be and failing to achieve that desire may still mean achieving a high level of satisfaction. The process followed to fulfill a desire may be as important or more important than getting what one desires. However, failure to achieve some aims can have devastating implications, many lives have been ruined through dedication to an unachievable goal. In short, the choice of desires can have implications for the state of well-being in cases in which desires are not fulfilled. Finally, there is some ambiguity in the definition about the type of desire under discussion. Throughout this manuscript I have been using desires to mean relatively short-term objectives, as opposed to aims and goals that are longer-term in nature. If that categorization is accepted, then the phenomenon of weakness-of-will raises obvious implications for this theory in that people often desire things that make their lives go worse.

And, once again, there is a similarity between desire-fulfillment theories and preference theory. Preferences, in common speech, are not the same as desires. For instance, a person can prefer a quick death to a long drawn out death, other things more or less equal, but not desire either. In addition, preferences are a relative concept, whereas desires are not. A person can desire an object without reference to another object, although in practice we often desire one thing more than another. These qualifications aside, in general, the critiques mentioned in relation to desire-fulfillment theories appear to equally apply to preference theory.

The third theoretical aim of self-interest that Parfit discusses is objective list theories. The idea is that certain things are good or bad for us, whether or not we want to have the good things or to avoid the bad things. We should, according to

this theory, attempt to aim at those objects that are good for us and avoid those that are bad. This theory obviously avoids the problem of the subjectivity of the two previous theories, but presents the problem of deciding what goes on the list and of who makes that decision. Although I believe it is possible to find agreement on a list of certain basic aims that most mature, reasonable people would agree are good for everyone in most circumstance, and another list of those things that are bad for most people most of the time in most contexts, the mechanics of developing a complete list that covers all possible things that are good and bad for people in all circumstances is too odious to contemplate--although I will mention a few general aims in the next paragraph. The complete guide to the goodness or badness of an act would, for instance, need to include the future repercussions of each alternative under discussion and have a mechanism for comparing alternatives. An alternative possibility, no less ominous, would be to develop a list of procedures for deciding what is good and bad for a person, but this too suffers from intractability. Another alternative, a sort of anti-list tack, is to use a particularist approach which doesn't try to create a list of goods and bads or procedures for determining what is good or bad but, instead, recognizes the uniqueness of many situations. A particularist approach explicitly recognizes the belief that general rules cannot be applied to every situation. Such an approach, by definition, cannot be modeled and so is of limited value to economists and business scholars. In any case, objective list theories have some appeal, but at the cost of the loss of individual sovereignty that marks many economic approaches. Whether that price is worth paying will depend on the context under consideration; for instance, while some behavior is influenced by an explicit or implicit "list" of approved objectives, other contexts allow a more open, liberal pursuit of preferences and desires. In practice, there are very few situations in which preferences are given totally free reign. Agent preferences are internally constrained by, for example, the norms of the community in which they matured. They may additionally be constrained by external forces in the form of laws that are not necessarily unanimously agreed upon and thereby create objective constraints on the preferences of some individuals.

### Some General Aims

I will now turn to the task of listing some more specific, but still very general aims and goals as compiled by C.D. Broad (1930). The generality of many aims can be seen by their repeated appearance in history. For instance, the Hellenistic philosophers typically referred to the goals of wisdom, temperance and justice, all of which are likely to appear on modern lists of worthwhile goals (Cicero, On

Duties, translated by Grant, 1971). While all of the aims discussed in this subsection can be pursued for self-interested reasons, they are not necessarily pursued for that reason. The pursuit of happiness, for instance, doesn't exclude performing benevolent acts in that people can derive happiness from performing such acts. While the standard, *prima facie* goal of self-interest is happiness or a pleasant feeling, other goals and aims have been suggested. For instance, while Sidgwick (1988) took happiness as a *prima facie* end, he also accepted excellence of human nature as a strongly supported *prima facie* end. Where excellence is "not primarily superiority to others, but a partial realization of, or approximation to, an ideal type of human Perfection." Under excellence falls aims such as doing god's will, self-development or realization, and living one's life according to nature. Few of these goals are necessarily associated with feelings of pleasure and fewer still are made more accessible by increases in wealth. However, all of them can rest safely under the umbrella of enlightened self-interest. The point, once again, is to argue that even if the claim that agents are exclusively self-interested is accepted, that claim alone doesn't significantly limit behavior.

An important point of consideration when modeling aims is their reducibility. C.D. Broad (1952) argues that there are many aims, some of which are irreducible including: pleasure, pain, wealth, power, security, liberty, glory, possession of particular objects, fame, health, longevity, status, self-respect, self-development, self-assertion, reputation, honor, and affection. For Broad then, wealth, power, etc., are not reducible to pleasure or happiness or anything else. Rather, they provide unique, separable goals, each of which can be in a person's interest. Broad supports his claim that not all aims are reducible to a common value or good by observing that some aims are relative in that they require the presence of other people, and some are absolute in that they can be realized in isolation. Power, for instance, is a relative concept, while self-preservation can be an aim in isolation. Therefore, following this argument, at least two types of distinctive aims can be identified. Another argument that some of the aims of agents are distinct rests on feelings. The argument is that, for instance, the feeling one has while acting honorably is fundamentally different than the feeling one has, for example, upon winning the lottery.

If Broad is correct in that there are different distinct aims, then any theory based on the motive of self-interest should specify the aims the modeler has in mind for the agents he is modeling because different aims will likely be pursued for fundamentally different reasons. For instance, agents aiming at the maximization

of liberty will likely behave differently than those aiming at the maximization of wealth. These two aims might be correlated, but not necessarily. To the degree that aims are unique and uncorrelated and have different implications for behavior, they need to be specified by a modeler hoping to analyze or predict said behavior. In short, modelers need to be explicit about the aims of agents to the degree that different aims lead to significantly different behavior.

What is the aim that economics specifies? As we've seen, axiomatic preference theory says nothing about aims, but the historical aim of economics has often been taken to be the maximization of wealth. In other words, what makes an economic agent's life go economically best, according to many economic models, is the maximization of wealth. J.S. Mill's definition of political economy recognized wealth as a key variable, but also its limits as a motivating force. Political economy is, according to Mill: "The science which traces the laws of such of the phenomena of society as arise from the combined operations of mankind for the production of wealth, in so far as those phenomena are not modified by the pursuit of any other object" (J.S. Mill, 1994, p. 54). Mill was also aware that: "... the conduct of mankind in the pursuit of wealth is under the collateral influence of any other of the properties of our nature than the desire of obtaining the greatest quantity of wealth with the least labour and self-denial, the conclusions of Political Economy will so far fail of being applicable to the explanation or prediction of real events, until they are modified by a correct allowance for the degree of influence exercised by the other causes" (J.S. Mill, in Hausman, 1994, p. 54). Mill's message is that reliance on the assumption of wealth as the principal economic objective may have to be modified to account for the influence of other motives to act if it is to allow for accurate explanations and predictions.

Wealth as an aim is not as restrictive as it would appear to be at first glance. It too doesn't, for instance, exclude the possibility of other-regarding motives. Except in the case of misers who acquire wealth as an end in itself, economic agents acquire wealth as an instrument to realize, to the extent that it is possible, other objectives including, power, prestige and security. Wealth can also be pursued to improve the well-being of others such as family members and friends, but it can also be pursued to improve the lives of people unknown by way of donations to charities and other aid organizations. Wealth doesn't necessarily restrict the description of economic man to egocentric self-interest as described or presumed in many economic models, it only restricts agents to those objects that can be realized through wealth. Although wealth may contribute to better health, status, and self-respect it, beyond

a certain minimum amount, is not a necessary component to acquiring these objectives. In addition, agents can be self-interested and egocentric without necessarily being driven by wealth alone. In short, the pursuit of wealth, the historical aim of agents in economic analysis, is not equivalent to self-interest.

#### Regard for Others

One of the more important external elements is the degree to which people consider the interests of others within their own interests. I am considering regard for others to be the aim of influencing the well-being (for better or for worse) of others. Economic models of self-interest that portray agents as isolated and non-interacting will likely yield a different set of results than those in which people's interests are taken to interact. This is not a new idea in either economics or philosophy, both of which recognize that people often consider the well-being of others in their decisions (see e.g., Francis Edgeworth, Alfred Marshall, David Hume). Rather than offering what can only be a cursory review given the immense amount of research conducted in this area, I will attempt to define cases that explore the boundaries of interests among people. The false dichotomy that is often portrayed of either absolute, isolated, self-interest or self-destructive altruism isn't subtle enough to account for the many different points at which the interests of people can interact.

The emphasis in this subsection is on the manner in which people incorporate the interests of others into their own decisions. I am not claiming that the interests of others are the only consideration when making a decision, however, the approach will allow me to concentrate the discussion on demarcating the different sorts of self-interested relationships. I have mentioned some of the following examples in previous chapters, my purpose in this section is to provide a more systematic categorization of those and other examples.

a. The first case, an admittedly extreme one, says that the interests of others are of no concern; the interests of people are completely independent of one another. For such agents, the focus is exclusively on realizing their own interests which don't include the interests of others. People are completely isolated, they do not interact and they expect to remain isolated. They choose what is best for themselves without reference to the impact that their choices will have on others and the impact that the choices of others will have on their own well-being. This is the most easily modeled assumption, but clearly a poor description of any realistic



situation economic or otherwise. This assumption is perhaps a good description of Robinson Crusoe's situation before Friday's arrival and before he was aware of any potential interaction with other, hostile, invaders. Crusoe's aims probably included self-preservation and happiness although he wouldn't have had to be concerned with issues like power, esteem, or wealth, which are relative concepts. Crusoe also seems to have been concerned with the aims of living a virtuous life and self-development, although these had to take a secondary position to the goal of survival. This is an extreme situation, later in the story, even Crusoe had to be concerned with others to the degree that he expected that people might invade his island.

b. The second case says that an agent includes the interests of others only to the degree it affects their own interests, they are indifferent to the interests of others as ends in themselves. Again, the focus of this model is exclusively on the aims of the agent and not on the interests of those who are impacted by his actions. The sole focus that an agent is taken to have is on realizing his well-being. In a competitive economic setting the reasoning might run something like: My well-being is inversely dependent on your well-being, so I'll try to reduce your well-being in so far as it promotes my well-being, otherwise I am indifferent to the impact that my actions have on you. This is a strictly instrumental approach and under most definitions of ethics such behavior would be considered wrong. The model has some disturbing implications that would test the resolve of even the most fervent instrumentalists. For instance, the relative degrees of gains and losses is not important, any small gain by the agent outweighs the interests of others regardless of the size of those losses. If, for example, a person stands to gain one dollar by causing others to lose everything they own, then such behavior is not ruled out under this model. There is room for right and wrong behavior in the model, but only as it has implications for the interests of the agent. For instance, if acting against the norms of a community damages the ability of an agent to achieve his interests, for instance because people disapprove of an act and therefore refuse to provide the things that the agent needs to achieve his aims, then the norms won't be broken.

c. A third case says, once again, that an agent includes the interests of others because those interests are expected to have implications for the agent's well-being, but in this case he's no longer indifferent to the well-being of others as he was in the second case. However, it is important to clarify the focus of the agent's actions. This is a very large category and for definitional purposes, namely to exclude the

case of altruism, I have set a limit on the 'net-interest' that an agent expects to receive. For instance, if the aim is happiness, the net amount of happiness that an agent expects to receive from a given act must be positive. For example, I am willing to promote your happiness because your happiness makes me happy, but only as long as I expect to receive a net increase in my happiness from making you happy; I am not willing to accept a net loss in happiness to make you happy.

In all of the following cases an agent expects his net-happiness to increase, therefore, others are used instrumentally. To make the examples clearer I will assume that the aim of the agent is his happiness.

i. In the first case, the focus of the action is mainly on the implications for the agent's interests and not on the overall net-happiness of others. For example, I am happier because you're happier, but I am focused exclusively on my own happiness. In this case a person gives to a charity because it gives him pleasure, he's glad the beneficiaries are happier (their happiness may be the source of his happiness), but that's not the main motivating factor behind his charitable contribution. The relationships between people may now become important. If the relationship between people is valued, a person may be willing to give to those with whom he has a special relationship despite the possibility that others may have a greater need. Simply, it brings the agent the most pleasure to give to those with whom he has a special relationship (friends, family, etc.). The overall net-happiness obtained for the community as a whole is not relevant in this case, the focus is on the happiness of the agent and he acts accordingly. If relationships were not important, then if a person could get the same amount of pleasure by giving the same amount or less to someone else or some other group, then he would do it.

ii. In the next case, the focus is on the interests of the other party (or parties), but the agent's interests are also important in that he still requires a net increase in his expected happiness in order to act. For example, an agent acts primarily because he expects the act will make others happier. Although he still expects to receive a net increase in his happiness, the focus is on the happiness of others. For example, an agent gives to a charity primarily because it makes others happier, but he won't give unless it makes him happy as well. The relationship between the agent and others is less important in this case. Given a choice between conducting an act that is expected to provide an equal or greater happiness to those with whom he has a special relationship and an act that provides an equal amount

of happiness to strangers while providing him with the same expected net increase in happiness, the agent will still chose the first option, otherwise he's indifferent. In short, the relationship among agents is less important than the overall well-being achieved.

iii. In the final case, the agent attempts to maximize the total well-being of others, given the constant stipulation that his own expected net happiness increases. The relationship between the agent and others is not important, the goal is to increase overall net happiness.

d. The case of altruism can be defined as those acts in which there is an expected net reduction in the agent's well-being for an expected net increase in the well-being of others. As previously argued, this is a strong form of altruism in that it makes altruistic acts easier to identify. Following Kavka's understanding of predominant egoism, it is possible to create gradations of altruism by specifying limits of the expected net-interest loss that an agent is willing to accept. For instance, a person might be more willing to accept a net reduction in his own well-being if the payoffs to others are expected to be large than if the expected gains to others are expected to be small. The relationship between losses of an agent and gains to others is probably not linear or continuous. For example, there may be a discontinuity if the loss to an agent includes the possibility of bodily injury or death. He will not, in that case, be willing to risk bodily injury even if acting would greatly promote the well-being of others. The problem of determining net payoffs is particularly acute in this category. For instance, does a person give money to a charity because it makes the benefactors better off and even though it means he expects to suffer a net loss in happiness or because his own net happiness is increased by the good deed? I have discussed this problem at the end of chapter 2 and won't repeat it here.

e. A non-instrumental model says that the agent cares about the interests of others regardless of, or perhaps without reference to, his own well-being. In this model of behavior there is no place for a net-interest calculation. An example might be the case of a parent's actions when the well-being of his child is in danger. The parent largely acts to save the child without reference to his own well-being. Alternatively, in the case of revenge, a person might be willing to damage another person's interests regardless of the impact that such actions have on his own interests. This category must be adjusted to account for a wide range of behaviors. At some point, in most situations, some sort of net-interest calculation will

probably come into play. There are cases, revenge comes to mind as do certain core beliefs, when calculations of net-interest seem to be ignored for good or bad. However, these are the exceptions that prove the general rule that most activities involve some sort of calculation of net-benefits or losses.

f. In another extreme position, the interests of people are taken to be completely dependent on one another. In a sense, there is no difference between people's interests--people share a common interest or common mind. An agent measures his own well-being completely with reference to the sum of well-being. For instance, an agent is happy when others are happy. Both extreme positions are practical impossibilities in most contexts, however, recent work in the field of philosophy, and particularly philosophy of mind, has explored the possibility of a common mind. That work has provided a methodology that helps to clarify some of the issues linking individual agents to groups (Tuomela, 1995; Gilbert, 1989; Pettit, 1993). Pettit (1993), in particular, makes an important distinction between the influence of aggregate or structural factors and relations between individuals themselves, what he calls vertical versus horizontal factors. The emphasis in all of these models is to show how the pluralistic "we" can replace or complement the atomistic "I". Unfortunately, I will have to leave that discussion to another occasion.

## Conclusions

In the first section of this chapter I reviewed the structural dimension of self-interest. I provided several categories of reasoning or rationality that agents use to pursue their interests. There are many sorts of self-interest that can be used to model agent behavior, each with its own limits. For instance, while the parsimony of psychological egoism is appealing, the extent to which it is a good description of all motives on all occasions, as it claims to be, is doubtful. Predominant egoism has some interesting characteristics that make it a good candidate for economic models, but it is far less parsimonious than psychological egoism. It explicitly recognizes that although people may act altruistically, such behavior is taken to be the exception both across people and within a person. A researcher that chooses this description of motivation needs, at a minimum, to specify the conditions under which altruistic behavior may be present. Rational egoism, a position taken by some economists, is usually taken to rest on psychological egoism. If it doesn't, then another basis needs to be specified on which to support the claim that only egoism is rational. Although ethical egoism still finds some adherents, claiming

that an act is ethically right only if it maximizes one's self-interest hopelessly confounds the meaning of the terms ethics and self-interest. The type of egoism that best describes an agent's motives will depend on the modeler's beliefs about the role that self-interest plays in a person's psyche at a moment of decision. The message of the section on the external dimensions of self-interest is that the aims or goals of an agent are linked to the type of rationality used to achieve a set of aims. In short, the successful pursuit of different aims may influence the type of rationality an agent uses. An agent can believe that his interests are best served by pursuing his own happiness, or the fulfillment of his desires, or living his life according to an objective list of aims. General aims, in turn, may give a clue as to the rationality used by the agent to achieve his given set of aims. I use the term *may*, because, as I have previously argued, the same aim can be pursued by those with very different sorts of rationality. In short, it is necessary for a modeler to explicitly specify both the structural and external dimensions of the assumed motivation of agents. Once the dimensions are specified, it becomes possible to argue whether they are good assumptions or whether some other rationality or aims could better explain observed or predicted behavior. The final message of this chapter is that the general assumption of self-interest leaves open the possibility of a nearly endless variety of behaviors. Its value as an assumption is limited by its generality and it may need to be supported by additional behavioral assumptions.



## CHAPTER FIVE: TRACING THE FOUNDATIONS OF MODERN UNDERSTANDINGS OF SELF-INTEREST

---

### Introduction

The philosophers I discuss in this chapter set the foundations of modern theories of self-interest. Thomas Hobbes and Bernard Mandeville initiated a discussion among philosophers about the role that self-interest plays within human nature and the benefits and costs that it brings to society. Some of the basic concepts advocated by Hobbes ([1651], 1996) and Mandeville ([1732], 1988) are still recognizable in modern economic discussions of self-interest, for instance, the role that self-interest plays in making society possible and prosperous. In addition, Hobbes is often taken to promote rational egoism and Mandeville can be read as promoting psychological egoism, positions which find adherents in modern economic theory. A careful analysis of these two authors will therefore allow a better understanding of the possible sources of modern economic interpretations of self-interest. Joseph Butler ([1725], 1986, 1983), responding to the ideas of Hobbes and Mandeville, wrote a rejoinder that can only be described as brilliant; any modern theory of self-interest needs to respond to Butler's critiques. Butler's work is the focal point of this chapter and his ideas stand behind many of the concepts I have been discussing in previous chapters. His most important contribution was to place the motive of self-interest within a greater psychological framework. In short, he believed that in order to understand the role of self-interest in human nature the entire psychological framework needs to be examined. The last author whose work I will discuss is Adam Smith. Smith, in part responding to the previous authors and generally inspired and influenced by the debates on the benefits and disadvantages of self-interest for society, launched modern theoretical economics by publishing *An Inquiry into the Nature and Causes of the Wealth of Nations* (WN) in 1776. The relationship between Smith's economic theory and his ethical theory, as found in *The Theory of Moral Sentiments* (TMS), has been a source of confusion because the works appear to some readers to contain diametrically opposed descriptions of human nature and the role of self-interest therein. Some authors have incorrectly claimed that WN should be read as a defense of egoism and TMS as a defense of altruism. Others, including Ronald H. Coase whose position I will discuss at the end of this chapter, see complementary messages in the two books. In particular, Coase claims that an important message that can be drawn from the two books is support for narrow economic self-interest in markets. I believe that the approach of Coase is partially correct, the two books

should not be read in isolation and are complementary accounts of human nature. However, I will offer an alternative account of Smith's works that challenges Coase's specific arguments and suggests that the purposes of the two books were quite different and the ideas represented should only be mixed with care. Self-interest plays a role in both accounts, but the types of interests under consideration are very different.

Volumes, particularly in the cases of Hobbes and Smith, have been written about the views of the authors under discussion. Fortunately, I can concentrate my efforts on a relatively small topic, namely, my analysis will be limited to the questions of how each of the authors viewed the role of self-interest within human nature and how they derived those views. My aim is to show how an understanding of the motive of self-interest has developed from a relatively simplistic, all-encompassing, explanation of behavior, to a more complex and subtle concept existing within a wider psychological framework.

#### Bernard Mandeville

The idea that economic self-interest by individuals leads to general benefits for the wider community as a whole, via a mechanism which is sometimes called the invisible hand, is often incorrectly attributed to Adam Smith, it more accurately represents the views of the benefits of self-interest as extolled by the Dutch born Bernard Mandeville. I begin with Mandeville because his colorful book, *The Fable of the Bees*, along with the earlier but more refined works of Thomas Hobbes, elicited such a strong reaction from later writers and initiated an important and still ongoing debate. Mandeville's views are not philosophically sophisticated, there are many questions his approach leaves unanswered and his argumentation is often not strong, but he forcefully articulated a view that struck a chord with advocates and critics of a certain narrow form of self-interest. The discussion which Mandeville started eventually helped to clarify what it means to be self-interested, lessons that are still relevant today given current economic and popular conceptions of human nature. In the following paragraphs I will offer an account of Mandeville's position and brief previews of the critiques of that position which I will discuss in detail in the sections on Butler and Smith.

The thesis of Mandeville's book, *The Fable of the Bees* ([1732], 1988), is that vice, which for Mandeville is a self-interested desire for luxury and happiness, benefits society by creating a higher standard of living by offering means of employment



for people who would otherwise be idle (ed. Beauchamp, p. 20 in Hume, 1998). Mandeville's view can be summarized as: since we are wicked (self-interested) by nature in any case, we might as well accept our wickedness and be honest and thankful for the prosperity it brings. It is important to note that Mandeville wasn't advocating self-interest as a normative position--he didn't claim that it was good that people are self-interested; rather, self-interest should be seen as a necessary evil of human nature that should be accepted for what it is and celebrated for the benefits it provides.

Mandeville's evidence that people are primarily or exclusively motivated by self-interest rests on what he takes to be a realistic observation of human behavior. He tells the reader that his concern is telling things as they really are rather than as they should be. He, like Ayn Rand, is concerned that moralizers are preaching the wrong message when they tell their audiences that they should sacrifice their own interests for the interests of others, an often repeated message of dubious worth, but one worth examining because of its prevalence. Mandeville prefaced his book with the following charge directed at moralists: "One of the greatest Reasons why so few People understand themselves, is, that most Writers are always teaching Men what they should be, and hardly ever trouble their Heads with telling them what they really are" (Mandeville, 1988, p. 39). On its surface, this is an unusual charge, ethicists and those from many other disciplines including economics, do tell people what they should be or do, and I suppose that that is precisely their task. A more interesting charge, the one I believe Mandeville had in mind, would be that ethicists and others tell people that they should be something which they can't be. Mandeville, like so many before and after him, takes the empirical high ground and accordingly proclaims that people are self-interested and those who preach otherwise are naive or self-serving. In this regard, Mandeville followed in what Joseph Butler described as a secret prejudice and open scorn of all talk of public spirit and real goodwill in our fellow creatures. However, as I will show, other observers of human behavior often come to very different conclusions about what is real in human nature.

Mandeville performs the following thought experiment to demonstrate his view that self-interest is a beneficial motive: "If all actions were to cease except those due to unselfishness, the pure idea of good, or the love of God, trade would end, the arts would be unnecessary, and the crafts be almost abandoned." All these activities, (trade, arts and crafts), exist only to supply purely mundane wants, which, according to Mandeville's analysis, are all at bottom driven by selfish

desires for happiness and luxury (1988, Mandeville, p. xlviii). In short, his claim is that unethical individual behavior leads to a better material world and greater happiness for all. Accordingly, he implies that ethical behavior would imply living in a poorer material world with less happiness. These claims are unsubstantiated by Mandeville, it is not obvious, except for in the case of ascetics, how living an ethical life necessarily excludes the desire for trade, arts and crafts. It is even possible to imagine sustainable cases in which non-self-interested behavior leads to economic prosperity, although in reality such cases appear to be the exception rather than the rule. Despite such glaring defects, and perhaps because of the simple manner in which they were proclaimed, Mandeville's views meet with popular approval at the time that they were written and they formed a central point of attack in the discussions which followed.

In Mandeville it is possible to see two important themes which I have been emphasizing throughout this manuscript, the first, in his favour, is that Mandeville has explicitly defined his view about the objects of self-interest (happiness and luxury goods), and the second is that he is clear about the context over which the motive of self-interest is expected to prevail, namely, he claims that self-interest is a part of human nature in all contexts. Although his concern is explaining the benefits of self-interested behavior in economic markets, he rests his conclusions on the theory of psychological egoism applied generally.

An interesting feature of Mandeville's argument on the prevalence of self-interest, and one later addressed by Adam Smith, is the link he attempts to make between ethical behavior and self-interest via the principle of pride. He claims that moralists appeal to our pride to make us better people: "Thus Sagacious Moralists draw Men like Angels, in hopes that the Pride at least of Some will put 'em upon copying after the beautiful Originals which they are represented to be" (ed. Beauchamp, Hume, 1998, p. 52). But pride, Mandeville tells us, is in fact a form of self-interest, "we may discover no small Symptoms of Pride, and the humblest Man alive must confess, that the Reward of a Virtuous Action, which is that Satisfaction that ensues upon it, consists in a certain Pleasure he procures to himself by Contemplating on his own Worth..." (ed. Beauchamp, Hume, 1998, p. 57). So pride, the weapon that moralists use to make us more like angels is, according to Mandeville, a self-interested motive ultimately based on the desire for pleasure. Pride is an important motive in the lives of most people, but the claim that people are virtuous only because it serves their pride is too bold a claim. The argument, as I understand it, is that people seek pleasure and they receive pleasure

when they experience pride. Therefore, by appealing to pride, moralists can get people to act according to an ideal. A belief in the virtuousness of an act, according to Mandeville, offers no reason to perform the act, only hedonistic reasons are accepted as reasons to act.

There are several things that should be observed in Mandeville's reasoning on self-interest and pride. The first is that pride can be taken in performing either self-interested or non-self-interested actions. A feeling of pleasure is felt by the agent, but the action may be directed towards others and may be benevolent or altruistic. Therefore, there is nothing in Mandeville's explanation that prevents feeling pride for performing virtuous acts whether moralists approve of them or not. Second, suppose that people perform virtuous acts only for the reason of feelings of pleasure as Mandeville suggests they do. Does the feeling of pleasure detract from the virtuousness of a virtuous act? The answer to that question closely parallels previous discussions on the difficulties of properly identifying altruistic acts versus, say, pseudo-altruistic acts, and I won't repeat those discussion here in their entirety. The basic idea is that if a person performs acts that have virtuous consequences but is motivated to do so for non-virtuous motives, then such acts should be called pseudo-virtuous. Truly virtuous acts are motivated for virtuous reasons; the motives behind such acts should therefore be non-self-interested. This doesn't imply that virtuous acts have to cause displeasure in order to be virtuous, although they may. The ultimate question, one that Mandeville doesn't address, is why a person receives pleasure from performing virtuous acts. It is possible that the agent does so for the parsimonious reason that he enjoys the effects that his virtuous actions have on the well-being of others rather than for his own pleasure. Certainly there are other, perhaps easier ways to derive pleasure other than by acting virtuous?

The shortfalls in Mandeville's work shouldn't discourage reading his texts. What Mandeville lacks in rigor he more than compensates for with bravado. I include the following section from Mandeville, (ed. Beauchamp, Hume, 1998, p. 43), if only to show the righteous indignation and accompanying enjoyment he must have felt writing his book:

"To introduce, moreover, an Emulation amongst Men, they [Moralists and Philosophers of all Ages] divided the whole Species into two Classes, vastly differing from one another: The one consisted of abject, low-minded People, that always hunting after immediate Enjoyment, were wholly incapable of Self-denial, and without regard to the good of others, had no higher Aim than their private Advantage; such as being enslaved by Voluptuousness, yielded without Resistance to every gross desire, and made no use of their Rational Faculties but to heighten their Sensual Pleasure. These vile grov'ling Wretches, they said, were the Dross of their Kind, and having only the Shape of Men, differ'd from Brutes in nothing but their outward Figure. But the other Class was made up of lofty, high-spirited Creatures, that free from sordid Selfishness, esteem'd the Improvements of the Mind to be their fairest Possessions; and setting a true value upon themselves, took no Delight but in embellishing that Part in which their Excellency consisted; such as despising whatever they had in common with irrational Creatures, opposed by the Help of Reason their most violent Inclinations; and making a continual War with themselves to promote the Peace of others, aim'd at no less than the Publick Welfare and the Conquest of their own Passions."

In this quote it is possible to see the straw man argument in all its glory. The contrast Mandeville draws between the 'vile grov'ling Wretches' and 'lofty, high-spirited Creatures' is a false dichotomy, one that is not defended by any scholar that I know of, but it is also clear as well that Mandeville is reveling in hyperbole rather than engaging in a serious debate. As a philosophical treatise, Mandeville's work is weak. There is little argumentation about rival positions and his evidence that people are self-interested amounts to utterances. We would have to, as previously stated, accept Mandeville's important claim that people act virtuously only because moralists and philosophers appeal to their pride. Yet, Mandeville set the stage for an important debate which was to follow. Mandeville represents a position that does have a certain appeal in terms of its simplicity, although we have to look to Thomas Hobbes for a more thoughtful position.

Thomas Hobbes

Thomas Hobbes is interesting for the purposes of this manuscript for three reasons: first, he is often taken to defend rational egoism, a position that still finds support among some economists including, as I have shown, Meckling and Jensen. However, in contrast to many modern advocates of rational egoism, Hobbes bases his position on a clearly stated assumption, namely, that the fundamental aim of

people is self-preservation. Hobbes's starting point is that, "the natural state of man is non-moral, unregulated; moral rules are means to the end of peace, which is a means to the end of self-preservation." Human nature, according to Hobbes, needs to be viewed as ultimately deriving from the need for self-preservation. His view of human nature is closely aligned to that described by Brennan and Pettit in a previous quote, namely, economic agents as envisioned by economists are selfish, isolated and unethical. Fortunately, for the people in Hobbes's imagined world, their rationality saves them by allowing them to see the necessity of forming a powerful central government which will force them to control their base instincts. Second, Hobbes holds the position that human conduct results from whatever desire or aversion is strongest in a person at a given time (Darwall, intro. to Butler, 1986 p. 1), a position which can in some respects be likened to a naive version of desire-fulfillment or preference theory. Third, Hobbes offers a sophisticated model of human nature which, from his understanding, justified his belief in the central role of egoism within the human psyche. Hobbes, in contrast to Mandeville, took a strong position that required a sophisticated response rather than an off-handed dismissal.

Hobbes, as I have noted, is often taken to advocate a position known as rational egoism, although there is some debate on this matter (see Shaver, 1999). Rational egoism is a normative position that claims that it is necessary and sufficient for an action to be rational that it contributes to the well-being of the agent. This is an understandably controversial position that causes a nearly visceral response by opponents, as I have previously mentioned, one such opponent is Amartya Sen who writes: "Universal selfishness as actuality may well be false, but universal selfishness as a requirement of rationality is patently absurd" (Sen, 1987, p. 16). However, I believe that Sen's outright dismissal of rational egoism may be too harsh and needs to be examined. First, a small qualification needs to be made to Sen's remark. By stipulating selfishness rather than self-interest Sen has taken an unnecessarily restrictive view of rational egoism, rational egoism is based on self-interest not selfishness, but Sen's intention is clear and I'll leave this point aside. Sen's comment exposes an important link between psychological egoism and rational egoism. I believe that Sen is referring to psychological egoism when he mentions universal selfishness. He leaves open the possibility that people may be psychological egoists in which case they are, in a sense, bound to act only from the motive of self-interest--to act otherwise would be unreasonable or unnatural. Clearly, psychological egoism and rational egoism are linked; claiming that people can only act from self-interest and then claiming that people may rationally base

their actions on some other motive is inconsistent. In short, by leaving open the possibility that people may be psychological egoists I believe that Sen has left open the possibility that people can only behave as rational egoists. However, if psychological egoism is refuted, then rational egoism is likely to fall as well, because allowing motives other than self-interest to motivate actions would seem to allow the possibility that it is reasonable to act from those other motives. For instance, accepting that altruistic reasons are good reasons to act would mean that self-interest is no longer a necessary condition for a rational act. Although I would agree with Sen's comment that it does seem absurd to claim that someone should be considered rational only if they motivated by self-interest, as I will show, Hobbes's position is not absurd given his assumptions regarding human nature.

Hobbes builds his position upon empirical observation and an important thought experiment meant to describe man in a primitive state of nature before the development of institutions and other social organizations. What makes Hobbes's analysis so powerful is that the parts of his system all appear to be self-evident. His argument rests on six features, the first five of which are: an aversion to death, concerns for reputation, forward-lookingness, conflicting desires among people for limited resources, and rough equality among people (Kavka, 1986). I will concentrate on the sixth assumption of egoism, the most controversial assumption he makes. Hobbes grounds his egoism on the view that people have a fundamental right to self-preservation. According to Hobbes, people have good reason to protect, and, importantly, to do what they can to make their lives livable. Because all people are taken to be essentially equal, competition for limited resources will often lead to violent conflicts that reasonable people would like to avoid. Of course, Hobbes tells us, if someone is significantly more powerful than the rest then he would have no need to compromise and he could just take what he wanted. According to Hobbes, a reputation for aggressiveness is helpful because, among other things, it can deter potential attackers. Fortunately, because people are forward looking, they can see that there is a high probability of an unwinnable state of war of all-against-all developing. That state can be avoided if people submit to a powerful government that is able to enforce order. In short, people use their reason, in the form of enlightened self-interest, to justify to themselves that submitting to a powerful government is in their interest. Ethical concerns for Hobbes, by which I mean disinterested concerns, are merely tools of the state, there is no inherent reason to be ethical in his model.

Hobbes's argumentation doesn't always support a rational egoistic interpretation-- in some passages he argues that self-interest is sufficient to make an act rational, but not necessary for an action to be rational (see Shaver, 1999, p. 7). However, many of the phrases for which he is most famous do confirm him as a (limited) rational egoist; for instance, "A Law of Nature, (Lex Naturalis), is a Precept, or generall Rule, found out by Reason, by which a man is forbidden to do, that, which is destructive of his life, or taketh away the means of preserving the same; and to omit, that, by which he thinketh it may be best preserved" (Leviathan, 14/91). People, according to Hobbes, are forbidden by reason from doing what will detract from their self-preservation and, importantly, they are forbidden from doing what they think will detract from what best preserves their lives. The importance of this last idea is that many actions can be and have been filed under the rubric of "best preserving". For instance, best preserving has been interpreted by some rational egoists as advocating a general form of rational egoism in which any activity that improves one's life is rational. This strikes me as an overly broad interpretation, Hobbes's chief concern was the motive of self-preservation, and unless motives can be closely linked to that motive, then they are not a part of the rational domain that Hobbes was defending. For instance, it would be difficult, I believe, to link the general desire for wealth or power with the need for self-preservation. In other words, although people need a certain amount of wealth or income or some other economic aim for survival, it is difficult to argue that anything beyond that small amount is necessary for the preservation of life. However, the ambiguity in Hobbes's remarks has left open the possibility of giving "best preserving" a broad interpretation that supports rational egoism.

Another point of interest is how Hobbes arrived at his egoist view of human nature. Sidgwick (1988, p. 163), explains that it is a result of Hobbes's materialism, by which he means that for Hobbes the sensations, imaginations, thoughts and emotions of people are mere appearances of motions in the interior parts of the body. This perception leads Hobbes to give special attention to bodily wants (Sidgwick, 1988, p. 164). It is not a large step for Hobbes to then conclude (correctly or not) that appetites and desires are thereby naturally directed to preserving life and making life bearable through pleasure and the avoidance of pain. In addition, because Hobbes doesn't distinguish between reflective and instinctive desires, all motives are self-interested and directed at preserving life and making it bearable. In short, Hobbes bases his rational egoism on the belief that the material requirements of people direct their behavior. Fortunately, according to Hobbes, man's reason leads him to realize that he'd better form compacts in order

to survive with his equally egoistic, unsociable, and materialist fellow humans and, given that others are willing to submit their sovereignty to a powerful government, he too is willing to do so. Hobbes's material views strike the modern reader as odd, but they are not necessary to the rational egoist argument. In other words, a person can be a rational egoist without holding a materialist perspective, nor does holding a materialist position necessarily imply rational egoism, it just makes rational egoism appear to be a natural choice. In short, arguing against materialism doesn't necessarily argue against rational egoism--modern rational egoists, for instance, don't follow Hobbes on this point.

The natural law of preservation as outlined by Hobbes can be seen as a sort of contract that people voluntarily enter into with others. The laws require: a. That an individual should not render unreciprocated obedience to ethical rules in the interest of others, and so "make himself prey to others"; and, b. That they should refuse to observe such rules when there is insufficient security that they will be observed by others, and so "seek not peace but war" (Sidgwick, 1988, p. 167). The first point is that ethical rules are at best instrumental; a person should obey them only if it helps their cause and only if there is a strong guarantee that others will do likewise, otherwise a person is under no obligation to follow them. In fact, it would be irrational to do so because doing so would be intentionally harming one's interests. Self-preservation in Hobbes's model is prior to ethics and ethics only serves to secure self-preservation. The second point is that there are no ethical obligations to act, the only obligation that needs to be observed is doing what is best to preserve one's life. For example, if others will provide for a common good, then I am under no ethical obligation to assist.

Similar natural laws are expressed in some economic models of the relationships among agents with the key exception that Hobbes was referring to self-preservation, while economists are presumably referring to economic survival. For instance, Oliver Williamson's idea of opportunism makes essentially the same claim as the first natural law above: if we can't agree to a contract then I should assume that you are actively going to seek advantage over me. The market, for Williamson, is a place of war of all against all, agents should, accordingly, expect others to behave opportunistically. Recall that for Hobbes ethical behavior is an instrumental convenience to be followed only as a second best solution. Similarly, for Williamson, economic contracts are an expensive alternative to the market; it would be preferable to use a well regulated market and most preferable to have the capacity to exercise some sort of advantage over partners to get them to behave as



one sees fit. In a similar vein, the logic of Mancur Olson (2000) parallels and extends the second law of preservation in that it predicts that public goods will be difficult to provide in large groups where there is no possibility of coercion. Olson's main argument is that it does not *logically* follow that even when all of the individuals in a collective realize that they would gain if they were to achieve a common objective that they should pursue that objective. The essence of his argument is contained in the following two sentence: "...it is now generally understood that if the firms in an industry are maximizing profits, the profits for the industry as a whole will be less than they might otherwise be" (Olson, 2000, p.10). From which he argues: "The important point is that this [the conclusion in the previous quote] is true because, though all the firms have a common interest in a higher price for the industry's product, it is in the interest of each firm that the other firms pay the cost-in terms of the necessary reduction in output-needed to obtain a higher price" (Olson, 2000 p. 10). By the same reasoning, namely, it follows that although agents may realize that by working together they could be better off, it does not necessarily follow that cooperation will occur. Olson extends his argument to include other types of organizations that have a strong economic component, such as countries, labour unions and political parties. Economic reasoning in these models, in part, parallels Hobbes's war of all against all. Nothing is prior to one's self-interest and everything needs to serve that interest or be rejected.

To summarize, Hobbes's materialist understanding of human nature leads him to an acceptance of rational egoism. The argument is that behaviors can be traced to some physical need that requires satisfaction. It is rational to seek to fulfill those needs in order to preserve one's life and to do what one can to make it livable, all other considerations are irrational. They are irrational because any other course of action would lead to a diminishing of one's right to self-preservation. Therefore, it is necessary and sufficient that an act contribute to agent's subjective understanding of his own well-being, all other acts are irrational under this reasoning. Hobbes arrives at rational egoism via a psychological egoism based on his materialistic understanding of human nature.

Bishop Joseph Butler

Psychological egoism, if true, would make the assumption of self-interest redundant; it would simply be the case that agents are naturally, at all times, whether in economic or some other context, motivated by self-interest, they can't

reasonably act otherwise. There might still be differences between economic and non-economic self-interested motivation, but those differences can't be based on motives other than self-interest since all reasonable motives are self-interested. However, is psychological egoism a good description of the motives behind actual behavior, and if not, why not? Joseph Butler, writing a short time after Hobbes and Mandeville, mounted forceful arguments against psychological egoism and self-interest in general. C.D. Broad, perhaps slightly over-stating his case, wrote: "As a psychological theory it [psychological egoism] was killed by Butler; but it still flourishes, I believe, among bookmakers and smart young business men whose claim to know the world is based on an intimate acquaintance with the shadier side of it" (Broad, 1930, p. 55). Whether one agrees with Butler's arguments depends critically on whether his view of human nature is accepted. In the next few paragraphs I will summarize several key features of Butler's complex argument against psychological egoism, while paying special attention to the role that self-interest has in Butler's understanding of human nature.

#### Butler's Hierarchy of Motives

Butler's conceives human nature as consisting of a system composed of sub-units which, like the pieces of a watch, cannot be fully understood unless the sub-units are put together or imagined to be working together. There is nothing particularly unique about this aspect of Butler's theory, Mandeville, for instance, also imagined people to be composed of various passions. However, Mandeville viewed the governance relationship among the different motives as random in that dominance of a passion at a given moment was determined by outside forces: "... that all of them, as they are provoked and come uppermost, govern him by turns, whether he will or no" (Mandeville, 1988, [1732]) Introduction, "Enquiry Into the Origin of Moral Virtue"). Under Mandeville's system there is no room for reflection, humans are at the mercy of their various passions that are triggered by outside forces. Butler, in contrast, envisioned a hierarchy among the motives and an ability to exercise reason that allows people to reflect and adjust their behavior. For instance, in the case of reflection Butler wrote: "It is from considering the relations which the several appetites and passions in the inward frame have to each other, and, above all, the supremacy of reflection or consciousness, that we get the idea of the system or constitution of human nature" (Butler, 1986, 14, 14). The human ability to reflect upon actions means that, unlike inanimate systems and animals, people are responsible for any disorder in their "system". For Butler, in contrast to Mandeville and many others, people are not passive vehicles at the mercy of

external forces that activate and manipulate various internal passions; they have rationality or reason that allows them to reflect upon and adjust their motives. Butler's approach has obvious implications for the rationality of Hobbes's materialist understanding of human nature. Under Butler's understanding, behaviors, whether or not they are physically motivated, must pass through the filter of reflection. Behaviors therefore reflect not only immediate physical needs, but, at times, can be modified by reflection as well.

People have the rational ability to reflect, and part of what they reflect upon is their "lower" or more basic motives, what Butler calls particular motives. In Butler's system there are, roughly, rational or regulating motives and non-rational, regulated motives. The regulating, rational motives are conscience (reflection), benevolence and self-interest. The non-rational, regulated motives are desires, passions and appetites. They include such things as habits and instincts which are automatic and operate without reference to reason. They also include "animal" instincts that affect our intentions, but don't suppose any operation of reason (Sidgwick, 1988, p. 226-229). Appetites are periodical, e.g., hunger, while desires are things like power, esteem, knowledge and superiority. Taken together, desires, passions and appetites are called particular motives because they are directed at particular external objects, for example, food or persons. "All particular affections whatever, resentment, benevolence, love of arts, equally lead to a course of action for their own gratification, that is, the gratification of ourselves; and the gratification of each gives delight; so far then it is manifest they have all the same respect to private interest" (Butler, 1986, p. 132).

The importance of the non-rational, particular motives in Butler's system can't be over-stated, his insight was to recognize that they can be directed either towards self-interested or social (benevolent) objectives. According to Butler, we have particular inclinations that compel us to perform social as well as self-interested acts. For example, when we see somebody or hear of somebody in need many people experience a natural reaction to help. For instance, many feel an immediate reaction of sympathy for someone in pain. It is difficult to explain sympathy in terms of self-interest. It could be argued that we sympathize with others because we have an expected hope to receive sympathy in return when we require it, a sort of sympathy account that we can call upon when necessary. This explanation fails to explain sympathetic feelings towards people and causes with whom we can in all likelihood never expect to repay our sympathy. Take for instance the case of feelings of sympathy towards animals other than pets, no one reasonably expects to

receive repayment from baby sea lions or whales, yet our instincts are to help suffering animals. Not all sympathy can be explained by reference to the happiness or pleasure it brings because feelings of sympathy are often accompanied by feelings of sadness--when we see or hear of a person suffering a terrible loss we don't experience happiness, but sadness. The point is that it appears to be a natural instinct that motivates us to help, just as natural as those instincts that lead us to perform acts of self-interest. Of course, not everyone feels sympathy and people have different levels of sympathy towards different objects, but the same can be said for self-interest. What is important is that sympathy, like self-interest, is a natural reaction to some events for a large majority of people. From this observation, Butler argues that people are as naturally benevolent as they are self-interested. Particular affections can result in behaviors that have interested as well as disinterested outcomes.

In contrast to the particular affections, self-love (self-interest), according to Butler, is a cool, rational principle. For Butler, self-interest, as with conscience and benevolence, is a regulating motive; it is a general desire that most people have for their own happiness. People don't aim at happiness or pleasure, rather, they have particular desires, passions, appetites, etc., which are directed at external objects. Self-interest acts as a regulator of these more basic affections, passions and appetites: "Self-love does not constitute this or that to be our interest or good; but, our interest or good being constituted by nature and supposed, self-love only puts upon obtaining and securing it" (1986, p. 48, 9). Therefore, the particular motives are more basic than self-interest because if they didn't exist, then there would be nothing for the general desire of self-interest to aim at. Pleasure is not the primary aim of self-interest, but is a result of achieving the objects which are the aims of particular desires, passion and appetites. My interpretation of the foregoing is that there is nothing outside of us that we would call happiness. We don't pursue happiness, rather, our happiness is realized when our particular passion, etc. are realized. Regulated or specific motives are directed at obtaining those external objects. Thus, e.g., the object of the appetite of hunger is the eating of food, not the *pleasure* of eating it; hunger is therefore, strictly speaking, no more "interested" than benevolence. This means that our particular appetite for food to satiate our hunger is separate from our motive of enjoyment. Pleasure comes from eating the food, not the appetite for that food.

Butler offers a compelling piece of evidence to demonstrate that particular desires, passions, and appetites are not always in our self-interest; he observes that

particular motives can often make our lives go worse. He identifies the problem as, "the confusion of calling actions interested which are done in contradiction to the most manifest known interest, merely for the gratification of a present passion" (Butler, 1986 19, 35). The pursuit of external objects is no otherwise interested than every action of every creature must from the nature of the thing be; for no one can act but from a desire or choice or preference of his own (Butler, 1986, 19, 35). In this paraphrase Butler does two things, first, he points out that our particular desires can be directed at objects that are in our interest, but also at objects that are not in our interest. Second, he addresses the position known of tautological egoism. As previously mentioned, this is the idea, still heard today, that because people make a choice, i.e. prefer one thing to another, it must be their interests that are being served. This reasoning confuses the decision maker with the motivation. Simply, a person chooses what he prefers, but what he prefers isn't necessary motivated by self-interest. For example, I can be motivated by benevolence to help someone in need despite whatever impact it has on my own interests. I prefer benevolence, and choose it, but benevolence isn't necessarily in my self-interest. By not understanding the idea that self-interest is composed of many diverse motivations, authors have fallen into the trap of identifying every action as self-interested. Interpreting desires or preferences in too broad a sense so that they refer to whatever motivational feature within the agent produces an action, means that such desires and preferences become truisms without empirical content. Of course agents prefer whatever it is they choose, but their reasons for choosing what they choose are not necessarily based on their self-interest (see Butler 1986, Shaver 1999; Kavka, 1986 for a discussion of these topics). By not distinguishing between particular motives that are directed at specific objects and the general motive of self-interest which seeks happiness or pleasure, previous authors, according to Butler, confound specific desires, passions appetites with the general motive of self-interest.

Butler is not claiming that people are not motivated by self-interest, rather, he observes that people frequently act against their self-interest: "Men daily, hourly sacrifice the greatest known interest to fancy, inquisitiveness, love or hatred, any vagrant inclination. The thing to be lamented is not that men have so great a regard to their own good or interest in the present world, for they have not enough; but that they have so little to the good of others." (Butler, 1986, 21, 40). Therefore, the particular passions can be directed at objects that are good or bad for us and therefore, according to Butler, they are not necessarily in one's self-interest. In modern terms this would be identified as weakness-of-will, a concept that I have

already discussed extensively in previous chapters and won't repeat in detail here. The point is that immediate passions for food, sex, etc., can lead people to actions that are clearly not in their best interests. Subjectively, upon cool reflection, the individual may know that what they desire is not in their best long-term interest, but at the moment of action, such desires, passions and appetites can lead people to perform acts which are self-destructive.

If Butler's hierarchy of motives is accepted, then it is possible to understand how benevolent passions and self-interest can be correlated. The argument is that particular passions can work against our best interests, but they can also be benevolent or what Butler calls social or community acts. Just as people can be guided by their particular desires for food, they can also be guided by other-regarding desires. For instance, sympathetic or compassionate desires can be just as natural as the desire for food. Therefore, benevolence, according to Butler, is not necessarily incompatible with self-interest and, in fact, it may be in one's interest to be benevolent. Butler argues that a part of man's constitution, like those of 'brutes', is made-up of instincts and principles of action, those that lead to the good of the community and those that are mostly directed to private good (Butler, 1986 p. 16, 18-21). He notes: "Now there have been persons in all ages who have professed that they found satisfaction in the exercise of charity, in the love of their neighbor, in endeavoring to promote the happiness of all they had to do with, and in the pursuit of what is just and right and good, as the general bent of their mind and end of their life; and that doing an action of baseness or cruelty would be as great a violence to their self as much breaking in upon their nature as any external force" Butler, 1986, (p. 52, 15). Butler thereby concludes that social and community acts can be aligned with one's self-interest: "Therefore it is not a true representation of mankind to affirm that they are wholly governed by self-love, the love of power and sensual appetites; since, as on the one hand, there are often actuated by these, without regard to right or wrong, so, on the other, it is manifest fact that the same persons, the generality, are frequently influenced by friendship, compassion, gratitude; and even a general abhorrence of what is base, and liking of what is fair and just, takes its turn amongst the motives of action" (Butler, 1986, p. 16, 21).

My interpretation of the above quote is that people are composed of a mix of motives, some of which are called unreflective or instinctual (sensual appetites) and others which might be called intuitions (fairness and justice). The unreflective motives can be self-interested or social, implying that people are as "naturally"

social or benevolent as they are self-interested. It could be argued that Butler has simply assumed away the problem of psychological egoism, but that would be incorrect. Butler has, as I have shown, separated self-interest from particular passions, he then claims that those more basic motives can lead to self-interested acts which are at times self-destructive, but they can also lead to benevolent acts. He summarizes his claim: "The most intelligible way of speaking of it seems to be this: that self-love and the actions done in consequence of it (for these will presently appear to be the same question) are interested; that particular affections toward external objects, and the actions done in consequence of those affections, are not so" (Butler, 1986, 51, 11). Since affections toward external objects can lead us in either a good or bad direction, they are not necessarily in line with our interest.

If Butler's claims are accepted, then people are as naturally benevolent or social as they are self-interested. In addition, there is no necessary conflict between benevolence and self-interest. Butler makes a similar argument in favor of the idea that self-interest and moral behavior can be correlated. "The chief design of the eleventh discourse is to state the notion of self-love and disinterestedness, in order to show that benevolence is not more unfriendly to self-love than any other particular affection whatever" (Butler, 1986, 18, 35). His argument is that if his idea that the general motive of self-interest has a controlling function is accepted, then it is possible to conclude that self-interest doesn't exclude goodwill or love of the arts or reputation or any other particular motive. "The short of the matter is no more than this: happiness consists in the gratification of certain affections, appetites, passions, with objects which are by nature adapted to them. Self-love may indeed set us on work to gratify these, but happiness or enjoyment has no immediate connection with self-love, but arises from such gratification alone. Love of our neighbor is one of those affections. This, considered as a virtuous principle, is gratified by a consciousness of endeavoring to promote the good of others; but considered as a natural affection, its gratification consists in the actual accomplishment of this endeavor" (Butler, 1986, 53,16). Therefore, motives directed towards benevolent actions and the pursuit of the public good are as much a part of self-love and the pursuit of private good as any other particular passions and their respective pursuits. In short, there is no necessary contradiction between other-regarding motives such as goodwill and self-interest because conducting acts of goodwill can also yield pleasure and happiness.

There needs to be an authoritative voice in Butler's system that allows people to determine right from wrong behavior. Butler has shown that self-interested and social inclinations are equally natural, but he also wants to show how people decide what is the right thing to do. Butler summarizes the problem as: "Now all this licentious talk entirely goes upon a supposition that men follow their nature in the same sense, in violating the known rules of justice and honesty for the sake of a present gratification, as they do in following those rules when they have no temptation to the contrary. And if this were true, that could not be so which St. Paul asserts, that men are 'by nature a law to themselves'. If by following nature were meant only acting as we please, it would indeed be ridiculous to speak of nature as any guide in morals, nay, the very mention of deviating from nature would be absurd; and the mention of following it, when spoken by way of distinction, would absolutely have no meaning" (Butler, 1986, II; 4). The observation that we sometimes violate, for example, rules of justice, and sometimes don't, and that we know that the first is a bad thing to do and the second a good thing, means that there must be some mechanism within us telling us what is bad or good. This, according to Butler, is the principle of reflection or conscience:

"But there is a superior principle of reflection of conscience in every man which distinguishes between the internal principles of his heart [the natural disposition to kindness and compassion] as well as his external actions, which passes judgment upon himself and them, pronounces determinately some actions to be in themselves just, right, good; others to be in themselves evil, wrong, unjust, which, without being consulted, without being advised with, magisterially exerts itself, and approves or condemns him the doer of them accordingly; and which, if not forcibly stopped, naturally and always of course goes on to anticipate a higher and more effectual sentence which shall hereafter second and affirm its own. It is by this faculty, natural to man, that he is a moral agent, that he is a law to himself; by this faculty, I say, not to be considered merely as a principle in his heart, which is to have some influence as well as others, but considered as a faculty in kind and in nature supreme over all the others, and which bears its own authority of being so" (Butler, 1986, II: 8).

At this point in the argument Butler appears to have created a problem for himself in that if the faculty of reflection is supreme and directs us to do the good, why do people do bad things? In fact, if it is a superior motive, then there could be no wrong actions. Butler clearly understood the dilemma: "Man may act according to that principle or inclination which for the present happens to be strongest, and yet



act in a way disproportionate to, and violate, his real proper nature" (Butler, 1986 II: 10). "Passion of appetite implies a direct simple tendency toward such and such objects, without distinction of the means by which they are to be obtained. Consequently, it will often happen there will be a desire of particular objects in cases in which they cannot be obtained without manifest injury to others. Reflection of conscience comes in, and disapproves the pursuit of them in these circumstances; but the desire remains. Which is to be obeyed, appetite of reflection?" (Butler, 1986, II: 13). Butler uses an analogy to suggest a solution; desires, passions and appetite may on occasion win out over the good, but reflection is always the superior in kind. The distinction Butler draws is between mere power and authority and what is possible versus what is lawful. It is possible to break the law, but the law remains the law. Where does the law come from and how do we recognize it? For Butler recognition is not the problem, he claims we have an internal gauge that lets us determine right from wrong. That gauge, like the law itself, is a natural part of ourselves: "Conscience does not only offer itself to show us the way we should walk in, but it likewise carries its own authority with it; that is our natural guide, the guide assigned us by the Author of our nature; it therefore belongs to our condition of being, it is our duty to walk in that path and follow this guide, without looking about to see whether we may not possibly forsake them with impunity" (Butler, 1986 III: 5). Incidentally, it is not necessary to believe in a higher being to share Butler's viewpoint, some Intuitionists, for example, claim that some moral truths are non-inferentially known and do not rely on the existence of a god to supply that knowledge.

In his thin but dense book Butler raises several other issues of relevance to the topic of this manuscript; I will mention just a few of them that still have resonance in modern discussions of self-interest. The first relates to the form of egoism that Butlers adopts, namely, veto egoism. A complete discussion of this topic would lead me too far off-topic, so I will present an abbreviated discussion of the relevant issues. The second topic concerns the paradox of hedonism, which is an argument against an exclusive concern for self-interest which is still heard today.

Butler's view of the role of happiness or pleasure in human nature is complex but informative. In the Eleventh Sermon he writes: "Let it be allowed, though virtue or moral rectitude does indeed consist in affection to and pursuit of what is right and good, as such; yet that when we sit down in a cool hour, we can neither justify to ourselves this or any other pursuit, till we are convinced that it will be for our happiness, or at least not contrary to it." H.A. Prichard argues that Butler is

maintaining that, in the final analysis, an action is only right if it is in an agent's interest (H.A. Prichard, 2002 [1928]). Duties then, according to this interpretation, have no moral authority, only our interests make an action right. A more problematic interpretation is that even if a person knows an act is right, his interests may determine another course of behavior; an agent won't be motivated to do it unless it serves his interests. Butler has already shown that benevolent acts are not necessarily contrary to happiness and therefore a person's interests can coincide with the interests of others. However, the deciding factor in his system is which act will promote one's self-interest. Butler admits reasons for action grounded in something other than self-interest, yet gives self-interest a veto over those reasons. In short, Butler appears to be what is called a veto egoist, which is a normative position that claims that it is necessary but not sufficient for an action to be rational that it contributes to the well-being of the agent (Shaver, 1999). In other words, an act has to be motivated by self-interest to be rational, but it also needs other things as well, namely, it has to be right or good.

Butler anticipates an important paradox that Sidgwick (1981) later identifies as the Fundamental Paradox of Hedonism. This is the idea that a person can be so fixated on his own pleasure that the fixation hinders him from achieving his aim. It says that a person can be so obsessed on, say, his own happiness, that he is less happy than he would be if he had another aim. For instance, having the sole fixation of being the most popular person in a room is likely to result in the failure to achieve that aim. The result, according to Butler, is that: "People may love themselves with the most entire and unbounded affection, and yet be extremely miserable" (Butler, 1986, 48, 9). A similar, but more general idea was expressed by the Epicureans who, as I will show, taught that the pursuit of all but the most basic desires and needs only leads to frustration. Greater happiness, according to the Epicureans, can thereby be achieved by intentionally restricting one's desires. The Fundamental Paradox is a further strike against psychological egoism because it means that an exclusive reliance on the motive of self-interest can be self-defeating. In other words, it may be in one's best interest to have some other motive than self-interest.

Finally, I wish to reemphasize that Butler never claims that people are not self-interested, rather, he claims that self-interest is a rational choice (or a choice about rationality) that people make. Butler's argument is that there are a variety of particular affections, passions, and appetites that are directed towards particular external objects. These particular motives are directed at specific objects,

particular things, and are not directed at obtaining feelings of pleasure or happiness. Simply put, psychological egoism fails because it assumes that all motives are self-interested which is clearly not the case because particular motives can work against one's interests. Furthermore, there is no necessary incompatibility between self-interest and benevolence because self-interested pleasures can be derived from the pleasure of others.

### Adam Smith

In Adam Smith's work it is possible to see the refinement of the approach first articulated by Butler. Smith's view of human nature is complex and can't be understood solely in terms of self-interest, however that term is understood. I will begin this section by examining one of the more misunderstood aspects of Smith's work if not one of the most misunderstood ideas in all of economics as it is popularly understood, namely, Smith's view of self-interest and the so-called Smith problem. I will then turn to the issue of Smith's general view of human nature and the role that self-interest plays within it.

Adam Smith is often misrepresented as advocating the position that self-interested behavior by the individual leads to general benefits for all; a belief, as I have shown, that accurately reflects Mandeville's position. Smith, in fact, rejected Mandeville's restricted definition of self-interest and argued that a proper understanding of that term would undermine Mandeville's position (Fleischacker, 2004, p. 46). As I will show, Smith, a professor of moral philosophy at Glasgow University, had a rich and complex understanding of the relationship of people to one another. In addition to having an expansive view of human nature, Smith also warned against making overly simplistic assumptions and attempting to explain too much with one theory. Both Smith's view on the diversity of human nature and his warning against theories that are overly parsimonious have been largely disregarded in the economic and business literatures in regards to the topic of self-interest.

I suspect that some of the inclination towards self-interest as a default position among economists and others may be due, in part, to the false belief that economists have always made the assumption that people are by nature self-interest. The quote at the beginning of chapter 3 by Milton L. Myers emphasizes the importance that self-interest has been given in modern conceptions of economic man. However, self-interest hasn't always been taken as an absolute description of

economic man. For instance, many of the early economists were Utilitarians, the central tenant of which is the greatest good for the greatest number of people. Utilitarianism thereby implies that individual interests may in some circumstances need to be sacrificed for the good of the community at large; in other words, the interests of the community, in some cases, outweigh the interests of the individual. Utilitarianism is not without its fair share of problems, for instance, modern Utilitarians and other consequentialists are still struggling with issues such as when and for whom we should be willing to sacrifice our interests (Mulgan, 2001; Smart and Williams, 1973; Griffin, 1988).<sup>12</sup> Modern economics largely avoids these and similar issues by placing an extraordinarily heavy weight on the value of individual autonomy. Under the approach of Utilitarianism, individuals were not seen as sovereign beings existing outside society, but members of a greater whole. Smith, although generally not considered to be a Utilitarian, expressed an appreciation of at least the sacrificing aspects that society may place on an individual: "man ought to regard himself, not as something separated and detached, but as a citizen of the world, a member of the vast commonwealth of nature", and "to the interest of this great community, he ought at all times be willing that his own little interest should be sacrificed" (Smith, 1993, p. 140). Smith, as I will show, may have been influenced in this regard by the Stoics who expressed similar sentiments. Smith then, doesn't accept the very basis of Hobbes's argument that people are forbidden by reason from doing what detracts from their lives. There can, according to Smith, be good reasons to sacrifice one's interests for the good of the community.

However, there is room for different interpretations of Smith's ideas on human nature as it appears in his two great works. The "Smith Problem" has occupied scholars for some time and largely concerns attempts to reconcile the largely non-moral view of human nature found in *The Wealth of Nations* (WN), and the message found in *The Theory of Moral Sentiments* (TMS). A cursory look at the two books appears to confirm a fundamental incompatibility; after all, one of the

---

<sup>12</sup> The difficulties I have touched on here are just a small sample of the issues that consequentialists have been facing ever since Utilitarianism was first proposed as an ethical theory. While many such issues in consequentialism concern how to balance the rights of individuals with the need for the greater good, many theories of self-interest have the dubious advantage of denying that the needs of others place any demands on us at all. Still, individual egoists have to be able to find a balance among their various interests or face the consequence of failing to achieve their best interest. See, for instance, Tim Mulgan's (2001) *The Demands of Consequentialism* or Darwall (2003) for a review.

most famous quotes in all of economics is: "It is not from the benevolence of the butcher, the brewer, or the baker, that we expect our dinner, but from the regard for their own interest" (WN, 1993: 22). While the first sentence in the TMS is, "How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it" (TMS I.i.1.1). The second quote is somewhat ambiguous, people undoubtedly derive pleasure from making others happy, but the by now familiar question could be raised: Why does it make people happy to see others happy? Whose interests are being served by making others happy? As I have attempted to show on numerous occasions, this question is really at the heart of the topic of self-interest. I suggest that in the quote above Smith leaves open the possibility that there is some principle in our nature that gives us pleasure from seeing others happy which doesn't rely on self-interest. In other words, the happiness of others gives us pleasure as an end in itself and not because we derive pleasure from performing the act. I believe this is the correct interpretation because in the quote Smith seems to be contrasting selfishness and interests with something else--I claim that the something else is disinterested or non-self-interested motives. However, my interpretation is admittedly open to debate, the quote leaves open the possibility that people won't make others happy unless their own happiness is increased as well. In other words, making others happy is a self-interested, instrumental, endeavor. In addition, even accepting my view of this quote, we're still left with the problem of reconciling it with the "butcher" quote.

Samuel Fleischacker, in what he claims is the first philosophical analysis of the *Wealth of Nations*, convincingly argues that there is no Smith Problem. His analysis of the butcher quote is typical of Fleischacker's approach, which is to place Smith's ideas in their proper context. Reading Smith is a pleasure partly because of his meticulous organization. For example, at the beginning of each chapter in his books, Smith clearly specifies the structure and the main points he wishes to make. He then proceeds to meticulously discuss each point in the order in which they were originally raised in his introduction to the topic. A cursory reading of Smith can lead to problems, because he often introduces counter-examples which challenge his own views. However, if thoroughly read, there is little room to misinterpret the main point of each chapter and the reasoning Smith uses to support his position.

Keeping Smith's methodology in mind, it is possible to interpret the importance that he attaches to the story of the butcher. The first thing to notice about the quote is that no one expects the butcher, brewer or baker to provide their products out of kindness. The butcher expects to be paid for the goods and services he renders and likewise, we expect to be paid for the goods and services we provide. Only a beggar, whom Smith mentions in his discussion of the butcher, brewer and baker, expects or hopes to receive things from kindness alone--but he's the exception to the rule which Smith mentions precisely to emphasize the point that people expect to be rewarded for the goods and services they provide. Smith is addressing a very specific issue in the second chapter of the book from which the butcher quote is taken, namely, the division of labor. In the first paragraph of the chapter, Smith writes that the division of labor is not a matter of considered human design, but: "It is the necessary, though very slow and gradual consequence of a certain propensity in human nature which has in view no such extensive utility; the propensity to truck, barter, and exchange one thing for another" (WN, I.ii; p. 21). Rather than appealing to benevolence, Smith argues that people will be more successful achieving their aims if they can rally the interests of others in their favor.

Why, Smith asks, does the situation occur that we need to appeal to the self-interests of others? He tells us in the same chapter that it is not due to the natural talents of people, which: "...is, in reality, much less than we are aware of; and the very different genius which appears to distinguish men of different professions, when grown up to maturity, is not upon many occasions so much the cause, as the effect of a division of labour" (WN, I.ii; p. 23). Smith explains that the difference "between a philosopher and a common street porter arises not so much from nature, as from habit, custom and education" (WN, I.ii; p. 23-24). Just as with Hobbes, for Smith people begin with more or less the same talents but, in contrast to Hobbes, Smith believes that as they mature, habits, custom and education make each of them efficient producers in different areas. Our specialization and the fact that others are specialized as well means that we are required to truck, barter and trade in order to get the things we need and want. The critical point is that although we expect to pay and be paid for the things we buy and sell, this doesn't require us to be exclusively self-interested. The division of labor means that people have to buy and sell their goods and services in the market, and they will do so as their interests dictate. Nevertheless, there is no requirement in Smith's explanation that people have to be exclusively self-interested in all aspects of life. As I will begin to show in the next paragraph, even within the market, there is room for the expression of interests other than that which is commonly called self-interests; in what should

now be a another familiar refrain, interests need to be interpreted within a given context. For Smith, the context is clearly economic. Finally, it is important to keep in mind Smith's reasons for publishing the *Wealth of Nations*, which were to describe the general workings of a large economy and to argue against the policy of mercantilism. For those tasks he didn't need a complete model of human behavior; for that we need to examine his *Theory of Moral Sentiments*.

What motivates an individual to act in Adam Smith's system? It is clearly not just the accumulation of material goods or wealth. In the following quote Smith argues, perhaps again reflecting his Stoics influences, that once basic material needs are met, the reasons that people desire material goods shifts from one of needs to superfluities: "For to what purpose is all the toil and bustle of this world? what is the end of avarice and ambition, of the pursuit of wealth, of power, and preeminence? Is it to supply the necessities of nature? The wages of the meanest labourer can supply them. We see that they afford him food and clothing, the comfort of a house, and of a family. If we examined his oeconomy with rigour, we should find that he spends a great part of them upon conveniences, which may be regarded as superfluites, and that, upon extraordinary occasions, he can give something even to vanity and distinction" (TMS, I.iii.2.1). Smith further argues that once basic needs are fulfilled, anything beyond that amount will not make people happier or better, in fact, experience shows, according to Smith, it will make them worse off. Rather than goods or wealth, he argues the following: "It is the vanity, not the ease, or the pleasure, which interests us" (TMS, I.iii.2.1). Wealth, for Smith, is a means to acquire the attention, sympathy and approbation that in turn feeds our vanity; it is not an end in itself once basic needs are satisfied. Vanity, in turn, "is always founded upon the belief of our being the object of attention and approbation" (TMS, I.iii.2.1). Simply, we care about what others think of us, we want them to sympathize with our plight where sympathy is used by Smith to denote, "... our fellow feeling with any passion whatever" (TMS, I.i.1.5).

If I have properly understood Smith's argument, people want to acquire wealth because it feeds their vanity and vanity is a part of a need for sympathy from others. If Smith's argument is accepted, a narrow, economic interpretation of self-interest as an isolated individual in pursuit of wealth is a poor description of human motivation. In the terms defined by Broad, sympathy is a relative concept and its value requires the presence of others. This is because sympathy, according to Smith, affects both the giver and receiver; it is not a one-way transaction, a person benefits or suffers because others benefit or suffer. A person cannot sympathize

alone or with an imaginary person, sympathy requires some sort of connection with another person. Through sympathy, "we enter as it were into his body, and become in some measure the same person with him, and thence form some idea of his sensations, and even feel something which, though weaker in degree, is not altogether unlike them" (I.i.1.2.). "Sympathy, however, cannot, in any sense, be regarded as a selfish principle. When I sympathize with your sorrow or your indignation, it may be pretended, indeed, that my emotion is founded in self-love, because it arises from bringing your case home to myself, from putting myself in your stances. But though sympathy is very properly said to arise from an imaginary change of situations with the person principally concerned, yet this imaginary change is not supposed to happen to me in my own person and character, but in that of the person with whom I sympathize" (TMS, VII.iii.1.4). Smith's theory of sympathy, his view of human nature, is clearly expressing an entirely different sentiment than a simple reliance on a self-interested pursuit of wealth and happiness.

Perhaps sympathy plays the same role that pride plays in Mandeville's scheme, namely, perhaps it is just disguised self-interest? Smith explicitly argues against this interpretation, speaking of "Those who are fond of deducing all our sentiments from certain refinements of self-love..." (I.i.2.1), he writes, "Man, say they, conscious of his own weakness, and of the need which he has for the assistance of others, rejoices whenever he observes that they adopt his own passions, because he is then assured of that assistance; and grieves whenever he observes the contrary, because he is then assured of their opposition" (I.i.2.1). Smith is here arguing against the instrumental position that we have a positive response to the reactions of others because we desire their support or a negative response because they oppose us. His argument against this understanding of sympathy is: "... pleasure and pain is felt so instantaneously, and often upon such frivolous occasions, that it seems evident that neither of them can be derived from any such self-interested consideration" (I.i.2.1). For Smith, our instantaneous reactions, (the desires, passions and appetites of Butler) represent our true response, not a response due to a calculation of the net benefits or harms that the reactions of others perhaps imply. While this argument isn't entirely convincing because it doesn't account for the fact that many of our reactions are learned responses or habits, to the degree that instantaneous reactions represent unchecked motives, they do indicate an ability to sympathize with both the plights and pleasures of others. Although some of our sympathetic responses are undoubtedly learned responses, some may not be. For instance, very young children often appear to sympathize with the sufferings of



others, well before they have been taught to do so. Individuals often react in extraordinary circumstances to what they perceive as acts of injustice or when heroic acts are called for. The point, for the purposes of this manuscript, is that while there is room in Smith's system for instrumental self-interest, there is also a place for selfless acts.

Like all of the authors I have discussed, Smith believed that people are motivated by both self-interest and benevolence. Self-interest plays an important although limited role in Smith's system. To explain its role Smith employs an approach to which I alluded to before, a methodology which can cause some problems for unwary readers, namely, he often describes an ideal (read unlikely) state and then presents a more realistic position which he in fact advocates. Using that methodology Smith first describes a world based purely on beneficence, "It is thus that man, who can subsist only in society, was fitted by nature to that situation for which he was made. All the members of human society stand in need of each other's assistance, and are likewise exposed to mutual injuries. Where the necessary assistance is reciprocally afforded from love, from gratitude, from friendship, and esteem, the society flourishes and is happy" (TMS, II.ii.3.1). So, contrary to Hobbes and Mandeville, Smith argues that, at least in an ideal state, benevolence could provide the means to a flourishing society. However, Smith explains that even a less than ideal state, one without the guarantee of mutual love and affection, will not necessarily collapse. This comes about because, according to Smith as with Hobbes, members of a society realize that for their own well-being they are dependent on each other. Where people in Hobbes's system rationally concluded that they have to submit their will to an all powerful state in order to escape mutual destruction, people in Smith's system use their rational faculties and decide that they have to depend on one another to prevent society from collapsing. Enlightened self-interest allows people to conclude that they require the cooperation of others in order to fulfill their own aims. While narrow self-interest can be overcome through enlightened self-interest, misplaced self-interest can be destructive: "Society, however, cannot subsist among those who are at all time ready to hurt and injure one another" (TMS, II.ii.3.3). It is not beneficence that holds the society together, that feeling is too weak, but justice that, "is the main pillar that upholds the whole edifice" (TMS, II.ii.3.4). Our naturally given conscience drives us to seek justice, a drive that compels us to take care of the needs of others. For Smith, the need of people to be worthy of sympathy drives them to seek their fortune, while justice keeps them from harming others upon pain

of conscience. Human conscience, as in Butler's model, plays a reflective role allowing people to choose between good and bad actions.

Unmitigated self-interest is contained by a stronger force: "It is a stronger power, a more forcible motive, which exerts itself upon such occasions [of restraint]. It is reason, principle, conscience, the inhabitant of the breast, the man within, the great judge and arbiter of our conduct. It is he who, whenever we are about to act so as to affect the happiness of others, calls to us, with a voice capable of astonishing the most presumptuous of our passions, that we are but one of the multitude, in no respect better than any other in it; and that when we prefer ourselves so shamelessly and so blindly to others, we become the object of resentment, abhorrence, and execration" (Smith, TMS, III.iii.3.4). Smith believed that people are motivated by both self-interest and benevolence and that they are capable of reflecting upon and controlling their actions and taking account of the feelings of others through, in Smith's case, the faculty of sympathy. Smith, like Butler, saw a hierarchy among the motives. For Smith it is the intellectual virtue of prudence through which one judges the appropriateness of an action. Finally, Smith believed that an over-reliance on self-interest could lead one to go astray of what is in his best enlightened interest.

How then is it possible to account for the idea of the invisible hand? The invisible hand is a powerful, gripping image; it is the idea that it is economically good that everyone pursues their own narrow self-interests because doing so has the unintentional side-effect of increasing the well-being of the group as a whole. The results of the invisible hand are not a good from the subjective perspective of any person--people don't pursue the good of the group as an aim--but good results as viewed by an objective observer. The motives of people in capital markets are focused on increasing their own wealth: "Every individual is continually exerting himself to find out the most advantageous employment for whatever capital he can command. It is his own advantage, indeed, and not that of the society, which he has in view" (WN, IV.ii: p. 289). Smith repeats the claim that in the capital market people are motivated by wealth: "But it is only for the sake of profit that any man employs a capital in the support of industry; and he will always, therefore, endeavour to employ it in the support of that industry of which the produce is likely to be of the greatest value..." (WN, IV.iii: p. 291). We're not surprised when the butcher asks us to pay for his products, likewise, we shouldn't be surprised that when investors choose to attempt to maximize their profits. Profits are a part of one's self-interest, not the only thing, and certainly not the major aim of leading a

good life if we take TMS as a guide to Smith's understanding of human nature. Smith's point expressed in the above quotes is the simple idea, just as it was with the example of the butcher in the meat market, that people try to do what is best for themselves in the narrow confines of the marketplace--no one, except perhaps a beggar, expects any other sort of motivation in the market. I will provide evidence in the eighth chapter that raises doubts even about this last claim; an economic context doesn't necessarily exclude the possibility of benevolent motivation. Smith recognized, like Butler, that the pursuit of private advantages can often lead to outcomes that coincide with what is advantageous for the group as a whole. Often, but not always, a person, by pursuing what is in their best interest promotes the interests of the group without intending to do so; the invisible hand guides him to do what is potentially good for the group. While the invisible hand can yield benefits for the group as a whole, Smith also recognized that it can also lead to decreases in the well-being of the group (WN, Smith, II.ii). That the invisible hand is not always benevolent and sometimes malevolent should be obvious. Modern markets are highly regulated entities that curtail the interests of individuals for the benefit of the group as a whole. Smith often expresses beliefs that go against the idea of giving self-interest carte blanche, "The wise and virtuous man is at all times willing that his own private interest should be sacrificed to the public interest of his own particular order or society" (TMS, p. 235).

An important feature of Smith's argument for the purposes of my manuscript is that when he refers to the butcher or the invisible hand, he is referring to a specific economic context. He is not, as is so often the case modern economics, implying that such reasoning extends to non-market situations. Economics, for Smith, is a clearly delineated aspect of human life, it is not, as in much modern economic theory, any situation in which resources, broadly defined, are used to fulfill human needs. *The Wealth of Nations* concerns a specific context in which specific, narrow, self-interested motives are appropriate, while *The Theory of Moral Sentiments* is a more general account of how people arrive at ethical judgments of their own and other's behavior.

#### Ronald H. Coase's Analysis of *The Theory of Moral Sentiments*

Ronald Coase's (1994 [1976]) analysis of what he takes to be Adam Smith's beliefs about human nature and the role that self-interest plays in that nature is typical of Coase in that it is incisive and novel; however, I believe it maintains a bias towards self-interest that is not contained in either of Smith's two books. Coase draws two

conclusions from his analysis of *The Theory of Moral Sentiments* (TMS); the first is that there is no essential difference between the views of human nature expressed in TMS and those found in *The Wealth of Nations* (WN). His second conclusion is that Smith's views of human nature as expressed in TMS strengthen the advantages of the market. Coase concludes, correctly I believe, that self-interest is "everywhere evident", but that is not a unique observation, Smith, like Hobbes, Mandeville, Butler and virtually every other scholar recognizes that self-interest is everywhere or nearly everywhere present (1994, p. 111). However, not all interests are necessarily reducible to economic interests. I will explore this basic idea as I discuss each of Coase's arguments.

#### Self-interest in TMS and WN

Although I believe that Coase is correct in concluding that self-interest is an important element in both TMS and WN, I believe that in the case of TMS, Coase has, despite his numerous qualifications, over-emphasized its importance. Nowhere is that over-emphasis more apparent than in Coase's analysis of Smith's key concept of sympathy. I will address some of the limits of Coase's analysis, but before doing that, it is important to understand the general difficulties of comparing TMS and WN.

That self-interest plays an important role in both books is not surprising given the importance self-interest in human psychology; I have expended quite some effort in the preceding chapters demonstrating the prevalence of that motive. Therefore, it shouldn't be surprising that whether the issue is explaining how people form moral judgments, as is the purpose of TMS, or explaining the functioning of an economy, the main purpose of WN, self-interest is an important component of Smith's analysis. However, the two very different purposes for which the books were written means that applying the conclusions of one book to the other must be done with care. There are several methods that could be used to bring together the findings of both books. One method would be to argue that moral judgments play an important role in the functioning of an economy.

Although Smith doesn't address the influence of moral judgments on economic activity in WN, he does discuss attempts to set a limit on the economic benefits of self-interest in TMS. The overall message is that a proper regard for one's economic self-interest is expected; self-interest in terms of profits motivates economic agents, but: "To be anxious, or to be laying a plot either to gain or to

save a single shilling, would degrade the most vulgar tradesman in the opinion of all his neighbors." A proper regard for one's interests is expected and respected, but needs to be controlled: "Those great objects of self-interest, of which the loss or acquisition quite changes the rank of the person, are the objects of the passion properly called ambition; a passion, which when it keeps within the bounds of prudence and justice, is always admired in the world, and has even sometimes a certain irregular greatness, which dazzles the imagination, when it passes the limits of both these virtues, and is not only unjust but extravagant" (WN, p. 173). Self-interest, within bounds, is clearly an important element of WN, but what isn't clear is the degree to which moral judgments are a part of the functioning of an economy; in other words, it is not clear how moral judgments influence economic behavior and drawing any such conclusions from either text on this issue would be speculative. Coase uses another approach to draw his conclusion that self-interest plays an important role in both TMS and WN. He correctly takes for granted that it is an integral part of WN, but then argues that it is an important element in making moral judgments. That conclusion is controversial and one that needs to be defended; I do not believe Coase has succeeded.

Adam Smith wrote the TMS in order to explain how people arrive at moral judgments; the book concerns the issue of how we judge our own behavior and that of others. Sympathy is an important component in Smith's approach where it is defined as the sharing of any fellow-feeling with any passion whatever (TMS, p. 10). Sympathy is not confined to benevolence, we can, for instance, share feelings of anger. In short, TMS is not exclusively about benevolence or altruism. According to Smith, our imagination allows us to put ourselves in the place of others, wherever that place may be. The match between people with whom we sympathize can never be perfect, our understanding of the motives of others is limited and it follows that our ability to sympathize, to imagine ourselves in the place of others, is correspondingly limited. Sympathy "enlivens joy and alleviates grief" (TMS 14). It enlivens joy because we are pleased when others share our feelings, and it alleviates our grief by allowing us to share the burden of our misfortunes with others. A central component of Coase's (1994 [1976], p. 100) argument of the prevalence self-interest in TMS is his contention that Smith argues that people form their moral sentiments for self-interested purposes: "It will be observed that Adam Smith's account of the development of our moral sentiments is essentially self-centered. We care for others because, by a sympathetic response, we feel as they feel, because we enjoy the sharing of sympathy, and because we wish to appear admirable in our own eyes; and we conform to the rules of conduct

accepted in society largely because we wish to be admired by others." Coase is only partially correct, although at first glance his claim seems entirely reasonable given some of Smith's comments. For instance, at the beginning of the third part of his discourse, Smith (TMS, p. 109-110) summarizes what he's done to that point in his book:

"The principle by which we naturally either approve or disapprove of our own conduct, seems to be altogether the same with that by which we exercise the like judgments concerning the conduct of other people. We either approve or disapprove of the conduct of another man according as we feel that, when we bring his case home to ourselves, we either can or cannot entirely sympathize with the sentiments and motives which directed it. And, in the same manner, we either approve or disapprove of our own conduct, according as we feel that, when we place ourselves in the situation of another man, and view it, as it were, with his eyes and from his station, we either can or cannot entirely enter into and sympathize with the sentiments and motives which influence it. We can never survey our own sentiments and motives, we can never form any judgment concerning them; unless we remove ourselves, as it were, from our own natural station, and endeavour to view them as at a certain distance from us. But we can do this in no other way than by endeavouring to view them with the eyes of other people, or as other people are likely to view them."

By observing the conduct of others we acquire an understanding of that conduct with which we sympathize and which we can then use to scrutinize our own behavior. By means of imagination we learn, as it were, to disengage ourselves from our actions and observe those actions from an impartial perspective. However, Coase has only partially followed Smith's argument when he claims that "we conform to the rules of conduct accepted in society largely because we wish to be admired by others." While it is true that Smith does argue that our conscience is a product of social relationships, it is more than that. Our behavior is judged by others, but we have an ability to reflect upon our behavior as well. Reflection is the capacity to see our actions not from the perspective of others, but from a detached, idealized person within. Neither view is perfect, the view of an external person is obscured by a lack of information about someone's true motives, while our internal view can be obscured by self-deceit in which we see ourselves in a better light than reality judges us.

Although Coase and others can be forgiven for interpreting the beginning of the third part of Smith's discourse as an argument for the position that we form or adjust our sentiments based on how our conduct will be interpreted by others, it is more difficult to understand how later sections of that chapter can be similarly misinterpreted. For instance, on (TMS p. 116-117) Smith argues:

"Nature, when she formed man for society, endowed him with an original desire to please, and an original aversion to offend his brethren. ... But this desire of the approbation, and this aversion to the disapprobation of his brethren, would not alone have rendered him fit for that society for which he was made. Nature, accordingly, has endowed him, not only with a desire of being approved of, but with a desire of being what ought to be approved of; or of being what he himself approves of in other men. The second was necessary in order to render him anxious to be really fit. The first could only have prompted him to the affection of virtue, and to the concealment of vice. The second was necessary in order to inspire him with real love of virtue, and with the real abhorrence of vice. In every well-formed mind this second desire seems to be the strongest of the two."

And later on the same page Smith writes: "To desire, or even to accept of praise, where no praise is due, can be the effect only of the most contemptible vanity." The wise man, "stands in need of no confirmation from the approbation of other men. It is alone sufficient, and he is content with it. This self-approbation, if not the only, is at least the principal object, about which he can or ought to be anxious. The love of it, is the love of virtue."

Smith explicitly moved away from the view that conscience is formed exclusively from popular opinion in later editions of his book because he was acutely aware that if conscience was taken to be only a product of society, then it couldn't be better or worse than that of given society at a point in time. Rather, it is Smith's view that it is our imagination that allows us to build a more objective, impartial view of our behavior. Therefore, while the judgment of the real, outside spectator depends on the desire for actual praise, that of the impartial spectator depends on the desire for praiseworthiness (Raphael, 1976, p. 16). According to Smith, enlightened people seek to be praiseworthy rather than praise: "Some splenetic philosophers, in judging of human nature, have done as peevish individuals are apt to do in judging of the conduct of one another, and have imputed to the love of praise, or to what they call vanity, every action which ought to be ascribed to that of praise-worthiness" (TMS p. 127). The two can be in concert, but, Smith argues,

people are unsatisfied if they are praised for unpraiseworthy intentions and actions. Likewise, praiseworthy intentions can be misinterpreted. (TMS, p 130) "But though man has, in this manner [from the "man without"], been rendered the immediate judge of mankind, he has been rendered so only in the first instance; and an appeal lies from his sentence to a much higher tribunal, to the tribunal of their own consciences, to that of the supposed impartial and well-informed spectator, to that of the man within the breast, the great judge and arbiter of their conduct."

Coase's argument that a person shows sympathy because it gives them pleasure to perform virtuous acts is also at best partially correct and applies more to David Hume's position than Smith's. Smith's argument is that people derive pleasure from the observation that there is correspondence between their feelings and the feelings of others. We're happy when people share our pleasures and displeasures. However, Smith clearly states, and it is certainly true, that feelings of sympathy can be either pleasurable or painful. According to Smith, we first respond to the pleasure or pain of the situation, and then we feel the pleasure of discovering that our sentiments coincide. In short, the first reaction can be either pleasure or pain so our sentiments are not necessarily triggered by feelings of pleasure.

I end this section with a final quote from Smith. In it, Smith clearly states that sympathy is not based on a regard for one's selfish desires (TMS p. 317).

"Sympathy, however, cannot, in any sense, be regarded as a selfish principle. ... When I condole with you for the loss of your only son, in order to enter into your grief I do not consider what I, a person of such a character and profession, should suffer, if I had a son, and if that son was unfortunately to die: but I consider what I should suffer if I was really you, and I not only change circumstances with you, but I change persons and character. My grief, therefore, is entirely upon your account, and not in the least upon my own. It is not, therefore, in the least selfish. ... That whole account of human nature, however, which deduces all sentiment and affections to self-love, which has made so much noise in the world, but which, so far as I know, has never yet been fully and distinctly explained, seems to me to have arisen from confused misapprehension of the system of sympathy."



### Self-interest and the Advantages of the Market

Coase's second conclusion is that the arguments in TMS strengthen the advantages of market. Markets, according to Coase, function primarily because people are self-interested, those interests: "...offset the weakness and partiality of benevolence, so that those who are unknown, unattractive, or unimportant will have their wants served" (p. 115). This strikes me as a very cynical conclusion for three reasons. The first is that it rests on the belief that benevolence is a partial motive; people are only benevolent towards those they know, find attractive and important. Coase has chosen a non-standard definition of benevolence which is usually understood to be a general desire to perform charitable acts. Benevolence would have us treat everyone, no matter what his characteristics, equally. The second reason is that it assumes markets are impartial in their consequences for agents. To the degree that market prices approach the ideal of being sufficient statistics, I suspect Coase is correct. In those situation prices contain all of the information necessary for an agent to decide whether to buy or not; buyers and sellers have no reason to interact other than through the price so there is no room for prejudices. If this is the situation which Coase has in mind, then he is closely in line with the conclusions found in Albert O. Hirschman's (1997) justly famous *The Passions and the Interests*. However, to the degree that prices are not sufficient statistics, I would suggest that the motive of self-interest is just as likely to admit prejudices as benevolence. The final reason that I find Coase's explanation unconvincing is more fundamental. Coase appears to believe that TMS supports benevolence alone, however, Smith was adamant in his belief that sympathy is the sharing of any fellow-feeling and is not confined to benevolent or altruistic feelings. Therefore, Coase may be correct that markets overcome the prejudices of benevolence, but benevolence is not the only reason people sympathize with others and benevolence is not the message of TMS. Coase's conclusion may be correct, but his reasoning is wrong.



## CHAPTER SIX: SELF-INTEREST IN HELLENISTIC PHILOSOPHIES

---

### Introduction

How can studying the systems of the Hellenistic philosophers help to clarify the two general goals that I have set for myself in this manuscript? Namely, how can they help to clarify the dimensions of self-interest and how do their various approaches help in building a coherent, systematic framework of economic self-interest? I'll first offer a general answer to these questions before presenting specific examples from Hellenistic philosophers. Two widely accepted reasons to develop economic models are that they allow researchers to make more accurate predictions about economic events and that they provide descriptive insights into the workings of an economy. The two key economic assumptions, self-interest and rationality, are presumably meant to add to the predictive and explanatory capacities of such models. In previous chapters I have argued that the economic understanding of self-interest is underdeveloped, while the extent of economic rationality has long been a topic of debate. In this chapter I attempt to narrow in on a better understanding of what an effective, practical, definition of self-interest should entail. By examining historical examples of models that assumed self-interest and rationality I can, in a loose sense, "test" the effectiveness of these assumptions in applied models. In other words, I can examine how the assumptions of self-interest and rationality have been used previously in models of human behavior and practice, and assess how they affect the explanatory and descriptive capacities of those models. Taking a narrow definition of self-interest, as do the Hellenistic philosophers, it is possible to ask whether self-interest, together with the assumption of rationality, provides the basis of a model of human nature that helps to predict and describe behavior.

There are several particular reasons that recommend the Hellenistic philosophers for helping to answer the general question posed above. The Cyrenaics, Epicureans and Stoics were all hedonists and supremely confident in the rationality of individuals to understand, with effort, what it takes to live a good life. The initiators of each of these philosophical positions all lived at around the same time in the city of Athens and they or their advocates competed with one another for adherents. In a sense, Athens was a market place of ideas in which these and other approaches competed for a limited audience. As a direct consequence of the interactions and reactions of the various positions, it is possible to see progress in the understanding of what it means to be self-interested. All of these groups were

concerned with how to live in order to have a happy life as individuals rather than being good parts of a greater social order. Anyone, from any social class including slaves, could follow, for example, the Stoic ideals of a good life. In addition, these philosophers actually lived the lives that they recommended to others and disseminated their views in language that was clear and precise. For instance, Epicurus, probably in response to earlier, idealistic, philosophical systems, asserted that words should be used in their immediate and most obvious sense (Sharples, 1996, p.6). A review of the Hellenistic philosophers gives a good indication of the basic characteristics that a theory of self-interest should address. Their views on what it means to be self-interested anticipate many of the modern debates or potential debates on that topic. The strong emphasis placed on the individual and a common nature shared with everyone is characteristic of these philosophers, and, I would add, is apparent in some of the writings of Joseph Butler and Adam Smith (see Long, 1986 2nd ed.). The Hellenistic writers also emphasized the power of rationality and the idea that inner resources alone can provide the foundations for a happy, stable life. A proper discussion of the entire systems of each of these philosophical approaches is not possible given limited space, therefore, I will limit my discussion to two topics directly related to the main research question of my manuscript. The first topic is an examination of the various Hellenistic definitions of self-interest or happiness given that all of the systems were based on hedonism. The second is an examination of how these philosophers arrived at their various definitions of self-interest; what, in other words, were their assumptions about human nature. I will discuss each school in the order of their historical appearance in Athens.

### The Cyrenaics

The Cyrenaics were very clear in regards to what I have labeled the external and structural dimensions of their definition self-interest. They, along with the other Hellenistic philosophers I will discuss, were hedonists, meaning that they believed that the pursuit of happiness or pleasure provides a good reason to act as opposed to some other aim such as the acquisition of knowledge or pursuit of justice, although pursuing one doesn't necessarily exclude pursuing the other. The focus of their behavior was egoistic and centered on promoting their own happiness and not, for instance, promoting the overall level of happiness in the wider community. However, the Cyrenaics had a very specific form of happiness in mind. Aristippus of Cyrene (c. 435-355 B.C.E) is famous for making the intriguing claim that a state of pleasure, specifically bodily pleasure, is the goal of life. While both the external

and structural dimensions of the Cyrenaic system were clear, an important and telling question that needs to be answered is what prompted the Cyrenaics to conclude that bodily pleasure should be the goal of life.

The Cyrenaics were a loosely knit group of philosophers who taught in Athens. They didn't start a school, such as Socrates, Aristotle, the Epicureans or the Stoics, and their teachings don't form a uniform body of work. However, although there were many splits and sects among the Cyrenaics, they all agreed that the supreme goal of life was bodily pleasure: not the accumulation of pleasure over a lifetime, but the bodily pleasure that a person experiences at present (Tsouna, 1998: 15). In short, they accepted a rather unusual form of hedonism in that they believed that people should follow their immediate desires. By immediate I mean that they did not consider the historical causes or future implications of their acts, only their current desires were of importance; clearly, the issue of weakness-of-will wouldn't present a problem for the Cyrenaic approach. However, because of their modest approach to life, the desire for immediate satisfaction didn't lead to an intensely happy, but short life. Their approach may appear strange, even irrational by today's standards, but they had what they believed were good reasons for their beliefs. Given their overall epistemology, which I briefly outline below, they made reasonable conclusions about how to lead a good life. The Cyrenaics were hedonists whose view of the best manner to live one's life can be summarized as, live for the moment, but live modestly. How did they arrive at that view?

The Cyrenaics were the intellectual heirs of the Sophists (Watson, 1895). The Sophists, through rhetoric and casuistry, challenged popular Athenian belief in divine authority and were critical of the unchangeability of morality. They based their challenges on the diversity of customs that they observed while teaching in various Greek city-states. Laws, customs and morality varied across these cities, and appeared to the Sophists to be based on conventions that were agreed upon for mutual advantage; Hobbes and many others would undoubtedly agree with this part of their analysis. The newly found mobility of these scholars meant morality and religious beliefs were not taken as absolutes but, for the first time, open to comparison and critique. The diversity of morality and religion which they observed led the Sophists to the belief that knowledge was impossible and that morality is nothing but laws imposed by rulers on others for their own selfish needs. According to the Sophists, there were no absolute rules of right and wrong because right and wrong couldn't be known. Other conclusions could obviously have been drawn from what the Sophists observed, for instance, that knowledge is

a relative concept, but the Sophist reasoned that knowledge itself was impossible. They didn't take the next step and build a new model of human behavior based on their observations, rather, it was the Cyrenaics who made that leap by assuming that there had to be one single end to which all human actions refer.

Both the Sophists and Cyrenaics recognized that the societies of Greek city-states were changing due in part to the increased mobility of people; individuals too were believed to be constantly in a state of flux. The epistemological framework of the Cyrenaics held that individual experiences are isolated in the sense that what one person felt couldn't be entirely understood by anyone else or even necessarily by the same person at another time in their lives. To take a common example, what feels hot to me may feel cold to you and may even feel cold to me at another time, what is hot or cold depends on the context. Our sensations, then, determine what is real at any given moment. Aristippus therefore argued that what we call knowledge is reducible to the immediate convictions or feelings of the individual at the time it is experienced. His argument rests on the belief that language is unable to convey general ideas. I cannot, for instance, prove to you or vice-versa that we both mean the same thing when we describe how we feel about something. We can both agree that a pot of coffee tastes bitter, but we cannot prove to one another that we are referring to the same thing when we use the term bitter. Since we cannot understand the feelings and experiences of others, it is impossible, according to Aristippus, to study nature. There are some obvious problems with the Cyrenaic idea of knowledge, for example, although some sensations like color, taste, sound, etc. are relative, other properties, like weight, length and motion are objective. The problems in the Cyrenaic approach won't be addressed here, it is their solution to what they saw as the limits of knowledge that are interesting for the purposes of this manuscript.

The rejection of empirical knowledge and of knowledge of other minds caused the Cyrenaics to look internally in search of certainty. The solution that Aristippus arrived at was to reduce knowledge to each person's immediate understanding of his own feelings of pleasure. In other words, the Cyrenaics concluded that hedonism was the one thing at which we can all aim because it is the only thing we can know, but what sort of pleasure? They distinguished between mental pleasures that ultimately depend on bodily pleasures and those that don't, and believed that pleasures having to do with the body pertain exclusively to oneself, in contrast to say pleasures that are independent of the body and which focus on objects other than one's self. An example of pleasure derived from an object according to

Diogenes Laertius is when, "... we feel joy merely on account of the prosperity of our country as if it were our own prosperity" (Tsouna, 1998: 15--quoting Diogenes Laertius). In order to maintain the certainty that they were looking for, the Cyrenaics expressly omitted the feelings of pleasure derived from external objects and meant by pleasure only those feelings which are personally experienced. In short, they argued that people are unable to understand the feelings of others or the nature of things and their solution was simply to fall back onto the one thing that we can all know, namely, our own experiences. The experiences we want are, according to the Cyrenaics, pleasant experiences and so maximizing the pleasant, personal, experiences should be the goal of life.

In contrast to the Epicureans, a school to which I will soon turn, the Cyrenaics claimed that the absence of pain cannot be a part of the moral end for the reason that pleasure and pains are related to motions and since the absence of pain is not associated with movement it cannot be a part of the moral end (Tsouna, 1985: p. 13). Movement is associated with an "active" engagement with a sensation, as opposed to a passive or indifferent state. The basic idea is that the pleasure that people experience is unrelated to other times present or future: it has no prospective or retrospective value, and can only be enjoyed while it is actually occurring. If we accept the Cyrenaic view, then we can ask, what is the best way to obtain maximum happiness? This is where the Cyrenaic view takes what from a modern perspective is an unusual but, in light of their empirical stance, reasonable twist. We aim, they claim, not at our overall happiness through time, but at those things which we know for sure, namely, bodily pleasures which are immediately apparent. Their message was to forget the past because it is gone and forget the future because it cannot be accurately predicted. Immediate pleasure is always and for everyone a good and always desirable.

There are two things to note about the Cyrenaic's approach. The first is that individualization in the Cyrenaic model is explicit and absolute; others are to be used as instruments to promote one's pleasure. There are obvious ethical implications for the Cyrenaic thesis; for example, a person shouldn't steal or lie, not because it is intuitively or rationally wrong, but only because it would or could detract from their pleasure if they were caught and punished. The Cyrenaic position towards others is an example of the perfectly instrumental approach I discussed in chapter 4 as it applies to the aim of happiness. Actions are only motivated by the desire to attain agreeable feelings and others are mere instruments useful in achieving those feelings. The second thing to note is that the Cyrenaics

posited pleasure as the supreme good of human life (see Tsouna, 1998). Positing pleasure as a supreme goal is a problematic approach for the reasons I have mentioned in previous chapters, but what makes their position particularly interesting is that they were not motivated by the accumulation of pleasure over a lifetime, but immediate bodily pleasures. They didn't, in other words, consider their actions in light of their overall happiness, but only in light of their immediate interests. Their reasons for doing so may appear odd to us, but that is precisely what makes it such an interesting approach for the purposes of this manuscript in that it emphasizes the importance of understanding and making clear one's underlying assumptions about human nature. The Cyrenaics had good reasons for promoting their own interests, which they, at least, found convincing. By explicitly stating their assumption we are in a position to argue the merits of their approach.

#### Epicurus

"The beginning and root of every good is the pleasure of the stomach. Even wisdom and refinements are referable to this" (Uncertain Fragments, 59 in Inwood, Gerson and Hutchison, 1994). In modern terminology an epicurean is someone devoted to the pursuit of pleasure, he is someone with refined tastes especially for food. The modern understanding of the term epicurean is ironic given that Epicurus wrote: "Plain dishes offer the same pleasure as a luxurious table, when the pain that comes from want is taken away. Bread and water offer the greatest pleasure when someone in need partakes of them. Becoming accustomed, therefore, to simple and not luxurious fare is productive of health and makes humankind resolved to perform the necessary business of life" (translated by O'Connor, 1993, p. 66). It was the opponents of Epicureanism who vilified the founder as a libertine and voluptuary, but those descriptions are inconsistent with both his teachings on pleasure and with his own lifestyle (Long, 1986: 16). For Epicurus and his many followers, leading a tranquil, good life meant fulfilling the basic desires rather than aiming at luxuries which could lead to unnecessary exposure to potentially disruptive experiences.

Epicurus, although a citizen of Athens, was born on the island of Samos in 341 B.C.E. He taught in Colophon, Mytilene and Lampsacus, in Asia Minor, and eventually settled with his students in Athens in 306 B.C.E. and began to teach in his kitchen-garden to anyone, whether men or women, slaves or free persons. The Garden, as his movement was known, was a society of friends who were devoted to Epicurus and his idea of living according to a common set of principles including



tranquility and living life without attracting attention (Fragment 86 in Inwood, Gerson and Hutchison, 1994). Epicurus wrote over three hundred books on topics ranging from science to love and the gods, unfortunately, few of his writings survived. However, two secondary sources are particularly helpful when read with care. The first is the famous poem, "On the Nature of the Universe", by Lucretius (ca. 100-55 B.C.E.). Lucretius was an Epicurean whose didactic and epic poem of around seven thousand five hundred lines praises Epicurus and the Epicurean system of life with an almost religious fervor (translator M. Smith, 2001). The second source is probably one of the more unique acts of devotion to any philosopher. In the town of Oenoanda, in the middle of modern Turkey, Diogenes Flavianus (Long cites Diogenes of Oenoanda) erected a large stonewall inscribed with some of the teachings of Epicurus. The degree to which these secondary sources represent the ideas of Epicurus or the authors of these works is at times difficult to discern, so whenever it is possible I will keep to Epicurus's original works.

The range of topics that Epicurus wrote about is impressive, but it is important to keep in mind the purpose to which all of his studies were directed. Epicurus explains his goal in a letter written to Pythocles. In the letter, Epicurus prefaces his remarks on natural science with the following: "First of all, we must not think that there is any other aim of knowledge about the heavens, whether treated in connection with other doctrines or separately, than peace of mind and unshakeable confidence, just as it is our aim in all other pursuits" (85-86). By unshakeable confidence Epicurus may be referring to his belief that nearly everything can be known through observation and use of the senses; if modern terminology is appropriate, Epicurus, like the Cyrenaics, was generally an empiricist. His message to Pythocles is that all other pursuits are secondary to the pursuit of a peaceful, tranquil, mind.

Epicurus was also a hedonist, and his evidence that people are hedonistic was, unsurprisingly, empirically based: "Therefore, we declare that pleasure is the beginning and the goal of a happy life. For we recognize pleasure as the first good and as inborn; it is from this that we begin every choice and every avoidance" (translated by O'Connor, 1993, p. 65). Epicurus' starting point is therefore different from that of, for example, Hobbes who, as we saw, believed that the fundamental principle guiding people is self-preservation. However, the message of Epicurus is slightly more complicated than the simplistic idea that people should begin searching for those things that will make them happiest, his message is that

only a wise person will be able to determine those things that are truly in his interest. In a "Letter to Menoeceus", Epicurus makes a distinction between natural and idle desires: "of the natural desires, some are necessary while others are natural only" (128 in Inwood, Gerson and Hutchison, 1994). "Of the necessary desires, there are those that are necessary for happiness, those that are necessary for the body's freedom from disturbance, and those that are necessary for life itself." Understanding these differences allows people "to refer every choice and avoidance to the health of the body or the calm of the soul, since this is the goal of a happy life." According to Epicurus, we have an ability to reflect on the choices we make and decide according to how they promote the goal of life, which is to be happy, but happy in terms of correctly selecting to act upon those modest desires that produce a calm soul. The approach thereby anticipates the concept which is now called enlightened self-interest.

In contrast to the common modern understanding of happiness, for Epicurus happiness meant the removal of all pain and fear. This subtle difference has profound implications for the model of the good life that Epicurus constructs. Once the removal of pain and fear is achieved nothing else is needed. The idea is not to then search for things that will make one happy, but to remove extraneous desires. Experiencing pleasure, for Epicurus, is an indication that something was previously lacking: "For we have need of pleasure at that time when we feel pain owing to the absence of pleasure. When we do not feel pain, it is because we no longer have need of pleasure" (translated by O'Connor, 1993, p. 65). Take the example of a hungry person, the initial pain that a hungry person feels on account of their hunger, according to Epicurus, is satisfied when they eat from which they derive immediate pleasure. However, from the satisfaction of eating a second pleasure arises, a sort of static pleasure of a complete absence of pain (Long, 1986). Once a certain (minimum) level of desires have been fulfilled, desires should be curtailed: "For it is not continuous drinking and revels, nor the enjoyment of women and young boys, nor of fish and other viands that a luxurious table holds, which make for a pleasant life, but sober reasoning, which examines the motives for every choice and avoidance, and which drives away those opinions resulting in the greatest disturbance to the soul" (131, in Inwood, Gerson and Hutchison, 1994). "We esteem self-sufficiency not so that we may always prefer the cheap and frugal, but so that we may feel no dread regarding them" (Ep. Fragments 29 in Inwood, Gerson and Hutchison, 1994). We've previously seen this idea of the advantages of living moderately also expressed in the works of Butler and Smith.

It is not that Epicurus was opposed to luxury and physical indulgences because they are inherently bad, he wasn't a prude, rather, the problem is that they may lead to a net increase in pain. For example, he accepts the pleasurable experience of drinking alcohol, but says that the pleasure should be balanced with the pain of the morning after. Just because something is pleasurable doesn't mean it should be pursued, and just because something is painful doesn't mean it should be avoided. We often pass over small pleasures, "whenever greater difficulty follows from them" (translated by O'Connor, 1993, 65). Sometimes pains are better than pleasures, "since a greater pleasure will attend us after we have endured pain for a long time." "Every pleasure, therefore, because of its natural relationship to us, is good, but not every pleasure is to be chosen. Likewise, every pain is an evil, but not every pain is of a nature always to be avoided." This is a familiar sort of hedonistic calculation that has to be conducted in order to determine the net happiness of an act and clearly conflicts with the Cyrenaic view that immediate pleasures are to be preferred.

Epicurus' view of happiness was most certainly influenced by his belief that people should live a trouble free, tranquil life. In fact, his primary concern was to show how a tranquil life could be obtained (Long, 1986, p. 62). It is informative to compare Epicurus' views with those of Plato and Aristotle, to whom he may have been replying. For Plato and Aristotle, some pleasures are good and contribute to happiness while others are bad (Long, 1986, p. 62). Good and happiness don't always correspond in the systems of Plato and Aristotle, while for Epicurus pleasure is always good, since the good is defined as that which is or causes pleasure. However, as I wrote in the immediately above paragraph, just because something is good doesn't mean that it has to be pursued in the Epicurean system. Regardless, the Epicurean position that happiness and good are different names for the same thing might be interpreted as a form of ethical egoism. Under this interpretation, a person should (in order to be ethical) perform those acts that make them happy, acting for any other reason would be non-ethical or unethical. But in contrast to Ayn Rand's understanding of happiness, happiness for Epicurus had a much more modest meaning. Although Epicurus can be understood to a hedonistic, ethical egoist, the sort of life he recommended was the very modest aim of leading a tranquil life.

What does Epicurus say about relationships among people? By all accounts Epicurus was much loved and, according to Diogenes Laertius, his friends were so

numerous "that they could hardly be counted by entire cities" (Laertius, "Lives of Eminent Philosophers"). However, it is difficult to imagine that an Epicurean would want to become entangled in personal relationships because of the potential disruption that those relationships might cause to the Epicurean's carefully constructed equilibrium. Nonetheless, if Diogenes Laertius can be believed (10.117-121 in Inwood and Gerson, 1997), a wise man, according to Epicurus, is reasonable but feeling. The basic message is to maintain an equilibrium that allows the Epicurean to use his reason unhindered to contemplate those things that really matter. The Epicurean view of relations reflects this belief and can be seen in the following maxims. The wise Epicurean will be affected by his feelings as long as they don't hinder his progress toward reason. He will not fall in love or pretend to be in love, presumably because doing so would lead to entanglements that might upset his tranquil state. His reason allows him to overcome the hate and envy of his enemies and prevents him from taking the opposite dispositions nor feign it willingly. "Sexual intercourse", the Epicureans said, "never helped anyone, and one must be satisfied if it has not harmed." However, the wise man will marry and father children (despite the harm), but only if it is indicated by the circumstances of his life. He will earn money when he is in dire straights, but only by exploiting his wisdom. While friendship comes into being because of its utility, it is also an end in itself. Although he may have to make a preliminary sacrifice to gain friends, it is then sustained by shared reciprocated pleasures. He will even sometimes die for a friend.

The list of do's and don'ts collected by Diogenes, of which the previous were just a small sample, is a rather odd mixture. However, the overall message seems to fit well with Epicurus's recommendation to satisfy one's minimum, natural desires which reason tells us are in our best long-term interests. Some advice, such as not to be a good public speaker, not to rant and rave when drunk or write poetry (whether drunk or not), but to erect statues and leave written treatises, strikes the modern reader as a curious mixture, some of which even appear to be self-serving. For instance, the rich members of his movement were expected to give money to the poorer members; Epicurus received a stipend. However, the overall message is one of self-sufficiency. Don't, for instance, depend on the gods, they exist but they don't care about people because to do so would disturb their state of happiness. Epicurus regarded self-sufficiency as "a great good, not that we may always have the enjoyment of but a few things, but that if we do not have many, we may have but few enjoyments in the genuine conviction that they take the sweetest pleasure in luxury who have need of it, and that everything easy to produce is natural while

everything difficult to obtain superfluous" (131 in Inwood, Gerson and Hutchison, 1994).

It is also helpful to understand Epicurus' ideas in the context of other approaches of the time. Diogenes Laertius writes the Epicurus disagreed with the Cyrenaics on the question of pleasure. "For they do not admit katastematic [stationary] pleasure, but only kinetic pleasure, and he admits both types in both body and soul ..." (136 in Inwood, Gerson and Hutchison, 1994). "For freedom from disturbance and freedom from suffering are katastematic pleasures; and joy and delight are viewed as kinetic and active" (136 in Inwood, Gerson and Hutchison, 1994). The katastematic pleasures are those that are not actively related to a feeling of happiness. This idea is expressed in English when we say that someone enjoys good health, where we don't mean to say an active state of pleasure, but a neutral state that we wouldn't want to change for the worse. This happy, neutral, state is defined as pleasurable by the Epicureans and reflected in their desire for tranquility. Epicurus also disagreed with the Cyrenaics about which pains were worse, he believed "pains of the soul are worse since the flesh is only troubled by the present, but the soul is troubled by the past and the present and the future" (137 in Inwood, Gerson and Hutchison, 1994).

Epicurus' views can also be seen as a response to the epistemological and ontological positions of Plato and Aristotle. Where Plato had his ideal forms, Epicurus accepted atomism. Atomism is a complex position, here I will note only that, for an atomist, all things are in a state of flux, moving from one state to another and thereby negating the possibility of ideal forms and opening room for the position that the search for pleasure, rather than ideals, was the way to live a good life. Where Plato and Aristotle used logical analysis, Epicurus relied on immediate sensations, particularly the sensation of touch, again implying the importance of pleasure in his system. He was generally an empiricist but not everything in his system could be observed, for example, the existence of the soul was argued by induction. Finally, he rejected the view that linguistic analysis by itself can tell us anything relevant to achieving a happy life. Rather than words alone, sensations were relayed upon to get to the truth of the matter. All of these positions emphasized the need to seek pleasure, but the Epicurean view of pleasure strikes the modern reader (at least this one) as somewhat sterile, overly cerebral and standoffish; all of which reinforces the need to specify the general aims and assumptions in order to assess the reasonableness of specific actions within a model.

## Stoics

Like the Cyrenaic and Epicureans, the Stoics too were hedonists, but that label doesn't do justice to their sophisticated ethical system. Stoic philosophy developed over hundreds of years, starting at around 300 B.C.E. when Zeno of Citium began to pace the Painted Colonnade (Stoa) in Athens and officially ending in 529 when Justinian closed four philosophical schools in Athens. Early and middle Stoic philosophy is marked by its rigor; studies on the topics of logic, nature and ethics intertwined to form a complex picture of the place of humans within the whole of nature. A well-known quote celebrates this cohesion: "The remarkable coherence of the system and the extraordinary orderliness of the subject-matter have made me prolix. Don't you find it amazing, in heaven's name? ... What is there that is not so linked to something else that all would collapse if you moved a single letter? But there is nothing at all which can be moved" (Spoken by 'Cato' in Cicero, *Fin.* iii 74). Within the limits implied by Cato's quote, I will attempt to draw out three relevant points. The first is the practical form that hedonism takes in the Stoic system. When reading the Stoics it is easily possible to confuse the ideal state which the Stoics set for sages, (the Stoics themselves doubted whether such an all-knowing person had ever existed), and the stages to which a mere mortal should aspire to live according to the Stoic system (Hankinson, 2003, p. 59). Examining the more modest goals of the Stoic system gives an indication of what leading a good, happy, life meant to them. Happiness, for the Stoics, is broadly defined and emphasizes the diversity of aims and goals that people can have. I will then discuss an important component of the external dimension of self-interest, namely, I will address the Stoic view of the interrelations of human interests. Finally, I will briefly note the Stoic influences in the works of some of the authors I have previously discussed.

According to the Stoics, all adult humans have the rational capacity to be happy, unfortunately, that same rational capacity means they can make unwise choices, that is, choices that will make them unhappy. Happiness for the Stoics meant to follow one's nature, but the happiness that results from following nature doesn't necessarily correspond to modern ideas of happiness. According to the Stoics, plants, animals and children all follow their respective natures and are therefore "happy" without reference to higher forms of reason. On the other hand, adult humans are the only creatures capable of using reason to understand their proper place in nature as a whole, and the only ones with the capacity to conform their

actions to what nature prescribes. Superhuman sages have complete understanding of nature and always act according to what nature prescribes and are, as a result, always happy. A sage's understanding of nature and his place in it is absolute, bad things (or what non-sages would call bad), can happen to him, but everything in nature serves some greater purpose and so all things are inevitable and therefore nothing really deserves the title of good or bad, things are just the way nature intended them to be. This view of life produces a sort of comforting detachment in the sage. The goal that normal, fallible humans should strive towards is to constantly add to their understanding of their place in nature and reform their actions to fit their evolving understanding of nature. However, the ability of adults to reason is a two-edged sword in that although they are the only creatures capable of adjusting their activities to fit nature, the freedom that reason gives means that adults can work against their nature. In contrast, plants and animals are in a sense programmed to follow their respective natures, while humans, by using bad reasoning, can go "off-course". The role of nature is critical in the Stoic system. All particular things can be analyzed in terms of their specific natures--the nature of a plant, the nature of an animal and so forth (Long, 1986: p. 189): "The nature of anything is simply that structure and pattern of behavior which universal Nature has ordained as appropriate or in the interests of the creature concerned" (Long, 1986: p. 189). The value of anything in Stoicism is defined by reference to nature in its universal sense. As a human being acquires rationality their nature prescribes new modes of appropriate behavior. Rather than following base instincts the function, the nature, of mature humans is to follow their reason.

Given this background, it is now possible to explain how the Stoics and Epicureans clashed on the important point of pleasure. Epicurus claimed that there was no need to argue or discuss why people pursue pleasure and avoid pain, he felt that it was sufficient merely to point it out as an obvious fact of human nature. The Stoics, to the contrary, argued that it wasn't pleasure, but self-preservation that is the primary impulse that people pursue. Therefore, happiness for both schools, in contrast to its modern usage, is not necessarily equivalent to pleasurable feelings. As I have shown, for Epicureans, freedom from pain is a happy state, while for Stoics it is living or attempting to live according to one's nature or, what amounts to the same thing, to living virtuously. Just as with Hobbes, the Stoics believed that all creatures begin life with the primary impulse to preserve themselves. However, the goal of self-preservation for Stoics is variously described as: life according to reason, life according to virtue, happiness and the attainment of happiness (Long, 1996, chap. 6). All of these ideas are linked for the Stoics, which

can lead to some confusion in regards to why they deserve the modern label of hedonistic.

The Stoics claimed that happiness is wholly dependent on virtue (Long, 1996, p.184). For the Stoics, the state of the soul constituting virtue is treated as an essential instrument of happiness, but not as something desirable as such or as part of the content of happiness. There are base instincts which compel plants and animals (including humans) to pursue some things (including the appropriate foods, shelter, parental affection) and avoid other things in order to preserve life. However, humans are capable of developing affiliations towards a wider range of things than irrational animals. "In a list of things according to and contrary to nature we find technical competence, health, beauty, wealth, high repute, nobility of birth and their opposites" (SVF iii 127 in Long, 1986: p. 190). It is these wider affiliations that can lead people astray. The point to remember is that although these affiliations are preferable to their opposites, they are not a part of Stoic virtue or happiness. For example, health is preferable to sickness, but neither is a requirement of virtue; a sick or poor person can be just as virtuous as a healthy and rich person. This idea is expressed in several of Seneca's Letters (II, 4). Seneca, referring to "the enemy" Epicurus, writes: " 'A cheerful poverty,' he [Epicurus] writes, 'is an honourable state.' But if it is cheerful it is not poverty at all. It is not the man who has too little who is poor, but the one who hankers after more." In Seneca's letter we can see the idea that a person can and probably will prefer wealth to poverty, but a wise person would see that neither really matters to living a virtuous life or, what amounts to the same thing, living according to one's nature. Later writers, including Adam Smith, repeat the idea that unconstrained desires are a source of unhappiness which need to be controlled. An exclusive fixation on fulfilling one's desires or preferences is not the means to living and achieving a happy life.

How, according to Seneca, do we know what to do in order to follow our nature and be virtuous and thereby happy? The short answer seems to be that we need to selectively look around us and see what wiser people are doing. The Stoic Epictetus's repeatedly stressed the importance of examining one's impressions and recognizing what is and is not up to us, his advice should be understood in the light of Stoic view of nature. Epictetus uses the following formula to help the novice towards a virtuous life: (1) reexamine the overall goals of one's desires; (2) adjust impulses to action and views of one's social commitments in the light of thought about goals; and (3) aim at complete consistency in belief, attitude, and state of



mind (Gill, 2003 p. 43). Epictetus wrote that happiness is within our reach (Long, 1996 p.196):

God [another term for nature] made all human beings with a view to their happiness, their good condition. To this end he gave them means, giving each person some things that belong to himself and others that do not. The things that are liable to frustration, removal, and compulsion are not his own, but those which are not liable to frustration are his own. As was right in one who cares for us and protects us like a father, he included the essence of good and evil among the things that are our own. (Epictetus, III.24.3).

The practical advice directed at people who want to become more sage-like is centered on determining what types of actions are, indeed, "appropriate" and in determining in one's life the right relationship between gaining preferable advantages and acting virtuously or at least making progress toward virtue (Gill, 2003 p. 41).

The Stoics used the prevailing concepts of virtue to define the things that are good, a list which included, prudence, justice, courage, and temperance. Things that are bad are the opposites of the good, accordingly, imprudence, injustice and so forth. "Prudence is knowledge of what one is to do and not to do and what is neither; or knowledge in a naturally social (and rational) animal of good things, bad things and what is neither (and they say that this [definition] is to be understood [to apply] in the case of the rest of the virtues too. Temperance is knowledge of what is to be chosen and avoided and what is neither. Justice is knowledge of the distribution of proper value to each person. Courage is knowledge of what is terrible and what is not terrible and what is neither" (John Stobaeus Anthology 2, 5b1 in Inwood and Gerson, 1997). So far, the Stoics list of goods and bads seems reasonable if somewhat difficult to implement in practice. However, the Stoic idea of indifferent needs some explanation. Things that are neither good nor bad are, for the Stoics, those things which neither benefit nor harm, such as life, health, pleasure, beauty, strength, wealth, good reputation, noble birth, and their opposites death, disease, pain, ugliness, weakness, poverty, bad reputation, low birth and such things (Diogenes Laertius 7.102 in Inwood and Gerson, 1997). How is it possible to reconcile this list of indifferent objects with the idea that the Stoics were hedonists? A clue to the answer can be found in the following: "For just as heating, not cooling, is a property of the hot, so benefitting, not harming, is a property of the good; but wealth and health do not benefit any more than they harm; therefore,

neither wealth nor health is good" (Diogenes Laertius 7.103 in Inwood and Gerson, 1997). The reasoning behind this argument is similar to that found in Butler. Whereas Butler was referring to desires, passions and appetites, the Stoics applied similar reasoning to the objects or aims that people may have. The objects are not important because they can lead to either a good or bad life, it is how a person views those objects that is important.

The Stoics use similar reasoning to argue that performing an appropriate act doesn't imply that it is right according to one's nature. It is rational to perform an appropriate act, but rationality admits both 'good' and 'bad' acts. For the Stoics there is more to being a good person than following one's rationality. Rationality, in short, is a necessary but not sufficient condition to being good. The good man is in complete, continuous agreement with nature: "The ever-growing consistency of his selection of natural advantages (performance of appropriate acts) brings a recognition that there is something of far greater worth than any of these objects, singly or collectively" (Long, 1986: 192). For example, the sage will make good use of poverty if it comes his way while a foolish man may use wealth badly. This does not undermine the objective fact that wealth is preferable to poverty, but wealth is not a constituent of virtue. The moral value of selecting wealth depends upon the agent's principles and manner of acting" (Long, 1986: 192).

According to Stoics, people are mistaken to follow their emotions and passions. These are understood in Stoicism as products of a specific kind of error; namely, that of treating merely "preferable" advantages as if they were absolutely good, which only virtue is (Gill, 2003, p. 41). Intense reactions (passions) constitute a disturbance of our natural psychophysical states which should be seen as disturbances and treated as "sicknesses" that need to be "cured" by analysis of their nature and origin and by advice. Seneca offers the following threefold pattern of ethical guidance which is similar to that of Epictetus listed above: (1) assessing the value of each thing; (2) adopting an appropriate and controlled impulse toward objects pursued; (3) and, achieving consistency between impulse and action (Gill, 2003, p. 42). The satisfaction or frustration of desires leads to production or avoidance of the passions so controlling desires will lead to fewer disturbances and a more balanced life. Unlike Plato and Aristotle, the Stoics thought of emotions as a part of a person's rational faculty which makes them responsible for actions based on emotional responses (Cooper and Procope, 1995). Contrary to common understanding, the Stoics did not wish to eradicate all emotions, only those

emotions which are excessive, after all, emotions are a part of rationality and therefore necessary in order to make rational decisions.

The issue of the relationships among people is typical of the Stoic approach. The social principles of the Stoics derive from the impulses implanted by nature to form familial and extra-familial relationships (Long, 1986: p. 191). "But the principle determining such behavior is not regarded as different in kind from that which prompts more obviously self-regarding actions" (Long, 1986: 191). This principle can be traced back to the Stoics belief that self-love is the primary, natural motivation. Equally innate to the primary motivation of self-love, but developing later, is the secondary motive of appropriation to a creature's offspring, which the Stoics took to be the foundation of human sociability (Long, 2000 in Inwood, 2003). The two instinctual motives that Butler makes primary, self-love and benevolence, are prefigured in the self-directed and other-directed objects of Stoic motives. Both motives, according to the Stoics, are equally natural, suitable and mutually compatible. For the Stoics, instincts towards self-preservation lead to natural inclinations to care for our offspring which then serves as the foundation upon which wise people can understand the unity of mankind. The Stoics saw the world as a single great community in which all men are brothers, ruled by a supreme providence which could be spoken of, almost according to choice or context, under a variety of names or descriptions including divine reason, creative reason, nature, the spirit or purpose of the universe, destiny, personal god, even the gods (Seneca in Campbell, 2004, p. 15). These same views are echoed by Adam Smith in a quote I previously mentioned: "man ought to regard himself, not as something separated and detached, but as a citizen of the world, a member of the vast commonwealth of nature", and "to the interest of this great community, he ought at all times be willing that his own little interest should be sacrificed" (Smith, 1790, p. 140).

The Stoic position has been attacked along two lines. The first is that it leads to emotional impoverishment. This can be seen in the following advice given by Epictetus (III 24: 84-87): "Whenever you are getting attached to someone, don't let it be as though they're something undetachable--but more as if you had a jar or crystal goblet, so that when it breaks, you'll remember that it is that sort of thing and not be upset... ." It is this disengagement with life that marks the Stoic system. The argument is that Stoicism leads to a sort of fatalism. Imagine, for instance, a wise Stoic walking along a street and noticing a burning house. He runs into the house and does everything he can to save some children trapped inside but is

unsuccessful. He has done everything he can to help and even risked his own life, but without success. A proper Stoic will reflect and conclude that the death of the children was an inevitable event ordained by deterministic nature. An emotional response by the Stoic would be out of place and unnecessarily troubling; he has no reason to feel bad, stronger still, he would be wrong to feel bad. For a non-Stoic, the Stoic's indifferent reaction would appear cold and heartless. Although reason might tell the Stoic that an emotional response doesn't do anything to improve the situation, many would find that the Stoic's limited understanding of what reason excludes unacceptable. The second attack claims that happiness is used by Stoics in a very loose, disingenuous manner. On this interpretation, the Stoics are not really offering a method to happiness, but rather moral guidelines that are quite independent of happiness. Certainly many modern readers wouldn't find much happiness in living the disengaged, over-intellectualized life of a would-be sage--frankly, it would seem to be quite boring. Life might go smoother under such a regime, but actively experiencing the highs and lows of life are part of what makes life interesting.

Along with their influence on Adam Smith, the Stoics influenced the views of Joseph Butler. Butler writes in the preface of the "Fifteen Sermons" that "the ancient moralists had some inward feeling or other" corresponding to his thesis, which they expressed by saying that "man is born to virtue, that it consists in following nature, and that vice is more contrary to this nature than tortures or death" (Butler, 1986, Preface p. 8). According to Long (2000), Butler and the Stoics agreed on two things: 1. Nature, with respect to human beings, is a term that has multiple references. One meaning is that following nature means that humans have the capacity to reflect upon their actions and adjust them to what is good; 2. There is no basis in our given nature for any necessary conflict between self-love and benevolence. Following nature for Butler meant: 1. Acting according to any psychological propensity; 2. Following whatever passion happens to be the strongest; and, 3. Following the principle of reflection which, in term of his teleological argument, is superior to all our other faculties (Long, 2000: 384 in Inwood, 2003). In both the Stoics and Butler, reflection is an important component separating base instincts from more the developed, better, desires. As I have shown, Butler claims, like the Stoic Cicero, that animals start with instincts of self-preservation and proceeds to the thesis that honourableness is the goal and fulfillment of a fully mature and rational human being. Butler would presumably fully agree with the following sentiment expressed by Seneca: "Those who make pleasure supreme hold that the good is perceptible, but we on the other hand

attribute it to the mind and hold that it is intelligible. If the senses made judgments about the good, we would not reject any pleasure; for no pleasure fails to attract us and every pleasure pleases us; conversely, we would not willingly suffer any pain; for every pain hurts our senses" (Seneca, in Inwood and Gerson, 1997, p. 249). Reason makes a decision good or bad: "But [the hedonists] let the lowest part [of man] make the decisions about what is better, so that judgment is pronounced good by sense perception, which is blunt and lazy, and slower in men than it is in beasts" (Seneca, in Inwood and Gerson, 1997, p. 250). In the previous quote, Seneca seems to be rejecting the hedonistic roots of Stoicism, but what he is actually rejecting is the idea that base, animal, instincts should be allowed to determine what will make one happy. He and Butler would both argue, I believe, that a wise person would see that what makes someone truly happy is revealed through reason, and doesn't depend on sense perceptions alone.

### Conclusions

Although all of the schools of thought I examine in this chapter were hedonistic and believed in the rationality of people, they arrived at very different conclusions about how one should lead a good life. For the Cyrenaics, happiness means to follow immediate desires because it is the only thing which can be known for certain; for the Epicureans, satisfying the basic desires will lead to the greatest happiness, while for the Stoics, happiness is the realization of one's proper place in the universe. The key differences in the theories relate to the hedonistic rather than the rational aspects of the approaches. The conclusion I draw is that even accepting a limited form of self-interest, one that might be called informed hedonism based on rationality, can imply very different conclusions about the best way to live one's life. Social scientists who model agents as self-interested and rational have left open a large range of motives and, consequently, behaviors; so large in fact as to be nearly useless as an aid to describing or predicting behavior. The final lesson of this chapter is that in order to understand the recommendations that each of these approaches made about how to live one's life it was necessary to understand their assumptions about human nature. A modeler attempting to represent human behavior needs to be clear about the contexts under consideration, the aims and goals of agents, and how those aims and goals are assumed to be linked in the psyche of the agents under consideration.



## CHAPTER SEVEN: FRAMEWORK FOR ECONOMIC SELF-INTEREST

---

### Introduction

Previous chapters have been concerned with developing both a theoretical and practical understanding of the elements of self-interest as they might apply in economic contexts. In this chapter I use those elements to build a basic framework that can be used by researchers to begin to correctly specify the type of self-interest that they believe is appropriate to describe the motives of economic agents under consideration.

I will begin this chapter by describing the purview of economic analysis as it relates to self-interest. The discussion is intended to demonstrate that an important aspect of defining agent motivation is taking a clear stance on the scope over which an analysis is meant to apply. Regardless of whether the assumed motive or motives of an agent can be described as self-interested or non-self-interested or a combination of the two, a researcher needs to specify the context in which agents are operating; that is, given that the assumption that different contexts call forth different motives is accepted--a point that advocates of psychological egoism would deny. That complication aside for the moment, in regards to the point at hand, a researcher should be unambiguous regarding what (if anything) distinguishes an economic context from a non-economic context or which elements of a context are economic and which are not. If a researcher decides that self-interest is an appropriate description of the motivation behind observed economic behavior in a given context, then the following framework is intended to help them to refine that motive into a more manageable, less ambiguous motive. The discussion will largely follow along the lines of the structural and external dimensions of self-interest that I have been developing and using throughout this manuscript. I won't repeat in detail the arguments that I used to support each of the elements I mention here, rather, the purpose of this chapter is to provide a helpful summary of previously discussed issues.

One qualification that I should mention before proceeding is that in the following sections I will largely ignore the possibility of non-self-interested motives in economic contexts. As previously shown, this category of motives includes duties, special relationships and altruism. In general, these are motives with which people are willing to sacrifice their net interests for the interests of others. I make the

simplifying assumption that these motives and other non-self-interested complications such as weakness-of-will are not present in order to concentrate the analysis on the topic of self-interest, which I take to be the default position in economic analysis. By making this assumption I am not claiming that these other motives are not present in economic contexts, I take quite the opposite stance as I will demonstrate in the chapter to follow, rather, I do so from a desire to keep the discussion manageable and concentrate on the main topic at hand--a practical framework for self-interest.

### The Purview of Economic Analysis

An issue that has been somewhat in the background for much of this manuscript is the purview or competence of economic analysis. What defines an economic context and, more specifically, what does it mean to be economically self-interested? I don't intend to offer a detailed description of what defines an economic context; my approach will be to attempt to demonstrate that there is a close relationship between the choice of a context and the type of self-interest that describes agent motivation and consequently behavior.

If it is accepted that self-interest and rationality are "the" economic assumptions, then the purview of economic analysis is indeed very broad.<sup>13</sup> Self-interest is clearly an important motive that stands behind many if not most actions whether those actions can be described as economic or not. Likewise, some level of rationality is necessary or useful to function in nearly any imaginable circumstance. Given the prevalence of self-interest and rationality in all human activities, it is not much of a stretch to then conclude that economic analysis is the appropriate tool to describe and predict most human behavior; a position explicitly taken by some of the authors mentioned in previous chapters. But the very prevalence of these two assumptions in human activities, along with a belief that behavior is significantly variable across contexts, should be a warning that if the assumptions of self-interest and rationality are accepted as describing all such activities, then they are

---

<sup>13</sup> Following Pettit (1993, 55), I use the term rationality in a practical sense to mean: "that an agent's belief should be such that, under accepted criteria of evidence, they stand a better chance of being true than available alternatives. More specifically, they should be well supported inductively or deductively. They should be inductively rational in the light of observation or they should be inductively or deductively rational in the light of rational beliefs."



exceedingly malleable terms. This remains true even if the position of psychological egoism is accepted. Recall that a psychological egoist contends that all behaviors are, at base, motivated by self-interest, which is different than arguing that all contexts are the same. I do not believe anyone argues that all social contexts are equivalent in that they involve the same or similar aims, and therefore I will ignore that possible line of reasoning. The important point is that I have argued that to the degree that contexts are taken to differ, and that different contexts call forth different aims and, additionally, different aims imply that different sorts of rationality are applicable, so accordingly should the sorts of motives that are used to describe agents in those various contexts. This remains true even if self-interest is taken as the only motive. In fact, rationality and self-interest are so widely applicable that they can better be described as necessary rather than sufficient descriptions of motivation.

What if we begin by narrowing the focus to exclusively economic contexts, can we then conclude that self-interest and rationality are good assumptions to describe motives? The answer to that question rests on whether it is possible to find a convincing link between what is meant by an economic context and an unambiguous subset of self-interest. In other words, are there purely economic interests? The task, under this approach, is to identify the distinguishing features of an economic context in which desires or preferences are economically inclined or directed. I contend that the intention of an act in such a context is, presumably, to improve one's economic well-being. The question becomes, can we restrict or interpret the definition of those preferences and desires that promote economic well-being to such a degree that the interests represented become manageable? Two possible approaches come to mind, one is to restrict the aims of agents to exclusively economic desires or preferences, and the other is to argue that all aims, economic or otherwise, are in reality closely related.

Both approaches entail limiting the external component of self-interest. I'll assume for the moment that the structural component is some form of pure self-interest (as opposed to mixed forms) as that term was previously defined. The first approach attempts to constrain preferences and desires to strictly economic preferences and desires. Part of coming to an agreement that a context is economic might include the idea that the goals and aims of an agent, the external dimension of self-interest, are best described as economic. Here I am simplifying the argument by equating preferences and desires with an agent's aims and goals at the point of making a decision. An agent's preferences therefore represent the aims and goals that an

agent has at a point of making an economic decision--a decision about his economic well-being. I am further assuming that a rational agent will order his desires or preferences in such a manner as to maximize his well-being.

Traditionally, economic agents were assumed to be attempting to maximize their wealth, income or some similar, easily accessible, assessable, monetary measure. Wealth, etc., was taken to be the external aim at which self-interest was directed in order to promote economic well-being. It is important to note that economic well-being wasn't necessarily equated with overall well-being, although it may be a part of well-being--a point Adam Smith repeatedly made and J.S. Mill reiterated. I do not know of any scholar who argues otherwise, rather, the argument is that all human activities are rationally pursued in order to maximize one's overall, economic and non-economic, interests. Traditional aims would then appear to place severe constraints on an agent's preference set, limiting their choices to those acts which maximize their wealth. Such a limit clearly contrasts with the primary value that modern preference theory is intended to support, namely, personal autonomy in all of its many aspects. However, perhaps the historical approach is not as restrictive as it first appears. Certainly many aims can be acquired or realized via wealth. After all, wealth, with the exception of misers, is not pursued as an end in itself, but as a means to acquire other aims or goals. Assuming wealth as the ultimate preference of agents shifts the issue to the reasons why agents pursue wealth, only some of which may be described as economically self-interested. Wealth isn't necessarily pursued for reasons of self-interest, and if it is pursued for self-interested reasons then those reasons would appear to vary widely as well. In short, the assumption that the aim of an agent is to maximize wealth or income doesn't necessarily restrict the motivational set which stands behind agent behavior. We'll have to find a more limiting set of preferences if we want to ensure that agents are directed at economic aims and goals.

This is not a difficult task, it is quite easy to construct hypothetical scenarios in which agents are restricted to sets of products or services which can generally be accepted or assumed to be of purely economic importance. For instance, we can conclude that an agent who decides to buy more widgets at a given moment does so because he expects to improve his economic well-being; after all, what else are widgets good for? However, as one moves further away from textbook examples and laboratory conditions, the (revealed) preferences of agents become more difficult to link with specific motives and therefore more difficult to identify as exclusively economically self-interested. The success of this first approach will

depend on the ability of a researcher to build a convincing case that the products and services under consideration are pursued strictly in order to promote one's economic interests.

The second approach is a more general maneuver that allows researchers to maintain the claim that economic analysis is everywhere applicable and avoids the issue of defining uniquely economic preferences. The approach is composed of two parts; the first is to take the position of psychological egoism so that self-interest is the only motive behind behavior. The second part of the argument is to argue that although contexts and consequently motives differ, there is an underlying equivalency among preferences and desires across contexts; in short, that there really is no difference between economic and non-economic aims and goals. The two parts taken together may imply that economic analysis is appropriate over a wide range of circumstances. I will demonstrate this approach by linking it with modern utility theory.

Modern utility theory avoids the issue of defining economic aims and goals by taking the position that preferences can be anything, no stance needs to be taken in regards to what an agent prefers and the motives that stand behind those preferences other than whether preferences remain rational in the limited sense that they are consistent through time. If this broad understanding of preferences is accepted, and self-interest is taken to be the only motive behind all acts, economic or otherwise, then the approach under consideration is to argue that all preferences, despite different contexts, can be collapsed to one general, underlying preference. There are two things to notice about this maneuver.

The first, as I have previously argued, is that assigning agents the motive of self-interest is a significant departure from the ideal of modern utility theory. Self-interest is an addition to, and not a part of, preference theory, one that constrains agent preferences. The second point involves the attempt to avoid the problem of varying contexts, and therefore varying motives, by finding a common defining characteristic underlying all aims in all contexts. In other words, the position is that although self-interest may take different forms in different contexts, those objects pursued under the rubric of self-interest essentially boil down to the same thing. That "thing" has traditionally been taken to be (net) happiness or pleasure. For instance, agents maximize their wealth or the number of widgets they acquire in order to maximize their happiness or pleasure, wealth and widgets are not an end, but a means to happiness. Such an approach might be appropriate in some

contexts, but certainly not all contexts in that, as I have shown, agents often have good reasons to act against their own happiness and pleasure. In short, the argument would have to be made that economic agents, no matter the context, always act to increase their net happiness and all happiness's can be compared. That may be a valid simplification in some contexts, but one that runs counter to arguments presented in previous chapters. Promising alternatives to net happiness or pleasure might include some definitions of well-being as argued by Griffin (1988) in his case for modern Utilitarianism. This isn't the place to begin a discussion of the pros and cons of Utilitarianism, that discussion has more articulate advocates and detractors. The shift in approach would involve a much more fundamental shift in economic theory, that of re-entering the head of homo-economous to paraphrase Hicks.

The choice faced by researchers is either to severely restrain preferences and thereby ease the ability to link economic behavior with a small subset of economic interests, or, as with modern utility theory, leave preferences unrestrained and lose the ability to link behavior with any specific motive. If the choice is to restrain preferences, then the problem of linking self-interest to economic objects remains. The alternative of claiming that all objects are pursued for a common, underlying, goal remains open, but then that goal and the motives for pursuing it need to be specified.

The wide purview of economic analysis is a two edged sword. Self-interested or, better said, selfish behavior, narrowly defined and restricted to traditional economic aims, might be a good description of behavior in well-functioning markets in which price is a sufficient statistic. However, as economic analysis is extended to other social endeavors a narrow interpretation of self-interest becomes less tenable without the additional assumption that all objects are somehow linked in terms of some common measure of desirability. People in many contexts are not necessarily focused on maximizing their own economic interests, they may be motivated to live a good, well-balanced, life that might involve sacrificing some of their interests for the well-being of others. For example, contrary to the assumed motives behind market behavior, living a good life may include a desire to reduce one's wants and desires in order to live a less complicated life. One final point of consideration raised by Adam Smith and others before him is the relationship between economic interests and other interests. Smith, following his Stoic influences, recognized that once basic human needs are satisfied, a quest for ever more material possessions often leads to unhappiness. Greed may be good in

markets, it may serve an agent's narrow economic interests, but it may also lead to a reduction in overall well-being. It should be enough to remind the reader that self-interest has been used as a rational justification to live in abject poverty as well as extravagant luxury.

The assumption of self-interest alone, or in combination with the assumption of rationality, is not precise enough to either explain or predict economic behavior unless preferences are severely limited as well. Without a restriction on preferences, and regardless of whether it is meant to reflect real behavior or act as a simplifying assumption made in order to build more parsimonious models, self-interest allows a much too broad set of potential aims and goals to be a description of motives in all contexts. The question of importance is, given that the assumption of self-interest is appropriate in a given context, what sort of self-interest should be assumed?

#### Structural Dimensions

Let's assume that we can agree to restrict the external elements of economic self-interest to a specified set of aims and goals and, furthermore, we agree to call a context economic, have we thereby significantly restricted self-interest to a meaningful assumption? Again, I do not believe so, unless we can further agree on a common understanding of the structural dimension of self-interest. Recall that there are many approaches to modeling how an agent orders preferences, including purely egoistic approaches such as a psychological, rational or ethical egoisms, and mixed egoistic approaches that allow various types of self-interests and non-self-interests to be included within a preference set. Different means of ordering internal preferences imply different sorts of behavior. For instance, an economic agent who believes that he should or ought to act egoistically--an ethical egoist--will presumably have an added impetus to act with vigor to achieve his goals, more so than perhaps a rational egoist who believes that it is rational to act egoistically. A related issue is the degree to which the interests of others are incorporated into one's own interests. An economic agent who integrates the interests of future generations within his interests is likely to behave differently than an agent whose concerns are restricted to his own narrow interests.

Recall that the structural dimension of self-interest can be thought of as the rationality that agents use to order their preferences or desires. In previous chapters I have drawn important distinctions based on the type of rationality that

self-interest is assumed to take in an agent's psychology. To prepare for the discussion to follow, I will briefly review the two broad categories of structural self-interest which I discussed in detail in chapter 4, namely, positive and normative approaches, and I will include a very brief review of the external dimensions. The first category is called psychological egoism, a descriptive position that claims that people are in all circumstance self-interested. This approach has the advantage of parsimony, but the disadvantage of being open to empirical refutation. In contrast, the second category, normative approaches including rational, instrumental and ethical egoism, claim that people should or ought to act from the motive of self-interest. These approaches require justification or leave themselves open to the charge of arbitrariness. For instance, if the claim that ethical egoism will make one's life go economically better is accepted, it is reasonable to ask whether it will make one's life go better in general. Normative approaches required, at a minimum, that researchers specify how they will measure the "good" they advocate and range over which their "shoulds" and "oughts" apply. For instance, in the case of rational egoism, the basis upon which rationality is assessed should be specified. Hobbes, for instance, is often taken to have based his assessment of rationality on the concept of self-preservation. Accordingly, all acts that damage one's ability to survive are irrational in his system. Similarly, if an economist wants to call an act irrational, then he needs to specify what he's using as a measure of rationality. If wealth is taken as the only aim, then all acts that intentionally reduce the net wealth of an agent are irrational--a claim that stretches the meaning of rationality to the breaking point. Other structural theories discussed include tautological egoism, an approach which misidentifies an act with a reason for performing an act, and predominant egoism, an approach that claims that most people in most circumstances act egoistically.

Predominant egoism is a mixed theory, one that includes interested and non-interested motives. Mixed approaches are probably the most promising sorts of egoism to describe practical situations, but they come with the stipulation that a researcher is required to explicitly state when and which motives other than self-interest are engaged and under what conditions interests and non-interested motives dominate. Mixed theories can quickly become prohibitively complex by sacrificing parsimony for realism. Some of the theories discussed are not compatible, for instance, psychological egoism isn't compatible with normative approaches if it is accepted that you can't tell somebody what they should do if they don't reasonably have a choice to act otherwise. In addition, there is a clear contrast between psychological and normative egoisms, and mixed theories. The

first two approaches are meant to define motivation or what should or ought to be an agent's motivation at all times and all circumstances, while mixed theories allow people to be motivated by a diverse range of motives at a given moment. The incompatibilities among the various forms of self-interest emphasize the need to get this part of the assumption of self-interest right.

### External Dimension

I will briefly review the main features of the external dimension of self-interest before turning to the task of providing a general framework of self-interest. The standard answer that modern microeconomics gives to the question of what the aims of an agent are is that agents attempt to maximize their utility. Utility, in turn, is a function for mapping preferences to a real number, preferences can be anything. Modern axiomatic utility theory, to remind the reader, doesn't make the assumption that people are self-interested: "All that is assumed is that people's preferences conform to a number of axioms: roughly, they simply need to be consistent. They can conform to the axioms without being self-interested. Yet the muddle over 'utility' leads many economists to forget this important discovery." (Broome, 1999: 22). In short, utility theory says nothing about the aims of an agent, and if self-interest is assumed to be behind those aims, the justification for making that assumption needs to come from outside the model.

I have expended some effort in this chapter and chapter 4 arguing that a proper understanding of self-interest needs to specify the aims and goals at which the assumed motive of self-interest is directed. As Joseph Butler made clear, people don't aim at increasing their self-interest, they aim at those objects that promote their interests. Many different aims have been suggested through time. Adam Smith, for instance, claimed that once basic material needs are fulfilled, people aim towards objects which feed their vanity. J.S. Mill suggested that the pursuit of wealth can be used as the aim of economic agents as long as that aim is not modified by the pursuit of any other object. C.D. Broad argued that there are many, irreducible, aims including: pleasure, pain, wealth, power, security, liberty, glory, possession of particular objects, fame, health, longevity, status, self-respect, self-development, self-assertion, reputation, honor, and affection. The overall aim of the Hellenistic philosophers was to live a tranquil life, which they practically achieved through the pursuit of wisdom, temperance and justice (Cicero, *On Duties*, translated by Grant, 1971). Mis-specifying the presumed aims of agents will have negative repercussions for the predictive and descriptive value of a model

to the degree that behaviors induced by different motives are not correlated with one another. An example that I will explore in great detail in the final chapter is the case of open source software. There is much evidence to suggest that many of the original open source software developers were motivated by duties such as the belief that information should be freely available, rather than the belief that a commodity should be sold to the highest bidder. These programmers started a community that has had a large influence on the software market. Although their interests don't conform to the traditional, narrow, interpretation of economic self-interest, they are pursuing their interests in a highly competitive economic context, and to the degree that their aims differ from profit maximizing aims, their behavior may differ as well. An advocate of preference theory can rightly claim that their theory accounts for both types of motives. He cannot however, without additional assumptions, claim that the behaviors are motivated for reasons of self-interested.

Theories of self-interest that portray agents as immoral, isolated and non-interacting are fundamentally different than those in which the interests of people are taken to interact. In the second half of chapter 4 I outlined one possible manner with which to illustrate how interests may interact. At one extreme, people's interests can be completely independent of one another and thereby implying a self-defeating type of behavior, while at the other extreme there is the notion of the common mind in which the interests of people are completely dependent. Between these two practically unlikely extremes there exists a spectrum over which interests interact. An important concept used to distinguish among the different ways in which interests interact is the idea of the focus of the motive in question. For example, if the focus is primarily on an agent's own interests, even if those interests include the well-being of others, then such motives should be called self-interested. An important point to note is that a person can be self-interested and still be concerned with the well-being of others. Obvious examples are cases in which there are special relationships among agents including friends and family members; in which case by promoting the interests of these others an agent is promoting his own interests. Within the spectrum along which interests interact, I drew a distinction between cases in which the focus of an act is on the interests of the agent, and cases in which an agent is willing to shift that focus to another and sacrifice his own net interests for the net benefit of others. In other words, there is an important difference between promoting the interests of others as a means to promote one's own interests, and a willingness to promote the interests of others even, perhaps, at the cost of one's own interests--what I call an altruistic act. I concluded the chapter by observing that the often drawn false dichotomy of, on the



one side, isolated selfishness, and on the other, self-destructive altruism, isn't helpful; they are straw-men. In general, it is important to remember that self-interest isn't taken to exclude the happiness we derive from others. Self-interest that incorporates the interests of others has many variations that may be important for correctly modeling and interpreting economic models.

### General Framework

In this final section I introduce a framework which will allow researchers to systematically address the important features of self-interest. The idea being that if a researcher decides that self-interest is the motive that he wants to use to describe agent motivation, then the following characteristics of that cumbersome assumption need to be considered and, whenever they provide a potential source of misunderstanding, explicitly stated. The breadth over which self-interest describes motivation and consequently behavior means that a common understanding of this fundamental assumption can only be achieved by further delineating its many components. The method I will use in this section is to ask the types of questions that a researcher needs to consider when he describes the motivation of an agent as self-interested. When helpful, I will include the parameters between which the answers to the questions posed should fall. I will not here provide detailed justification for the positions taken, the reader is referred to previous chapters for discussion.

### Identifying motivation

Identifying and then categorizing motives even broadly, is difficult, sometimes even for actors themselves (chapter 1). Doing so will never be an exact science, but clues can be derived from several sources. There is, for instance, no substitute for close study of actions over a long period; particularly actions at important moments in a person's life under circumstances encouraging reflection.

### Does it matter?

The first question to answer is, naturally, are the motives of the agents under consideration important to the analysis? Economists typically describe agents as self-interested; many describe self-interest as one of the key assumptions of economic analysis. However, self-interest is a very general description of intentions and behavior, and should be more narrowly specified if a researcher has

a specific set of intentions in mind. For instance, it is often convenient in applied models to represent agents as maximizers of their own short-term financial gains.

What is the context under consideration?

It is vital to understand the context in which an actor is operating because it can give clues to the motives of actors. For instance, can the context be described as economic in the sense that profits can be taken to be the main aim of participants? How complex is the environment in which the actor is operating? A complex environment allows for the possibility of many different motives. Economic markets are complex in that many types of motives can operate. Consequently, in general, the more realistic a context under consideration, the less appropriate a simplistic definition of self-interest. All of which emphasizes the importance of picking a subject of study carefully. When making claims about interests, it is necessary to understand the context in which actors are operating. A well-documented, transparent, context will help when divining motives.

An important question that researchers should ask themselves is whether self-interest is the only motive operating in the situation under consideration or whether some other, non-self-interested, motive is prompting agents to action. Not all motives can be or should be reduced to that of self-interest; I have discussed several non-interested motives including altruism and duties, but there are many more as well, perhaps nearly as many as there are people (chapter 2). That point aside, self-interest is the predominant motive in most human activities including economic activities.

If the decision is made to describe motivation as self-interested, then both the structural and external components of that motive should be specified; where the structural dimension is the type of rationality that an agent uses to order his interests and the external component is comprised of the aims and goals of the agent. I extensively discussed various forms of structural reasoning including psychological egoism, ethical egoism and rational egoism, and refer the reader to chapter 4 for a list of some of the rationalities that have been identified in the philosophical literature. By the aims of an agent I don't necessarily mean a detailed list of the specific objects of desire of an agent, rather, I mean the more general categories under which aims and goals can be collected as discussed in chapter 4. For instance, aims can be described as hedonistic, desire fulfilling or found on an

objective list. It is important to note Broad's observation that aims may be irreducible, thereby emphasizing the need for a researcher to explicitly state what it is, generally, at which agent's aims. Finally, a researcher should be aware that many aims that have consistently been cited as worthy of pursuit such as wisdom, temperance and justice, are difficult to categorize as strictly self-interested (chapter 6).

Rather than being wholly self-interested or non-self-interested, perhaps agents are motivated by a combination of both types of motives, what I have previously called mixed motivation. If such is the case, then a researcher should state the conditions under which the various motives are actuated or risk sending a confusing message about the various aims and motives under consideration, and by implication, agent behavior. Mixed motivation is probably a more accurate description of practical motivation than either extreme, but one that invariably complicates a model. That stipulation aside, self-interest is in all likelihood a broad enough category to describe most economic activity in most contexts, non-self-interested motives should only be considered after careful consideration.

What is the time period over which an agent's interests are measured?

This question is closely related to the external dimension of self-interest mentioned above, but needs to be emphasized because of the potential complications it can add to the issue of interpreting the motives behind observed behaviors. Some interests are immediate, while others act as reference points that extend throughout mature lives. The nature of immediate aims and goals can be very different and may conflict with those we set for the long-term. Joseph Butler's analysis (chapter 5) of self-interest rests on the understanding that there is a hierarchy of interests. According to Butler, immediate interests, expressed as passions and appetites, can lead people to perform acts that are not in their best interest. Importantly, Butler also effectively argued that passions and appetites can lead to benevolent and altruistic acts--non-self-interested motives are thereby just as natural as self-interested motives. In addition to immediate passions and appetites, people have reasoned medium and long-term aims and goals. Agents may decide to sacrifice short-term interests in order to achieve medium or long-term interests; at any point in time people are bundles of interests that have different moments of fruition. A researcher who unwarily plucks an action out of time without due consideration of the act's place within an agent's overall set of aims and goals may misinterpret the motives inciting the observed action. The sort of inter-temporal reasoning that I

am referring to is common in finance and the rest of economics, similar sorts of reasoning needs to be applied when deciphering motivation. Actions that may appear to be self-sacrificing in the short-term (and therefore under some interpretations, irrational), may, for instance, be conducted in order to achieve a more valued, self-interested, long-term goal. As a practical matter, an important method of addressing this problem is to examine actions over a period, a method that I will demonstrate in the next chapter.

Why isn't preference theory enough?

There are several problems with arguing that by pursuing preferences agents are necessarily pursuing what is in their interest. The first is the case of weakness-of-will, in which agents prefer objects that are not subjectively in their informed interest. Weakness-of-will is a well-documented phenomenon, one that breaks any necessary link between preferences and interests. The second problem exists at a more theoretical level, it says that by making a claim about motivation we are making a claim about what informs a set of preferences, not about the preferences themselves. Motives are prior to preferences, so knowing an agent's preferences doesn't necessarily give insight into motivation. The third problem is the danger of tautological reasoning that goes along the following lines: agents prefer what is in their interest, therefore all preferences are self-interested (chapter 3). The bottom line is that interests and preferences shouldn't necessarily be equated with one another.

Is the objective to describe how agents act or how they should act?

Positive models of motivation such as psychological egoism are open to empirical refutation. The claim that agents are in general self-interested is a simplification, clearly there are acts motivated by altruism, duties and other forms of non-self-interest. Researchers may decide that non-self-interested motives can be ignored for reasons of parsimony, but they can't reasonably argue that self-interest is the only possible motive.

Normative forms of self-interest, those that tell people that they should, for whatever reason, be self-interested, raise ethical issues. A researcher who makes normative claims should be able to defend those claims against other norms. Why, a person might ask, is it better for me to act from self-interest as opposed to some other motive? Will self-interest make my life go better than some other

alternative? Perhaps, for instance, it is better to live life according to the Stoics and maximize my understanding of the world? Clearly, the issue of what makes a life good is linked to the external dimension of self-interest and to the wider topic of ethics. Greed in the marketplace might be tolerated, where some argue that all is fair (as in love and war), but a proper defense of the motto greed-is-good in wider contexts would have to address the basic issue of what good one has in mind.

Finally, as discussed in chapter 4, it is generally believed that there is a close relationship between self-interest and ethics. Therefore, to the extent that researchers claim that agents should be self-interested, they are making normative statements which they should be prepared to defend.

Is a subjective or objective view taken?

Preference theory takes autonomy (subjectivity) as its starting point, and by doing so severs economic analysis from many forms of objective evaluation. By subjective I mean that an agent's interests are not open to evaluation by other parties. A subjective approach means that a researcher is excluded from, for instance, claiming that certain interests are more worthy or valuable than others. Subjective models dominate modern economic analysis, but that hasn't always been the case. Early economists, including the Utilitarian J.S. Mill, drew a distinction between lower pleasures, those that bring contentment, and higher pleasures, those that are the goals of a developed human being. An idea which Mill summarized as: "It is better to be a human being dissatisfied than a pig satisfied; better to be Socrates dissatisfied than a fool satisfied." Modern economists have intentionally barred themselves from making similar pronouncements by sanctioning personal autonomy over every other value.

Objective models, for example, ones that evaluate acts on the basis of an agreed upon list of criterion, are generally anathema to most modern economists. Economics, particularly neoclassical economics, celebrates the individual's right to make good or bad decisions, based on his preferences, as he sees fit. However, as I have shown, self-interest is not necessarily a normative free position, and to the degree that economic models are normative, they imply the possibility of objective evaluations; a rational egoist, for instance, claims it is rational to be self-interested and irrational to act otherwise. However, the rational egoist presumably bases his definition of rationality on some objective choice. All rational people, Hobbes said, would agree that the preservation of life is a basic goal from which other

goals derive, while others have suggested the basic aims as the avoidance of pain or happiness. The point that I would like to make is that economists, by assuming self-interest, may unintentionally already be making objective evaluations. Perhaps their norms need to be made explicit and thereby subject to assessment?

To what degree are the interests of agents interconnected?

I have extensively discussed this theme in chapter 2 and throughout the manuscript as a major part of the external dimension of self-interest. The extremes of pure-egoism and pure-altruism are ideals and, as a practical matters go, they are largely self-defeating. However, between these two extremes there are many possible ways to model the interaction of agent interests. The fundamental message is that the interests of agents often include the interests of other agents. An explicit position on this point will help to resolve a major source of confusion in regards to what it means to be economically self-interested.

At what intensity are interests pursued?

By intensity I mean the vigor with which agents pursue their interests. The intensity of self-interest parallels the discussion within the economic literature on the rationality of agents. The agent who maximizes his self-interest in every instance is an illusion. Such an agent would be required to constantly perform rigorous calculations of expected net changes in interests across all potential outcomes; an unrealistic and probably unworkable description of human behavior. A more realistic, but consequently more difficult to implement description of behavior, is that agents satisfice across interests. Any number of intensities can be imagined and the context in which an agent finds himself will surely influence the energy with which he pursues his aims and goals. Therefore, even within a positive position such as psychological egoism, there is room for the expression of a large range of different intensities of self-interest. A maximizing egoist will behave differently from a non-maximizing egoist, therefore, a comment on the intensity with which interests are presumed to be pursued might be helpful to reaching a common understanding of agent behavior.

## Conclusions

The common theme running throughout this manuscript and emphasized in this chapter is that the assumption of self-interest, without qualification, doesn't significantly curtail potential reasons to act. Claiming or assuming that agents are self-interested without specifying characteristics such as the context in which agents are operating, how aims and goals are ordered, and what those aims and goals are, does little to constrain motivation or interpretations of resulting behavior.

Researchers need to take a consistent and coherent position about what they mean when they use the term self-interest. This will allow them to make meaningful statements about the behavior of agents and will allow others to contest those statements. This is a big step for economics which has been able to largely shield itself behind preference theory. In contrast to that theory, the approach I have outlined requires taking a position on what is going on inside the heads of agents. However, if a researcher's goal is to attach a label to an agent's motives, whatever the motive, I do not see that it is possible to avoid describing how that motive operates or should operate within an agent's psychology. The other part of the story are the objects at which self-interest is directed. Generally, the more identifiable aims in a context, the more complicated the motivations need to be in order to describe observed behavior. My objective has been to broaden the understanding of what self-interest entails and, particularly, to refute the idea that self-interest excludes other-regarding behavior. The result is a rough framework that is designed to get researchers to ask the right questions about the motive of self-interest.





## CHAPTER EIGHT: THE CASE OF OPEN SOURCE

---

### Introduction

Many explanations have been offered as to why people choose to contribute to open source projects. Explanations range from narrow, self-interested motives as found in many accounts of standard economic theory to non-interested, altruistic motives. The previous chapters aimed at building a theoretical overview of the complex topic of self-interest, this chapter aims to apply that overview to the case of open source. Using the theoretical framework developed over the previous chapters, I will attempt to add a coherent structure to the issue of determining the type of motivation exhibited by members of the open source community. The central question addressed is whether the motives of open source contributors are primarily economically self-interested.

### Practical Relevance

As a practical matter, does it matter whether the motives of open source contributors are properly identified? Open source communities, and open sources production methodologies, are a potentially rich resource for profit seeking firms. Indeed, many proprietary firms are searching for ways to use and incorporate the advantages of open source and other community based systems into a structured business environment, see, for example, the work of West and O'Mahony (2008); Dalle and Jullien (2003); West and Dedrick, (2001); Jeppesen and Frederiksen, (2006), von Hippel, (2005); Franke and Shah, (2003); Lakhani and von Hippel, (2003). If the arguments of Frey (1997) and Ryan and Deci (2000) are correct as I have outlined them in chapter one, then appeals to self-interested motives when non-self-interested motives are determining behavior will be counter-productive. As a consequence, it is important for potential users of open source to understand the motives of the open source developers. However, it is not just software developers that should be of interest to firms, rather, it is the entire community existing around an open source software project; including active members who use the software, provide documentation, offer help to other users, and suggest "wish lists" to developers of features they would like to have, which make open source software a product worth emulating.

Open source communities have proven to be robust and vibrant, for example, the Linux, Perl, Emacs and R communities continue to provide high-quality software

which is free to download, but which also comes with common open source license requirements as to rights and obligations regarding the use and redistribution of software derived from, and in some cases connected to, the software. Even lesser known projects such as those supported by the GFortran and Lisp communities continue to provide helpful support for these relatively old, yet powerful, software programs. Getting the most out of these opportunities will depend on properly identifying the motives of open source software community members, while getting it wrong may damage an important potential resource.

The remainder of this chapter is divided into several sections. All sections aim at determining the motives of open source programmers with assistance from the framework and general concepts considered in previous chapters. In general, evaluation of the social situation or context in which open source programmers operate is a critical feature in determining or estimating motivation. Where the social situation is broadly understood to include the circumstances in which the open source community exists and operates. It includes, for example, the institutions, history and important personalities which have contributed to its existence and viability.

The first section is a general discussion about open source and asks what makes open source an interesting case to discuss motivation. It is not a coincidence that the motives of open source programmers have been the subject of speculation since open source became academically interesting (Lerner and Tirole, 2001). Academics have been drawn to the case precisely because it appears to be an exception to the general understanding of motivation in the marketplace. I will then, briefly, discuss the basic concepts of open source. There are many detailed guides to open source software, several of the more interesting written by developers themselves (Raymond, 2001, Torvalds and Diamond, 2001), but a brief overview of the basics of open source, with particular attention to concepts which are of importance to my argument, will allow me to avoid potential areas of confusion. Over simplification of what it means to be open source has led to both over and underestimation of its academic and economic importance. Next, the general economic context will be discussed, specifically in relationship to the motives of firms that have lost in the race to become the dominant design. A review of the history of open source, particularly as it applies to licenses, follows. I will argue that licenses are an expression of the long term motives of their issuers. Open source has a long and well-documented history that provides further clues as to the intentions of the writers and users of open source licenses. Furthermore, I

will provide an overview of the main categories of open source licenses; seemingly small differences in licenses have large implications for the types of activities they allow and support. The differences in licenses are often neglected, causing some scholars to draw general conclusions about open source software when, in fact, their conclusions only apply to a subset of those licenses.

The overview and analysis will allow me to be more precise about the motives of open source contributors using the framework developed in the previous chapter. I would like to stress that the overview and details presented in this chapter are not peripheral, but, rather critical to identifying the motives of open source contributors. The social situation in which the open source community has operated and operates, provide guidelines to determine the intentions of open source users.

#### Open Source and motives

The open source software development process is a significant economic phenomenon with both practical and theoretical lessons for economic and business studies. The process brings together, in a sort of loose coalition, the talents of individuals and, more recently, firms, to produce high quality software that is, in many product categories, able to compete with functionally equivalent proprietary products. The open source software development process has been made more likely by the confluence of several technical developments including the advent of cheap, ubiquitous communications technology and modular software languages that can run on various computer platforms, e.g., the C computer programming language. The combination of these technical factors allows remotely located programmers to independently design and develop software which can then be rapidly and inexpensively exchanged and recombined to form viable software programs. While technical characteristics define important pre-conditions that have made the development of open source more likely, they alone are not sufficient to explain its uniqueness given that proprietary software firms have access to, and use, the same techniques. The additional factor, and the one that is the most significant and theoretically interesting, is the motives of open source software programmers and how those motives are expressed.

Lerner and Tirole (2001) were the first academics to ask what motivates highly skilled programmers to contribute their efforts and resources to developing software which is then made available to anyone under very generous licensing

agreements. More specifically, the question which I address in this chapter is whether the economic conception of the motive of self-interest, as understood by Lerner and Tirole and other economic writers, can explain the willingness of programmers to contribute to open source projects. The task at hand is not to attempt to dissect all of the various reasons that programmers have to contribute to open source projects as in the Floss (2002) study, rather, I am primarily interested in answering the much more modest question of whether economic self-interest can explain such behavior.

Open source software is also an important economic phenomenon in a highly competitive, billion dollar market, particularly in markets for the software that runs on servers and behind Internet applications. It is expected to play an increasingly important part in other software markets such as data base management (MySQL) and the software for portable telephones (Android). Large hardware firms such as Sun Microsystems, Hewlett-Packard and particularly IBM, have made large software contributions to the open source community, and many other firms, including Oracle, now offer servers running open source software. Microsoft was quick to see open source as a potential threat and wrote a famous position paper outlining the threats open source posed to its business model (see Halloween Documents, and Young, Chap. 8 in DiBona, 1999). The economic significance of open source alone requires a coherent and consistent response on the part of economic and business analysts.

The motives of open source programmers have been a subject of speculation and debate since its first appearance. For instance, I have previously mentioned the important work of Raymond and Lerner and Tirole, furthermore, von Krogh and von Hippel (2003), editors of a special issue on open source, were careful to categorize different sorts of "individual motives" in their overview article in the special addition.<sup>14</sup> The individual motives mentioned by von Krogh and von Hippel closely coincide with Raymond's (2001) earlier, and important "musings" on the ethos' of many of the original computer programmers, some of whom later contributed to open source projects. I believe that Raymond's views shaped the popular impression of the motives of open source programmers. According to Raymond (2001), there is a strong sense of community among programmers and the feeling that they need to work together to solve issues. Von Krogh and von

---

14 I believe that the use of the term individual motives is particularly helpful, at their most basic level, all motives are individual motives, there is no group mind.

Hippel also mention the "non-individualistic" motive of protecting open source software, which is certainly a concern of many in the open source community, but one that I would suggest is a secondary motive in that it does not explain what motivates programmers to contribute their code to open source projects in the first place. In other words, programmers might be more willing to contribute to open source projects that offer their software protection against, say, copyright infringement, but it does not explain why programmers decide to contribute to open source projects rather than proprietary project. Copyright laws protect both proprietary and open source software (Jacobsen v. Katzer (2008)).

In more general terms, arguments have been made that contributors have motives ranging from ideological concerns for issues such as the freedom of information and fairness (Stallman, 1999), to economic strategies designed in response to barriers resulting from network externalities which are prevalent in software markets (Rossi, 2004). Although some of these motives appear to correspond to standard economic assumptions about self-interest, others stretch that definition beyond reasonable recognition. Open source contributors, like everybody else, are almost certainly motivated by their interests, but the question is whether those interests should be classified and understood to be primarily economic or something else entirely. Lerner and Tirole (2002) present the clearest defense of what they take to be economic interests, their position will therefore be the focal point in determining the degree to which open source contributors are economically motivated. In keeping with a main theme of this thesis, I will suggest that much of the confusion on the issue of whether open source programmers are motivated by economic self-interest or some other motive is due to a misunderstanding of what the motive of self-interest entails.

One of the principal practical lessons of the preceding chapters is that assigning motives to actions is a difficult and an invariably controversial undertaking, one that requires a detailed and deep analysis of the context and activities under consideration. Open source software has several characteristics which make it a particularly interesting case to explore issues raised in previous chapters, I will briefly mention several of its more important characteristics.

Open source licenses themselves provide a reliable and consistent source of evidence of the motives of open source programmers. Software licenses represent clear, long-term, commitments to specific goals on the part of their issuers. I would argue that such commitments by open source licensors are every bit as

important as the commitments described by Ghemawat (1991). Issuing a software license for an important piece of code is not something that is taken lightly given the expectations on the part of users that are incurred once a product has been released under a particular license. The open source initiative (OSI) is tasked with determining whether a license submitted to them is an open source license. They offer advice and guidance to potential license issuers via an active mailing list and freely available literature so that both the positive and negative implications of using open source licenses can be clearly understood. In addition, there is wide variation in the types of open source licenses available, with a license to meet any goal. Its economic importance and the significance of the licensing discussion allow me to claim that open source licenses provide a reliable indication of the underlying intentions of licensors.

Finally, open source is an interesting case because it is going through a rapid development process. Given its open nature, researchers can watch the development of authority structures, rules and norms of behavior, and are thereby privy, via newsgroups and other on-line forums, to the internal stresses that the open source community experiences as new participants join and new open source licenses are issued. These stresses, as represented by the different interests of open source contributors, are an important reoccurring theme that I will be addressing. I will argue that the community was long ago transformed from a loosely connected group of idealistic computer programmers, highly motivated to change the way information was disseminated, to a much more diverse groups of participants including larger, profit motivated, computer firms.

That said, the uniqueness of open source software can also be seen as a drawback to its usefulness as an object of study. I know of no other commercial product that is developed and sold in a competitive market with characteristics similar to open source software. There are many examples of individuals volunteering and performing services or providing products for disinterested reasons or for reasons other than promoting their own economic interests, but no other service or product is produced and supported by groups of organized, skilled, workers who then make their products freely available in a highly competitive market under liberal licensing conditions. Many people in modern economies choose to work for reasons other than profit maximization (Florida, 2002), that happy occurrence is not the issue I am addressing, my concern is whether narrow economic interests are being represented or some other interested or perhaps non-interested motivation is at work in open source communities. The strength of the

case of open source is that it is a possible indication of the diverse range of motives that have always been present in economic contexts. People, whether acting as economic agents or not, are motivated by a complex mixture of motives, the case of open source, I believe, is an indication of that complexity within an economic environment. At a practical level, a firm or economic policy that could tap into such motives would yield potentially large economic benefits. The uniqueness of open source refers to its ability to provide evidence that non-self-interested motives play an important role in a highly competitive market. It is its transparency that makes it of interest. Whether contributors to open source are more or less self-interested than other economic agents is not addressed.

Determining the motives behind actions is not easy, but, as I have mentioned, several indicators provide evidence that helps to decipher the motives of open source contributors. I will use three indicators to assess such motives. The first is the structure of the software market. While not an obvious choice to determine motivation, it turns out to be a crucial determinant of the motivation of profit maximizing firms to contribute to open source projects. The second feature is the history of the UNIX operating system; that history led to many important lessons being learned and later applied to Linux, the most important open source project. Finally, and most importantly, I will examine the open source licenses themselves because they provide a good indication of the intentions of open source contributors. By this means I hope to arrive at a better understanding of the structural and external dimensions motivating open source programmers. However, before discussing these items, I will present a brief introduction to open source.

### Basics of Open Source

Before examining the details of open source, it is helpful to begin by giving a very brief description of how a typical open source project gets started and how open source licenses differ from their proprietary equivalents. A typical software development project begins with a computer programmer needing to solve a problem (Raymond, 2001). A programmer writes a piece of code that will make, say, some repetitive task such as data manipulation, easier. Perhaps the project begins to grow too complex for the programmer to complete alone or the programmer simply loses interest in the project. The programmer may then decide that he needs help in order to complete the project. By posting a message on freshmeat.net or some other public site that caters to software development

projects, the programmer may be able to interest others to contribute to the project or taking it over entirely. The programmer uploads his piece of code to the site where it is typically classified by name, type of product, operating systems on which it runs, etc. . The project will later be rated in terms of its popularity within the community, turnover, and other measures of its success. At this point a programmer faces a critical decision, as the copyright holder of the code he has to make a decision about the conditions under which he wishes to license his code. The choice of license determines the degree to which programmers and other users can examine, change and add to the software project, as well as any other conditions the programmer would like to specify. The most liberal of these licenses are called open source licenses.

How do open source licenses compare to proprietary software licenses? Closed or proprietary software licenses typically attempt to keep the code hidden from users so that it cannot be examined or changed, copying the code is forbidden or limited. In addition, licensees are often required to pay a fee to use the software; users do not own the software they use, they are essentially renting it. The proprietary model of software development is currently the primary model of software distribution. Firms like Microsoft hire programmers to code software that the firm owns and licenses to users for a fee. Microsoft retains ownership of the software. All major traditional research streams in economics and business support the basic idea that firms should retain and protect their valuable resources. The assumption is that the rent generating assets of a firm must be owned by that firm (Penrose, 1995; Teece, 1980, Wernerfelt, 1984, overview in Foss, 1997). In other words, to generate rents, the assets in question must be owned and controlled by the firm. However, the traditional links between outputs, ownership, and control of assets have been questioned, where the term assets is widely defined to include both physical and non-physical value generating (Child and McGraph, 2001, Hennart, 1994)

The contrasts between different open source licenses and proprietary licenses are critical to understanding the diverse set of motives of open source developers. Open source licenses, in contrast to proprietary licenses, often include provisions that allow users to copy, change and redistribute the program's source code. The Free Software Foundation lists four basic principles of open source software: 1. The freedom to run the software for any purpose; 2. The freedom to study how the software works and adapt it to fit one's needs; 3. The freedom to redistribute copies of the software, and; 4. The freedom to improve the software and distribute



the improvements to the public (FSF). Licenses in general provide a mechanism for firms to fine-tune the degree of ownership that they wish to exercise over an asset, making ownership an important strategic variable. By using open source software users are entering an agreement, whether they have entered into a formal contract is debatable, but the common understanding is that users must abide by certain rules (ProCD, 1966). The main stipulation of some open source licenses is that if users exercise their right to add to or modify the code, then that altered program must also be released under the same or similar open source license. This requirement prevents people from taking open source software and then re-releasing it under a proprietary license unless *all* of the code released is covered under the open source license. Note that it is possible, and quite common, to use open source code without contributing anything to the community. The obligations are generally believed to come into effect only if the code is modified or linked to open source software and then redistributed.

The choice of a license can be particularly important in an industry exhibiting network effects because there is a tendency for the market to "tip" in favor of one, eventually dominating, standard. By varying the degree of ownership of an asset, a firm can potentially influence the direction that a market tips or survive in a market that has already tipped; a topic to which I now turn.

#### Economic Context: Networks

Software is an example of a product that exhibits network externalities and the single most important feature of such markets is that they can be tipped in favor of a particular product or system of products. I will argue that networks play an important role in the motives of firms to contribute to open source projects. Due to clever foresight on the part of Bill Gates, the market for the desktop operating system tipped in Microsoft's favor while the lucrative server market is beginning to go the same way (Frieberger and Swaine, 2000).

The computer software market, and many other markets, exhibit what are known as increasing returns on the demand side. For products in such markets the utility that a user derives from consumption of a good increases with the number of other agents consuming the good (Katz and Shapiro, 1985). The properties of increasing returns are: market instability, unpredictability, the possible predominance of an inferior product, and immense profits for winners (Arthur, 1994). Many strategic issues arise under such conditions and must be incorporated into a firm's strategy,

for example: (1) Should a firm move first to exploit a market with network externalities or follow the leader?; (2) Should a firm form alliances to exploit markets and when should it go at it alone?; (3) Should a firm work within a given standard or create a new standard? (Farrell and Saloner, 1992; Katz and Shapiro, 1985, 1994; Besen and Farrell, 1994; Church and Gandal, 1992). The second and third items are of particular concern to firms considering using or contributing to open source products. I claim that for proprietary software firms, open source is a loser's response to the Microsoft dominance. The reactions of Microsoft officials to the gains of Linux as well as the professed strategic intent of Linux distributors confirm this assumption (see Halloween Documents, Young, Chap. 8 in DiBona, 1999).

The network characteristics of the software market mean that compatibility issues play an important role in consumer decisions. Once a market tips in favor of one product switching costs become an especially important factor. Switching costs result from a consumer's desire for compatibility between current and future purchases and previous investments. For instance, Microsoft's dominance of the desktop computer means that the great majority of software can run only on the Microsoft system; switching to a new operating system requires current Microsoft users to buy new software as well. Therefore, any standard that hopes to overcome Microsoft's monopoly will have to offer users considerable benefits, financial or otherwise. The strategy that the open source community, and specifically firms, has chosen to challenge Microsoft is similar to the strategy that Sun Microsystems used at one time to establish a dominant design in the computer chip market (Khazam and Mowery, 1993). This strategy consists of adopting a liberal licensing policy that effectively concedes control over intellectual property in order to gain market share. In addition, to the degree that liberal licenses require that the licensor make his code publicly available, the ability to realize profits on those particular products correspondingly falls.

Many of Microsoft's competitors are following a loser strategy, by which I mean the market has tipped in Microsoft's favor and these firms have decided to support open source software as a low cost alternative to Microsoft products in order to break through the barriers caused by Microsoft's dominance in the market. Licenses give copyright holders, in this case firms, the ability to finely adjust the amount of control over the products they release to the public. By making their software more open they make it easier for others to copy and change the product and, thereby, encourage rapid product dissemination. The downside of an open

strategy is that it severely constrains the ability of a firm to control the development process and realize profits from their products. Some firms, for instance Netscape, have released their products using a dual licensing approach that means that they release both an open source and a proprietary version. Open source strategies, if consumers accept such products, may force Microsoft to change its strategy, perhaps through releasing some of its software under liberal, open source like licenses. However, addressing all of the pros and cons of the different open source business strategies would take me too far away from my main theme which is motivation (Meeker, 2008). Microsoft, before Linux, was faced with a choice between setting a lower price for its products in order to capture or maintain its market share and setting a higher price to harvest profits by exploiting current locked-in customers Klemperer (1995). Open source, by removing or reducing the importance of the issue of purchase cost, may force Microsoft to lower its prices or face losing its dominant position.

I take it for granted that firms are motivated by the need to maximize profits within legal constraints; their stockholders demand it and the law in many countries supports those demands (Millstein and MacAvoy, 1998). I strongly suspect that the nature of network externalities in these markets means that many firms have decided that they either won't or can't profitably compete directly against the dominance of Microsoft and have therefore decided to support open source software in order to provide an alternative to Microsoft software. Releasing proprietary software to the open source community has the possible advantage of overcoming or circumventing the dominant design. By helping to create a pool of freely available software, these firms shift the competitive playing field away from software towards providing services or hardware. The idea, presumably, is that by making a piece of software free or nearly free, the price of that particular piece of software will no longer be a consideration in the decision as to which computer system to buy. For example, by helping to create a competitive open source alternative to Microsoft's operating system, hardware firms and software firms which make complementary products that work with the open sourced operating system will be able to take away or reduces the advantages of Microsoft's lead in that area.

The battle to control the desktop and server software markets is worth billions of dollars, while the networked nature of the computer market means that the standard has, for the moment, tipped in favor of Microsoft's operating systems. Whether operating systems based on the Linux kernel will be able to challenge the

dominance of Microsoft operating system is a matter of intense debate and speculation. Much of the debate involves the minutia of comparing the total cost of operating Linux versus Microsoft. Licenses that encourage proprietary firms to contribute their efforts to open source projects, while allowing these firms to reach a wider audience, might provide a powerful enough combination to overcome the immense switching costs involved in changing operating systems. I won't concern myself with issue of operating cost versus total costs, a topic that is best left to systems administrators and chief information officers, however, what is not debated is that open source products provide a cheap, up-front, alternative to Microsoft products and are therefore an important component for profit seeking firms hoping to challenge the dominant design.

## History of Open Source

### Learning from the Fragmentation of UNIX

The underlying network structure of hardware and software markets means that the history of the computer sector is relatively repetitive and therefore partially predictable, similar opportunities and constraints often reappear. Therefore, it is not surprising that the issues facing firms in the personal computer and servers markets today are very similar to those that firms in the mainframe market faced more than thirty years ago. At that time, IBM had near monopolistic control of the mainframe sector, while the "seven dwarfs" (Sperry Rand, Burroughs, NCR, RCA, Honeywell, General Electric and CDC) struggled to survive by finding safe niches from which to compete (Campbell-Kelly and Aspary, 1996). General Electric, in particular, decided to build a mainframe computer that would use a time-sharing system, a market in which IBM was weak (Campbell-Kelly and Aspary, 1996). They joined with MIT and AT&T to build the Multics (Multiplexed Information and Computing Service) system, although Multics was a failure because it was too technically complex, it led to the development of the smaller, more elegant UNIX operating system. The structure of UNIX is simplistic, almost everything in UNIX can be thought of as a file which gives the entire system a modular form that, in the jargon, allows easy "scaleability", a feature that would only later show its true value (see Kernighan and Pike, 1984). Another important innovation was that UNIX was written in the language C, which allows it to be portable to computers from different manufacturers. Previously operating systems were written for specific computers, a time-consuming, difficult task.

However, there arose a fundamental contradiction in AT&T's strategy that would later repeat itself in other computer markets. Network externalities mean that the more people used UNIX the better it was for AT&T but, in order to draw away users from IBM, AT&T would have to offer UNIX under license conditions attractive enough to make it a viable alternative to IBM's systems. Although the precise role that AT&T's UNIX played in breaking the domination of IBM is not entirely clear, (IBM's fate was probably more influenced by the charges of monopoly it was facing from the U.S. Government), what is clear is that UNIX became the standard for large mainframes. However, while UNIX became the standard, AT&T did not profit from its dominance. In short, AT&T lost control of its proprietary UNIX system in part because it used a license that was too liberal. I will explore the details of this story in the following paragraphs because of its direct implications for open source software.

UNIX is written in the C programming language, a language especially developed for the purpose of building UNIX, but one that is still used today because of its speed and flexibility. C is special because it is system-independent; in other words, it provided programmers the possibility to write software that can run on multiple systems or platforms as opposed to having to write programs specifically for a particular computer system. Because UNIX was written in C, it too could be used across platforms. For the first time, and in contrast to IBM's operating systems, an operating system was written in a language that could be "ported" to different computers. Independent, hobby programmers began using, sharing and improving tools that could be used on computers running UNIX, especially when Usenets began operating in 1980. These communities of programmers would later form the basis of the hacker culture that would help support the development of Linux (Raymond, 2001). The development phase of UNIX was marked by a spirit of cooperation between AT&T's Bell Laboratories and universities, each benefiting from the other's expertise, Bell in terms of the insights gained from the application of their operating system on a large scale and access to some of the best programmers in the world, while the universities gained in terms of the advanced research and the administrative role of Bell Laboratories (McKusick, Chap. 2 in DiBona, 1999).

The spirit of cooperation between Bell Labs and the University of California at Berkeley was of particular importance. That relationship continued throughout the 1970's until the distribution of UNIX was taken away from the Labs and commercialized by another group from AT&T (at which time Berkeley took over

Bell's role of chief researcher). The split between AT&T, which predictably and justifiably wanted to profit from its years of investment in the operating system, and researchers at Berkeley, whose primary concern was research, led to conflict. The result was that UNIX split into two versions, a commercial version sold by AT&T and a Berkeley version that was licensed under very liberal conditions; essentially anyone could get a copy of the program who was willing to take the time to download it from an Internet FTP site (McKusick, Chap. 2 in DiBona, 1999). However, the slitting or "forking" of the operating system was not over. The Berkeley version later split into the NetBSD (Berkeley Software Distribution) group, which emphasized porting the system across as many platforms as possible and continuing research, and FreeBSD, which supported the PC architecture. OpenBSD later spun off from NetBSD with the aim of improving the security and ease of use of the system. In addition to these free distributions, Berkeley's BSDI began selling a commercial version of its UNIX-like operating system, which AT&T would try, largely unsuccessfully, to sue for copyright infringement. The settlement in that case resulted in yet another split of the BSD system, one version which required an AT&T source license, and another version which allowed BSDI to freely redistribute source and binary forms subject only to the condition that the university's copyrights remain intact and the university receive credit when others used the code (the last condition was later rescinded by the university).

In addition to the forking of AT&T's and Berkeley's versions of UNIX, commercial hardware companies such as Hewlett-Packard, IBM and Sun Microsystems would eventually issue their own, largely incompatible versions of UNIX. By this time AT&T had lost control of its operating system. Attempts to recombine this jumble of UNIX versions resulted in the creation of a single, common UNIX specification, which was organized by the IEEE in 1986 under the name POSIX. Its administrative efforts have produced an evolving set of standards that provide a basic, common set of UNIX code which in principle can run across all UNIX based computers. Despite these efforts, the fragmentation of UNIX into small proprietary domains continues today as firms search for a balance between market share and rents. The fragmented nature of the market eventually gave Microsoft the opportunity to try to tip the market in its favor. Microsoft's NT and Windows Servers operating systems are now strong competitors to the different UNIX operating systems in part because they offer users stability, in other words, buyers know that Microsoft will maintain a single, proprietary standard. The story of UNIX is largely one of forking and fragmentation. Both AT&T and Berkeley lost control of their versions of the system, allowing Microsoft easy access to the

market. However, Microsoft would soon come under attack from a distant cousin of UNIX, but this time from a largely united version of the operating systems, a unity that derives from the license under which it is distributed.

To summarize this section, the main lesson that the open source community was to learn from and apply to latter projects was to avoid the damaging fragmentation that came about because of the different types of UNIX on the market. What became the Linux community, and later the open source community, realized that it needed to create one standard operating system that would be available to everyone and work across all major computer systems. On the positive side, the software community centered on the efforts of the Berkeley group had shown that it could organize itself to such a degree that it could develop and maintain even a very complex software project, the portability of UNIX was a great help in that respect. However, the community had also begun to realize the limits of its governance forms and its relationship with proprietary firms, both limits would be addressed with the development of Linux (Raymond, 2001).

#### History of Linux

While the major computer manufacturers such as Sun, IBM and Hewlett-Packard, were fighting over an increasingly divided computer server market, the personal computer market was beginning to develop as an important sector. The vision of Alan Kay, a computer scientist at Apple Computers, to build an easy-to-use computer for the masses was beginning to be realized in the middle 1980's (Hiltzik, 1999). Although Apple was successful in building a series of personal computers running its own operating system, by a series of lucky coincidences and great foresight and initiative by Bill Gates, MS-DOS came to dominate the desktop system (Frieberger and Swaine, 2000). Network externalities then helped Microsoft to maintain its dominance of the desktop. In the middle 1990's, Linus Torvalds would attempt to bring the UNIX and personal computer worlds together through the creation of a smaller, UNIX-like, operating system for the desktop. Torvalds was not the first to attempt to port an UNIX-like operating system to a desktop computer; in the early 1990s, Andy Tanenbaum at the Vrije Universiteit in Amsterdam was developing Minix as a learning tool for his operating systems class, Richard Stallman was developing the HURD operating system as a free alternative to the expensive UNIX, and FreeBSD was porting a version of the Berkeley UNIX to the desktop. Each of these other attempts failed, why did Linux succeed?

The Linux operating system is the most famous, popular, example of a successful open source project. The origins of the Linux operating system was simplicity itself; Linus Torvalds, a Finish computer science student, wanted to develop a compact, UNIX-like version of the operating system for use on his desktop computer. Versions of UNIX were then the standard operating systems for large, mainframe computers used by many businesses, universities and research centers. One of Torvalds's first steps was to request a copy of the POSIX standards so that his kernel (the core of an operating system) would work with the programs that were currently available, in other words, to use the knowledge base that had collected around the UNIX operating system. The ability to re-use previously written programs was made possible because, as mentioned previously, UNIX had been written in the universal language C. While Torvalds's request for the POSIX standards resulted in the disappointing answer that he would have to pay to get them, a sympathetic administrator decided to place Torvalds's code on an FTP site. As a result, his code was placed on a publicly accessible Internet site that would eventually act as a clearinghouse for improvements and additions to what was to be called the Linux operating system. The growth of Linux was enormous, within three years, with the help of a large development community, an official release of Linux was issued.

There are three factors accounting for the success of Linux versus the other UNIX-like desktop operating systems. First, governance by meritocracy; Torvalds was willing to release large chunks of development code to competent developers (Moody, 2001: p. 84). Second, Torvalds chose for a simpler, older, operating system design rather than a state-of-the-art design. Thirdly, Linux was eventually released under the GNU open source license. I will discuss each point in turn because together they informed many latter open source projects. Governance by meritocracy guaranteed that the best and most enthusiastic programmers took responsibility for major parts of the coding of Linux. Linux was the first project for which a conscious and successful effort was made to use the entire on-line world as its talent pool (Raymond, 1999: p.51). This is in contrast with many other projects that were run "... in a carefully coordinated way by a relatively small, tightly-knit collective of people (Raymond, 1999: p. 16)." An example of the type of project to which Raymond is referring is Perl, an open source project whose management is closely controlled by Larry Wall. Although Torvalds maintained control of the Linux kernel, he knew when to cede control to trusted associates. The issue of control plays a surprising role in the way open source projects are run.



Computer code is released into the community upon which an informal review process takes place. For instance, Linux kernel questions are handled by Torvalds while others supervise other parts of the operating system. The Linux kernel forms the center of the Linux system, in other words, other programs must interact with the kernel and therefore Torvalds theoretically has final say about what gets added to the system in the sense that it must be able to interact with his kernel. In fact, it would be impossible for Torvalds to micro-manage the thousands of projects currently under development to such an extent. Therefore, major projects, such as building a movie player for Linux or a security system, are managed by others. The supervisor of a project does not have absolute power because of the liberal nature of the open source license. Open source licenses guarantee that everybody has equal access to the same basic code base of a project. The ultimate power therefore rests with the programmer's satisfaction with a project. For example, if a programmer doesn't like the direction that a project is heading, he simply takes the freely available common code base and begins his own project in a process known as forking. The GNU licenses makes projects easily susceptible to forking by requiring that the source code of a project is readily available, i.e., a rival project has access to the entire inner workings of a project and faces no legal restrictions if it wants to change the program, provided that it too provides access to source code with its new program. A precarious balance is thus maintained, no project manager can afford to alienate a workforce that can so easily start a competing project. In this manner, projects and project leaders informally compete with one another for the programming resources of the community.

Torvalds' second important decision was to use an established operating system design known as a monolithic kernel rather than a more modern micro kernel (Raymond, 1999). Monolithic kernels were considered obsolete because they place too much burden on the kernel in terms of the number of jobs it must do. Micro-kernels perform a much smaller set of operations and are so designed as to be less hardware specific, in other words, they are easily portable to other computer systems. When Torvalds first issued his request for help on the FTP site there ensued a now famous series of emails between Torvalds and supporters of the mono-kernel on one side, and the operating systems academician Andy Tanenbaum, the developer of the Linux rival Minix, on the other (debate in Appendix A, DiBona, et al., 1999). The intricacies of the debate, while interesting, are not important for the questions in this manuscript, it is enough to realize that the simplicity of Linux has helped to make it the most widely ported operating system available for the PC. While the Minix kernel was technically superior to

Linux, its failure to become widely used was, in part, due to its complexity. The HURD kernel project by Richard Stallman, also suffered in part because of its over-complexity (see Stallman, chap. 4 in DiBona, 1999).

While the governance structure and technical characteristics of Linux were important to its success, the license under which Linux was eventually released guaranteed its success by finding the proper balance between openness and proprietary concerns. Torvalds initially rejected the GNU-GPL license in favor of the following abridged conditions: 1. Full source must be available (and free), if not with the distribution then at least on asking for it. 2. Copyright notices must be intact, meaning that the copyright notices must accompany the code. Small partial excerpts of code may be copied without bothering with the copyrights. 3. The software may not be distributed for a fee, not even "handling" costs (See Moody, 2001: p. 44). The third provision was the major exception to the GNU-GPL, Torvalds was probably, in part, responding to the fact that Tanenbaum had charged a fee for Minix. Although Torvalds initially found the GNU-GPL too restrictive, he eventually changed his mind because he realized that by denying fellow Linux developers the right to charge a fee he was restricting the distribution of his program (Moody, 2001: p. 45). Distribution, of even free software, costs time and resources, accordingly, Torvalds decided to release his code under the GNU-GPL, allowing companies such as Redhat, Caldera and SuSE to eventually mass produce the Linux kernel along with other GNU-GPLed software. In contrast to popular understanding of open source software, the GNU-GPL explicitly allows license holders to charge a fee for their products. However, the condition that the full source code is available to anyone who asks for it sets a limit on how much a programmer can reasonably expect to receive for his contributions.

## GNU-GPL

As a starting point for further discussion of open source licenses I will summarize the GNU-GPL open source software license (GNU-General Public License, GNU is an apparently clever, self-referencing definition which stands for GNU is not UNIX). The GNU-GPL was one of the first open source licenses, and it serves as the blueprint for many of the open source licenses followed. The GNU-GPL is only around seven typed written pages long and, as far as legal documents go, it is easy to understand because it is written in a simple, non-legalistic, style. Its simplicity, or perhaps over-simplicity, has meant that the document is not as thorough as it should be given the increasing importance of GNU-GPLed software.

For instance, issues such as whether the GNU-GPL can be considered a contract even though no formal acceptance procedure is required to use the software, the details of the difference between a new piece of code versus derived works, and the degree to which pieces of non-GNU-GPLed code are allowed to work with GNU-GPL code, have all been points of contention in the open source community. The result is that the GNU-GPL is an evolving document that is continuously being re-written to incorporate new concerns and resolve certain ambiguities. In addition, the GNU-GPL is more than just a legal document, it is also a political statement with what amounts to a manifesto defending the public's right to change and share software. Richard Stallman and Eben Moglen, the authors of the GNU-GPL, wrote in the preamble to the GNU-GPL that their primary objective in creating the license was to address the following issue: "The licenses for most software are designed to take away your freedom to share and change it. By contrast, the GNU General Public License is intended to guarantee your freedom to share and change free software-to make sure the software is free for all its users." The political manifesto has created controversy, many open source users and contributors resent its proselytizing and anti-business tone; these issues were addressed in later open source licenses, however, before discussing those, I will first briefly discuss the difference between free or public domain software and GNU-GPL software and provide some important examples of open source software.

The ambiguity of the word "free" in the GNU-GPL has led to one of the biggest misconceptions about open source software, a misconception that is still committed by some scholars writing about open source. The fact that many of these licenses allow developers to charge for "free" software has caused confusion, and arises out of limits of the English language. The open software community never tires of reminding people, that the "free" in free software refers to the freedom to be able to use, understand and change a piece of software, and should not be confused with the meaning of free as in, for instance, the phrase "free beer". In fact, free software can be, and often is sold, as I have previously suggested, Torvalds's acceptance of this provision was one of the reasons that Linux has become such a success. One of the most important lessons learned from the history of UNIX by the open source community and addressed in open source licenses is that free software needs to be protected by licenses that forbid others from taking code that is freely available and "hijacking" it for proprietary purposes. To do that, many open source licenses are explicitly not "free" in that they require licensees to use open source approved licenses rather than proprietary licenses if they wish to redistribute modified code-- I will return to this important issue shortly. The company Redhat, for instance,

collects open source software, adds some software that eases installation and update procedures, places the software in easy to use computer packages, and then sells to the public. They, and many other companies, make money from software that can be downloaded for free by making it easier to use or providing support services to open source users. Of the fifty most popular down-loadable programs on the CNET site there are many programs listed under the heading of "free" software. However, there the term refers to the price, at least the initial price of using the software. In general, the free software that one finds on the Internet often comes with strings attached. The software may be a diminished version of the proprietary program, for instance, or it may have fewer functions or less storage space. In general, either the quality or the amount of time over which the product can be used is severely limited. Much of the software is cost free only if it is used as an add-on to a product for which the customer has already paid. For instance, free software available on CNET from Microsoft can only be used in conjunction with the proprietary Microsoft operating system. Of the fifty "free" software programs I recently examined on the CNET site which were available to download, all but three appear to have some hidden charge. For example, a large percent, around a third, mainly act as platforms for advertisers, i.e., users are subjected to a large and constant amount of advertising when using the program, and many of the ads refer to costly upgrades of the program in use. Of the very few programs that are free, perhaps three, two can be used to avoid annoying pop-up advertisements, and one is issued by CNET itself. Many of the programs are also very difficult to dispose of once installed. In short, nearly all of these free software programs have some sort of severe limitation. However, the biggest limitation, from the open community's perspective, is that none of the programs provides the source code necessary to understand and change the program.

Finally, before turning to the subtleties of open source licenses, I think it is important to briefly remind the reader of the role that first free and then open source software has played in the development of computing and, as an example, the Internet. Anyone who has used the Internet has used open source technologies. For example, when your modem connects to your Internet server, it is probably using the Transmission Control Protocol/Internet Protocol (TCP/IP), network protocols that determine how packages of information are split into pieces and distributed; both protocols were developed openly. The Domain Name System (DNS), which translates the names you type (e.g., Amazon.com) into IP addresses, was also developed through open source processes. The most critical component of the Internet, the Berkeley Internet Name Daemon (BIND), the piece of software

that runs the DNS, is also open source. Even browsers by Netscape and Microsoft are based on the open source protocols developed by Tim Berners-Lee (see, O'Reilly, chap. 12 in DiBona, 1999). Finally, the web server that you connect to is probably running open source Apache software and the scripts that it in turn runs are likely to be written in an open source scripting languages such as Perl, Python, and Tcl.

The GNU-GPL license contains thirteen clauses, the most important of which concern issues relating to the copying, modification and fees of open source software. The essence of the copying clause is that a person is allowed to modify or copy licensed programs with the condition that recipients receive the same rights. A person can copy all or parts of a program provided copyright notices are included and the changes are noted and dated. Each time a program is distributed the recipient automatically receives the same rights and no further restrictions can be added. In contrast to the original license that Torvalds used, users are explicitly allowed to charge a fee for the program. One of the stricter requirements of the GNU-GPL is the condition that when the program or parts of the program are attached to other programs, the new, derived work must also be covered by the GNU-GPL license. However, mere compilations of works which include a GNU-GPLed program are not required to be under the same license. What constitutes a derived work is not entirely clear in the GNU-GPL, but this issue has been addressed in newer open source licenses. For example, Torvalds modified the part of the GNU-GPL related to derived works as they apply to Linux; he allowed programs that use kernel services by normal system calls, in other words, programs that pass messages to the kernel, to be exempt from the requirement of falling under the GNU-GPL. One of the primary lessons that Richard Stallman learned from the history of UNIX was the importance of preventing hijacking. Hijacking occurs when public code is made private. For instance, a person writes an interesting piece of code, places it in the public domain, and a firm copies it and places it in its entirety or by combining it with pieces of their own propriety code under a proprietary license thereby removing it from the public domain. The GNU-GPL overcomes this problem by first licensing software with the GNU-GPL instead of issuing the code without a license and requiring software that works with GNU-GPL software to be licensed as GNU-GPL software. While the GNU-GPL has been very good at preventing proprietary firms from hijacking open source projects, some members of the open source software community were concerned that the political manifesto and the perceived anti-business tone of the GNU-GPL were retarding the growth of open source software. In reaction, they decided that

they could do better and wrote their own definition of open source and called it the Open Source Definition (OSD). In keeping with the basic spirit of the open source community, the OSD includes many of the same ideas of the GNU-GPL and is considered a derivative of the ideas contained in that license (Perens, Chap. 11 in DiBona, 1999).

The OSD is not a license, it is a set of guidelines clarifying the necessary provisions that are required to earn the designation of open source software. The OSD essentially broadens the definition of open source beyond the restrictive conditions of the GNU-GPL. In addition to the GNU-GPL, the BSD license, the X Consortium, artistic licenses, etc., are all certified to be open source licenses. Attempts to broaden definitions of what it means to be open source have raised concerns among some members of the community that safeguards built into the GNU-GPL to prevent fragmentation of the code base will be circumvented. As previously shown, fragmentation was the issue that severely damaged the portability of UNIX and led to the development of several competing “flavors” of UNIX. On the other hand, the new open source licenses covered under the Open Source Definition have undoubtedly been the cause for an increase in the participation of proprietary firms, including, IBM, Apple, Netscape and HP, in open source development projects. The clauses of the Open Source Definition cover the same issues as the GNU-GPL, namely, copying, modifications, and fees. Modifications, or derived works, are allowed, but rather than being required to use the same license as the original program, the derived work can be licensed under another set of terms. For instance, a company such as Netscape can take open source software, modify it, and then release it as proprietary software, something that was not allowed under the GNU-GPL, although it still must make the code available to anyone upon request. In fact, a program can be issued under multiple licenses, some of which are proprietary while others are open source. The idea is to encourage the participation of those firms who want to use and contribute to open source projects and allow them to profit from their efforts. The approach appears to have been successful, for instance, Apple Computers, a company notorious for its closed architecture, announced in 1999 that it would release parts of the Mac Open Source X Server as an open source BSD system called Darwin. IBM chose the open-source Apache web server to support some of its most important products, as has Sun Microsystems, Netscape and many other corporations ([www.opensource.org/docs/products.html](http://www.opensource.org/docs/products.html)).

For the purposes of completeness, I will list some of the general provisions of open source licenses. By modifying or distributing a program covered by the GNU-GPL, it has been argued that the user is indicating an acceptance of the license. In short, users do not have to sign a license agreement or actively agree to the terms of the contract by, say, clicking on an "I Accept" button. This passive type of acceptance has not yet been fully challenged in court and its validity is a cause for concern within the community. Any attempt to copy, modify, sub-license, or distribute a GNU-GPLed license except as expressly provided by the license, terminates the rights under the GNU-GPLed license. No warranty is included with GNU-GPLed software programs, although Lloyd's of London recently announced it would insure open source software (OSRM). Licensing software under the GNU-GPL is as easy as adding a copyright notice and a statement of copying permission to the software, along with an explanation that the program is distributed under the terms of the GNU General Public License. Although there is no legal requirement to register the copyright, it is considered safer to register it with the US Registry of Copyrights for legal protection purposes. The license covers the program itself, translations of the program, and any derivative of the program. Such clauses are designed to prevent others from simply rewriting the program or parts of the program in, say, another computer language, and then licensing the result as a proprietary product. Activities other than copying, distributing and modifying are not covered in the license, they are outside its scope. Therefore, for example, running the program is not restricted and the contents of the output are also not covered by the GNU-GPL license (OSI).

#### Licenses and the Motives of Open Source Programmers

In the following sections I will discuss important differences among modern open source licenses. My intention is to show that the choice of license provides a good indication of the motives of licensors and, through the conditions of the chosen open source license, the motives of licensees. There are several possible ways to classify open source licenses which make their basic characteristics clearer. Bruce Perens (1999), for instance, distinguishes among the following four categories: (1.) Whether the licensed software can be mixed with other software; (2.) Whether modifications to the software can be taken private and not necessarily returned to the original copyright holder; (3.) Whether the software can be re-licensed by anyone; and, (4.) Whether the software contains special privileges for the original copyright holder. Public domain licenses, licenses which are not open source licenses but often confused with them, allow the first three rights, namely, it is

possible to combine the software with other software, it is possible to make modifications to the software that are not returned to the original licensor, and the software can be re-licensed by anyone. The GNU-GPL, on the other hand, doesn't allow any of the above rights. Finally, many proprietary firms issue software with an additional fourth condition which gives them the right to retain special privileges, but denies users the other three rights.

Perens' categorization is helpful for analyzing the business implications of the various open source licenses; however, Larry Rosen's (2004) simplified classification system is particularly informative in regards to the intentions of copyright holders. Rosen draws a broad distinction between *reciprocal licenses* such as the GNU-GPL, and *academic licenses* such as the Berkeley BSD, MIT and Apache licenses. The GNU-GPL is considered business unfriendly because it requires that, putting it simply, works that come into contact with GNU-GPL licensed software must also be released as GNU-GPL software. For instance, if a person downloads a piece of GNU-GPL software and changes it or incorporates it into their own software, then any new software deriving from the original work must also be licensed under the GNU-GPL. This, in a sense, creates an "island" of GNU-GPL software that can only be accessed if a person is willing to make their own software GNU-GPL. Firms which are dependent on the secrecy of their software for their profitability will therefore generally be unwilling to incorporate or use GNU-GPL in any way because doing so will require them to allow others to study, adapt and change their software. The bargain created by the GNU-GPL is essentially: You may have this free software with the condition that any derivative works that you create from it and distribute must be licensed to all under the same license. On the other hand, software released under academic licenses, such as the very successful Apache server software, have essentially the same characteristics as GNU-GPL software except licensees are free to distribute copies of the work and derivatives *under any license they choose--even proprietary licenses*. Excluding the standard warranty and liability statements, the MIT or X11 licenses is: "Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions: The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software."



Academic licenses are not an insignificant category of open source license. Many of the commercial UNIX systems, including Sun Microsystems, Hewlet-Packard and Apple, use modified variants of the X11 system in their proprietary software. Ubuntu, a very successful Linux based operating system, also uses X11. The Berkeley Software Distribution, largely regarded as the first open source license, makes no legal requirements to return anything to the common pool of software. Despite this lack of obligation, companies and individuals continue to contribute code to the common pool of software created by these licenses. In short, free-riding hasn't constrained the supply of code. In contrast to reciprocal licenses, academic licenses have no attribution clause. In the case of reciprocal licenses such as the GNU-GPL, the names of copyright holders are contained in the body of the license. But there is no requirement that all contributors to the software receive credit for their contributions. The first condition of the second paragraph of the GNU-GPL states: "You must cause the modified files to carry prominent notices stating that you changed the files and the date of any change". However, the intention of the condition is to protect the reputations of creators of work against injury arising from flawed derivative works (St. Laurent, 2004, p.39). The legal distinction Rosen's makes between academic and reciprocal licenses also provides a critical distinction for assessing the motives of open source licensors as I will explain below when discussing individual motives for contributing to open source.

When assessing the motives of open source programmers a distinction has to be drawn between open source projects that are initiated by individuals and loose groups of individuals, and those instigated by large software developers such as IBM, Sun Microsystems, Hewlet-Packard--for simplicity and historical reasons I will call the first group individual open source and the second group business based open source. Much of the popular and academic commentary referring to open source is directed at individual open source programmers. These programmers are often represented as a large community of volunteers who give their time and expertise to develop free software. It is the presumed motivation of this group that I will primarily be addressing in later sections. The motives of the other group, open source software issued by firms, is easier to understand.

### Business Motives

Many firms have become rich re-packaging open source software in convenient to use media and others have been successful supporting firms using open source products or selling products which use open source software. I classify these

ventures as supporting or secondary open source in that they depend on a large base of available open source software. While many of these firms do provide code to the open source community, their primary task is to support rather than to add significant amounts of code to the community. This same reasoning is true of all those firms who use open source resources for various reasons ranging from testing their software to keeping informed on the latest software developments. The more interesting category is the one that I mentioned earlier in this chapter, namely, that of firms who wish to build an alternative to proprietary software. The desire to be independent of any one supplier is a standard theme in the industrial organization literature. Following that line of reasoning, Lerner and Tirole (2002) suggest that firms may be motivated to support open source projects as a way of achieving independence from the pricing and licensing policies of software firms. However, open source comes with significant “costs” in terms of the openness that firms need to provide in order to join the open source community. Many things including power, reputation, firm value, profits, etc., motivate firm owners, and traditional software licenses serve those needs. However, the primary objective of those who issue reciprocal licenses is to increase the amount of open source software available in the public domain, while the objective of those who issue academic licenses is to increase the use of open source software.

As I have claimed in previous paragraphs, the software market exhibits network externalities and some important market categories have tipped or are tipping towards Microsoft. The remaining firms are in a “losers” position; they are faced with the prospect of either being locked-in to Microsoft as the sole supplier of software or creating a viable alternative. A low cost alternative such as that provided by open source software is an enticing alternative. Creating a successful open source alternative means that server software, for instance, acquires the characteristics of a commodity and shifts the focus of competition to hardware or the software that runs on top of the operating system. I therefore suspect that the nature of network externalities in these markets means that many firms have decided that they either won't or can't profitably compete against the dominant design of Microsoft. Releasing proprietary software to the open source community then has the advantage of overcoming, or postponing, a dominant design. Space considerations do not allow me to address the many strategic issues that are raised when proprietary firms use open source licenses, for instance, the sort of open source licenses to use or the trade-offs involved in issuing software under more than one license. However, these firms are probably more inclined to use software released under academic licenses because they allow firms to re-release software

under proprietary licenses, as opposed to reciprocal licenses that would require them to release their proprietary software using the same conditions. For instance, the IBM open source license explicitly states that users can re-license IBM code using their own, i.e. non-IBM, license agreements.

### Individual Motives

As mentioned in previous chapters, the motives which prompt an act are not always easy to identify. Introspection and empirical investigation may not be very useful by themselves since motivation is often obscured by distorting factors such as rationalization, repression, dissembling and delusion. In short, simply asking open source participants what motivates them to provide their valuable resources free to open source projects may yield answers of dubious value. It may be the case that observers are in a better position to assess the motives of agents than the agents themselves. Motivation is also a theoretically complex concept that combines beliefs, desires, rationality, contexts and a host of other issues, each of which would require writing volumes to adequately locate a position within these dense literatures. My task is somewhat simplified in that my interest is the relatively specific issue of whether open source can be understood within the narrow confines of the behavioral assumptions of standard economic and business models. In other words, I won't try to assess all of the specific motives and combinations of motives that drive individual open source contributors to contribute; rather, I will concentrate my efforts on building a better understanding of the meaning and implications of self-interest in general and economic self-interest in particular.

While identifying the myriad of potential reasons why individual programmers give their time and contribute their valuable skills would be a futile undertaking, the history of UNIX and Linux, the statements of the license writers and, most importantly, the licenses themselves all provide clues as to the motives of contributors. My approach in the following sections is to examine these indicators in relationship to a few of the many explanations that have been offered on the motives of open source contributors. My emphasis, naturally, will be on the economic explanations that are offered and particularly those of Lerner and Tirole.

### The Analysis of Lerner and Tirole

Lerner and Tirole's approach to open source is unabashedly economic in that self-interest and rationality play major roles in their approach. They were the first

academics to see the significance of open source and I find their explanation of the motives behind what I have called firm open source as outlined above convincing. However, their explanation of what I am calling individual open source is less convincing in that economic self-interest is inadequately distinguished from non-economic self-interest. In Lerner and Tirole (2001), the authors propose a research program for the study of open source, part of which includes the motivation of programmers who contribute their code to open source projects. In that article, they defend their assumption that open source programmers are narrowly self-interested and respond to charges made by other researchers that open source software challenges the economic paradigm. Lerner and Tirole ask and respond to the following questions:

1. Individual incentives: Why do top-notch programmers choose to write code that is released for free? Is this "gift economy" (Raymond, 1999) consistent with the self-interest-economic-agent paradigm?

2. Corporate strategies: Why do commercial companies allocate some of their talented staff to open source programs? Why do software vendors initiate open source projects?

3. Organizational behavior: Is the apparently anarchistic process of open source production, in which no one tells anybody else what to do, a new model of business organization?

4. Innovation process: How does the new process fit with the conception of the innovation process driven by intellectual property rights (patents, copyrights and trade secrets) that we have inherited from Arrow (1962) and Schumpeter (1975)?

I have addressed the second point already and the fourth point is outside the boundaries of this manuscript, which leaves the first and third issues.

Before proposing their own explanation to the question of why open source programmers contribute to public goods, Lerner and Tirole are keen to reject altruism as an explanation. They tell us that: "economists (and indeed several open source leaders) are suspicious of the altruism hypothesis: the view that contributions are driven by pure generosity, or a sense of duty to give back to a community that has provided a useful piece of code" (2001, p. 822). Lerner and Tirole explain that the reason economists are wary of altruism is because: "Of

course, altruism exists and can do marvels in certain circumstances, but the altruistic hypothesis fails to explain why programmers do not focus their generosity on more needy beings and why free riding would be less pervasive than in biotechnologies or other industries" (p. 822). The first charge assumes, like that previously seen in Jensen and Meckling's agency theory (chapter 3), that altruists must maximize the value of their contributions in order to be considered altruistic. They must find, in other words, the best use for their altruistic efforts or, according to Lerner and Tirole, they are not altruistic. If I have properly understood their argument, then it is unconvincing. If a firm can be profit seeking without maximizing profits, why can't a person be altruistic while not maximizing their altruistic efforts? And, more generally, just as a self-interested person can be self-interested without maximizing their interests. The altruist faces the charge that he can do better and that he is not a "pure" altruist, but not the charge that he is acting from other than altruistic intentions. Finally, not all economists are as wary of altruism as Lerner and Tirole. Although most of the evidence supporting altruism comes from behavioral experiments, altruistic behavior is regularly observed in everyday life, for example, as in firm wage negotiations (Bewley, 2000), tax compliance (Andreoni, Erard, and Feinstein, 1998), and cooperation in the protection of local environmental public goods (Acheson, 1988; Ostrom, 1990).

It is not entirely clear what the authors mean by the second part of their charge when they ask "why free riding would be less pervasive in software markets than in biotechnologies or others industries." One interpretation of the remark is that an act can't be altruistic if free riding is not a significant problem or put more simply, open source is possible in the software market because free riding is a problem. It is not clear to me what free riding has to do with altruism. On the one hand, free riding is an issue in open source; in fact, it is an accepted "cost" of those choosing to issue their code under academic licenses. People contribute their resources to academic licenses without restrictions, without, in other words, concern that others will free ride upon their efforts. Yet the academic license works, people continue to contribute despite receiving no guarantees that others will contribute and despite the fact that others are allowed to take their intellectual property and use it in their own proprietary products. On the other hand, reciprocal licenses are effective in part because they restrict others from using licensed products unless they submit their products to the same conditions. If free riding were a problem that couldn't be overcome, then there would be no reason to use reciprocal licenses. In addition, if free riding were a big problem then there would be no sense in using proprietary software either, all software, in a sense, would be a public good. Another

interpretation is that the authors wonder why open source or its equivalent isn't prominent in other industries. I suspect that one of the reasons is the nature of the software market (externalities), the history of the market, and the nature of the product itself. Software, unlike biotechnology and many other industries, is a product that is easy to transport and manipulate and the costs of reproduction are nearly zero. Software, like many other information goods, has well-known characteristics that make it substantially different from physical goods. Perhaps the issue in Lerner and Tirole's second point is that an often given reason to use open source licenses is that they prevent free-riding by firms or individuals that use the software but contribute nothing to the developer community. That would certainly be a reason not to use "free" licenses and why licensors might decide to use reciprocal licenses, but it doesn't explain why they would use academic licenses. However, all of that said, I hold open the possibility that I may simply have misunderstood the point that the authors are attempting to make about altruism.

Why then, according to Lerner and Tirole, do programmers contribute their skills to open source projects? The authors give two reasons, the first is that programmers "know UNIX" so they "find programming straightforward". The explanation that programmers find it easy to code and are therefore more likely to use open source licenses is questionable. I wouldn't categorize programming as straight-forward no matter what the platform--UNIX is a platform upon which many computer languages run, and is itself not a computer language--given such things as the rapid proliferation in the number of computer languages and a corresponding increase in compatibility issues, let alone the new platforms on which software is expected to run, but others may. Regardless of the ease of programming, the explanation offered doesn't seem to fit into even a liberal understanding of self-interest. Many occupations may appear "easy" to outsiders, but finding something easy to do is generally not considered a good economic reason to give away valuable economic resources. Therefore, the explanation that it is easy for some people to write and reproduce software is not a convincing reason for why programmers would voluntarily give software away when they can charge for those services. A more important question in such cases is what value others place on the software. For example, it may be inexpensive for a company like IBM to release important code under open source licenses, but its value to the open source community is immense in that it saves programmers and users from having to write code that serves the same function.

The second reason given by the authors is more interesting, they write, "...many contributors are sophisticated users who need to remove a bug or tailor the code to their specific applications. Having done so, they turn the code back to the community. They may then even see others improve on their modifications, increasing their private benefit further"(p. 822). If I have properly understood the argument, programmers give away their code because they expect or hope that others will make improvements to their code and, presumably, release those improvements back to the community so that the original issuer can access the improved code. There are some problems with this reasoning if a narrow, economic understanding of self-interest is used. First, it seems to conflict with Mancur Olson's well known claims about the logic of collective action. In particular, the leap of faith that programmers have to make that others will return the favor contradicts the basic concept of selfish behavior. It would be better, according to economic "logic", to wait for others to foolishly release their code using open source licenses. Of course, no one wants to be foolish and so no one, according to the logic, will be the first to use open source licenses. The fact that people do use open source licenses, and especially the fact that they use academic licenses, seems to point to a flaw in this reasoning. Secondly, the assumption of Lerner and Tirole seems to be that improvements in open source software will be returned to the community, however, returning software to the community is not a requirement of any open source license. In other words, a person can use open source software without ever contributing anything back to the community. In fact, the vast majority of people use open source code without ever giving anything back. The willingness of programmers to contribute their code to the community with the hope that an improved version will later reappear rests on the assumption that others may be motivated by something other than a narrow definition, profit maximizing, definition of self-interest. In short, if the assumption is that people are narrowly self-interested is accepted, then open source wouldn't be the success that it is. Programmers would have no good reason to contribute their code.

Another benefit to open source programmers that Lerner and Tirole mention is the "non-negligible" delayed benefit of contributing to open source projects in terms of an improved reputation. This reason, I believe, is at the heart of the author's argument against a non-self-interested interpretation of open source; it is the belief that by contributing to open source projects programmers add to their reputations. This is because contributions are taken to be public and scrutinized by potential employers and peers as well. Reputation effects have a long history in economic analysis, recall the important role that sympathy and vanity play in Smith's system,

rather than goods or wealth. Smith argues: "It is the vanity, not the ease, or the pleasure, which interests us" (TMS, I.iii.2.1). Wealth, for Smith, is a means to acquire the attention, sympathy and approbation that feeds one's vanity; it is not an end in itself once basic needs are satisfied. Vanity, in turn, "is always founded upon the belief of our being the object of attention and approbation" (TMS, I.iii.2.1). Reputation doesn't lead to wealth in Smith's system, but people acquire wealth to increase their vanity. Reputation plays a large role in Mandeville's system as well: "The Greed we have after the Esteem of others, and the Reputes we enjoy in the Thoughts of being liked, and perhaps admired, are Equivalentents that overpay the Conquest of the Strongest Passions..." (Mandeville, [1732], 1988, i. 68).

However, Lerner and Tirole do not explain why contributing to open source is an especially good way, let alone the best or maximizing way, to further one's reputation. Many, firms, as the authors point out, also base rewards on the contributions of their employees, so why should a programmer work without pay on an open source project instead of getting paid while working for a standard software firm and perhaps earning a bonus for good work? Another alternative to raise one's reputation would be to start or join an on-line software project that uses proprietary licenses rather than open source licenses. Alternatively, a programmer could contribute to projects that use academic licenses and reissue the software under a proprietary license. Another related issue that the authors don't address is why a person's reputation is important to them. The assumption seems to be that people desire a good reputation because it leads to potential economic benefits. There is no doubt that this is one reason that some people are concerned with their reputation, but there are many others as well. For instance, to name just a few occupations, judges, religious leaders and researchers may be concerned with their reputations for a variety of reasons none of which necessarily involve economic rewards. Similarly, I suspect that open source programmers are at least as interested in their reputations as excellent programmers as they are in lining their pockets. So, while some open source programmers, for example, Larry Wall and Linus Torvalds, have made money from reputations gained through participation in open source projects, there is nothing to suggest that their financial gains are solely based on their participation in such project or that any of the thousands of other often nameless open source participants contribute in order to maximize their economic gains. Open source, as a means to financial success, is of dubious worth.



The third question raised by Lerner and Tirole is something of a straw-man, I know of no one who claims that open source projects are anarchistic nor does anyone believe that no one tells anyone else what to do in open source projects. Leadership in open source projects is often quite authoritative (e.g., the Apache project and the Linux kernel), the difference with standard firms is that anyone can start a new project when they want to. One of the remarkable features of open source projects is that they exist independently of their original developers. As I have mentioned, some types of open source licenses create a pool of free software from which anyone can draw. The relevance to the issue posed by Lerner and Tirole is that anyone can start a new open source project based on the original, freely available base code. For instance, if the leader of a project is abusing their authority, dissatisfied programmers can simply start their own projects based on the commonly available source code. The most successful open source projects, such as Apache and Linux, have well-defined structures that have kept these projects active for years. The newness of open source is not in the structures, but in the ability of "employees" of projects to become managers if they choose and that it allows projects to be developed independently of their original authors.

I now turn to Weber (2004). Weber's approach is interesting because of his critique of the proposed motives of open source contributors, and because of his view of open source as a complex public good; an explanation which has become popular among some researchers.

Weber (p. 130) writes: "First I want to dispose of two myths about open source that get in the way of moving toward analytically rigorous answers". The first myth, according to Weber, is that writing open source code is basically an amusement for enthusiasts; a game of fun among like-minded hobbyists. "Imagine this analogy: If all the model train enthusiasts in the world could join their tracks together through the Internet, they would build a train set just as elaborate as Linux. Nobody has to tell them how to do this and surely nobody has to pay them; it is a labor of love. And because everybody feels the same way, there is nothing to argue about."

Weber argues that this "macro", as he calls it, explanation of amusement, can't alone be the motivating force behind open source because open source programmers have disagreements: "The macro part of this story is either unarticulated or naively wrong. Like-minded or not, participants in the open source process argue more or less continuously, and about both technical and organizational issues." And later Weber writes: "If open source software were

simply the collective creation of like-minded individuals who cooperate easily because they are bound together by hobby-love or semi-religious beliefs, there would be little disagreement in the process and little need for conflict resolution among developers."

Amusement may not be a primary motive of open source, although it has often been mentioned by open source programmers themselves as a motive (Torvalds and Diamond (2001), Raymond (2001)). A more relevant question is whether programming for an open source project is more amusing than programming for a proprietary firm. Survey evidence from the FLOSS (2002, part IV) study suggests that the two most popular reasons for participating in open source projects are to share knowledge and skills and to learn and develop new skills; the question of whether people participate because it was amusing was not directly addressed in the survey. However, a question about the perception of the reasons why others participate in the community was asked, and 20 percent answered that they thought that other participants contributed because they derive pleasure from doing so, a relatively small percentage compared to the 65 percent who said that it enables more freedom in software development and the 40 percent who said that open source provides more variety in software. I would, however, add that coding for open source projects is an overwhelmingly voluntary affair, so one can reasonably assume that participants receive some amount of "net-amusement" from participating, and even those programmers who are paid to program for contributing to open source projects may enjoy participating. It is important to realize as well, that coding is only one aspect of participating in an open source project. Many participate by answering user questions posed in newsgroups or other online forums. For instance, the R and Perl groups, but many smaller groups as well, have active mailing lists that cater to the needs of new users. Experienced users and programmers voluntarily answering novice questions is a vital service provided to developing open source communities. One incentive for participating, often voiced by these experienced programmers, is the pleasure they receive from helping others.

The "micro" part of Weber's story attacks and then evades the assumption of altruism in open source. He argues that if there were no mechanism for conflict resolution, the "hobbyists" would have to be following motivations other than personal satisfaction. It follows, according to Weber, that because there are means to resolve conflicts within open source communities, altruism cannot be a primary motivating factor. Participants would have to be acting in an explicitly pro-social

way, and doing things that others want, for the sake of the other and the act itself. This, according to Weber, would be altruism. "To act selflessly in this setting would be to write and contribute code for no apparent compensation other than the personal gratification that comes from doing something that helps someone else." Clearly, Weber is arguing against the motive of altruism: "But the evidence confounds any straightforward version of this argument. If altruism were the primary driving force behind open source software, no one would care very much about who was credited for particular contributions. And it wouldn't matter who was able to license what code under what conditions." Weber makes several possible errors in his claim, the first is an overly restrictive understanding of the requirements of an altruistic act, the second is caused by a misunderstanding about open source licenses.

With regard to the first possible error, and as discussed in the first chapter, altruistic acts do not necessarily require a sacrifice on the part of the altruist, similarly, they do not require that an altruist receives anything in return for performing such acts. Practically, most altruistic acts, like any acts, do "cost" something, say, for instance a net-loss in well-being. But if the reason for performing an action is altruistic, and there are benefits to others, then an act is altruistic. Contributors to open source projects that use licenses that stipulate that their names must be attached to the software in some way are not necessarily less altruistic for doing so. Besides, no open source license requires that the names of all contributors are listed in the license, such a requirement would be unworkable. The name of the copyright holder is generally available, but the nature of many open source projects means that many hundreds and in some cases, thousands, of programmers have contributed their code and expertise; such contributors remain anonymous. In addition, most open source licenses only require that the names of the license holders be attached to the source code, something few users ever see.

The sentence: "And it wouldn't matter who was able to license what code under what conditions.", is confusing. Weber seems to imply that because there are restrictions on who can license code and determine license conditions open source licenses cannot be altruistic. Altruism, for Weber, appears to require that a contribution be an unconditional gift to everyone, as I have shown in chapter 2, this would be a very restrictive form of altruism. Weber reinforces his point in the footnote that accompanies the above sentence: "Popular media often portray the open source community in this light, but fail to account for the fact that many "beneficiaries" of this altruism (apart from the developers themselves) are major

corporations that use Linux software." It is true that proprietary firms can benefit from open source software, but many who benefit are not associated with proprietary firms, for instance, those who use open source browsers. A contribution to the needy that happens to help those that don't need it, can still, clearly, be an altruistic act. Weber also seems to have missed that a cornerstone of all open source licenses is that there are no restrictions on who can use open source code and no restrictions on the purposes for which it can be used. So, in fact, it doesn't matter who uses the licenses for what proposes, as long as the rights and obligations of the license are obeyed.

Weber adds one last, as he puts it, pragmatic reason to steer clear of altruism as a principal explanation. Altruism, according to Weber, and I would agree, is a highly loaded term; I, incidentally, would add that self-interest is a highly loaded, ambiguous, term. Weber (2004, p. 131) writes: "For better or worse, arguments about altruism tend to invoke an economics-inspired intellectual apparatus that places altruistic behavior in opposition to self-interest". He goes on to say that this can quickly go on to become an unproductive discussion about whether or when it makes sense to redefine self-interest so it can accommodate a desire to do something solely for someone else. "In other words, should individual utility functions include a term for the welfare of another?" Weber chooses to "sidestep" this issue. Weber is correct, the question of how to combine or even compare individual utility is difficult and one that is still debated at a fundamental level (Elster and Roemer, 1991). However, as I have explained, questions about what self-interest and altruism entail are of practical importance in regards to how to motivate people. The problem is, as Weber correctly identifies, when researchers attempt to subsume altruism under self-interest. The two concepts involve different intentions and should not be conflated. Self-interest is one of the most important assumptions in economic theory and often, correctly, placed in contrast to altruism. Finally, no reputable theory of self-interest suggests that helping others cannot be a component of one's interest (chapter 4).

The positive account of open source motives that Weber's (2004) offers frames open source as a common pool resource following Ostrom (1990) (see also Smith and Kollock, 1999). Weber (2004; p. 16) believes that: "Open source radically inverts the idea of exclusion as a basis of thinking about property. Property in open source is configured fundamentally around the right to distribute, not the right to exclude." The position in its purest form is found in van Wendel de Joode (2005), which I will presently discuss. Van Wendel de Joode (2005) claims that his main

behavioral assumption comes from Ostrom (1990, pp. 33-40); he writes: "It is important to note that Ostrom's findings do not depart from the basic principles of economic thought. The individuals in communities described in Ostrom's book are not altruistic. Instead, they want to maximize their utility and act in their own self-interest" (van Wendel de Joode, 2005, p. 23). I believe that a closer reading of these same pages in Ostrom's book provides a much more complex view of behavior than is common in economic thought as understood by Van Wendel de Joode, in fact, much of Ostrom's work would seem to support altruism as a prime motivator (Ostrom and Gardner, 1992). In particular, Ostrom discusses the importance of norms of behavior in regards to collective activities. For instance, she writes that norms of behavior reflect valuations that individuals place on actions or strategies in and of themselves, not as they are connected to immediate consequences. When an individual has strongly internalized a norm related to keeping promises, for example, the individual suffers shame and guilt when a personal promise is broken. Norms of behavior therefore affect the way alternatives are perceived and weighed (Ostrom, 1990, p. 35). Following norms may be in one's self-interest, but they are not required to be in one's interest and in many cases may work against those interests. Telling the truth, to take Ostrom's example, if taken as a categorical imperative, is a rule which has force regardless of its influence on interests. In the same quote above, van Wendel de Joode seems to confuse utility maximization with self-interest. A point he makes again on page 15: "How are communities of utility-maximizing individuals able to overcome the problems facing common pool resources?" (van Wendel de Joode, 2005, p. 15). And again, on page 194: "The individual behavioral rules identified are based on the assumption that individuals want to maximize their utility. This is in line with state-of-the-art research on people's motives to participate in open source communities" (van Wendel de Joode, 2005, p. 194). Open source may not be non-self-interested, but Ostrom's argument and method cannot be used to support that assumption. In addition, as I have shown, utility maximization remains silent on the issue of motives; claiming that someone follows his utility is a descriptive statement, not a statement about motives self-interested or otherwise.

Much of van Wendel de Joode's (and Weber's) argument, and his reason for using Ostrom, rests on the idea that open source is a common pool resource, but is it? Van Wendel de Joode argument (p. 51) goes: "Usage of software in general and of open source software in particular does not affect the amount of source code available to others. To use software developed in a community implies making a copy of the source code and installing it on a local computer. This does not

diminish the original stock of source code, and therefore using the resource itself does not diminish the amount available to others." The inability to exclude others is critical to van Wendel de Joode's (p. 5) argument: Open source communities generally lack such exclusion mechanisms: the culture is to attract as many people as possible to participate. For this reason the boundaries of open source communities are much less formal and are arguably highly permeable. Membership in open source communities is claimed to be fluid; current members can leave the community and new members can also join at any time.

However, I am not entirely convinced by van Wendel de Joode's arguments because open source, like proprietary software, is exclusionary; while the argument that software is costless to reproduce applies to both proprietary as well as open source software and so adds little to the discussion. A user needs to agree to the conditions of a license in order to use it or, as is frequently the case in open source and other software, by using it the user is implicitly agreeing to abide by its terms. *Open source licenses do not remove the software from protection by the copyright laws.* The original author of the software, which can be a company, doesn't need to undertake any formal act to obtain a copyright; any original software that is recorded on some format is automatically protected by copyright laws. As discussed previously, these rights include the exclusive rights to make copies, prepare derivatives works and distribute copies of the original work and derivatives. Rights can be added and subtracted, but basic copyright law applies automatically to all original software.

As discussed, open source academic licenses are more liberal than reciprocal licenses, but both categories are protected under international copyright laws. Van Wendel de Joode is correct in claiming that the open source community wants to attract as many people as possible to participate, so do proprietary firms, but in order to participate in either open source communities or use proprietary software, users must accept the legally enforceable terms of the copyright attached to the software. Open source software is more "open" than proprietary software because the source code is made easily available, but that openness shouldn't be confused with unrestricted. Again, a distinction should be made between academic as opposed to reciprocal license. Proprietary software has an additional barrier in that it is generally not available in source form, but techniques to reverse-engineer binary code means that such protection is fleeting at best.

Open source software is exclusionary by design; that was one of the lessons to be drawn from the history of UNIX outlined above. Free software, by which I mean software for which copyright holders explicitly renounce their rights, was replaced in favor of the GNU-GPL in order to prevent proprietary firms from exploiting the goodwill of programmers. The GNU-GPL message is, if you want to use this software, then you have to play by our rules. A close analogy can be found in any public library, a library card holder can check-out books for free and read them, but the copyrights of the original authors are still in tact, the author has the exclusive rights for, essentially, the life of the author plus seventy years. As discussed above, academic licenses are more liberal in that they allow license holders more flexibility to distribute open source software, but this category of open source license still protects licensors under copyright law.

In short, Weber's claim that open source radically inverts the idea of exclusion as a basis of thinking about property and fundamentally revolves around the right to distribute is only partially true. In fact, the power of reciprocal open source licenses rests firmly on copyright laws and explicitly excludes those who don't want to contribute to open source projects.

#### Other Motives

It is now possible to return to the motives outlined in the introductory chapter and attempt to categorize them according to the structure I have been developing throughout this manuscript. The question is whether the motives offered by researchers concerning why people contribute to open source projects can be categorized as self-interested. If not, attempts to appeal to interests will fail and perhaps be counterproductive.

#### Self-interest

Recall Lerner and Tirole's (2001, p. 3) original question quoted in the introduction: "Why should thousands of top-notch programmers contribute freely to the provision of a public good? Any explanation based on altruism only goes so far. While users in less developed countries undoubtedly benefit from access to free software, many beneficiaries are well-to-do individuals or Fortune 500 companies. Furthermore, altruism has not played a major role in other industries, so it would have to be explained why individuals in the software industry are more altruistic than others." Whether open source programmers and programs are "top-notch" is a

subject of debate (Rusovan et al., 2005; Neumann, 2005), but the point is taken. Is open source a public good? As argued above, it does not fulfill a basic condition of the definition of a public good in that it is possible to exclude others from using it. The GNU-GPL, for instance, is very exclusive, requiring software with which it interacts to be licensed under the GNU-GPL as well.

Lerner and Tirole use a restrictive definition of altruism which requires that altruistic acts can only, exclusively, benefit the needy. In order to be altruistic an act needs to be motivated by the intention to be of benefit to others, but that doesn't exclude the possibility that those in less need may receive benefits as well. The fact is that many people benefit from open source software, even some Fortune 500 firms. The final question that Lerner and Tirole ask is whether individuals in the software industry are more altruistic than others. This is an empirical question which can be tested, although I do not believe it has. I suspect, as I believe Lerner and Tirole intend to suggest, that open source contributors are not any more altruistic than workers in other sectors. However, the conditions of dissemination and the incremental process of writing software make it especially easy to be altruistic, in other words, barriers to being altruistic are low in the software market. Dissemination of software is virtually costless, while the process of coding or helping others online can take place anywhere and anytime and doesn't require large capital investments. The Internet and mail servers make it easy to coordinate efforts, further reducing the costs of contributing code and expertise that benefit others. It might be the case that altruism is a very common motive for actors in many situations including marketplaces, only the means of expressing that motive are less readily available than in the software market. Contributions to open source projects are altruistic, albeit perhaps they are a form of qualified, low-cost or costless altruism.

#### Internal motivation and self-interest

I will now shift to a topic first presented in chapter one of this manuscript, namely, the categorization of motives used by Ryan and Deci (2000, p.70). My purpose is to demonstrate the confusion in the literature concerning motives in general, and the motives of open source programmers in particular.

The focus of internal motivation as defined by Ryan and Deci and used by many others would appear to be on well-being. Recall that Ryan and Deci define internal motivation as: "... the inherent tendency to seek out novelty and challenges, to



extend and exercise one's capacities, to explore, and to learn." The authors claim that the focus of this sort of motivation is non-economic, subjective, and centered internally on the needs of the actor. However, I believe that Hars and Ou's (2001, p. 2), who use Ryan and Deci's interpretation of internal motivation, have misunderstood its meaning. Recall from chapter 1 that Hars and Ou make several claims about the motives of open source programmers. In particular, they claim that advocates of the position that open source programmers are motivated by internal motives implies that open source programmers are not motivated by monetary incentives. Rather, according to Hars and Ou, open source programmers are motivated by the enjoyment they derive from practicing their hobbies and their general preferences or they receive rewards from increasing the welfare of others. There are several points of confusion in Hars and Ou's explanation, in particular, they mix self-interested and altruistic motives and they accord too much power to the economic conception of preference.

The internal, external, categorization made by Ryan and Deci is presumably made in order to emphasize the source of the motivational force, but motivation is ultimately always internally determined. The point of the categorization is to presumably show that external motivators, such as an employee receiving a check once every two weeks from an employer, and working primarily for the enjoyment of working, are two different, sometimes incompatible ways to get work done (Frey, 1997). But the starting points of internal and external motives are the same, the employee chooses to react based on his own set of preferences or desires. However, in order to understand those preferences, one is required to take one step "backwards" and attempt to understand the mechanisms that help to form those preferences. Therefore, Hars and Ou's suggestion that altruism is an internal motivator is confusing because both altruism and self-interest are "internally" motivated. Consequently, Hars and Ou's categorization of altruism as an internal motivator doesn't distinguish it from the motive of self-interest. An interesting question is whether altruism can be motivated by an external force, for instance, a governing body. It is clear that benevolent behavior can be so motivated, for example, tax deductions for charitable contributions, but altruism requires intention, which is a much deeper concept and one that is not so readily malleable. Hars and Ou (2001) also name community identification as a variant of altruism, which, they claim, is the need to be loved. This too is a questionable use of the term altruism because the need to be loved may be a self-interested matter. To use the language of chapter 4, it is an external dimension that can be pursued for self-interested or non-self-interested reasons.

Note that my argument is not that the motives which Hars and Ou identify are not motivating contributions to open source, rather, my critique concerns their identification of self-interested acts as altruistic acts. The problem can also be seen in empirical studies. For instance, in a question designed to test whether the motives of open source contributors are altruistic, Hars and Ou ask survey respondents whether they strongly agree or disagree with the following statements:

1. I don't care about money. (strongly agree/strongly disagree)
2. You can always trust an open source programmer. (strongly agree/strongly disagree)
3. Recognition from others is my greatest reward. (strongly agree/strongly disagree)
4. Open source programmers should help each other out. (strongly agree/strongly disagree)
5. I deeply enjoy helping others - even if I have to make sacrifices. (strongly agree/strongly disagree)
6. Open source programmers are a big family. (strongly agree/strongly disagree)
7. I am proud to be part of the Open Source Community. (strongly agree/strongly disagree)

Some of the questions do not, under the definition of altruism that I have provided, identify altruistic motives, namely, answers to questions one, two, and seven would not help to identify the motive as altruistic as, I suppose, is the intention. Briefly, an altruist can care (or not) about money; not trust other open source programmers (or do trust them); and not be proud to be a part of the open source community (or be proud). Question three is interesting because it would be difficult for a programmer to be both primarily motivated by altruism and have the aim of recognition as the greatest reward. It is difficult to imagine that the intention to help others is closely correlated with receiving the greatest amount of recognition for one's self. A person could strive to be the greatest provider of open source

software, but if the intention is to become the greatest provider in order to maximize one's recognition, then the person would not be altruistic. In short, question three is probably not a good identifier of altruistic motives. An affirmative answer to question four would be ambiguous for what should now be obvious reasons; the question is why should programmers help each other, not do they help each other. For instance, open source programmers motivated to help each other because it harms proprietary firms have a different primary motive than those who help others because it is the right thing to do or because of a general intention to benefit others. For most reasonable cases, a negative answer to question four would be a sign of self-interest. Question five is similar to four; a person may deeply enjoy helping others because it is the right thing to do or because helping others increases his recognition as in question three or for many other reasons only some of which are altruistic. The intentions of asking six are clear, someone who sees the open source community as a big family is more likely to be altruistic towards family members.

Questions from the FLOSS (in particular part IV, 2002) study are more promising. The FLOSS Internet questionnaires and studies are important because they represent the most comprehensive attempt to assess contributor motives and the conclusions drawn from the studies have been used to ground theoretical claims about those motives. Ghosh (2005), for instance, a leader of the FLOSS study, uses what he calls a "selfish measure" or "altruism measure" to get at the motives of open source developers. Ghosh is aware that the measures he uses don't reflect intent as he, correctly, says they should, but outcomes, in which case the terms selfish and altruistic don't necessarily "fit". However, he goes on to conclude that, on four motive classes, namely, social/community, career/monetary, political and product related: "... the responses are consistent with self-interested participation and indicate that developers perceive a net positive value flow (Ghosh, 2005, p. 33-34)." But do the questions measure what they are intended to measure? The questions related to this issue asked respondents whether their own contributions were more or less than their perceived rewards; simply, "I give more than/less than/the same as I take." They found that 55.7 percent of all developers classified their relationship as taking more than giving, 14.6 percent take and give in equal measure, and only 9 percent give more than the take, i.e., were altruistic. The set of questions that Ghosh is referring to were not included in the FLOSS-US survey (FLOSS-US, 2003), which is smaller than its European and Asian counterparts, but very similar results were reported in the much smaller FLOSS-Asia (2004).

Given the question, it is surprising to me that anyone, except perhaps a handful of developers including Richard Stallman, Linus Torvalds and Larry Wall, could answer affirmatively to the question posed. Just by using Linux or Firefox most people are taking much more than they could ever give to the open source community in return. It is the nature of the product, which is built on the contributions of largely anonymous people over many years. For instance, I have contributed small amounts of code and much time to an open source projects related to the economics of fisheries and other topics, but that contribution, and any contribution I or anyone else could hope to make, is minuscule in comparison to the combined efforts that have gone into making Linux a functioning operating system. The question should not be whether there is a net positive or negative flow of value going to the contributor, that question automatically makes all contributors self-interested or selfish, rather, the question is what are the intentions of open source contributors. The empirical evidence designed to identify the motives of open source contributors, particularly those questions designed to identify altruistic motives, is only partially successful. Many of the questions asked don't adequately distinguish between altruism and self-interest because they confuse these two motives. A better understanding of what both motives entail may have led to better questions.

Several other internal motives have been identified in the literature, namely, the benefits derived from being a part of a community, the desire to support the ideals of the open source community, and the belief that open source software is superior to proprietary software. Von Hippel and von Krogh (2003) conclude that: "Contributors to open source software projects may also get private benefits from participating in the project "community" that are not available to free riders (Raymond 2001, Wayner 2000, Lerner and Tirole 2000, Moon and Sproull 2000)." Community identification is also apparent in "feelings of obligation to the open source software community" which Feller, et al. (2005) mentions, and can also be found in Barbrook (1998, 2005) when he writes about open source as a gift economy. Other authors include the intrinsic satisfaction of doing the work, fun, and learning as a source of pleasure (Lakhani and Wolf (2005), Torvalds and Diamond (2001), Feller et al. (2005)). In general, most of these motives can be traced to self-interested motives, although few of them relate to maximizing financial gains. However, references to feelings of obligations would suggest duties play an important part in some the motives of some open source programmers. This concern for duties is best exemplified by the views of Richard

Stallman, an important figure in the open source community, and to whose views I now turn.

The most articulate defense of the belief that software should be free and open source can be found in the preamble of the GNU-GPL, written primarily by Richard Stallman a respected leader of the open source movement. Stallman (1984 version of the GNU-GPL license) suggests that supporting open source is a question of an obligation or duty to support the freedom of information. His manifesto at the beginning of the GNU-GPL is a clear statement of his motives; even the terminology used to describe this type of software has been a subject of intense debate and is an important indicator of the motives of these programmers. Stallman, commonly regarded a key leader of open source software, suggests that this type of software should be called "free" software, where free refers to the freedom of program users to run, modify and redistribute software and not its price (Stallman, chapter 5 in DiBona et al. (1999). Stallman specifically rejects the labels of "liberated", "freedom" and "open" as misrepresenting what he regards as the key element of such software, keeping it free from proprietary constraints. The language he uses to describe his decision to develop such software is telling of his motivation. He refers to "a stark moral choice" he faced whether to join the proprietary software world and make "the world a worse place" by "building walls to divide people" or becoming a programmer that "could do something for the good". Such language is controversial, at least when compared to the average economic textbook used in business schools. The millions of people who use the GNU-GPL and other reciprocal licenses, in my opinion, are supporting Stallman's values, values which conflict with standard economic interpretations of self-interest in favor of creating a large island of free software for those willing to accept the conditions of the GNU-GPL. Contributors who use reciprocal licenses are, in part, supporting Stallman's intention to free information.

Lerner and Tirole and others have suggested that open source constitutes a form of gift economy. It is again helpful to distinguish between users of open source software and contributors when analyzing whether this is a good account of motives. Those planning only to use open source software are not required to give anything in return for using the software. Similarly, users of many proprietary products are not required to pay for the software they use, for example, the most popular proprietary Internet browser is free to download and use. Both groups of users, those using open source and those using proprietary software, are placed under no legal obligation to return a "gift" for using the software. Contributors to

open source software projects, specifically those who modify open source open source software, are obligated to follow the conditions as described in the software's license. To that extent, open source and proprietary licenses are the same, it is the license that defines the performance obligation. Perhaps the authors are referring to the sense of obligation as explained by Scheffler (1997) and outlined in chapter one of this manuscript? Users of open source software may feel a special responsibility to the open source community which they don't feel towards proprietary firms. The argument might be that by taking and using open source software, which is costless for most users, users acquire a relationship with accompanying responsibilities. This is an empirical question which needs to be explored. However, in general, the open source community can't be adequately described as a gift giving culture because the anonymity with which users can download, change, and use open source software means that there is no need to acknowledge the giving and particularly the receiving of gifts. A gift giving culture would seem to require that receivers of gifts be identifiable, this is not the case with open source software. The extent to which a user of open source feels responsible for the open source community as a whole rests solely upon one's own conscience and not a need to satisfy the expectations of others.

The three motives discussed above represent a mix of self-interested and non-self-interested reasons for contributing to open source. Identifying the reason for joining the open source community in terms of derived benefits would seem to place the issues squarely in terms of self-interest. People join because they expect to benefit, with the benefits usually expressed in terms of happiness. While feelings of obligation tend to suggest something other than self-interested, perhaps deontological reasons for contributing. While Stallman's and the Free Software Foundation's GNU-GPL license (GNU-GPL, 2007) is clearly expressed in terms of duty to an ideal. Finally, the idea that open source software is better than proprietary software provides a reason to use open source software, and is not primarily a reason to support it or contribute to its development. Again, in general, these three motives represent a mixed bag of self-interested and non-self-interested motives.

Finally, a variety of other motives have been offered to explain why people participate in open source projects. I will divide these explanations into self-interested and non-self-interested explanations. Self-interest is generally defined to include a wide range of aims, not just financial gains. Lakhani and Wolf (2003) suggest that enjoyment-based reasons drive open source participation and while

enjoyment can certainly be an self-interested reason to act, I have emphasized throughout this manuscript that pleasure can come from a variety of aims not all of which are self-interested. The issue that Lakhani and Wolf would have to address is whether participating in open source projects is more pleasurable than participating in proprietary projects and if so why. In the case of open source, I suspect that many people participate because it is fun to be a part of a group that is seen as offering resistance to the dominate Microsoft. Another reason mentioned is that of filling an unfilled demand (see e.g., G.L. Green 1999). Again, the question that these authors would have to answer is why filling an unfilled demand requires using an open source license. An alternative possibility, for instance, would be to use a proprietary license with very liberal conditions.

#### External motivation and self-interest

Once again, external factors are the rewards and punishments induced from outside the individual (Ryan and Deci 2000). For open source programmers, the rewards are generally taken to mean the direct and indirect benefits gained by increasing their marketability and skills base for the purpose of making themselves more valuable to others. Open source contributors, under this assumption, see open source primarily as a means to improve their job prospects. Whereas a contributor motivated internally does so because some aspect of contributing provides motivation, perhaps because of the pleasure of programming or the benefits of being a part of a group of like minded individuals.

The important difference between internal and external motivation is the immediate source of rewards (or punishments) facing an individual. The approaches of Lerner and Tirole (2002, 2005) and similar positions expressed by Bonaccorsi and Rossi (2003) and Lakhani and von Hippel (2003) all assume that working in an open source project provides prestige and visibility that gives developers a chance to be noticed by software companies. This point is emphasized by some advocates of open source (Raymond, 2001). Raymond goes so far as to name prestige as central to the hacker culture's reward mechanism (p. 89). I have discussed issues related to reputations above in a separate section on the approach of Lerner and Tirole and won't repeat them here.

However, in general, the external type of explanation follows standard economic theory, an efficient manager pays workers according to the value they add to the firm; open source is a means to demonstrate one's potential value. In the literature

on open source internal motives are often used as a foil against the rational behavior offered by standard theory. For instance, Lakhani and Wolf (2005, p. 3) write: "Many are puzzled by what appears to be the irrational and altruistic behavior by movement participants: giving code away, revealing proprietary information, and helping strangers to solve their technical problems." Lakhani and Wolf's position is that while external motivational factors such as better jobs and career advancement have been the main motivational explanations of contributions to open source used by other researchers, intrinsic, enjoyment-based motivations, namely creativity, are the strongest motivators. A more basic question is whether a person motivated by external motives is necessarily self-interested. However, the internal/external divide is made to show the source of a motivational force, not the motivations themselves. All motivation, at least the type of motivation of importance in this manuscript, is ultimately internal, by which I mean, motivation is how an individual chooses to react, if at all, to external stimuli. The problem, as explained by Frey (1997), is that attempting to motivate people motivated by creativity or some intrinsic characteristic by, for instance paying them more, might be de-motivating. If people contribute to open source in order to signal their value from which they would like to profit, then the crowding-out effect would not be an issue. Data from the FLOSS (2002) report show that a large minority (24 percent) join the open source community in order to improve their job prospects; a far higher percentage want to learn and develop new skills (79 percent) and share knowledge and skills (50 percent). An employer attempting to exploit open source directly or use open source production techniques within a firm will have to find a balance between these different motives or risk canceling-out potential benefits.

#### Applying the framework

I have discussed the complex social situation in which the open source community operates including relevant features of software markets, the history of open source software, and the legal framework in which software is distributed. I have also criticized the positions of others in regards to the motives of contributors to open source projects. I am now in a position to ask whether the economic assumption of self-interested as defined by Lerner and Tirole and others agents adequately explains the success of open source.

First, in keeping with my framework, I will clearly state my assumptions about the software market and human nature. The context or social situation under consideration is clearly a competitive, multi-billion dollar, marketplace. A specific



feature of the market that influences the analysis is an underlying network structure in which the winner can take all or nearly all of a market. The structure means that, in a broad context, market developments are predictable. Competition shifts heavily in favor of a dominant design and competitors must be innovative in order to compete. Many software firms, which I take to be profit seeking entities, have been forced to follow a "losers" strategy, meaning that they support open source projects because they are unable to compete against the dominant design. For these firms, open source provides a means to shift competition away from the software market towards their relative strengths. In addition, technical developments, including the Internet, have made communication easy, quick and cheap; thereby providing like minded individuals a forum in which to communicate. The history of the software market demonstrated the need to minimize "hijacking", proprietary firms taking free software and using it for proprietary purposes. The result is that open source uses legally enforceable licenses to protect software, with some of the licenses imposing strong conditions on licensees who want to modify and reissue the software.

My general assumption about human nature in the particular economic context under consideration is that people are generally, predominately, self-interested (not selfish), but that self-interest is widely inclusive to include non-monetary aims. Non-self-interested motives, such as duties and altruism, can also influence behavior, but there appears to be a very specific form of altruism at play. In addition, and in contrast to some of the authors quoted above, people, even people motivated by non-self-interested motives, are rational. By which I mean that they are generally able to incorporate the perceived desires of others into their own interests, thereby remaining self-interested, but that they are also capable of acting benevolently. The aims of open source contributors are diverse and include fun, fame and fortune. Finally, the interests of open source contributors are long-term in nature and informed by history, in other words, their actions are not a product of passions.

The key to understanding the motives of open source contributors are the licenses they use. Licenses help to identify the long-term intentions of licensors. Choosing, let alone writing a license under which to distribute software is not a trivial matter, it involves careful consideration and has long-term, often explicitly irrevocable, consequences (see, for instance, the Academic Free License and the Open Software License). Much of the confusion surrounding discussions of open source can be solved by using Larry Rosen's (2004) distinction between academic and reciprocal

licenses. Together these licenses cover the most popular and influential open source software, including Linux (GNU-GPL) and Apache (Apache license). However, before discussing the intentions represented in those two types of licenses it is important to make one further distinction.

When reading academic papers on open source, including those referenced above, it is often confusing as to which open source community authors are referring. An important point when discussing open source licenses is to understand who is affected by the conditions of a license. In practice, users of open source software who don't intend to modify and reissue open source software or link their software to open source software are free to use it without practical restrictions. For the vast majority of users of open source software, open source is a gift without practical restrictions. For instance, GNU Emacs, GFortran, Emacs Lisp, VI Editor, and many hundreds of other programs can and are used by people who never intend to contribute to the open source community or modify the code in any way; indeed, many of these programs can be used to code proprietary projects. Rather, the conditions of open source licenses are mainly directed at those who plan to modify, link and reissue open source software. These are the contributors that allow open source software to continue to develop.

The motives of those using academic licenses are primarily altruistic. Recall that academic licenses are very permissive, allowing licensees to distribute copies of the copyrighted work and derivatives under any license they choose--even proprietary licenses. Those choosing to use such licenses are clearly intending to be of benefit others, even at a cost to themselves. It could be argued, as discussed above, that these altruists could do better by giving only to the needing instead of making their code equally available to rich and poor alike. However, altruism doesn't need to be selective, as long as the intention to help needy others is present. In addition, the nature of the Internet means that, as a practical matter, it is difficult to exclude the less fortunate from the less needy. The license itself provides few barriers to who can use the code or for what purpose, nor does the Internet. It could also be argued that the costs of contributing code are small and so using academic licenses is not altruistic. The argument might follow that contributors only give away code that has little value, but this is clearly not the case with the Apache server system and many other open source software projects that run "behind" servers. In any case, altruism doesn't require a great sacrifice, only the right intent.

Are contributors who use reciprocal licenses primarily altruistic? Once again, for the vast majority of users who don't intend to modify and reissue the code, reciprocal licenses provide a generous gift with no strings attached. However, for those who plan to modify and reissue software covered by reciprocal licenses, they are a type of conditional or selective altruism in that only those who agree to abide by the same, strict, license terms benefit. However, conditional altruism, setting conditions on who benefits from a contribution, begins to stretch the definition of altruism rather thin. Licensors using reciprocal licenses make the deal that they will only perform an altruistic act if others are willing to do the same in the future. The condition can be a heavy, exclusionary, burden, particularly for profit seeking individuals and firms. Reciprocal licenses create an exclusive island of software, where like-minded individuals flourish. A further danger with this license is that it is possible to exclude those who really need the benefits because they can't or don't want to fulfill the required conditions. Some, including, Jensen and Meckling (1994), would argue that such conditional altruism isn't altruism at all because altruism should be applied in a general, unconditional, way. I would be inclined to agree that conditional altruism is less than the ideal, but there is still reason to believe that the licensors intend to be altruistic to a select, like minded, group of individuals.

The altruistic effects of open source licenses are long lasting in that once source code is made available and changeable, it can't be readily removed from the public domain. Modifiers of open source are obliged to follow license conditions, and as long as they do so the original information contained in the software remains and contributions can potentially accumulate over time. Note that I am not claiming that all members of the open source community are altruistic; there are clearly self-interested reasons to contribute to open source projects as many authors have suggested. However, the primary effect of these licenses, particularly academic licenses, are benevolent. There are open source licenses which attempt to balance proprietary interests with open source sensibilities. These licenses, for instance, those issued by such firms as Apple and Netscape, have been less successful in part, I believe, because they attempt to gain access to a free resource, without recognizing that a primary motivating force of open source programmers is altruism. The point is that open source would not be possible without the altruistic intents of coders who issue their code using either academic or reciprocal licenses. Academic licenses increase the popularity of open source software, while reciprocal licenses have created a strong core of valuable, robust, freely available, software.

The supply of new open source software depends on the continuation of altruistic acts. Altruism is an internally motivated desire, preference or disposition. If Frey's (1997) analysis is correct, appealing to altruistic individuals via their self-interest may crowd-out altruistic acts. For example, offering to pay individuals to produce open source software may make them less productive. In addition, if potential open source licensors believe that their benevolence is being exploited solely for self-interested reasons, then they may reduce their contributions.

### Conclusions Chapter Eight

The question that originally prompted me to write this manuscript was: Can standard economic theory, as understood by authors such as Lerner and Tirole, account for the open source software development process? To answer that question I have analyzed the basic assumptions of standard economic theory. I have tried to show that economics, as it is often used by economists and business scholars, uses an ambiguous definition of self-interests that often makes normative claims about behaviors that are not supported by the theory. My next step was to show that rationality and self-interest, as those terms are used by economists, are often narrowly interpreted as selfishness, however, the elastic quality of those assumptions means that there is probably room for open source even within the assumptions of the standard economic theory. In fact, the vagueness of these assumptions means that very few activities can be labeled anything other than self-interested. The standard economic model is in fact so general that it indicates that although rationality and self-interest may be necessary, they are unlikely to be sufficient conditions to explain or describe economic behavior. Rationality and self-interest are nearly always a part of actions, the crux of the problem therefore falls on defining the distinctive characteristics which make an issue an economic issue and then applying the appropriate definition of self-interest and, perhaps to a lesser degree, rationality to fit the issue. For the case of open source, a narrow form of self-interest or selfishness can't adequately explain the motives of some of the key contributors to open source projects. In other words, despite the acknowledged prevalence of the motive of self-interest, that motive is not behind important types of open source software.

I have tried to show that there are two principal motives at work in the open source community. The first motive is profit maximization, which has to be the recognized aim of the firms participating in the open source community. The

other, original motive, represented by Richard Stallman and embedded in reciprocal licenses, is the desire to keep software freely available for everyone. Again, reciprocal licenses create an island of freely available code for those who are willing to release their code under the GNU-GPL. Academic licenses are even more difficult to explain under the motive of self-interest because they create a pool of software from which anyone can draw without the requirement of having to give anything in return. Regardless of whether open source is a self-interested or altruistic pursuit, it points to a weakness in economic assumptions. Simply stated, a strict interpretation of standard economic theory is unable to account for this important economic phenomenon. Certain bold economists who accept the approach of rational egoism may go so far as to declare that the behavior of computer programmers who freely contribute their code to open source projects using academic licenses is irrational in part because those programmers are not following their own narrow, self-interests by maximizing or satisfying their financial interests. However, as has been recognized for thousands of years, motives other than financial concerns may serve one's interests, suggesting that some common definitions of economic self-interest may need to be broadened to include a greater range of motives, and combination of motives, including power, liberty, fame, status, self-development, self-assertion, reputation, honor, and affections.



## BIBLIOGRAPHY

---

Abrahamson, E. (1996). "Management fashion." *Academy of Management Review* 21: 254-286.

Abrahamson, E. and G. Fairchild (1999). "Management fashion: Life-cycles, triggers, and collective learning processes." *Administrative Science Quarterly* 44 (4): 708-740.

Acheson, J. (1988). *The lobster gangs of Maine*. Hanover, New Hampshire: New England Universities Press.

Andreoni, J., Erard, B., and Feinstein, J. (1998). "Tax compliance." *Journal of Economic Literature* 36 (2): 818-860.

Arrow, Kenneth (1962). "The Economic Implications of Learning by Doing." *Review of Economic Studies* 29 (3): 155-173.

Arthur, B. (1994). *Increasing Returns and Path Dependence in the Economy*, Ann Arbor, Michigan: University of Michigan Press.

Axelrod, Robert (1984). *The Evolution of Cooperation*, New York, NY: Basic Books.

Barbrook, R. (1998). "The hi-tech gift economy." *First Monday* 3 (12).

Barbrook, R. (2005). "The hi-tech gift economy." *First Monday, Special Issue* 3 (12).

Becker, Gary S. [1957] (1971). *The Economics of Discrimination*, 2nd ed., Chicago, Illinois: University of Chicago Press.

Becker, Gary S. (1968). "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76 (2): 169-217.

Becker, Gary S. (1974). "A Theory of Marriage: Part II." *Journal of Political Economy* 82 (2): 11-26.

Becker, Gary S (1976). *The Economic Approach to Human Behavior*, Chicago, Illinois: University of Chicago Press.

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

Bergstrom, Theodore C. and Oded Stark (1993). "How Altruism Can Prevail in an Evolutionary Environment." *American Economic Review* 83 (2): 149-155.

Besen, S. M. and J. Farrell (1994). "Choosing How to Compete - Strategies and Tactics in Standardization." *Journal of Economic Perspectives* 8 (2): 117-131.

Bewley, T. F. (2000). *Why wages don't fall during a recession*, Cambridge, Massachusetts: Harvard University Press.

Blaug, Mark (1996). *Economic theory in retrospect*, 5th ed., Cambridge: University of Cambridge Press.

Brennan, Geoffrey and Philip Pettit (2004). *The Economy of Esteem*, Oxford: Oxford University Press.

Brennan, Michael (1994). "Incentives, Rationality, and Society." *Journal of Applied Corporate Finance* 7 (2): 31-39.

Broad, C.D. (1930). *Five Types of Ethical Theory*, London and New York: Routledge and K. Paul.

Broad, C.D. (1952). "Egoism as a Theory of Human Motives." In *Ethics and the History of Philosophy*, London and New York: Routledge and K. Paul, pp 218-231.

Broome, John (1999). *Ethics out of Economics*, Cambridge: Cambridge University Press.



Butler, Joseph (1983). *Five Sermons*, Stephen L. Darwall editor, Indianapolis, Indiana: Hackett Publishing Company.

Butler, Joseph (1986). *Fifteen Sermons*, Charlottesville, Virginia: Lincoln-Rembrandt Publishing,

Campbell, Robin (2004). *Seneca: Letters from a Stoic*, 2nd ed., London: Penguin Books.

Campbell-Kelly, M. and W. Aspray (1996). *Computer: A History of the Information Machine*, New York, NY: BasicBooks.

Casson, Mark and Marina Della Giusta (2006). "The Economics of Trust", in *Handbook of Trust Research*, Edited by Reinhard Bachmann and Akbar Zaheer, Northampton, Massachusetts: Edward Elgar Publishing Limited, pp 332-354.

Child, John and Rita McGraph (2001). "Organizations Unfettered: Organizational Form in an Information-Intensive Economy." *Academy of Management Journal* 44 (6): 1135-49.

Church, Jeffrey and Neil Gandal (1992). "Network Effects, Software Provision, and Standardization." *The Journal of Industrial Economics* 40 (1): 85-103.

Coase, Ronald (1937). "The Nature of the Firm." *Economica* 4 (16): 386-405.

Coase, Ronald [1976] (1994). "Adam Smith's View of Man", in *Essays On Economics and Economists*, Chicago, Illinois: University of Chicago Press, pp 95-118.

Collard, David (1978). *Altruism & Economics: A Study in Non-Selfish Economics*, Oxford: Martin Robertson & Co. Ltd.

Cooper, John M. and J. F. Procope (1995), editors. *Seneca: Moral and Political Essays*, Cambridge: Cambridge University Press.

Daan, Niels (1997). "TAC management in North Sea flatfish fisheries." *Journal of Sea Research* 37: 321-341.

- Dahlander, Linus and Mats Magnusson (2005). "Relationships between open source software companies and communities: observations from Nordic firms." *Research Policy* 34 (4): 481-493.
- Dalle J-M and N. Jullien (2003). "'Libre' software: turning fads into institutions?" *Research policy* 32 (7): 1-11.
- Dancy, Jonathan (1997), editor. *Reading Parfit*, Oxford: Blackwell Publishers.
- Dancy, Jonathan (2000). *Practical Reality*, Oxford: Oxford University Press.
- Darwall, Stephen (2003), editor. *Consequentialism*, Oxford: Blackwell Publishing.
- Dawes, Robyn M., Alphons J.C. van de Kragt, and John M. Orbell (1990). Reprinted in "Cooperation for the Benefit of Us-Not Me, or My Conscious." in Jane Mansbridge, editor, *Beyond Self-Interest*, Chicago, Illinois: University of Chicago Press, pp 97-110.
- Dawes, R. and Richard H. Thaler (1988). "Cooperation." *Journal of Economic Perspectives* 2 (3):187-96.
- Dawkins, Richard (1989). *The Selfish Gene*, 2<sup>nd</sup> ed., Oxford: Oxford University Press.
- Deci, E. (1975). *Intrinsic Motivation*, New York, NY: Plenum Press.
- Diamond, Jared (2008). Annals of Anthropology, "Vengeance Is Ours," *The New Yorker*, April 21, p.74.
- DiBona, Chris, Sam Ockman and Mark Stone (1999), editors. *Open Sources: Voices from the Open Source Revolution*, Sebastopol, California: O'Reilly Media.
- Douma Sytse and Hein Schreuder (2002). *Economic Approaches to Organizations*, Harlow, U.K.: Pearson Education Limited.
- Edgeworth, F.Y. (1881). *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences*, London: Kegan Paul.

Elster, Jon (1989). *Nuts and Bolts for the Social Sciences*, Cambridge: Cambridge University Press.

Elster, Jon and John E. Roemer (1991), editors. *Interpersonal Comparisons of Well-Being*, Cambridge: Cambridge University Press.

Farrell, J. and G. Saloner (1992). "Converters, compatibility, and the control of interfaces." *Journal of Industrial Economics* 40 (1): 9-35.

Feller, J., Brian Fitzgerald, Scott Hissam and Karim R. Lakhani (2005), editors. Introduction to *Perspectives on Free and Open Source Software*, Cambridge, Massachusetts: MIT Press.

Fleischacker, Samuel (2004). *On Adam Smith's Wealth of Nations: A Philosophical Companion*, Princeton and Oxford: Princeton University Press.

Florida, Richard (2002). *The Rise of the Creative Class*, New York, NY: Basic Books.

FLOSS Asia, FLOSS-ASIA survey (2004). Available: <http://oss.mri.co.jp/floss-asia/>, was conducted by Mitsubishi Research Institute, Inc. (MRI), Japan. It is a part of the project on the "Survey on Human Resource Development for Open Source Software Developers", supported by Ministry of Economy, Trade and Industry (METI), Japan.

FLOSS FINAL REPORT (2002). International Institute of Infonomics available at: <http://www.infonomics.nl/FLOSS/report/>, University of Maastricht, The Netherlands.

FLOSS US (2003). Stanford Institute for Economic Policy Research. Survey available: <http://www.stanford.edu/group/floss-us/>, Stanford University, Stanford, California.

Foss, Nicolai, J. (1997). *Resources, Firms and Strategies: A Reader in the Resource-Based Perspective*, Oxford: Oxford University Press.

Frank, Robert (1988). *Passions within Reason*, New York, NY: W.W. Norton.

Frank, Robert, Thomas Gilovich, and Dennis Regan (1988). "Does studying economics inhibit cooperation?" *Journal of Economic Perspectives* 7 (2): 159-172.

Franke, Nikolaus and Shah, Sonali (2003). "How communities support innovative activities: An exploration of assistance and sharing among end-users." *Research Policy* 32 (1): 157-178.

FSF (2010). Free Software Foundation, <http://www.fsf.org/>, accessed, April, 2010.

Frey, Bruno S. (1992). *Economics as a Science of Human Behaviour*, Boston and Dordrecht: Kluwer.

Frey, Bruno S. (1997). *Not Just For the Money: An Economic Theory of Personal Motivation*, Northampton, Massachusetts: Edward Elgar Publishing, Inc..

Frieberger, Paul and Michael Swaine (2000). *Fire in the Valley: The Making of the Personal Computer*, New York, NY: McGraw-Hill Education.

Friedman, Milton (1953). "The Methodology of Positive Economics", in *Essays in Positive Economics*, Chicago, Illinois: University of Chicago Press, pp 3-44.

Ghemawat, Pankaj (1991). *Commitment: the dynamic of strategy*, New York, NY: Free Press.

Ghosh, R.A. (1998). "Cooking pot markets: an economic model for the trade in free goods and services on the Internet." *First Monday* 3 (3).

Ghosh, R.A. (2005). "Understanding Free Software Developers: Findings from the FLOSS Study." In Feller, Joseph, Brain Fitzgerald, Scott A. Hissam and Karim R. Lakhani. *Perspectives on Free and Open Source Software*, Cambridge, Massachusetts: MIT Press, pp 23-46.

Ghosh, R.A. (2005b). "Cooking pot markets: a balanced value flows." In Ghosh, R.A., editor. *CODE: Collaboration and ownership in the digital economy*, editor R. Ghosh, Cambridge, Massachusetts: The MIT Press, pp 153-168.

Gilbert, Margaret (1989). *On Social Facts*, London and New York: Routledge

Gintis, Herbert (2000). *Game Theory Evolving*, Princeton, New Jersey: Princeton University Press.

Gill, Christopher (2003). "The School in the Roman Imperial Period." In *The Cambridge Companion to The Stoics*, Cambridge: Cambridge University Press, pp 33-58.

GNU-GPL (2007). The license can be viewed and downloaded from:  
<http://www.gnu.org/licenses/gpl-3.0.html>

Grant, M. (1971), translator. *Cicero: On the Good Life*, London: Penguin Books.

Green, G.L. (1999). <http://badtux.org/home/eric/editorial/economics.php>., accessed June 24, 2005.

Griffin, James (1986). *Well-Being: Its Meaning, Measurement and Moral Importance*, Oxford: Oxford University Press.

Hankinson, R.J. (2003). "Stoic Epistemology." In Inwood, Brad, editor. *The Cambridge Companion to the Stoics*, Cambridge: Cambridge University Press, pp 59-84.

Halloween Documents (2010). Document found at:  
[www.opensource.org/halloween](http://www.opensource.org/halloween), accessed, April, 2010. Site links to:  
<http://catb.org/~esr/halloween/index.html>

Hars, Alexander and Shaosong Ou (2001). "Working for Free? Motivations of Participating in Open Source Projects." Proceedings of the 34th Hawaii International Conference on System Sciences.

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

Hausman, Daniel M. (1992). "Why look under the hood?" In Hausman Daniel M., *Essays on Philosophy and Economic Methodology*, Cambridge: Cambridge University Press, pp 70-73.

- Hausman, Daniel M. and Michael S. McPherson (1994). "Economics, rationality, and ethics" in Hausman, Daniel M.. *The Philosophy of Economics: An Anthology*, 2nd ed., Cambridge: Cambridge University Press, pp 252-278.
- Hausman, Daniel M. and Michael S. McPherson (1996). *Economic analysis and moral philosophy*, Cambridge: Cambridge University Press.
- Hennart, Jean-Francois (1994). "The 'Comparative Institutional' Theory of the Firm: Some Implications for Corporate Strategy." *Journal of Management Studies* 31 (2): 193-207.
- Hertel, G., Niedner, S. and Herrmann, S. (2003). "Motivation of software developers in Open Source projects: An internet-based survey of contributors to the Linux kernel." *Research Policy* 32 (7): 1159-1177.
- Hicks, John (1956). *A Revision of Demand Theory*, Oxford: Clarendon Press.
- Hiltzik, Michael, A. (1999). *Dealers of Lightning: Xerox PARC and the Dawn of the Computer Age*, New York, NY: Harper Business.
- Hirschleifer, Jack (1985). "The Expanding Domain of Economics." *American Economic Review* 75 (6): 53-70.
- Hirschman, Albert O. (1985). "Against Parsimony: Three Easy Ways of Complicating some Categories of Economic Discourse." *Economics and Philosophy* 1: 7-21.
- Hirschman, Albert O. (1992). "The Concept of Interest: From Euphemism to Tautology." In Hirschman, Albert O., editor. *Rival Views of Market Society*, Cambridge, Massachusetts: Harvard University Press, pp 35-48.
- Hirschman, Albert O. (1997). *The Passions and the Interests: Political Arguments for Capitalism before Its Triumph*, 12th edition, Princeton, New Jersey: Princeton University Press.
- Hobbes, Thomas [1651] (1996). *Leviathan*, Revised student edition, Cambridge: Cambridge University Press.

Hume, David [1751] (1998). *An Enquiry concerning the Principles of Morals*, edited and introduction by Tom L. Beauchamp, Oxford: University of Oxford Press.

Inwood, Brad (2003). *The Cambridge Companion to The Stoics*, Cambridge: Cambridge University Press.

Inwood, Gerson and Hutchinson (1994). *The Epicurus Reader: Selected Writings and Testimonials*, Indianapolis, Indiana: Hackett Publishing Company, Inc.

Inwood, Brad and L.P. Gerson (1997). *Hellenistic Philosophy: Introductory Readings*, 2<sup>nd</sup> ed., Indianapolis, Indiana: Hackett Publishing Company, Inc.

Jacobsen v. Katzer (2008). Reference found at:  
<http://www.cafc.uscourts.gov/opinions/08-1001.pdf>, accessed April, 2010.

Jensen, Michael C. (1994). "Self-Interest, Altruism, Incentives, and Agency Theory." *Journal of Applied Corporate Finance* 7 (2): 40-45.

Jensen, Michael C. and William H. Meckling (1976). "Theory of the firm: Managerial behavior, agency costs and ownership structures." *Journal of Financial Economics* 3: 305-360.

Jensen, Michael C. and William H. Meckling (1994). "The Nature of Man." *Journal of Applied Corporate Finance* 7 (2): 4-19.

Jeppesen, Lars Bo and Lars Frederiksen (2006). "Why Do Users Contribute to Firm-Hosted User Communities? The Case of Computer-Controlled Music Instruments." *Organization Science* 17 (1): 45-63.

Kahneman, D. J. L. Knetsch and R. Thaler (1986a). "Fairness as a Constraint on Profit Seeking Entitlements in the Market." *American Economic Review* 76 (4): 728-741.

Kahneman, D. J. L. Knetsch and R. Thaler (1986b). "Fairness and the Assumptions of Economics." *Journal of Business* 59: S285-S300.

Kahneman, D., Slovic, P. and A. Tversky, (1982). *Judgment under Uncertainty: Heuristics and Biases*, Cambridge: Cambridge University Press.

Kalin, Jesse (1970). "In Defense of Egoism." In Gauthier, David P., editor. *Morality and Rational Self-Interest*, Englewood Cliffs, New Jersey: Prentice-Hall, p 79.

Katz, Michael, L., and Carl Shapiro (1985). "Network Externalities, Competition, and Compatibility." *The American Economic Review* 75 (3): 424-440.

Katz, Michael, L. and Carl Shapiro (1994). "Systems Competition and Network Effects." *The Journal of Economic Perspectives* 8 (2): 93-115.

Kavka, Gregory S. (1986). *Hobbesian Moral and Political Theory*. Princeton, New Jersey: Princeton University Press.

Kernighan, Brian W. and Rob Pike (1984). *Unix Programming Environment*, Englewood Cliffs, New Jersey: Prentice-Hall

Khalil, Elias L. (2004). "Critical Commentary: What is altruism?" *Journal of Economic Psychology* 25 (1): 97-123.

Khazam, David C. and J. and Mowery (1994). "The Commercialization of RISC: Strategies for the Creation of Dominant Designs." *Research Policy* 23 (1): 89-102.

Klemperer, P. (1995). "Competition when Consumers have Switching Costs." *Review of Economic Studies* 62 (4): 515-539.

Kurz, Mordecai (1978). "Altruism as an Outcome of Social Interaction." *The American Economic Review* 68 (2): 216-222.

Lakhani, Karim R. and Robert G. Wolf (2003). "Why Hackers Do What They Do: Understanding Motivation Effort in Free/Open Source Software Projects." MIT Sloan School of Management, Working Paper 4425-03.

Lakhani, Karim R. and Robert G. Wolf (2005). "Why Hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects." In Feller, Joseph, Brain Fitzgerald, Scott A. Hissam and Karim R. Lakhani.



*Perspectives on Free and Open Source Software*, Cambridge, Massachusetts: MIT Press, pp 3-22.

Lakhani, Karim R. and E, von Hippel (2003). "How open source software works: 'Free' user-to-user assistance." *Research Policy* 32 (6): 923-943.

Lepper, Mark R. and David Greene (1978), editors. *The Hidden Costs of Reward: New Perspectives on the Psychology of Human Motivation*, Hillsdale, New York: Erlbaum.

Lerner, J. and J. Tirole (2000). "The Simple Economics of Open Source." Working Paper, Accessed at: <http://ssrn.com/abstract=224008>.

Lerner, J. and J. Tirole (2001). "The open source movement: Key research questions." *European Economic Review* 45 (4-6): 819-826.

Lerner, J. and J. Tirole (2002). "The Scope of Open Source Licensing." Harvard NOM Working Paper No.02-42 <http://ssrn.com/abstract=354220>.

Lerner, J. and J. Tirole (2002b). "Some Simple Economics of Open Source." *The Journal of Industrial Economics* 50 (2): 197-234.

Lerner, J. and J. Tirole (2005). "Economic Perspectives on Open Source." In Feller, Joseph, Brian Fitzgerald, Scott A. Hissam and Karim R. Lakhani. *Perspectives on Free and Open Source Software*, Cambridge, Massachusetts: MIT Press, pp 47-78.

Long, A.A. (1986). *Hellenistic Philosophy: Stoics, Epicureans, Sceptics*, 2nd ed., Berkeley, California: University of California Press.

Long, A.A. (1996). *Stoic Studies*, Berkeley, California: University of California Press.

Mandeville, Bernard [1732] (1988). *Fable of the Bees: or, Private Vices, Publick Benefits*, commentary, etc., by F.B. Kaye. Indianapolis, Indiana: Liberty Funds.

Mansbridge, Jane (1990), editor. *Beyond Self-Interest*, Chicago, Illinois: The University of Chicago Press.

Mansbridge, Jane (1990b). "On the Relation between Altruism and Self-Interest." In Mansbridge, Jane, editor. *Beyond Self-Interest*, Chicago, Illinois: The University of Chicago Press, pp 133-146.

Marwell, Gerald and Ruth Ames (1981). "Economists free ride. Does anyone else? Experiments on the provision of public goods. IV." *Journal of Public Economics* 15 (3): 295-310.

Maslow, A.H. (1943). "A Theory of Human Motivation." *Psychological Review* 50 (4): 370-96.

McGraw, K.O. (1978). "The Detrimental Effects of Reward on Performance: A Literature Review and a Prediction Model." In Lepper, M.R. and Greene, D., editors. *The Hidden Cost of Rewards: New Perspectives on the Psychology of Human Motivation* Hillsdale, New Jersey: Erlbaum, pp 33-60.

McKusick, Kirk (1999). "Twenty Years Berkeley Unix." In DiBona, Chris, Sam Ockman and Mark Stone, editors. *Open Sources: Voices from the Open Source Revolution*, Sebastopol, California: O'Reilly Media, pp 31-46.

Meeker, Heather J. (2008). *The Open Source Alternative: Understanding Risks and Leveraging Opportunities*, Hoboken, New Jersey: John Wiley and Sons.

Mitchels, Robert (1966). *Political Parties*. Glencoe, Illinois: Free Press.

Mill, John Stuart (1994). "Excerpt form On the Definition of Political Economy and the Method of Investigation Proper to It." In Hausman, Daniel M.. *The Philosophy of Economics: An Anthology*, 2nd ed., Cambridge: Cambridge University Press, pp 52-68.

Millstein, I and P. MacAvoy (1998). "The Active Board of Directors and Performance of the Large Publicly Traded Corporation." *Columbia Law Review* 98: 1283-1321.

Milo, Ronald, D. (1973). *Egoism and Altruism*, Belmont, California: Wadsworth Publishing Company, Inc.

- Moody, Glyn (2001). *Rebel Code: Linux*, New York, NY: Perseus Books
- Moon, J. Y. and L. Sproull (2000). "Essence of distributed work: The case of the Linux kernel." *First Monday* 5 (11).
- Moore, G.E. (1993). *Principia Ethica*, Cambridge: The University of Cambridge Press.
- Mulgan, Tim 2001. *The Demands of Consequentialism*, Clarendon: The Oxford University Press.
- Myers, Milton L. (1983). *The Soul of Modern Economic Man: Ideas of Self-Interest, Thomas Hobbes to Adam Smith*, Chicago, Illinois: The University of Press.
- Nagel, Thomas (1970). *The Possibility of Altruism*, Princeton, New Jersey: Princeton University Press.
- Nagel, Thomas (1986). "Agent-Relativity and Deontology." *The View From Nowhere*, New York: Oxford University Press, pp 164-185.
- Moll, Jorge, Frank Krueger, Roland Zahn, Matteo Pardini, Ricardo de Oliverira-Souza, and Jordan Grafman (2006). "Human fronto-mesolimbic networks guide decisions about charitable donation." *Proceedings National Academy of Sciences* 103 (42): 15623-15628.
- Neumann, Peter, G. (2005). "Attaining Robust Open Source Software." In Feller, Joseph, Brain Fitzgerald, Scott A. Hissam and Karim R. Lakhani. *Perspectives on Free and Open Source Software*, Cambridge, Massachusetts: MIT Press, pp 123-126.
- O'Connor, Eugene (1993), translator. *The Essential Epicurus*, Amherst, New York: Prometheus Books.
- Olson, Mancur (2000). *The Logic of Collective Action: Public Goods and the Theory of Groups*, 18th printing. Cambridge, Massachusetts: Harvard University Press.

O'Mahony, Siobhan (2005). "Non-Profit Foundations and Their Role in Community-Firm Software Collaboration.", In Feller, Joseph, Brian Fitzgerald, Scott A. Hissam and Karim R. Lakhani . *Perspectives on Free and Open Source Software*, Cambridge, Massachusetts: MIT Press, pp 393-414.

Open Source Initiative (2010). <http://www.opensource.org/licenses/gpl-license>, Accessed, April, 2010.

Osborne, Martin J. and Ariel Rubinstein (1994). *A Course in Game Theory*, Cambridge, Massachusetts: MIT Press.

OSRM (2010). <http://www.osriskmanagement.com/solutions.html>, Accessed April, 20, 2010.

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

Ostrom, E., Walker, J., and R. Gardner, (1992). "Covenants with and without a sword: self-governance is possible." *American Political Science Review* 86: 404-417.

Ozging, J. (1999). *Altruism*, Westport, CT: Praeger Publishers.

Parfit, Derek (1984). *Reasons and Persons*, Oxford: Oxford University Press.

Perens, Bruce (1999). "The Open Source Definition." In DiBona, Chris, Sam Ockman and Mark Stone , editors. *Open Sources: Voices from the Open Source Revolution*, Sebastopol, California: O'Reilly Media, pp 171-188.

Penrose, Edith (1995). *The Theory of the Growth of the Firm*, 3rd ed., Oxford: Oxford University Press.

Pettit, Philip (1993). *The Common Mind*, Oxford: Oxford University Press.

Popper, Karl R. (1985). "Popper Selections.", edited by David Miller, Princeton, New Jersey: Princeton University Press.

Popper, Karl R. (1994). "The Myth of the Framework." Edited by M.A. Notturmo, London and New York: Routledge

Prichard, H.A. [1928] (2002). "Duty and Interest." In MacAdam, Jim , editor. *Moral Writings*, Oxford: Oxford University Press, pp 21-49.

ProCD, Inc. versus Zeidenberg, 86 F. 3rd. 1447 (7th Cir. 1996).

Rand, Ayn (2007). *Atlas Shrugged*, London: Penguin Classics.

Raphael, D.D. (1976), editor. Introduction to: *The Theory of Moral Sentiments*, Oxford: Oxford University Press.

Raymond, Eric S. (1999). "Revenge of the Hackers." In DiBona, Chris, Sam Ockman and Mark Stone, editors. *Open Sources: Voices from the Open Source Revolution*, Sebastopol, California: O'Reilly Media, pp 207-220.

Raymond, Eric S. (2001). *The Cathedral & The Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary*, Sebastopol, California: O'Reilly Media.

Reid, Thomas [1785] (1969). *Essays on the Intellectual Powers of Man*, introduction by B. Brody, Cambridge, Massachusetts and London, England: The M.I.T. Press.

Robbins, Lionel (1935). *An Essay on the Scope and Nature of Economic Science*, 2nd ed, London: Macmillan.

Rogers, Kelly (1997), editor. *Self-Interest: An Anthology of Philosophical Perspectives*, New York and London: Routledge.

Rosen, Larry (2004). *Open Source Licensing: Software Freedom and Intellectual Property Law*, New Jersey: Prentice Hall.

Rosenberg, Alex (2000). *Philosophy of Science*, New York and London: Routledge.

Rossi, Maria A. (2004). "Decoding the 'Free/Open Source (F/OSS) Software Puzzle' a survey of theoretical and empirical contributions."

<http://opensource.mit.edu/papers/rossi.pdf>, Accessed Aug., 2005.

Rusovan, Srdjan, Mark Lawford, and David Lorge Parnas (2005). "Open Source Software Development: Future or Fad?" In Feller, Joseph, Brain Fitzgerald, Scott A. Hissam and Karim R. Lakhani . *Perspectives on Free and Open Source Software*, Cambridge, Massachusetts: MIT Press, pp 107-122.

Ryan, R.M., and E.L. Deci (2000). "Intrinsic and extrinsic motivations: Classic definitions and new directions." *Contemporary Educational Psychology* 25: 54-67.

Scheffler, Samuel (2001). *Boundaries and Allegiances: Problems of Justice and Responsibility in Liberal Thought*, Oxford: Oxford University Press.

Schlick, Moritz (1939). *Problems of Ethics*, in Rynin, David, translator., New York, NY: Prentice Hall, Inc..

Schumpeter, Joseph A. [1942] (1975). *Capitalism, Socialism and Democracy*, New York, NY: Harper.

Searle, John L (2001). *Rationality in Action*, Cambridge, Massachusetts: MIT Press.

Sen, Amartya K. (1978). "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory." In H. Harris, editor, *Scientific Models and Man*, London: Oxford University Press, pp. 317-344.

Sen, Amartya K. (1987). *On Ethics & Economics*, Oxford: Blackwell Publishers, Ltd..

Sen, Amartya K. (1999). *Reason before Identity*, Oxford: Oxford University Press.

Sharples, R. W. (1996). *Stoics, Epicureans and Sceptics: An introduction to Hellenistic Philosophy*, London and New York: Routledge.

Shaver, Robert (1999). *Rational egoism: A selective and critical history*, Cambridge: Cambridge University Press.

Sidgwick, Henry (1981). *The Methods of Ethics*, based on an unabridged and unaltered 7th edition from 1907, Indiana, Indianapolis: Hackett Publishing Company.

Sidgwick, Henry [1907] (1988). *Outlines of the History of Ethics for English Readers*, reprint, Indianapolis, Indiana, Indianapolis: Hackett Publishing Company, Inc..

Simon, Herbert (1961). *Administrative Behaviour*, 2nd ed., New York: Macmillan.

Simon, Herbert (1985). "Human Nature in Politics: The Dialog of Psychology with Political Science." *American Political Science Review* 79 (2): 293-304.

Smart, J.J.C and Bernard Williams (1973). *Utilitarianism: For and Against*, Cambridge: Cambridge University Press.

Smith, Adam (1976). *The Theory of Moral Sentiments*, Edited by D.D. Raphael and A.L. Macfie, Oxford: Oxford University Press.

Smith, Adam (1993). *An Inquiry into the Nature and Causes of the Wealth of Nations*, introduction and notes by Kathryn Sutherland, Oxford: Oxford University Press.

Smith, Marc A. and Peter Kollock (1999), editors. *Communities in Cyberspace*, London and New York: Routledge.

Smith, Martin F (2001), translator. *Lucretius: On the Nature of Things*, Cambridge: Hackett Publishing.

St. Laurent, Andrew M. (2004). *Understanding Open Source and Free Software Licensing*, Sebastopol, California: O'Reilly and Associates.

Stallman, Richard (1999). "The GNU Operating System and the Free Software Movement." in Chris DiBona, Sam Ockman and Mark Stone, *Open Sources: Voices from the Open Source Revolution*. Sebastopol, California: O'Reilly Media, pp 53-70.

Straw, Barry M. and Lisa D. Epstein (2000). "What Bandwagons Bring: Effects of Popular Management Techniques on Corporate Performance, Reputation, and CEO Pay." *Administrative Science Quarterly* 45 (3): 523-556.

Teece, David, J. (1980). "Economies of Scope and the Scope of the Enterprise." *Journal of Economic Behavior and Organization* 1, 223-33.

Torvalds, L. and D. Diamond (2001). *Just for Fun: The Story of an Accidental Revolutionary*, New York: HarperCollins Publishers.

Trivers, Robert (1974). "Parent-Offspring Conflict." *American Zoologist* 14: 249-264.

Tsouana, Voula (1998). *The Epistemology of the Cyrenaic School*, Cambridge: Cambridge University Press.

Tuomela, Raimo (1995). *The Importance of Us: A Philosophical Study of Basic Social Notions*, Stanford, California: Stanford University Press.

Van Wendel de Joode, R. (2005). *Understanding open source communities: an organizational perspective*, Ph.D. thesis. Enschede: Febodruk BV.

Varian, Hal (1999). *Intermediate Microeconomics: A Modern Approach*, 5th ed., Berkeley, California: University of California Press.

von Hippel, E. (1988). *The Sources of innovation*, New York: Oxford University Press.

von Hippel, E. (2001). "Innovation by user communities: Learning form open source software." *Sloan Management Review* 42 (4): 82-86.

von Hippel, E. (2005). *Democratizing innovation*, Cambridge, Massachusetts: MIT Press.

von Hippel E and G. Von Krogh (2003). "Open Source Software and 'Private-Collective' Innovation Model: Issues for Organization Science." *Organization Science* 14 (2): 209-23.



Von Krogh, von, Georg and Eric von Hippel (2003) "Special issue on open source software development." *Research Policy* 32 (7): 1149-1157.

Watson, John (1895). *Hedonistic Theories*, London: Macmillan.

Wayner, P. (2000). *Free for All*, New York: Harper Business.

Weber, Steven (2004). *The Success of Open Source*, Cambridge, Massachusetts: Harvard University Press.

Wernerfelt, B. (1984). "A resource-based view of the firm." *Strategic Management Journal* (5): 272-280.

West, Joel and Jason Dedrick (2001). "Open Source Standardization: The Rise of Linux in the Network Era." *Knowledge, Technology & Policy* 14 (2): 88-112.

West, Joel and Siobhan O'Mahony (2008). "The Role of Participation Architecture in Growing Sponsored Open Source Communities, forthcoming in Industry and Innovation." *Industry and Innovations* 15 (2): 145-168.

Williams, Bernard (1973). "Egoism and altruism." In Williams, Bernard . *Problems of the Self*, Cambridge: Cambridge University Press, pp 250-265.

Williamson, O.E. (1975). *Markets and Hierarchies: Analysis and Antitrust Implications*, New York: Free Press.

Williamson, O.E. (1985). *The Economic Institutions of Capitalism*, New York: Free Press.

Williamson, O.E. (1993). "The Logic of Economic Organization." In Williamson, Oliver E., and Sidney G. Winter, editors. *The Nature of the Firm: Origins, Evolution, and Development*, Oxford: Oxford University Press, pp 90-116.

Williamson, O.E. (1993). "Calculativeness, trust, and economic organization." *Journal of Law and Economics* 36 (April): 453-486.

Williamson, O.E. (1998). "Transaction cost economics: How it works; where it is going." *De Economist* 146 (April): 23-58.

Wilson, Edward, O. (1975). *Sociobiology: The New Synthesis*, Cambridge, Massachusetts: Harvard University Press.