GROUP SELECTION, ALTRUISTIC PREFERENCES AND ECONOMIC THEORY

by

Ruth Ellen Forsdyke

B.Sc.H., Queens University, 1989
B.Ed., Queens University, 1992
M.A., Queens University, 1997

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

In the Department of Economics

SIMON FRASER UNIVERSITY
August 2004

APPROVAL

Name: **Ruth Ellen Forsdyke**

Degree: **PhD (Economics)**

Title of Thesis: **Group Selection, Altruistic Preferences and Economic Theory**

Examining Committee:

Chair: Prof. Terence Heaps

Prof. Jasmina Arifovic
Senior Supervisor

Prof. Clyde Reed
Supervisor

Prof. Steeve Mongrain
Supervisor

Prof. Gordon Myers
Internal Examiner

Prof. B. Curtis Eaton University of Calgary
External Examiner

Date Approved: August 4th, 2004

ABSTRACT

Social scientists have long debated the extent to which humans are motivated by prosocial concerns versus material self-interest. While the 20th century focus on self-interest has received support from traditional Darwinian theory, there is currently renewed interest in prosocial emotions, due to factors such as work by experimental economists and a reconsideration of group selection within biology. The dissertation has two parts, related by the common theme of altruistic preferences.

Part 1 examines the debate on which units are subject to Darwinian selection. Do these units include genes, organisms and/or groups of organisms? Due to difficulty in explaining altruism as an organism level adaptation, biologists proposed that selection might act upon groups allowing for group level adaptations. During the 1960s and 1970s, group selection fell into disrepute. Various models suggested that it was unlikely to be biologically important and alternate explanations of altruism were provided. Currently, although group selection is receiving serious reconsideration, the "units of selection debate" remains unresolved. I argue that this failure is due to widespread acceptance that the largest replicators (heredity units of selection) correspond to genes. It is proposed here that replicators exist at a variety of levels, including genes, organisms and groups. Higher level replicators are constructed from lower level replicators via inheritable strategies for assortment into groups, allowing for adaptations at multiple levels.

Part 2 examines some potential effects of altruistic preferences for behaviour and information. When making decisions, altruists attempt to account for both private material net benefits and external consequences. High private material benefits may lead to asocial choices, but at the cost of negative social emotions, i.e. "cognitive dissonance". Dissonance can be reduced *ex post* by "selectively exposure" to (choosing or avoiding) information sources that report that one's behaviour is respectively prosocial or asocial. The aggregate effects of such "selective attention" affect information demand potentially biasing information sources, which can then affect behaviour in two-way feedback with potential for lock in to low welfare equilibrium. The dynamics of activity choices and information are modeled with application to 1) transition from feudal to market economies, and 2) environmental damages and policy.

DEDICATION

This thesis is dedicated to Dad, Mom, Sara, Polly, Charlotte, Thomas, Sophie, Eric, and Sweetpea, the cat, who have provided support and encouragement during the graduate school experience.

ACKNOWLEDGEMENTS

I would especially like to thank my senior supervisor, Jasmina Arifovic, for providing much encouragement and for reading and making many useful comments on the various versions of this thesis as it evolved. Also many thanks to my junior supervisors, Steeve Mongrain and Clyde Reed, and the examiners, Curtis Eaton and Gordon Myers, for reading the dissertation, asking excellent questions, and making suggestions for improvements and future work. I have also benefited from graduate course instruction from many of the faculty at SFU. Of particular importance to this thesis were Richard Lipsey's course on long run economic growth, Lawrence Boland's course on economic methodology, Curtis Eaton's course on industrial organization, which provided much practice in creative modeling, and Jasmina Arifovic's course on macroeconomics, which contributed to my background in evolutionary economics.

I would also like to thank Richard Dawkins for writing the incredible book *The Selfish Gene* (1976), which initially stimulated my interest in evolution. Particular thanks goes to my father, Donald Forsdyke, who years ago drew my attention to this book. Dad also read through the thesis in its entirety and provided comments, suggestions for improvements, and encouragement.

Lastly, thank you to all of the staff in the economics department, including Barb Clark, Tim Coram, Lorraine Cushie, Sherrill King, Gisela Seifert, Tracey Sherwood, Laurie Sutterlin, Dorothy Wilson, and Dorothy Wong, who have not only provided essential administrative and technical support during the graduate school experience but have also made a large contribution to the department's friendly environment.

Any errors in this dissertation are the sole responsibility of the author.

TABLE OF CONTENTS

LIST OF FIGURES

Part I

Part II

LIST OF TABLES

Part I

GLOSSARY FOR PART I

Allele The term "allele" is used to refer to the alternative forms of a gene that can exist at a particular gene locus. For example, at a gene locus for eye-colour in a diploid organism, there might be a blue-eye allele and a brown-eye allele.

Altruistic Action In this thesis, an action is referred to as "altruistic" if the donor of the action incurs a fitness cost as a *direct* result of the action, while conferring a fitness benefit on the recipient of the action, also as a *direct* result of the action. Sometimes altruistic actions are referred to as "apparently altruistic". For example, in the case of reciprocal altruism, the donor expects to be the recipient of an altruistic action at a later date and over time to gain more net fitness benefits than if he had not committed the initial altruistic action. Sometimes the term "helping" or "cooperating" is used instead of the term "altruism" because genes for actions that on balance lower their own fitness or the fitness of their organism carriers would not be expected to evolve.

Amino acids Amino acids are the "building blocks" of proteins. A protein is composed of a linear sequence of amino acids.

Assortative grouping Out of a larger population of entities that is subdivided into groups, where there are multiple types of entities which vary according to some characteristic(s), grouping is "assortative" if the within-group frequency of types differs from population-wide frequency. Assortative grouping either occurs due to chance during a random sampling process or is non-random, that is, correlated with variant type. If the entities reproduce and assortment is non-random (i.e. correlated with type) and if the source of variation between types is inheritable, then the hereditary information for type is correlated with the capacity to assort. It then follows that the ability to assort is contained in the hereditary information for type and is therefore inheritable (see inheritable assortative strategies).

Chromosome A cell's total genetic information (in the form of a chain of nucleotides, which constitute the DNA) is divided up into non-attached sections called chromosomes. In a haploid cell, there is one copy of each type of chromosome. In a diploid cell there are two copies of each type of chromosome, one of which is inherited from the father (paternal) and one of which is inherited from the mother (maternal). Some organisms such as yeast have more than two copies of each chromosome. There are 23 pairs of chromosomes in humans for a total of 46.

Codon A codon is a section of the DNA, which consists of three nucleotides in a row on an exon. Each codon corresponds to a specific amino acid. A linear sequence of codons on the DNA contains the information to instruct the cell's protein making apparatus to produce a specific chain of amino acids, which corresponds to a protein.

Coefficient of Assortment Conditional upon a given group member being type X, the probability that any other group member is also type X is equal to *f*, the coefficient of assortment. Hamilton (1975) illustrated that in the kin selection case, the coefficient of assortment corresponds to the coefficient of relatedness.

Coefficient of Relatedness (see Hamilton's rule)

Cross over During meiosis each pair of chromosomes is duplicated after which each set of four chromosomes line up in the centre of the cell. The DNA molecules are broken and material is swapped between chromosomes in an orderly manner such that the locations of specific gene loci are normally preserved.

"Darwinian" selection process In this thesis, a selection (or grouping) process is referred to as "Darwinian" if three conditions are satisfied as discussed by Lewontin (1970). These are 1) variation within the population of entities; 2) correlation between the variants and their survival and reproduction rates; 3) this variation is inheritable. Criterion 3 is used in a weak "traditional" sense, whereby offspring entities are required to resemble parents, and also the stronger "Dawkins- Fisher" sense, in which entities must produce perfect copies and are defined as "replicators".

Deme In genetics, a "deme" is defined as a locally interbreeding population. In a broader sense, as used in the literature on group selection, a deme is a group of entities which interact in some manner to affect each other's fitness.

Diploid A cell is diploid if it possesses two copies of each type of chromosome. Diploid organisms (organisms with diploid somatic cells) reproduce sexually such that one chromosome of each pair is inherited from the father and the other from the mother.

DNA (deoxyribonucleic acid) A DNA molecule is composed of two long chains of nucleotides, which twist to form a double helix. There are four types of nucleotides on a DNA molecule, C (cytosine), G (guanine), T (thymine), and A (adenine). Each of the two chains is related to the other because a given nucleotide pairs with a particular nucleotide on the other chain; an "A" always pairs with a "T" and a "C" always pairs with a "G". For example, if one chain can be represented as ATTGCTTG, then the other chain is TAACGAAC. This pairing capacity allows the DNA molecule to use one strand as a template to make a perfect double-stranded copy.

Domain of Actions In this thesis, the "domain of actions" refers to the "material" part of an entity that interacts with the environment, to affect the entity's fitness.

Williams (1992) refers to this domain as the "material domain" which includes Dawkins' "vehicles" and Hull's "interactors". These "actions" can be thought of as the "tools" by which the heredity information (see the domain of heredity) differentially survives and reproduces. For example, organisms are the "actions" of "genomes" and a blue-eye phenotype is the "action" of a blue-eye gene. It is useful to distinguish between the action and hereditary domains, because "mutations" are inheritable in the latter case but not in the former. The use of the term "actions" is consistent with standard game theory in which actions are the physical outcome of strategies of play that are used in the game. Correspondingly, in evolutionary game theory, genes can be thought of as "hereditary strategies" which provide the "blueprints" for phenotypic "actions".

Domain of Heredity In this paper, the "domain of hereditary" refers to the hereditary information component of an organism and is distinguishable from the "domain of actions" in that mutations in the hereditary domain are inherited by offspring, while those in the action domain are not. Williams (1992) uses the term "codex" or "codical domain" to refer to the hereditary domain.

Dominance One allele at a chromosomal locus in the heterozygote condition is dominant relative to the other allele (which is referred to as recessive) if it suppresses the phenotypic effect of this recessive allele. For example, an allele for the brown-eye phenotype will suppress the expression of an allele for the blue-eye phenotype if they are paired at a locus.

Evolutionarily Stable Equilibrium refers to the equilibrium concept used in evolutionary game theory in which the frequency of strategies in the population does not change. These strategies are referred to as evolutionary stable strategies. This requires that a small population of mutants or invaders from outside the population, which play a strategy (or strategies) other than the evolutionary stable strategies, will be selected against such that they will not be able to establish an equilibrium presence in the population.

Evolutionarily Stable Strategy A Strategy that persists and is not selected against in an evolutionarily stable equilibrium.

Exon The sections of a gene that code for the gene's protein. Each exon codes for a specific sub-unit of the complete protein. A gene's exons are sometimes separated by "introns", long sections of DNA that do not code for proteins.

Fitness Dawkins (1982) writes: "A technical term with so many confusing meanings that I have devoted a whole chapter to discussing them (chapter 10)." In this thesis, the "absolute fitness" of an entity refers to the number of offspring in the next generation. "Relative fitness" refers to the absolute fitness divided by the average absolute fitness in the population.

Gene (standard molecular biology) The section of a chromosome which codes for one chain of amino acids in a protein.

Gene (Williams, 1966; Dawkins, 1976) Smaller sections of a "parent" DNA molecule are less likely to be fragmented by meiosis than larger sections and hence have a higher probability of being represented in the gametes and therefore the next generation. Smaller DNA fragments therefore have higher copying fidelity than larger fragments. Copying fidelity is an important determinant of the relative fitness of any replicating entity (replicator), including DNA fragments in competition with other variant fragments. All else equal high copying fidelity fragments will be perfectly copied relatively often such that they win out in the competition to leave a lineage of identical descendents. Williams' "gene" corresponds to the largest fragment of DNA that has sufficient copying fidelity to potentially leave a lineage of such descendants.

Gene locus The specific location of a gene on a chromosome.

Gene pool "The whole set of genes in a breeding population. The metaphor on which the term is based ...de-emphasizes the undeniable fact that genes actually go about in discrete bodies, and emphasizes the idea of genes flowing about the world like a liquid (Dawkins, 1982, p. 294)."

Genome The genetic information of an organism in its entirety. For the human, this consists information is located on 46 chromosomes.

Genotype An organism's genotype is the set of genes that it possesses. Often, the term refers to the types of alleles at a specific locus or types of alleles at a set of loci in the organism.

Germ-line cell Germ-line cells can potentially copy and transmit their DNA to the next generation. They include both the gametes (sperm and eggs) and also the cells that produce these gametes. Compare with somatic cells.

Hamilton's rule (roughly $bf > c$), specifies the conditions under which altruism can evolve under conditions of kin selection, where c represents the cost to the donor of the altruistic act, b represents the benefit to the recipient of the altruistic act, and f represents the coefficient of relatedness, the degree of relatedness between the donor and the recipient of the altruistic act. For example, $f = ½$ for siblings and parents, and ¼ for cousins and grandchildren.

Haploid A cell is haploid if it possesses one copy of each type of chromosome. In diploid organisms (see diploid), the gametes are haploid. During sexual reproduction, two gametes, one from each parent fuse to form a diploid zygote.

In cases such as male bees, in which all of the somatic cells are haploid, the organism is referred to as haploid.

Heterozygous refers to the condition of having two non-identical alleles on at a chromosomal locus, where a chromosomal locus refers to the pair of corresponding loci on a chromosome pair.

Homozygous refers to the condition of having two identical alleles at chromosomal locus (see heterozygous).

Inclusive Fitness refers to the sum of the benefits and costs of a particular gene's actions that are conferred to identical genes in other organisms where each benefit or cost is weighted by the coefficient of relatedness (see Hamilton's rule) of the donor of the helping action to the recipient. Under conditions of kin selection, a special case of intrademic group selection, genes should evolve that behave as if maximizing their inclusive fitness. It is important not to ignore the "as if"—what is really happening is that genes that code for two pleiotropic actions, "helping kin" and "assorting with kin who will help them", have higher fitness than genes that code for "not-helping kin" and "not assorting with helping kin". Of course, genes that code for "not helping kin" and "assorting with helping kin" would have the highest possible fitness. However, this option is unlikely to be available because non-helping kin tend to inherit non-helping kin.

Independent Assortment If two genes, which code for different traits and are located at different loci on a diploid genome, are un-linked, (not on the same chromosome), or distantly linked, (far apart on the same chromosome), then due to fragmenting during crossover, they will "assort independently". In this case, there is a 50% chance that either one of the two genes will be represented in a given gamete for a 25% chance that both genes will be represented together in a given gamete. Recall that one gene from each gene pair at a diploid loci is represented in the haploid gamete. If the genes do not assort independently (due to close linkage), then, conditional upon one of the two genes being represented in a given gamete, there is a greater than 50% chance that the other gene will also be represented. Unconditionally, there is a greater than 25% chance that both will be represented together.

Indirect reciprocity Refers to the action of helping others who have not helped you but have helped others in the past. For example, people may help others who have established as reputation as a helper.

Inheritable assortative strategy An inheritable assortative strategy is a strategy for assortative grouping (see assortative grouping) that is inheritable.

Interdemic Group Selection (also known as traditional or "propagule-type" group selection). A process in which traits that are advantageous from the perspective of groups but not individuals are selected. The population structure of interdemic selection is as follows. Individuals interact in groups in a manner that affects individual fitness and then reproduce after which these "parent" groups produce "offspring" groups by group fission (the creation of propagules). Group-level traits (such as the frequency of a types of individuals within the parent group are inherited by the offspring groups).

Intrademic Group Selection A process in which traits that are advantageous from the perspective of groups but not individuals are selected. The population structure of intrademic selection is as follows. Individuals interact in groups such that the interaction affects their fitness, after which they reproduce and groups completely disperse such that individual offspring enter a "migrant pool". New "offspring groups" are formed by sampling from the migrant pool. Unless group formation is "assortative", traits that are advantageous from the perspective of groups but not individuals will not be selected.

Intron The sections of DNA between the exons that do not code for proteins. To make a protein a gene is first transcribed into RNA after which the RNA is translated into a protein. Some portions of the RNA transcript are cut out and are not used to produce the final protein. The DNA corresponding to these sections are called introns.

Kin Selection A process in which genes that code for altruistic actions that are directed towards near relatives are selected. Even though altruism is costly from the perspective of individual genes (and the individuals that house them), if altruism is discriminate in that it is directed towards near relatives, altruistic genes, (and the individuals that house them), can be selected because they are more likely to receive help than free rider genes (and the individuals which house them). In contrast, free rider genes save on the cost of helping but are less likely to be helped by relatives. Hamilton (1975) and Wilson (1975) each argued that kin selection is a special case of intrademic group selection. Roughly speaking altruism can evolve by kin selection if Hamilton's rule is satisfied.

Linkage If two genes are located on one chromosome, they are said to be linked. The closer the genes are on one chromosome, the more likely they will avoid fragmentation during cross-over and the more likely that they will be inherited as a unit, in which case the two genes will not assort independently.

Meiosis occurs in the diploid germ-line cells. Each pair of chromosomes duplicates after which the four chromosomes of each type line up in the centre of the cell and are subjected to crossover (see crossover). After crossover, the four chromosomes in each set are separated and the cell divides twice to create four haploid gametes, each of which contains a full set of chromosomes. The haploid gametes may later merge with the haploid gametes from another organism of the opposite sex (sexual reproduction) to produce a diploid zygote. Due to cross over, the genetic composition of the chromosomes is not conserved during meiosis. However, the basic structure of chromosomes and the genome as a whole, including the ordering and spacing of the loci for specific genes, is preserved.

Meiotic drive During meiosis normally the maternal and paternal gene for a particular locus each have a 50% probability of being represented in a particular gamete. Meiotic drive refers to the phenomenon in which a particular type of gene is able to bias meiosis to have a greater than 50% probability of finding themselves in a particular gamete. These genes may spread rapidly throughout a population even if they reduce the fitness of their carriers.

Meme The term "meme" was coined by Dawkins (1976, chapter 11). Dawkins (1982, p. 297) defines a meme as: "A unit of cultural inheritance, hypothesized as analogous to the particulate gene, and as naturally selected by virtue of its 'phenotypic' consequences on its own survival and reproduction in the cultural environment."

Mitosis The process in which the entire genetic information of a cell is duplicated after which, normally, the cell divides such that each new cell contains a complete copy of the parent cell's genetic information. This process occurs in somatic cells.

Nash Equilibrium refers to a set of strategies in game theory in which given the strategies of the other players, a given player cannot increase his payoff by switching strategies.

Neutral mutation A mutation in a gene which has no effect on the fitness of its carriers. The mutation may or may not change the organism's phenotype. For example, a change in the third nucleotide of a codon need not change the amino acid for which the codon codes. Even if the organism's phenotype changes due to a mutation, the organism's probability of survival and reproduction need not be affected.

Nucleotide Nucleotides are the building blocks of DNA and RNA that are joined in a linear molecule. There are four nucleotides in DNA, cytosine (C), guanine (G), adenine (A) and thymine (T). In RNA, the nucleotides are C, G, A and uracil (U), instead of T. In DNA, C pairs with a G on the other strand and T pairs with an A. The specific pairing allows the double strand of DNA (the double helix) to open up and use a single strand as a template to make a copy of itself. Sequences of three adjacent nucleotides on the exon segments of a gene are referred to as codons, each of which codes for a specific amino acid.

Phenotype The material expression (actions) of a specific genotype. The term may refer broadly to the aggregate "actions" coded for by an organism's genotype (all of its genes) or may be used more particularly, to refer to the "actions" due to particular alleles at a gene locus or set of gene loci, for example the blue-eye phenotype.

Pleiotropy literally means "many movements". In genetics, pleiotropy refers to a situation in which one gene codes for more than one phenotype. In this thesis, I have drawn attention to pleiotropic effects of genes that code for both standard phenotypes such as altruism and also assortative grouping.

Prisoner's dilemma This term refers to a class of games in game theory in which the aggregate group payoff is highest if all members cooperate. However, it is always optimal from the perspective of an individual to cheat independent of the actions of the other players. In the Nash equilibrium, all players cheat and each player receives a lower payoff than could have been attained if all players cooperated—this is the dilemma.

Random Drift In evolutionary biology, random drift refers to changes in the frequency of replicating entities that is due to chance sampling variation and not Darwinian selection. Random drift may have important implications for evolution. For example, in interdemic group selection, after group fission, some groups by chance may have a high proportion of altruists. Such groups may grow faster in turn producing more offspring groups with a high proportion of altruists. Chance sampling variation may allow isolated groups to reach the basin of attraction of different evolutionarily stable equilibrium than the population at large.

Recessive An allele at a chromosomal locus is recessive if it is not expressed in the heterozygous condition (see dominance).

Reciprocal Altruism refers to altruistic actions that are conditioned upon past altruistic actions of the recipient of the action towards the donor. In the case in which an individual is the first to initiate the altruistic action, continued cooperation depends upon whether the recipient of the altruistic action reciprocates by conferring an altruistic action upon the donor. In strong form (strong reciprocity), agents not only refuse to help non-reciprocators but are willing to inflict punishments where the act of punishing is costly. Also see indirect reciprocity.

Replicator The term was coined by Dawkins (1976) and is defined as "Any entity in the universe of which copies are made" (Dawkins; 1982, p. 300).

RNA (ribonucleic acid) A single stranded chain of nucleotides, which is similar to DNA except that the nucleotide Thymine (T) is replaced by Uracil (U) and the sugar molecule components are ribose sugars in comparison to the deoxyribose sugars in DNA. RNA serves as the genetic information in some organisms analogous to DNA. In organisms in which DNA is the genetic information, RNA transcripts play an important role in protein production.

Selection Another word for grouping or isolation. Selection may be either non-random or random. Darwinian selection is an example of non-random selection.

Somatic cells These are the cells of the body of an organism. They are "dead-end" in the sense that their DNA is not transmitted to the next generation.

Vehicle The definition given by Dawkins (1982, pg. 302) reads " Used in this book for any relatively discrete entity, such as an individual organism, which houses replicators (q.v.), and which can be regarded as a machine programmed to preserve and propagate the replicators that ride inside it."

PROLOGUE

This thesis is broadly concerned with the extent to which humans are motivated by concern for the welfare of others verses their own "material" self-interest. Part I considers the biological "units of selection" debate. Upon which units does Darwinian selection act? Do these include genes, organisms and/or groups of organisms? The debate has generated interest within the social sciences because selection at the level of the group may allow for the evolution of altruistic preferences. The literature on the levels of selection debate is reviewed in a form that is intended to be comprehensible to social scientists and a resolution to this long-running debate is proposed. Part II considers some possible implications of altruistic preferences for information about the social benefits and costs of behaviour, behaviour and policy. Although the evolutionary background provided in Part I provides support for the arguments in Part II, each part is a separate work and can be read independently.

The following briefly describes the history of how I came to work on these topics in hopes of providing a background context. The "path of investigation" that led to the development of Part I illustrates that my approach to the topic of group selection partially stems from the mathematical methodology of economics and evolutionary biology. Economists receive a level and type of mathematical training that more readily facilitates understanding the mathematical literature on group selection, than would a typical biologist. Although these mathematical approaches are useful in that they can allow for much clarity of focus, a mere training in economics or mathematical population biology would have made it difficult to understand the units of selection debate within its historical context. My research approach has also benefited due to an understanding of details of the chemical structure of the genetic information. This background was acquired during an undergraduate degree in biochemistry.

Biology is concerned with the competition between life forms for scarce resources while economics is the study of the allocation of scarce resources in human societies. Due to common subject material, there has been a great deal of "cross-fertilization" of ideas between the two disciplines. Evolutionary economists take the view that Darwinian principles have a role to play in modeling and understanding economic systems. My biological background played a role in the decision to focus upon evolutionary economics during the Ph.D. program.

Of particular importance was a paper by Eaton and Morrison, published in 2003, that considers the role of the emotion of revenge in ensuring credible commitments to punish opportunism. They consider an investment game in which player ONE has the choice of investing in either a risky but potentially lucrative project or a safe project. If ONE invests in the risky project, a second player, TWO, can either choose to behave opportunistically or not. In the former case, ONE can choose to retaliate or can allow the opportunism to go unpunished. Although the punishment is sufficient to deter opportunism, it is not credible because the act of punishing is costly to ONE. TWO infers that *ex post* it will not pay ONE to punish, and so, if ONE invests, TWO will behave opportunistically. ONE, similarly, predicts TWO's opportunism and therefore does not invest.

In this "sub-game perfect" (SGP) equilibrium, ONE does not invest so that TWO has no opportunity to "cheat". This equilibrium is not Pareto efficient because if ONE invests and TWO does not "cheat", ONE is strictly better off while TWO is indifferent in comparison to the SGP equilibrium payoffs. This latter efficient outcome might, however, be achieved if TWO believes that ONE will "irrationally" experience the emotion of revenge, in so inducing him to inflict a costly punishment in the advent of opportunism. Using an evolutionary game theoretic approach, Eaton and Morrison illustrate that the strategies for risky investment and irrational retaliation can be selected in equilibrium.

For a class paper, I extended Eaton and Morrison's evolutionary framework to consider games in which players can contribute to a public good at a cost to themselves (C) or free ride (F) by not contributing. Cooperators can either retaliate (C&R) or not (C). If a critical proportion of the population consists of cooperating retaliators, then free riders are selected against so that eventually the whole population comes to play one of the two cooperating strategies.

Although the system can remain in an all cooperating equilibrium for some time, eventually, given that C&R and C give identical payoffs, the proportion of cooperating retaliators may drift below the critical level that is required to select against free riders. The equilibrium cannot, hence, be maintained indefinitely and is referred to as "neutrally stable". Moreover, it is difficult to explain how cooperators could establish themselves from within a population that initially consists entirely of free riders. The second equilibrium, in which all players free ride, is "evolutionary stable"; a small population of C and C&R players that arise due to mutation could never achieve the critical level of C&R before being eliminated. It occurred to me that a solution to these two problems might be to partition the population into groups in a way that served to maintain or initiate the critical level of C&R types. Moreover, if groups are small, the critical level of cooperating retaliators might be achieved by chance. To consider such "grouping" possibilities from within the game theoretic setup, I chose to model agents as possessing two strategy loci, a locus for the strategies C, C&R, or F and a second locus for assortative strategies for group membership.

Meanwhile, I had started work on my thesis, which proposed to use an economic framework to investigate diversity in economic systems. While the topic of diversity has received a large amount of attention within each discipline, there has been a limited cross-disciplinary sharing of ideas. While economists have worked on product diversity, price discrimination, and the concepts of separating and pooling equilibrium, in evolutionary biology, the issue of diversity takes center stage with the problem of "the origin of species". A species, when taken as a whole, would have a greater chance of surviving over time if able to "hedge its bets" over many varieties in the face of environmental changes. However, within a species, individual variants are eliminated during competition with the fittest variant. The "species problem" is a subcategory of the altruism problem because differentiated variants are likely to sacrifice their own fitness for the good of the greater group—here the species. The maintenance of variety requires isolation, i.e. grouping, which can occur by means such as geographical location, and mating and breeding barriers. This interest in species was also stimulated by discussions with my father about his work on barriers to recombination (breeding) due to structural features of the DNA. Similar to the work on the public goods game, I approached the problem of the maintenance of diversity by considering assortative strategies for grouping. This work included a game theoretic model and also a computer simulation.

While this work was in progress, I attended a seminar. After making a comment, I was "accused" of invoking "group selection" in my explanation. I was puzzled by this and also felt somewhat insulted, given that I believed that group selection had been widely discredited. I had thought that I was simply reiterating an explanation given by Richard Dawkins in *The Selfish Gene* (1976), a book which focuses on selection at the level of genes and surely qualifies as one of the most unabashedly "anti-group selection" books ever written. As an economist, from a discipline that champions the emergent phenomenon approach, I had given little attention to the possibility that selection could act upon groups.

Due to my work on "assortative strategies" and also in response to the seminar comment, I began to think more seriously about the possibility of selection acting upon groups. It was relatively easy to alter the species program to examine "group selection" from the context of the simpler "public good game" altruism model. In a new computer simulation, players were assigned two strategy loci; the first could be occupied by either the free rider (F) or altruist gene (C), while the second locus specified a strategy for membership in one out of a larger number of groups. After interacting in groups to play a public good game, individuals receive fitness payoffs according to their own strategy and group composition. For example, an altruist in a group of free riders has a low fitness due to incurring the costs of helping while not being helped in return. A free rider in a group of altruists receives a high fitness payoff due to receiving help without incurring the costs of helping. After interacting, individuals reproduce according to their fitness payoffs. Offspring inherit the strategies of their parents except in the case of mutation.

Mutation can occur whereby C genes mutate to F genes or vice versa and also with respect to group membership.

Dependent upon parameters, the simulation illustrated that altruism could evolve and be maintained at high frequencies at the level of the population. However, within any given group, altruism is short-lived and cycles. Small groups consisting purely of altruists are periodically established by chance, after which they rapidly grow. Eventually, however, such groups are invaded by free riders, due to either within group mutations from C to F, or due to mutations of the group-strategy of free riders in other groups. More complex simulations allowed for an additional strategy, cooperate and retaliate (C&R). C&R types cooperate but punish free-riders at a cost to themselves. C&R and C types have identical expected fitness as long as there are no free riders in the group. If the within-group frequency of C&R types exceeds a critical level, free riders are selected against within the group. After many generations some group attains the critical proportion of C&R types allowing for within-group convergence to the "all-cooperating" neutrally stable equilibrium, which is sustained for some time. This group (and individuals within the group) rapidly out-compete other groups (and individuals) due to high levels of cooperation.

Encouraged by these results, I began a literature review during which I discovered a vast literature on group selection. Although there is a great deal of controversy surrounding the subject, I learned that group selection was not as disreputable as I had previously believed. The formal mathematical modeling of group selection had commenced in the 1960s with computer simulations being developed by the early 1970s. The cycling pattern, observed in my model, has recently been illustrated in other computer simulations such as Aviles (2002). The process of trying to understand the debate led me to propose a solution. This proposed solution, a theory of a hierarchy of replicators based upon inheritable assortative strategies, constitutes Part I of this thesis.

Part II, which is concerned with the implications of altruistic preferences for economic theory and policy, was devised while teaching environmental economics, a process which put me into regular contact with depressing facts about the state of the environment. Many of my students illustrated poor understanding and awareness of the seriousness of the environmental problem. In order to explain this ignorance, I proposed that my students might be altruists, who might experience guilt if they were to acknowledge the environmental and social impacts of their own private materially preferred activity choices. To alleviate such guilt, they might selectively avoid information that serves to link their own behaviour to environmental damages. The aggregate effects of such "selective attention" to information might affect the demand for "news" about the environmental problem, which might in turn affect the supply. In this way social consensus information about the morality of activity choices might come to support privately preferred but socially undesirable behaviour, which in turn reinforces the behaviour in two-way feedback. The altruistic behavioral assumption could potentially generate the same "self interested" behaviour that is to be predicted under the standard selfishness assumptions.

There is a growing literature on media bias with regards to environmental information (Part II, section 5). Most of this work focuses on the "supply-side" factors such as issue complexity in conjunction with media format, and the influence of vested interests. Part II contributes to this literature by considering demand-side bias. A separate application considers the interaction between moral information in the form of religion and the adoption of capitalistic activities during the transition from Feudal to market-based economies in Europe.

Part I

The Levels of Selection Debate: A Theory of a Hierarchy of Replicators based upon Inheritable Assortative Strategies

1 Introduction

The debate surrounding group selection has persisted in spite of a large number of publications on the topic. The subject of much criticism during the 1960s, group selection was to fall out of favour during the 1970s and 1980s. Nevertheless, empirical and theoretical work continued so that by the 1990s, group selection had made a comeback. Today, although there is broader acceptance within biology, there is still controversy. The initial criticisms leveled against group selection had not been based upon theoretical impossibility but upon its importance as a biological phenomenon and whether certain processes, such as kin selection, qualify. Today, there are still disagreements about the scope of the phenomena. Although the debate is to some extent semantic "what shall we mean by 'group selection' "? (Maynard Smith, 1987a), fundamental issues remain unresolved.

The group selection debate is relevant for the social sciences due to its potential to shed light upon human psychological characteristics. Until the onset of the agricultural revolution, some 12,000 years ago, *Homo sapiens* and our immediate ancestor species lived in small groups. Proponents of group selection argue that competition both within and between groups shaped human psychological traits, the latter facilitating the selection of prosocial behaviour towards group members and the potential for asocial behaviour towards outsiders. Darwin himself argued that competition between tribes might lead to prosocial preferences such as, "the spirit of patriotism, fidelity, obedience, courage, and sympathy". However, for the most part he focused upon competition at the level of individuals with adaptations evolving for the good of individuals (Darwin, 1871, p.166). The dominant latter focus has provided a biological basis for the methodological individualism of the social sciences, to which group selection offers a challenge.

In 1994, in order to alert social scientists to the reconsideration of group selection within the field of biology, Wilson and Sober titled a long review paper "Reintroducing Group Selection to the Human Behavioral Sciences". This was followed, in 1998, by a full-length popular book *Unto Others: The Evolution of and Psychology of Unselfish Behaviour*'. On the other side of the debate, Matt Ridley's (1996) book *The Origins of*

Virtue presents the argument that humans are distinctive in their remarkable degree of sociality. However, these prosocial psychological attributes can be explained due to selection acting upon genes, not groups. The differences of opinion expressed in these popular science books illustrate that there is still controversy within biology. Part I of this thesis reviews the controversy from within the historical-biological perspective and proposes a solution.

The group selection debate is a component of the "units of selection" debate, which revolves around the question; "For which units do adaptations evolve for the good of?" This question can be interpreted to mean; "What biological units will evolve through a "Darwinian" selection process?" For example, do these units include genes, organisms, and groups of organisms? All of these units may appear well adapted to their environments, but is this "unit-level functionality" a direct product of selection acting at that level or the product of selection acting at another level? To give the example of Williams (1966), a group of fast deer that out-reproduces a group of slow deer is well adapted. However, surely, it is fast individuals that are *directly* selected by the Darwinian process with fast groups being a mere byproduct; it would then appear that the group characteristic of "average group speed" did not evolve for the good of fast groups of deer.

The "units of selection debate" is about which units will evolve by the process of "Darwinian" or "natural" selection. There are three components as identified by Lewontin (1970); 1) there must be variation among units; 2) some variant units must be better at surviving and reproducing than others; and 3) there must be heredity such that units produce offspring or copies that resemble themselves.[1,2] Given these three conditions, the theory of natural selection predicts that given enough time, these units will become well adapted to their environments. In Darwin's time, the basis for heredity could only be speculated at, and adaptations were for the most part viewed as having

---

[1] Note that throughout this essay the term "Darwinian" will be used in Lewontin's sense and not merely to selection upon individual organisms in the sense of "traditional Darwinism".

[2] Selection eliminates variation. New sources of variation are ultimately generated by mutation.

evolved for the good of "organism units". With respect to criterion 3, offspring resembled parents and hence there was some basis for heredity at the level of organisms.

Yet, there are plenty of puzzling examples of biological traits that appear to be disadvantageous from the perspective of organisms but advantageous from the perspective of groups or even species. Examples include sexual reproduction, standard altruistic acts in which individual behaviour benefits the group at a cost to the individual, senescence, and speciation.[3]

To explain such at least "apparent" altruism two polar solutions were proposed.[4] The first was to argue that Darwinian selection could act upon groups leading to group level

---

[3] A brief explanation as to why sexual reproduction, senescence, and speciation can be regarded as examples of altruism follows. Mutation aside, in asexual reproduction, 100% of a given parent's genes are inherited by a given offspring, while in sexual reproduction, only 50% of a given parent's genes are expected to be inherited by a given offspring. The sexual species, however, produces more variants due to meiosis (see section 2.2.2). Variation increases the chance that some organisms will survive if the environment changes, or due to possible occupation of a variety of environmental niches, so increasing the chance of survival of the sexual species as a whole. Sex then appears to be good for the species but against the short-term interests of particular genes or individuals within the species. Similarly, in the case of senescence, if an individual lives longer, it can potentially produce more offspring. However, shorter life cycles allow for the creation of more variants potentially increasing the chance of survival of the species. Speciation, loosely refers to the process during which a sub-set within a species become reproductively incompatible with the rest of the species but is able to breed within the sub-set. Ignoring other mating isolation barriers, if members of a species randomly breed, then the statistical distribution of variants within the species should consist of some aggregate mean species phenotype about which there is some variation. Due to selection, the mean aggregate species phenotype may change over time but there will be no branching. Speciation allows for non-random breeding within isolated sub-groups, which allows for branching evolution, increasing the chance that some members of the two species (and hence some descendents of the initial ancestor species) will survive at later points in time. From the perspective of an individual organism, a reduction in reproductive compatibility should be fitness reducing due to a lower probability of successfully breeding.

[4] See Maynard Smith (1987a) for a discussion.

adaptations.[5] The second was to argue that not only did groups not qualify as units of selection, but neither did organisms. In this view, both organisms and groups are emergent phenomena due to selection acting at a lower level, that of the gene (Williams, 1966; Dawkins, 1976). The latter argument was supported by 20th century research, which had led to an understanding of the material basis for heredity, the DNA. While Darwin had no alternative other than to rely upon resemblance as evidence for a basis for heredity, it was now possible to be precise about criterion 3.

There are two possible domains for a hierarchy of units of selection. Firstly, there is the domain of hereditary information, including nucleotides (the basic chemical building blocks of the DNA), genes, genomes (the hereditary information of the organism as a whole), and the hereditary information of groups of organisms when taken as a whole. Secondly is the domain of "actions", the phenotypic expression of the hereditary information. For example, sequences of three nucleotides code for amino acids (the building blocks of proteins), genes code for proteins (the building blocks of organisms), and genomes code for phenotypic expression as embodied in the physical and psychological characteristics of the organism as a whole.[6]

Most evolutionary biologists accept that the Darwinian units of selection are in the domain of the hereditary information, not the domain of actions. To use Dawkins' (1982, p. 97) example, consider the genes that code for legs in stick insects. These genes instruct the insect to develop legs, which in turn help the insect to reproduce. Given that the genes for legs are inherited by offspring, the leg genes are instrumental in causing their own replication. Similarly, the phenotypic traits of legs (in the action domain) also play a causal role in ensuring their own replication; the leg phenotype helped the stick insect to reproduce and along with it the leg genes. These genes, in turn, provided leg development instructions to the offspring. It then appears that both domains are

---

[5] Note that group selection would be predicted to lead to altruism directed towards in-group members but to selfishness with respect to out-group members.

[6] Williams (1992, p. 10) refers to the domain of hereditary information as the "codical domain", while the domain of actions is referred to as the "material domain".

candidates for Darwinian units of selection. However, there is an important difference. If a stick insect looses its leg and then reproduces, this acquired characteristic is not passed on to offspring. To contrast, if a leg gene mutates, descendent copies of that gene will also contain the mutation. In this sense, it could perhaps be argued that the hereditary domain has a more fundamental effect upon causality.[7] It is useful to view the phenotypic actions as tools by which the hereditary information makes copies of itself. At the very least, if units in the domain of actions are to be accorded the status of Darwinian units of selection, there must be inheritable information underlying these units. It is simpler to focus upon units in the hereditary domain.

If it is accepted that the hereditary information is the relevant domain for Darwinian units of selection, the stage is set to argue that genes qualify as such units, but that genomes in sexually reproducing species do not. Proponents for the genic view have argued that since the offspring in sexually reproducing species do not inherit a whole copy of each parent's genome due to a process called meiosis, to be explained in section 2.2.2, then genomes lack the hereditary basis to serve as units of Darwinian selection, i.e. adaptations cannot evolve for the good of organisms. To contrast, smaller sections of the genome, which Williams (1966) referred to as "genes", do survive meiosis intact and hence there is hereditary at the level of the "gene".[8] According to this argument, put forth by Williams (1966) and popularized by Dawkins (1976), "genes" alone satisfy the criteria of a unit for which adaptations can evolve for the good of. Note that Lewontin's 3rd criterion (hereditary) has been strengthened from that of "resemblance" to "perfect copying" (to be discussed in section 2.2.1). If genomes don't qualify as units of selection, then making the case for groups of genomes is even more problematic.[9]

---

[7] That the order of causality goes from genes to phenotypic traits is referred to as the "central dogma". According to Dawkins (1982), to argue that other way is to invoke the Lamarckian inheritance of acquired characteristics.

[8] This meaning of the word "gene" was coined by Williams (1966) and is used by Dawkins (1976, 1999). In standard usage, a gene is the fragment of DNA that codes for one amino acid chain in a protein.

[9] These views of Williams and Dawkins are discussed at length in section 2.2.2.

Dawkins coined the term "replicator" to refer to anything in the universe of which copies are made. The concept of a replicator is a broad concept. For example ideas can be copied. In the biological realm replicators correspond to Williams' "genes". Phenotypic traits are regarded as tools by which genes copy themselves. Organisms and potentially higher level groups in the domain of actions are referred to as "vehicles", the means by which replicators interact with their environment in order to get copied.

Those on both sides of the units of selection debate have agreed that the largest organic replicators are Williams' "genes". Some biologists, such as Maynard Smith (1987a), have maintained that adaptations can evolve for the good of groups, based upon the weaker Darwinian criteria of resemblance. Wilson and Sober (1994, 1998) accept that the largest replicators are "genes" but argue that both organisms and sometimes groups of organisms satisfy the criteria of units of selection. Given that replicators cannot be identified at these higher levels, they argue that replicators are irrelevant to the question of whether adaptations can evolve for the good of a unit and that such units are in the domain of "actions" or vehicles.[10]

Throughout this essay it will be argued that Darwinian selection of a unit and adaptations at the level of that unit require heredity at that level. Furthermore, it is possible to do better than mere resemblance by utilizing the replicator concept at a variety of levels. Replicators are identified at the levels of nucleotides, genes, organisms, and groups of organisms. Each of these replicators can serve as a Darwinian unit of selection.

In addition to coding for standard phenotypic traits such as eye-color or altruism, lower level replicators (such as genes) also code for inheritable "assortative strategies", which allow lower-level replicators to aggregate in a non-random manner to construct higher level replicators.[11] These inheritable assortative strategies are not as obvious as strategies for standard phenotypic traits. Given that higher-level replicators are "constructed" by

---

[10] The arguments of Wilson and Sober are discussed more carefully in section 2.3.5. These include philosophical arguments regarding the nature of causality.

[11] When a gene codes for more than one trait, this is referred to as a pleiotropic effect.

lower- level replicators, they qualify as emergent phenomena, yet they also contain the information to make perfect copies of themselves and hence qualify as replicators in their own right.

Replicators are identified at the level of individual organisms and groups of organisms (for both the cases of interdemic and intrademic group selection).[12] A precise relationship between replicators and vehicles is identified, with a unique expected vehicle corresponding to each replicator. In section 4, it is argued that all forms of Darwinian selection depend upon inheritable assortative strategies, that is, strategies for non-random grouping. Whether selection qualifies as "Darwinian" depends upon whether these strategies are transmitted through hereditary information to descendants or whether they are "played" by chance. Only in the former case can adaptations evolve for the good of the selected units. Darwinian selection, in essence, corresponds to a process of non-random grouping, whereby the strategies for such non-random grouping are inheritable.

Before setting out to develop the argument, the major developments of the group selection debate are reviewed within the historical-biological context with a view to providing the background necessary to understanding the argument. Although most sections are descriptive and depend only upon basic mathematics, given that sections 2.3.2 and 4 are more mathematically involved, non-mathematical descriptions are provided in hopes that the material will be comprehensible without needing to go carefully through the mathematics.

2 Historical Overview of Group Selection

2.1.1 The Fable of the Bees Revisited

In 1732, Bernard Mandeville wrote a poem called *The Fable of the Bees* in which he suggested that the harmony and productivity of the beehive is not due to cooperation but to the pursuit self-interest by individual bees. Mandeville was not writing about

---

[12] Interdemic and intrademic selection, the two categories of group selection, are explained in section 2.1.1.

biological systems but instead used the beehive as an allegory for human society in order to illustrate how self-interested individuals interacting in markets might bring about a reasonably harmonious social order. The more obvious explanation for a harmonious social order, the fable, is that it is the result of planning and/or cooperation. Adam Smith (1778) further developed Mandeville's idea coining the term "invisible hand" to denote the illusion of guidance of the market by benevolent social planners.

Along these lines, prior to Darwin (1859), the harmony observed in nature was most easily explained as the workings of a benevolent divine planner, who, as pointed out by Gould (1990), was the biological equivalent to Smith's planner. Gould noted that Paley in his book *Natural Theology* maintained that from the overall harmony of nature, one could infer the existence of God—nature's planner or designer.[13] In contrast, Darwin, like Smith, argued that any functional harmony above the level of the individual (for example cooperation within groups of organisms) was an emergent phenomenon, the result of selection acting at the lower level of individuals.[14]

> The ...argument accepts the premise that, yes, most organisms are well designed, yes, ecosystems tend to be harmonious but-and here is the almost perverse turning of Paley on his head—that is true, it is only a side consequence. There is no law or principle that operates directly for it. It is an accidental side result, a consequence of something else that is working at a totally different level, namely, the only thing happening out there is that individual organisms are struggling for personal reproductive success. That is all. It sounds ultimately selfish and yet the good design of organisms and the harmony of ecosystems in Darwin's theory is nothing but a side consequence of something that has apparently an opposite moral force. (Gould, 1990)

As argued by Lewontin (1970) Darwin's original theory of selection required three conditions; 1) phenotypic variation among individuals, 2) different rates of survival and reproduction based upon 1), and 3) heredity as indicated by a correlation between the phenotypic traits of parents and offspring. In Darwin's time, the hereditary basis of the

---

[13] Paley's book *Natural Theology* was the standard "biology" book prior to Darwin.

[14] Gould (1990) wrote "...probably the strongest component of input into the theory of natural selection that Darwin developed in 1858 was his reading of Adam Smith's life by Dugald Stuart".

phenotype could only be speculated at and so Darwin based most of his arguments upon the selection of individuals. As offspring organisms resembled parents, individual organisms satisfied the three criterion and could be viewed as units of selection. The theory predicted that individual organisms should come to be well adapted due to natural selection. Adaptations were said to "evolve for the good of individuals". For the most part observations of higher level functionality were explained as emergent phenomena due to selection acting upon individuals, not upon units at higher levels.

Ironically, given Mandeville's allegory, the beehive (social insects) provided a puzzle to Darwin who had trouble explaining the harmony and coordination of the hive in terms of self-interested bees. He proposed that in special cases such as social insects and also human tribes, there might be differential competition between and selection of groups. Hamilton (1975) wrote:

> Returning to the problem of units of selection, Darwin himself vague about the process of heredity, based most of his arguments on considerations of the fitness of individuals. He made occasional exceptions, as for the social insects where he treated the "family group" as the unit of selection. I believe even these limited concessions were incautious (Hamilton 1972), and value his judgement more where, discussing the evolution of courage and self-sacrifice in man, he left a difficulty apparently unsolved. He saw that such traits would naturally be counter-selected *within* a social group whereas in competition *between* groups the groups with the most of such qualities would be the ones best fitted to survive and increase.

Gould (2002) also noted that for the most part Darwin stuck to his original insistence that organisms were the fundamental units of selection. Group selection, selection at the level of the group based upon the three criteria listed above, appears to run counter to the methodological individualism of "conventional" Darwinism and decentralized approaches used in biology and the social sciences, particularly economics. Yet, starting with Darwin (1871) and until the 1960s, group selection was increasingly invoked to explain biological phenomena not easily explainable as due to selection acting at the level of individuals (Maynard Smith, 1987a; Williams, 1966; Wilson and Sober, 1994, 1998).

A variety of phenomena including speciation, altruism, sex, senescence and population regulation were increasingly explained as group level adaptations.[15]

It has been well noted that such early assertions of group selection were indeed "incautious" (for example see Maynard Smith, 1987a; Williams, 1966; Wilson and Sober, 1994, 1998). Prior to the 1960s, there were limited theoretical efforts to illustrate whether such higher level selection was possible and indeed these efforts suggested that it was unlikely. Some of these efforts are briefly outlined below.

Fisher (1930, 1956, p. 49-50) discussed the possibility of selection acting upon groups, where groups are entire species with births and deaths respectively corresponding to speciation and extinction events. If some trait is adaptive from the perspective of individuals within the species, but maladaptive from the perspective of the species, the individually advantageous trait should win out because the life-cycle time of an individual is short in comparison to the life-cycle time of an individual species. Fisher does however leave open the possibility that sexuality is a species level adaptation.

Haldane (1932, 1966, p. 207-210) noted that groups with high proportions of altruists should grow faster than those with relatively low proportions. However, within groups, free riding individuals should have a fitness advantage ultimately leading them to takeover within groups. As a solution to this problem, Haldane suggested that the lifecycle duration of the group must not greatly exceed the lifecycle duration of individuals within the group. Shortening the lifecycle of groups could potentially be accomplished by group fission prior to within-group takeover by free riders. As groups of altruists grow faster than groups of free riders, altruistic groups should produce more offspring groups than are produced by free rider groups, potentially allowing altruism to increase in frequency at the level of the population. The process is more likely to get

---

[15] See footnote 3 (page 4) for a brief discussion of these arguments. The argument that population density regulation is a group level adaptation is discussed on page 13.

started if groups are small so that one group happens to have a high proportion of altruists by chance.

Similarly, Wright (1945) argued that parent groups might send out migrant colonies. If these propagules are small, then some "offspring" groups might by chance have a high proportion of altruists, allowing them to grow relatively fast. These in turn would produce a relatively large number of offspring propagules, each with a high proportion of altruists. Groups need to be sufficiently isolated such that free riders from other groups cannot invade. Both Haldane and Wright concluded that although group selection was theoretically possible, it was unlikely to be a frequent occurrence in nature as the parameter requirements needed to get group selection to work were stringent (Wilson and Sober, 1998).

These models later came to be referred to as traditional, propagule or interdemic models of group selection, where "demes" refer to the groups of individuals. The groups in interdemic models satisfy Lewontin's three conditions for Darwinian units of selection. Firstly, there is variation with respect to a group trait, the proportion of altruists. Secondly, if groups with more altruists grow faster and there is fission prior to within-group selection of free riders, then the group trait, the frequency of altruists in a parent group, is correlated with reproductive success in terms of offspring groups. Thirdly, this group trait, the frequency of altruists equals the expected frequency of altruists in offspring groups and so is inheritable.

Williams and Williams (1957) produced a mathematical model with similarities to those of Wright and Haldane (Wilson and Sober, 1998). After random mating within the population at large, each female produces an isolated colony of offspring, which interact with each other, either cooperating or free riding. After the interaction and before mating, offspring disperse into the population at large, and the process is repeated. Offspring of altruistic mothers are more likely to be altruists. Hence, although a given altruist incurs costs helping siblings, it is more likely to receive help from siblings. To contrast, free riders do not incur the costs of helping but are unlikely to receive benefits.

Altruists are hence potentially fitter than free riders. Under particular parameter specifications, altruism may increase in frequency at the level of the population.[16]

This model later came to be referred to as kin selection, a special case of what came to be known as intrademic group selection (Wilson, 1975). It has since been argued that intrademic models do not qualify as "group selection" due to the absence of group-level heredity, Lewontin's third criterion for selection of the Darwinian sort (Maynard Smith, 1976, 1987a).

In addition to the mathematical models described above, it was common for field biologists to explain phenomena as group level adaptations. Most famously Wynne-Edwards (1962) argued that groups of animals regulate their population density in order to prevent open access problems and furthermore that such regulation can be regarded as a group level adaptation, the product of group selection. To understand his argument, consider a population composed of groups of rabbits, where high and low fecundity rabbits can be thought of as free riders and altruists respectively. High fecundity rabbits over-exploit the resource base reducing the fitness of the group. Hence, *between* group selection should favour low fecundity rabbits. However, *within* groups, high fecundity rabbits are relatively fit potentially allowing them to take over the group. If selection between groups (for example due to extinction) outweighs selection within groups, then low fecundity might evolve by group selection.

Economists are familiar with the non-cooperative open access problem whereby selfish individuals take into account private net benefits while ignoring the external costs of their actions (Gordon, 1954; Hardin, 1968). The resultant overexploitation leads to an

---

[16] Most of the biological papers to be examined do not focus upon the proportion of types that will exist at equilibrium, but rather upon, whether, due to Darwinian selective pressure, altruists will increase in frequency given particular parameters and population structure. For example, in Hamilton's model (section 2.3.2), as long as there are initially enough altruists to form a group with a proportion (*f*) of altruists, and inequality [3-19] is satisfied, then altruists will be fitter on average than free riders such that a population consisting entirely of altruists is an evolutionarily stable equilibrium.

inefficient allocation from the perspective of the group. In the human case, some form of regulated access is required to prevent over exploitation and to maximize catch over time. Wynne-Edwards suggested that selection acting upon groups of animals could lead to population regulation, a group level adaptation, helping to solve open access problems.

At first sight, fecundity regulation in the interests of the group of rabbits appears analogous to invoking a planner to explain the organization of free-market economies. This flies in the face of the idea that higher level order is an emergent phenomenon due to selection of individual strategies which evolve to be best responses given the pool of other strategies in the environment—the game theoretic version of *The Fable of the Bees*. Not surprisingly, throughout the 1960s and 1970s, the ideas of group selection met with increased resistance from a group of mathematical biologists. Of particular importance was the work of Williams (1966).

2.2 The Arguments Against Group Selection

Williams (1966) argued that traditional group selection was theoretically possible but unlikely to be an important process in nature. Furthermore, most of the purported examples of group selection could be explained due to selection acting at lower levels. To establish his point, he provided well-reasoned examples. Moreover, Williams threw the traditional notion of organism level adaptations into question by arguing that the unit of selection was not the organism but the gene—organisms like groups were arguably emergent phenomena. This view was to move into the mainstream after the publication of Dawkin's popular book, *The Selfish Gene*, in 1976. Yet the seeds for the "genic" approach were planted far earlier. The next section outlines the historical development of the genic approach and also the arguments used to discredit group selection. Importantly, this stream of thought provides a precise meaning for the expression "a unit for which an adaptation evolves for the good of".

2.2.1 A Particulate Basis for Heredity

Although a particulate basis for heredity was hypothesized in the 19th century, and Mendel had collected evidence of discrete heredity elements that were transmitted from parents to offspring, Darwin and his contemporaries were unaware of Mendel's findings. For Darwin, resemblance between parents and offspring offered evidence of a basis of heredity that should allow individual organisms to evolve to become well adapted. Mendel's work was rediscovered in 1900 and geneticists set to work conducting breeding experiments in order to work out the rules of transmission of the genetic hereditary elements. Biologists hoped that analogous to atoms in chemistry, these hereditary elements could be used to understand biological phenomena from the ground up.

Fisher (1930, 1958), a pioneer in mathematical genetics was to integrate this new evidence of the hereditary genetic elements into Darwin's framework. Given evidence that the genetic elements were transmitted in discrete form, it was possible to be more precise about Lewontin's third criterion for a "Darwinian" selection process. Biologists no longer needed to depend upon resemblance as an indicator of heredity because they had evidence that, mutation aside, genes were perfectly copied from parents to offspring. The following hypothetical example employs a one-gene-one-trait focus in order to illustrate the necessary conditions for the "Darwinian" selection of units over one generation.

Consider a population that is initially composed of 50% of each of two types of units, A-types and B-types. A-types are twice as likely to survive to reproduce and hence produce twice as many offspring on average as B-types. For purposes of illustration, suppose that a mutation occurs with a biologically unrealistically high probability such that there is a 50% chance that an A-type will mutate to a B-type and vice versa. Although, A-types on average produce twice as many offspring as B-types, given that half of each type mutate to become the other type, the effects of mutation exactly offset selection. The process is illustrated in figure 1.

Figure 1: Selection and Copying Fidelity

Initial population:

A A A A A A B B B B B B

↓

Selection of more A's than B's due to differential survival and reproduction:

A A A A A A A A B B B B

↓

Mutation: (1/2 A's become B's and ½ B's become A's)

A A A A B B B B A A B B

The first two of Lewontin's criteria for Darwinian selection lead to selection of A-types; there is differential survival and reproduction based upon the phenotypic expression of the two variant genotypes. However, due to mutation, the third criterion, heredity, is too weak to allow the more adaptive trait to increase in frequency in the next generation. Due to low copying fidelity, A-types fail to increase in frequency in the population and the A-trait fails to evolve as an "adaptation for the good of the underlying unit".

The work on Mendellian factors and its integration with Darwin's work by Fisher and others, set biology on a historical trajectory that was to move its focus away from competition between organisms towards that of competition between genes for specific loci on genomes. The gene itself came to be regarded as a Darwinian unit of selection; genes that code for adaptive phenotypes and have high copying fidelity should tend to increase in frequency in a pool of genes competing for a particular locus on a genome.

2.2.2 Williams' "Gene" as the Fundamental Unit of Selection

Similar to Fisher (section 2.1.1), Williams (1966) argued that although group selection was a theoretical possibility, it was unlikely to be an important process in practice because the lifecycles of individuals are shorter than those of groups. In cases of conflict between individual and group interests, individual interests should overcome group interests unless particularly stringent conditions are satisfied.

> Selection has nothing to do with what is necessary or unnecessary, or what is adequate or inadequate, for continued survival. It deals only with an immediate better-vs.-worse within a system of alternatives, and therefore competing, entities. It will act to maximize the mean reproductive performance regardless of the effect on long-term population survival. It is not a mechanism that can anticipate possible extinction and take steps to avoid it (Williams, 1966, p. 31).

Williams devoted a large part of his book to providing lower level (individual or genic) explanations for adaptations that appear to have evolved for the good of the group. Although he argued that interdemic (traditional) group selection is theoretically possible, its importance as an evolutionary force was left open for further empirical work. He did, however, suggest that evidence of uneven sex ratios would provide support for group selection; substantive evidence of sex role bias was not available circa 1966.[17]

In addition to providing individualistic explanations for purported group level adaptations, Williams goes much further by arguing that *the* unit of selection for which adaptations evolve for the good of is not the individual organism, but the "gene". Building upon the work of Fisher, the stage was set to argue that the hereditary information of organisms as a whole does not qualify as a unit for which adaptations can evolve for the good of. In response to the group selectionist claims of Wynne-Edwards and others, Williams takes the argument the other way by arguing that not only do groups not qualify as units of selection, but neither do organisms. In this view, the genome, the

---

[17] Uneven sex ratios can be theoretically explained by group selection, while even sex ratios are to be expected if selection acts at the level of individuals. Ignoring the issue of open access problems (Section 2.1), consider a population consisting of groups of rabbits and conditions of interdemic group selection. A given group produces more offspring if the sex ratio is biased towards females; females can only produce a limited number of offspring while one male can potentially sire many offspring. Hence, competition *between* groups favours a sex ratio with a high proportion of females. If competition is between individuals and not groups, even numbers of each sex should be an evolutionarily stable strategy (Fisher, 1956). To see this, consider a population with more than 50% percent females. If males can mate more than once, individual males should produce more offspring than individual females. The situation is reversed if the population has more than 50% females. Hence, competition *within* groups favours an equal proportion of each sex. Later, Hamilton and others found conclusive examples of sex ratio bias in support of group selection hypothesis (Wilson and Sober, 1998, p. 38-43).

aggregate hereditary information of the organism as a whole, is an emergent phenomenon, the result of selection acting at the lower level of the genes.[18] The gene becomes the fundamental unit of selfishness.

In order to understand Williams' argument and also the non-standard sense in which he defines the word "gene" (also used by Dawkins, 1976), I digress to give a short lesson in genetics.

There are two types of cells in multi-cellular sexually reproducing organisms, somatic cells that constitute the body or soma of the organism and germ-line cells, which produce the gametes that potentially merge with a gamete from a member of the opposite sex in order to produce a zygote. The zygote can then grow into a new organism. Somatic cells reproduce during an organism's lifetime, in turn passing on their genetic information to daughter cells during a process called "mitosis". Copies of the hereditary information in somatic cells are not transmitted to the next generation of organisms. The transmission of hereditary information between generations of organisms is accomplished purely through the germ-line cells.

The genome is the hereditary information for the organism in its entirety. This hereditary information is contained in the linear sequence of nucleotides that constitute the building blocks of the DNA. Each somatic cell in organisms such as humans contains two complete sets of such hereditary information, one from each respective parent, and such cells are referred to as "diploid". The DNA is divided up into long linear fragments referred to as "chromosomes". In each diploid cell, there are two of each type of chromosome, one from the paternal parent and one from the maternal parent. There are

---

[18] Note that in the case of asexual organisms, the genetic information of the organism in its entirety is defined as a "gene" (in Williams' sense) and therefore, selection acts both upon genomes and genes.

23 types of chromosomes in each human diploid cell giving 44 chromosomes in total.[19] Shorter fragments of DNA called "genes" code for the production of particular proteins.

Diploid germ-line cells in the gonads (sex organs) undergo a process called "meiosis", which is illustrated in Figure 2. One pair of chromosomes is illustrated for simplicity. Firstly, each chromosome is duplicated (2) after which the four chromosomes line up in the centre of the cell (3). During a process called cross-over, the chromosomes are cut and sections are swapped such that the initial order of specific types of genes is maintained (4). The chromosomes are separated and there are two cell divisions such that the four new cells, the gametes (sperm and eggs) are produced (5 and 6). Each cell contains one copy of each type of chromosome and is referred to as "haploid" (for example 23 chromosomes in the case of humans). Note that due to cross-over none of the chromosomes in the gametes is a perfect copy of any of the original chromosomes. Supposing that these gametes are sperm, during sex, they may fuse with a gamete from another organism of the opposite sex to produce a diploid zygote, which contains 46 chromosomes in the human case (7). This zygote may grow into a new organism, which will possess its own germ-line cells.

Due to cross-over during meiosis, genomes are imperfectly copied so that offspring do not inherit perfect copies of parent genomes. Hence, it does not then appear that the hereditary information of sexual organisms (when taken as a whole) possess sufficiently high copying fidelity to qualify as a "Fisherian" unit of selection (section 2.2.1).

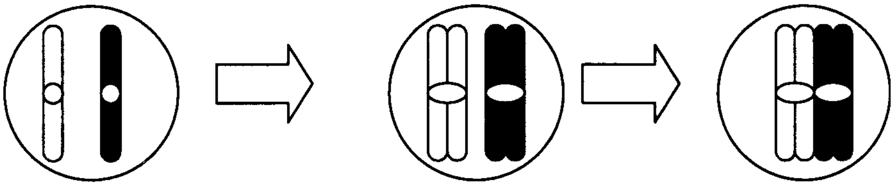
---

[19] Note that 2 out of the 46 chromosomes are sex chromosomes, which do not have identical structures in the case of males. There are two X chromosomes in female diploid cells and one X and one Y chromosome in male diploid cells.

Figure 2: Meiosis and Sex

Note that there are many pairs of chromosomes in each diploid cell. Only one pair is represented for simplicity.

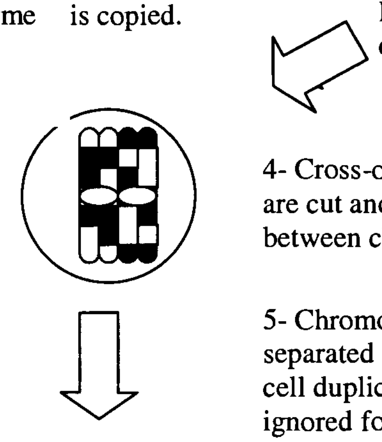
Meiosis:



1- Diploid Germ-line cell: Grey represents a maternal chromosome and white represents a paternal chromosome.

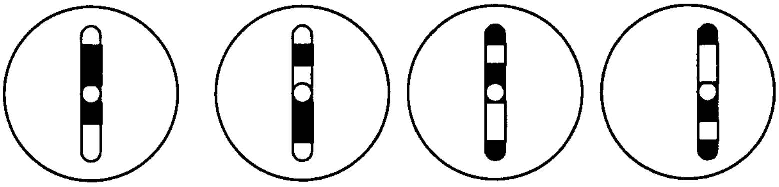
2- Each chromosome is copied.

3- Chromosomes line up in centre of cell.



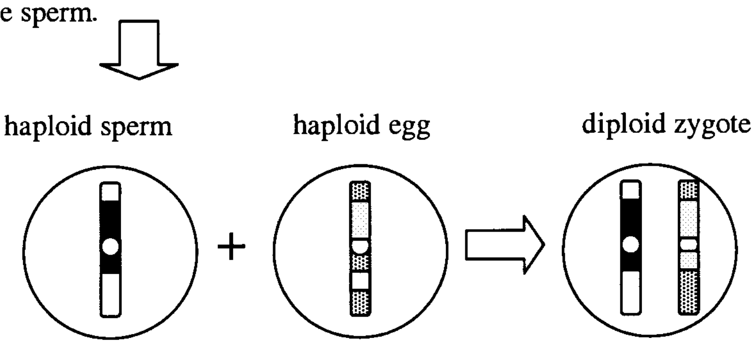
4- Cross-over: Chromosomes are cut and DNA is swapped between chromosomes.

5- Chromosomes are separated and there are two cell duplications. Details are ignored for simplicity.



6- These haploid gametes, either sperm or eggs, contain one complete set of genetic information of which approximately 50% is maternal and 50% is paternal DNA. "Maternal" and "paternal" are defined with respect to the original organism but correspond to "grand-maternal" and "grand-paternal" with respect to the gamete. Assume that the original organism is a male so that these gametes are sperm.

Sex:



7- A sperm created during meiosis (shown above) fuses with an egg from another organism of the opposite sex to produce a zygote, which grows into an organism.

Williams (1966) argued that due to this splintering by meiosis, sexual genomes lack sufficient copying fidelity to qualify as units of selection for which adaptations can evolve for the good of. However, smaller sections of the genome are more likely to survive the process of meiosis intact and hence have a higher likelihood of being perfectly copied to the next generation of organisms. If these smaller sections code for traits that increase the survival and reproductive success of their vehicle organisms, they can increase in frequency in the pool of genome fragments competing for a particular position on a species' genome. In this view, the units of selection "for which adaptations evolve for the good of" must be small enough to have a high probability of surviving meiosis intact, so that the traits that they code for can be transmitted to descendants. Williams refers to such DNA fragments as "genes" and argues that adaptations cannot evolve for the good of larger units. Note that Williams' definition of a "gene" is not the standard one in which a "gene" refers to the fragment of DNA that codes for a specific protein.

If meiosis is "fair", meaning that a given gene in a germ-line cell has a 50% probability of being copied to a given gamete, then a gene that increases the reproductive success of its organism vehicle should increase its own frequency in future generations. Hence, the interests of the organism and the gene tend to be synchronized. Genes that on average confer higher fitness to their host organism, meaning genes that work well given the average genetic environment and non-genetic environment of that organism type, will tend to increase in frequency relative to other genes coding for that locus.

Although individual sexually reproducing organisms can be differentially selected with respect to the number of offspring they produce, their hereditary information as a whole is not. When we talk about differential selection of blue eyed individuals, we mean that such individuals (holding all else constant on the genome) are selected in the sense that they have more offspring. In the view of Williams, the genome is an emergent phenomenon, due to selection acting at the lower level of the 'genes'.[20]

---

[20] The genome is asexual organisms is haploid and is not subject to meiosis, although there is some limited sex though DNA transfer between organisms. These genomes qualify as "genes" in Williams' sense.

Accordingly, if genomes do not qualify as units of hereditary information for which adaptations can evolve for the good of, then neither does the hereditary information for groups of organisms. Most biologists on both sides of the group selection debate have accepted that the largest units of selection in the hereditary domain are Williams' genes.[21] In spite of this many biologists allow that adaptations can evolve for the good of organisms and sometimes higher level groups. There are two basic arguments. Firstly, Maynard Smith (1987a) argues that that units of selection only need to satisfy the weaker Darwinian basis for heredity, that of resemblance between offspring and parents and not the stronger Fisher/ Williams' form in which the units of heredity are perfect replicates. The second approach of Wilson and Sober (1994, 1998) and Gould (2002) is to argue that the units of selection for which adaptations evolve for the good of is in the domain of actions or vehicles and not in the domain of hereditary information.

2.2.3 Hamilton and Inclusive Fitness/ Kin Selection

If genes are the units of selection, then how can we explain the selection of genes that instruct their vehicle organisms to undertake altruistic actions that are expected to reduce the reproductive fitness of the organism and also the particular altruistic gene within the organism?

Prior to the publication of Williams' book, Hamilton (1962, 1963) had provided a gene-based explanation for altruistic behaviour among kin. Hamilton sought to explain the common observation of altruism among kin, whereby organisms take actions that decrease their chances of survival, while increasing the chances of survival of near relatives. His theory of inclusive fitness illustrates that if the donor and recipient of the altruistic act are sufficiently related, then, genes for altruism can increase in frequency in the gene pool.

Consider a gene that induces the individual upon which its propagation depends to donate b fitness units to a sibling at a cost of c units of its own fitness. The probability that a

---

[21] For example, see commentary on Wilson and Sober (1994).

sibling possesses the same gene is ½. Both the donor organism and the particular gene that induces the donation serve to loose c units of fitness. With a probability of approximately 50%, the recipient sibling has a copy of the identical gene due to descent from common parents.[22] Hence with a probability of approximately ½, the altruistic sibling's donation is conferred upon a sibling, which also possesses the altruistic gene. Therefore the expected net fitness benefits of the altruistic act to the collective of altruistic genes is $b/2 - c$. Altruists should increase in frequency relative to free-riders if $b/2 > c$. This is Hamilton's rule for siblings.

The inclusive fitness method takes into account all of the effects of one gene upon identical genes in other organisms. To calculate inclusive fitness, add up all benefits and costs of a particular gene's actions upon all identical genes, including itself. Each benefit or cost is weighted by the coefficient of relatedness of the recipient of the action to the donor (example ½ for siblings). Genes will evolve to code for actions that maximize their inclusive fitness. Note that from the perspective of the *particular* gene that induces the donation and also the particular individual that "houses" this gene, the behaviour is altruistic. However from the perspective of the *group* of genes, where "group" refers to the set of identical genes in other organisms, the action is selfish. According to the theory of inclusive fitness, genes should be selected if they code for actions that maximize inclusive fitness (a group property).

Maynard Smith (1964) referred to this class of models as "kin selection". He argued that kin selection models do not qualify as group selection because in contrast to the traditional group selection models, new kin groups do not descend from one parent kin group but from two, one for each parent. Therefore, kin groups lack the requisite group level heredity to qualify as "Darwinian" units of selection.

---

[22] This description is a simplification, which roughly holds if, firstly, the gene for altruism is "dominant", so that if an altruistic gene is paired with a free rider gene at a diploid locus, then the expressed phenotype is altruism. Secondly, assume that q is small so that the second parent is likely to be a free rider.

2.2.4 Evolutionary Game Theory and Reciprocal Altruism

In addition to the work of Williams and Hamilton, the development of evolutionary game theory, in the early 1970s, was to contribute to the ascent of the "gene's eye view". Evolutionary game theory grew out of standard game theory, in which individual players are matched (usually at random) into groups of size n. Out of a set of strategies, each individual chooses the strategy that he believes will yield the highest payoff (his best reply) given the information available and the belief that other players will behave in a similar "rational" manner. A Nash equilibrium exists if no player has an incentive to change his strategy given the strategy choices of others. Each player's strategy determines the action taken, while the strategy's payoff depends upon the strategy choice of other players.

In evolutionary game theory, "strategies" correspond to "genes", "actions" correspond to "phenotypes", while "payoffs" are measured in absolute fitness units, which correspond to the number of offspring produced. For example, a gene for altruism can be thought of as a strategy or set of instructions that reads "donate b units of fitness to all other group members". The organism with the altruistic strategy (gene) then performs the altruistic action, which is referred to as its "phenotype". The expected payoff due to a given individual's genetic strategy depends upon the strategies of other players in his group, which depends upon the frequency of each type of gene strategy in the population at large, or "gene pool". Organisms are matched, usually by random sampling from the population with replacement, into groups of size n. Individual organisms interact and reproduce according to their genetic strategies, after which the offspring disperse into the population at large. These genetic strategies are inherited by offspring unless there is a mutation. The "game" is repeated for many generations. The genes that code for actions that give relatively high reproductive success increase in frequency over time. The system may evolve to converge to an evolutionarily stable (ES) equilibrium in which the frequency of gene strategies in the population does not change because each gene has an equal expected fitness, given the other genes in the population. Moreover, a small population of players using alternate strategies cannot invade the population.

Evolutionary game theory illustrates the difficulty of getting genes for altruism to evolve. For simplicity, suppose that players are randomly matched into groups of size two. There are two possible strategies, altruism (A) and free riding (F). An altruist that is paired with an altruist receives a payoff of $x+ b - c$, where b is the benefit received from its partner, c is the cost of the altruistic action and $x > 0$ represents the baseline fitness. Assume that $b > c > 0$. An altruist that is paired with a free rider receives $x - c$, while the free rider receives $x+ b$. In a pair of free riders, each receives only the baseline payoff of x. The payoffs are illustrated in figure 3. If q represents the population wide frequency of altruists, then under random matching, a given player meets an altruist with probability q and a free rider with a probability $1-q$. The expected fitness of a free rider is $x + qb$, which is higher than $x + qb-c$, the expected fitness of an altruist. Under random matching, $q=0$ is evolutionarily stable equilibrium and F is an evolutionarily stable strategy (ESS). Even if the whole population initially consisted of only altruists, free rider mutants would eventually invade and take over the population.

Figure 3: Individual Fitness Payoffs in Free Rider Game

		Player 2	
		A	F
Player 1	A	$X+b-c, x+b-c$	$x-c, x+b$
	F	$X+b, x-c$	x, x

The payoffs can also be examined from the perspective of groups, (here pairs), with AA groups producing the most "offspring", AF groups producing an intermediate number, and FF groups producing the fewest. Although the AA group is fittest from the group perspective, strategy A will not evolve under standard conditions of random matching. If instead groups compete as units, after which they produce offspring groups, (for example, an AA group would produce AA offspring), gene A would evolve to fixation.[23] Competition *between* groups favors altruism, however competition *within* groups favours

---

[23] "Fixation" refers to a situation in which the frequency of a gene in the population is 100%.

free riding. Evolutionary game theory nicely illustrates the problem of explaining the evolution of altruism; *within* group selection always outweighs *between* group selection under conditions of random matching.

Yet the framework is simplistic given that individuals often interact more than once in a lifetime. Trivers (1971) suggested that evolutionary game theory could support the evolution of strategies for cooperation if interactions are repeated allowing altruists to condition their strategies by cooperating or punishing those players who have previously respectively cooperated or taken a free ride. This theory of reciprocal altruism (conditional altruism) was further developed by Axelrod and Hamilton (1981).

Suppose that an altruistic retaliator cooperates on the first round but conditions his play in subsequent rounds upon the other player's previous behaviour. If the other player cooperated in the previous round, he cooperates, otherwise, he free rides. If paired with a free rider, he receives a payoff of $x - c$ in the first round and x in all subsequent rounds. If paired with an altruist, he receives $b - c$ in all rounds. If there are T rounds in total, the payoff if paired with a free-rider is $Tx - c$, while that if paired with an altruist is $T(x+ b - c)$. A free rider receives $b + Tx$, if paired with an altruist, and only Tx, if paired with a free rider. Under conditions of random matching, followed by repeated interactions, if there is a critical frequency of altruistic retaliators in the population, altruism can evolve.

Extensions include indirect reciprocity, whereby altruists establish a reputation increasing the chances of being helped by other altruists, even if they have not met before (for example, see Nowak and Sigmund, 1998). In order to remember previous interactions and keep score as well as to communicate this to others, sophisticated communication abilities and memory are required. It has been suggested that the benefits of reciprocity (both direct and indirect) may have provided a large selective force for the evolution of intelligence (Trivers, 1971; Ridley, 1996).

In the presence of reciprocators, altruism can be selected, however there is still a problem of achieving the initial critical frequency of reciprocal altruists out of a population of free riders. If the population is divided into small groups, the critical frequency of reciprocal altruists could potentially be achieved by chance. However it is difficult to explain how reciprocal altruism, a relatively complex strategy, could have evolved if not preceded by the evolution of simpler non-reciprocal altruism. Reciprocal altruism may have evolved subsequent to the evolution of altruism under conditions of kin group or group selection.

2.2.5 Summary of Early Arguments against Group Selection

By the early 1970's although there was a theoretical basis for traditional (intrademic group selection), mathematically inclined biologists had concluded that although theoretically possible, the required parameter restrictions were too unlikely for group selection to be an important process in nature. In addition to the earlier informal models of Haldane and Wright, this conclusion was supported by a later early round of mathematical models such as those of Maynard Smith (1964), Eshel (1972), and also computer simulations by Levin and Kilmer (1974). Although the argument of unlikely parameter restrictions was disputed by ecologists such as Wynne-Edwards (1964), the views of the mathematical camp were to win the day.

In addition, as argued by Williams (1966), models of selection acting at the level of individuals (or genes) were in many instances able to explain the purported incidents of group selection. Cases in which altruistic acts were directed towards relatives could be explained by kin selection (inclusive fitness) while reciprocal altruism illustrated how selection acting upon individuals (or genes) could explain altruistic acts in the context of repeated interactions among non-relatives.

The case that group selection was possible but unlikely was made using mathematics, while the initial proponents of group selection used descriptive arguments with support from field studies.[24] Due to the level of mathematics required to understand some of the papers of the formal camp, full comprehension requires a substantial time investment,

---

[24] Wilson and Sober, the current major proponents for group selection have referred to these arguments as "naïve", perhaps partially to distinguish their work from that of earlier proponents.

particularly for the less mathematically inclined. Of particular note with respect to difficulty are Hamilton (1964, part 1) and Eshel (1972). Hence, it is fair to say that the majority of biologists and social scientists interested in the topic of altruism received their information second hand through descriptive explanations. Although Williams' (1966) book and Trivers' (1971) paper are accessible to non-mathematical audiences, it was not until the publication of Dawkins' (1976) book *The Selfish Gene*, which has currently sold over a million copies, that the ideas of kin selection, reciprocal altruism, and the "gene's eye view" were communicated to both lay and broader biological audiences.[25] Gradually the view of the gene as the unit of selection moved into the mainstream, while group selection fell into disrepute.

2.2.6 Dawkins' Popularization and Extension of the "Gene's Eye View"

In *The Selfish Gene* (1976) not only did Dawkins bring the ideas of mathematical biologists including Fisher, Williams, Trivers, and Hamilton into the mainstream but he also developed the replicator / vehicle distinction.

2.2.6.1 The Ancient Replicators

Dawkins starts from life's beginnings and uses simple Darwinian principles to explain the gradual evolution of the complex life forms from non-living matter. Out of a collection of inorganic molecules in the "primordial soup", those with a high rate of formation net of destruction will tend to increase in frequency relative to other molecules. This is the "survival of the stable" and is the subject of chemistry. Now consider two molecules A and B, each with a similar stability but suppose that molecule B is able to serve as a template to make copies of itself. Molecule B should increase in frequency relative to molecule A. Dawkins coined the term "replicator" to describe these molecules that contained the information allowing them to duplicate themselves. In turn the copies inherited this information.

> At some point, a particularly remarkable molecule was formed by accident. We will call it the *Replicator*. It may not necessarily have been the biggest or the most complex molecule around, but it had the

---

[25] The 1989 edition adds two chapters and extensive endnotes.

> extraordinary property of being able to create copies of itself (Dawkins, 1976, p. 15).

A replicator's relative frequency in a population of replicators (the replicator pool) is determined by three properties: 1) stability (in common with non-replicating molecules) referred to as "longevity"; 2) the rate of replication; and 3) copying fidelity (replicators with high copying fidelity will become more common). Replicators can evolve by a Darwinian selection process if there are variants that differ according to stability and rate of replication (Lewontin's first two criteria for a Darwinian selection process). In the replicator context, Lewontin's third criterion for a Darwinian selection process, heredity based upon resemblance between parents and offspring, is strengthened to that of perfect copying fidelity.

The material structure of the early replicator molecules had two separate functions; firstly, to interact with the environment (the domain of actions) in order to differentially replicate and secondly to carry the hereditary information for self copying, which in turn is passed on to copies. It is the latter function that distinguishes replicators from non-replicators. The "actions" and "hereditary information" functions were to eventually become separated. Some replicator variants, created by mutation, came to possess chemical structures that allowed them to associate with other molecules such as amino acids, the building blocks for proteins. If such association increased the longevity, replication rate and or copying fidelity, these replicators would increase in frequency in the replicator pool, due to increased success in the competition for sub-units. In turn, descendent replicators inherited the information that coded for these more complex actions. As time progressed, these replicators evolved to code for increasingly more complex actions such as the walls of phospholipids found in single celled organisms and eventually multicellular organisms.

When the initial replicators attained the ability to code for actions other than the structure of the replicator molecule itself (the DNA or RNA), the action component of the replicator became for the most part *de-coupled* from the material structure of the replicator molecule. The essential element of such de-coupling can be understood by referring back to Dawkin's stick insect argument on page 5. A mutation in the DNA molecule is inherited by offspring, while if a stick insect looses its leg, any offspring will not inherit a missing leg. Dawkins describes the "action" component of replicators as *vehicles*, the tools with which the replicators interact with the environment. The hereditary information contained in the replicators may be viewed as blueprints providing information for the construction of vehicles.

> Replicators began not merely to exist, but to construct for themselves containers, vehicles for their continued existence. The replicators that survived were the ones that built survival machines for themselves to live in (Dawkins, 1976).

Replicators are not limited to molecules of DNA and RNA, but include any entity that possesses information allowing it to make copies of itself. For example, the bible is an example of an extremely successful replicator, which can be viewed as manipulating its environment in order to get copied. At the end of chapter 2, Dawkins writes;

> Four thousand million years on, what was to be the fate of the ancient replicators? They did not die out, for they are past masters of the survival arts. Now do not look for them floating loose in the sea; they gave up that cavalier freedom long ago. Now they swarm in huge colonies, safe inside gigantic lumbering robots, sealed off from the outside world, communicating with it by tortuous indirect routes, manipulating it by remote control. They are in you and in me; they created us, body and mind; and their preservation is the ultimate rationale for our existence. They have come a long way those replicators. *Now they go by the names of genes, and we are their survival machines* (my italics).

Note that in the last italicized sentence, Dawkins has equated replicators with "genes". This would go unnoticed on a first read, because Dawkins has not yet defined genes. At this stage in the book, a reader is likely to imagine a replicator as the linear structure embodying all of the heredity strategies to produce its aggregate actions (the vehicle), which allow it to make copies of itself. In other words, the replicator is a genome. In addition, there is a one-to-one relationship between the hereditary information of the replicator as a whole and its actions, the vehicle. The first multicellular organisms in the primordial soup (the environment described in chapter 2) might most simply be imagined as asexually reproducing organisms so that the *genome* in its entirety is a replicator given the assumption that meiosis has not yet evolved. A reader versed in the standard concept of a gene as the hereditary information coding for particular actions such as eye-colour or in Williams' definition of a gene might view the equation of replicators with genes as a rather large jump in the story. This shortcoming is however understandable as the details regarding the relationship between replicators, genes and sexual genomes would merely serve to confuse at this point in the argument. Defining a "gene" is the task of Dawkins chapter 3.

2.2.6.2 A Subtle Change in Meaning of the Word Replicator

In chapter 3, Dawkins defines genes according to Williams (1966). A "gene" is defined as any portion of chromosomal material that potentially lasts for enough generations to serve as a unit of selection (in Fisher's sense as described above). This requires a sufficiently high level of copying fidelity. Dawkins writes "Copying fidelity is another way of saying *longevity-in-the form-of copies* and I shall abbreviate this simply to longevity" (my italics, Dawkins, 1976). Note that the meaning of "longevity" is different from that in chapter 2 in which it referred to the stability or lifetime of an individual replicator.

The early replicators in the primordial soup were long chains of hereditary information, which did not get splintered by meiosis and so could be copied precisely. With the evolution of meiosis, whole genomes no longer qualify as replicators due to low copying fidelity. Although organism level vehicles can differentially survive and reproduce, the underlying information that codes for the vehicle is not transmitted as a unit to offspring. Dawkins argues that selection acts upon fragments of the genome, which are small enough to survive meiosis intact, i.e., Williams' "genes". These genes qualify as replicators on the basis of a high level of copying fidelity. Both genomes and their organism vehicles are emergent phenomena, due to selection acting at the lower level of genes. As such, in this view adaptations cannot "evolve for the good of" genomes or organism vehicles but only for genes.

A replicator can persist indefinitely through copies. More accurately, a lineage of a replicator can escape entropy to become potentially immortal. It is still the case, however, that a particular *individual* replicator in a lineage is as ephemeral as the genome and the vehicle organism in which it is contained. "Longevity", which initially referred to the lifetime of a particular *individual* replicator, now comes to refer to the lifetime of a particular gene replicator and all of its successors, the lifetime of a lineage. Genes are potentially "immortal" is another way of saying that gene lineages are potentially immortal. And so the selfish gene becomes a lineage—or a population (group) of copies of the same gene existing throughout time.

Genomes are then coalitions of gene replicators jointly coding for the actions or their vehicles of survival. The reproductive success of the vehicle given the environment determines whether the hereditary information will be transmitted, albeit fragmented, to offspring vehicles. The gene pool comes to be composed of genes that have proven to be fitter on average given the average environments, including the actions coded for by the company of genes with which they have found themselves with in the past.

In addition to employing Williams' concept of a "gene", Dawkins makes use of Hamilton's concept of inclusive fitness and Trivers' reciprocal altruism in order to explain the evolution of altruism from the perspective of selfish genes. A particular type of gene can increase in frequency in the gene pool if it codes for actions that increase the frequency of identical genes in the gene pool, even if the action lowers the fitness of the donor organism and also the individual gene that codes for the action.[26] Dawkins argues that genes will be selected to take actions that maximize their inclusive fitness (the fitness of all genes of the same type that exist in the gene pool). Selfishness then appears to be a group property and individual organisms and *individual* genes appear to be altruists.

> An animal's behaviour tends to maximize the survival of the genes "for" that behaviour, whether or not those genes happen to be in the body of the particular animal performing it (Dawkins, 1982, p. 233).

---

[26] Dawkins (1982, p. 294) defines "gene pool" as "the whole set of genes in a breeding population".

2.2.6.3 Dawkins on the Harmony and Coordination Observed in Organisms

Given this view that the genome is not a replicator, Dawkins next tackles the problem of explaining the functionality of the organism vehicles.[27] He employs the concept of "shared fate", which he explains by analogy to the process by which rowing coaches select crews. Coaches race two boats against each other, after which rowers are switched for specific seats. The coach switches seats until she finds the fastest combination of rowers. The outcome of each trial race depends upon the joint efforts of all the rowers; a free rider cannot take advantage of the other rowers because she will hurt herself along with the whole boat. Hence all rowers are in the same boat and have a common interest of winning the race—this is shared fate. If the rowing boats could have offspring at the end of each race according to their relative standing in the race, groups of rowers that row well together should tend to be selected over time.

Similarly, the genes (replicators) on a genome are in the same boat (vehicle), in that each has an equal expected fitness given that meiosis is usually fair.[28] Within the genome, a gene that does not well with the collective of genes reduces not only the fitness of the collective but also its own fitness. All genes on a genome have an equal expected fitness in most circumstances. Shared fate provides selective pressure for genes to work well together.

---

[27] This view is explained at length in the chapter 13 of the 1989 2nd edition of *The Selfish Gene* and chapter 14 of *The Extended Phenotype* (1982).

[28] In some cases, outlaw genes are able to bias the process of meiosis in their favour via processes such as meiotic drive. During meiotic drive, a gene biases meiosis such that it has a greater than 50% probability of finding itself in a successful gamete. For example, in mice, a gene called a t-element is able to bias meiosis such that it is inherited in 95% of gametes. If a given parent is homozygous (possesses two copies of a t-element at a diploid locus), it is sterile. As t-elements are fitter than normal genes in heterozygote form (only one copy of the t-element at a diploid locus) they can rapidly take over a mouse population leading to the ultimate extinction of the mouse population. They only survive if they are able to spread to another mouse population before the old population goes extinct. Modifier genes can evolve to combat the effect of such outlaws. For example, populations of mice with modifier genes will not go extinct in the presence of an outlaw gene. Strong evolutionary pressure for such modifier genes ensures that for the most part, meiosis is fair.

According to the view that genomes lack sufficient copying fidelity to pass on their strategies as a whole, the harmony and coordination of parts that is observed within organisms and sometimes in larger social groups is an emergent property. The harmony of vehicles is an emergent property due to the joint actions of individual gene replicators each of which has been selected due to being relatively fit given the average environment experienced by its ancestors. Note that the environment includes the other genes on the genome with which ancestors have on average tended to find themselves with. Dawkins (1989, p. 254) writes:

> ...the organism and the group of organisms are true rivals for the vehicle role in the story, but neither of them is even a candidate for the replicator role. The controversy between 'individual selection' and 'group selection' is a real controversy between alternate vehicles. The controversy between individual selection and gene selection isn't a controversy at all, for gene and organism are candidates for different, and complementary, roles in the story, the replicator and the vehicle.

Dawkins' next point is to argue that it is difficult for replicators to share a fate in groups larger than the individual because they are not forced through the same life-cycle bottleneck during reproduction and hence need not share a common fitness. Hence, except for situations such as kin group selection, individual organisms, not groups, are the relevant vehicles of selection. While organisms and potentially larger groups can serve as vehicles of selection, (in the biological context), the hereditary information that codes for these "actions" lacks the copying fidelity to be transmitted intact to the next generation. Only in the case of genes is the information for the coded phenotypic action, transmitted intact to the next generation. The remarkable harmony and coordination observed in organisms, and sometimes, higher level vehicles are emergent phenomena due to selection acting at the lower level of the gene. Wilson and Sober have accepted Dawkins' view that genes are the only important biological replicators and also that the "units of selection debate" is about vehicles. However, in disagreement with Dawkins, they argue that group level heredity is not required for group level adaptations. The views of Wilson and Sober will be examined in the following sections.

2.3 The New "Group Selection" Theories

In spite of group selection's fall into disrepute during the 1960s and 1970s, a number of biologists continued to work on group selection.[29] For example, Wilson and also Gould, have produced many papers arguing not only that group selection is theoretically possible but that it is an important process in nature.[30] Wilson and Sober have argued that both interdemic and intrademic (including kin group) selection are capable of generating group level adaptations. Group selection appears to have come back into popularity with the development of many formal mathematical and computer models. The work on group selection has not been limited to biology. Within anthropology, group selection models have been applied to the human context, including the evolution of culture (for example, see Richardson and Boyd, 1990; Boehm, 1993, 1996). Even economists, from a discipline that champions the emergent phenomenon approach, are contributing to this growing literature (Samuelson, 1993; Canals and Vega-Redondo, 1998; Hirshleifer, 1999; Hayek as discussed by Caldwell, 2000; Gintis, 2000; Bergstrom, 2002; Gintis and Bowles, 2004). An overview of the new group selection arguments is presented below with an emphasis upon models that will be important to my argument.

2.3.1 Simpson's Paradox

The following adapted from Wilson and Sober (1998) illustrates "Simpson's paradox" whereby an apparently individually disadvantageous trait can increase in frequency if the population is partitioned into groups at some stage in the organism's life cycle. Assume that individuals in a population of size n have a choice between two possible strategies, A for altruism and F for free riding. The population wide frequencies of altruists and free riders are respectively q and $1-q$. Absolute fitness indicates the number of immediate offspring of a particular individual. In example 1, it is assumed that there are no groups. An altruist incurs a cost of c units of fitness due to donating b units of fitness, which is

---

[29] Favorable overviews of the literature on group selection by persons other than Wilson, Sober, and Gould include Wade, 1978; Nunney, 1985; and Bradley, 1999.

[30] Some of the many works by Wilson are sited in the index. Gould's (2002) book *The Structure of Evolutionary Thought* provides a comprehensive, although somewhat wordy, overview of the group selection debate.

shared equally among members of the population. An altruist is assumed not to bear the fruits of his own gifts but receives an equal share *1/(n-1)* of the gifts from the *nq*-1 other altruists. Free riders receive a share of *1/(n-1)* of each gift but do not incur the cost *c* of contributing. Let *x* represent the baseline fitness common to each type. The payoffs in fitness units to an individual altruist and free rider are respectively represented as $w_A = x - c + \frac{b(nq-1)}{n-1} = x + \frac{bnq}{n-1} - c - \frac{b}{n-1}$ and $w_F = x + \frac{bnq}{n-1}$.

Assuming parameter values of n =100, q =0.5, x =10, b=5 and c=1 yields individual fitness values of w_A= 11.47 offspring for altruists and w_F= 12.53 offspring for free riders. To calculate the expected number of altruists in the population in the next generation, multiply the expected number of offspring by the number of altruists in the preceding generation and similarly for free riders. The expected values are nqw_A = 573.5 altruistic offspring and $n(1\text{-}q)w_S$ = 626.5 free rider offspring for a total of 1200 offspring. In the next generation, the frequency of altruists is $q' = 0.4779$ yielding a total of 47 altruists and 53 free riders if the population is normalized to 100. As free riders are always expected to be fitter, they eventually they evolve to fixation.

Consider a second example with the same parameters as above except that the population is partitioned into two groups, each of size 100. An altruist's donations are conferred only upon members of his own group. Again, let q = 0.5 represent the initial frequency of altruists in the entire population with q_1 = 0.2 and q_2 = 0.8 representing the respective frequencies of altruists in each group. The fitness of each respective type is now group dependent.

In group 1, on average, an altruist and free rider respectively produce $w_A^1 = 9.96$ and $w_F^1 = 11.01$ offspring. In aggregate, the 20 altruistic and 80 free rider parents are respectively expected to produce 199 altruistic and 881 free rider offspring for a total of 1080 offspring. In group 2, similarly, $w_A^2 = 12.98$ and $w_F^2 = 14.02$. In aggregate, the 80 altruistic and 20 free rider parents are respectively expected to produce 1038 altruistic and 280 free riders for a total of 1310 offspring.

The frequency of altruists in each respective group in the next generation is $q_1^{\prime} = 0.18$ and $q_2^{\prime} = 0.79$. Notice that, *within* each group the frequency of altruists has fallen, illustrating that within groups free riding is the fittest strategy. However, somewhat counter-intuitively, the frequency of altruists in the population at a whole increased from 0.5 to 0.5157. In example 1, in which there were no groups, the frequency of altruists fell from 0.5 to 0.4779. The difference of 0.5157-0.4779 = 0.0378 can be attributed to the presence of groups and has been referred to as the *group* component of selection (Wilson, 1975; Hamilton, 1975). *Within* group competition reduced the frequency of altruists while *between* group competition increased the frequency.

In example 1, a free rider is expected to receive the same number of gifts as an altruist and avoids the costs of donating. Hence free riders are expected to be fitter. In example 2, altruists have a higher probability of being members of group 2, in which there is a higher proportion of altruists. On average, altruists in the population receive more gifts than free riders, which provides sufficient fitness benefits to outweigh the costs of donating. Interactions are said to be "positively assortative" in contrast to the random assortment in example 1. In example 2, positive assortment into groups is sufficient to allow altruism to increase in frequency in the population. The population wide average level of fitness is 12.38 for an altruist and 11.62 for a free rider. Although the example does not illustrate that groups are Darwinian units of selection for which altruism can evolve for the good of, it illustrates that a group advantageous trait can evolve if altruists are positively assorted into groups.

2.3.2 The Price Equation and Intrademic Models of Group Selection

In 1975, three formal models of intrademic group selection were independently developed by Charnov and Krebs, Hamilton, and Wilson. Both Hamilton and Wilson used the term "group selection" to describe their models. Wilson supported his use of the term group selection by arguing that intrademic selection can generate group level adaptations. This claim has generated controversy because in comparison to interdemic group selection, in which offspring groups derive from a single parent group, in intrademic group selection, except in the case in which new groups consist of one individual, offspring groups derive from more than one parent group. Hence, there is apparently no group level heredity, the third criterion for selection of the "Darwinian" sort. In order to investigate the claim that intrademic selection qualifies, Hamilton's model of intrademic group selection is examined below.

Hamilton considered a population consisting of a mixture of two types of individuals, altruists (A) and free riders (F).[31] The population is subdivided into groups, indicated by the subscript *s*. Individuals reproduce asexually. The parameters for a given group and the population as a whole are defined in table 1.

Table 1: Parameters for Hamilton's Model

	Group	Population
Number of individuals of both types	n_S	$N = \sum n_S$
Frequency of individuals of type A	q_S	$q = \sum n_S q_S / N$
Mean fitness	w_S	$w = \sum n_S w_S / N$
Change in frequency of individuals of type A in one generation	$\Delta q_S = q_S' - q_S$	$\Delta q = q' - q$

Fitness represents the number of offspring. The superscript $'$ and "no superscript" respectively denote the next and current generation. The size of the population in the next generation is obtained by multiplying the original population by the average fitness of an individual entity in the population, such that $N' = wN$. Similarly, the population in each group in the next generation is the product of the initial group size and the average within group fitness, so that $n_S' = w_S n_S$.

Hamilton used Price's (1972) equation to partition the effects of selection into within and between group components as follows. The change in frequency of trait A at the level of the population is represented as $\Delta q = q' - q$. Substituting from Table 1 gives

$$\Delta q = \sum \frac{n'_s q'_s}{N'} - \sum \frac{n_s q_s}{N} = \sum \frac{w_s n_s q'_s}{wN} - \sum \frac{n_s q_s}{N}$$

$$= \sum \frac{w_s n_s (q_s + \Delta q_s)}{wN} - \sum \frac{n_s q_s}{N} = \frac{\sum w_s n_s q_s - w \sum n_s q_s}{wN} + \sum \frac{n_s w_s \Delta q_s}{wN}.$$

Multiplying both sides through by w gives

$$w\Delta q = \sum \frac{n_s w_s (q_s - q)}{N} + \frac{n_s w_s \Delta q_s}{N} = Cov(w_s, q_s) + E(w_s \Delta q_s).^{32} \qquad [3\text{-}1]$$

Equation 3-1, the Price equation, illustrates that the change of frequency of the trait A can be written as a sum of two components. The first term, the *between*-group selection component, is the covariance of the within-group fitness and a group characteristic, the within-group frequency of altruists. This term represents the effect of the within-group frequency of altruists upon the average group fitness. The second term, the *within*-group selection term, is the expectation of the product of within-group fitness and the within-group change in frequency of the altruistic trait.

If there is no between-group variation in average frequency of trait A then $q_s - q = 0$, so that the covariance term is 0. This corresponds to a situation in which there are no selective forces acting *between* groups. In order to have selection at the level of the group, groups must vary with respect to a group-level trait, i.e. q_S must differ between groups. If so, then groups satisfy Lewontin's first criterion for Darwinian selection; variation among units is required for Darwinian selection of units. The covariance term also illustrates Lewontin's second criterion for selection at the level of groups. If in addition to variation in a group-level trait, this group-level trait (here q_s) is positively or negatively related to group level fitness (here w_s), then the covariance term is respectively positive or negative. If this covariance term is 0, then group-level traits have a neutral effect upon group fitness.

---

[31] Note that traits A and F can be any traits of interest and are not limited to the altruism context.

[32] Alternately, $\Delta q = \sum \omega_s n_s (q_s - q)/N + \sum n_s \omega_s \Delta q_s / N$, where $\omega_s = w_s / w$ represents the relative fitness.

The second expectation term illustrates that if the frequency of trait A does not on average change *within* groups such that $E(w_S \Delta q_S) = 0$, then on average there is no positive or negative selection acting at the level of individuals *within* the groups. In the current example, there is selection of individuals within groups because free riders are by definition fitter within groups. Hence, the frequency of altruists is expected to fall within a given group. Note that the expectation term can be further decomposed by recursive use of equation 3-1, if there are smaller groups within each *s*-group.

If groups are very large and formed at random, then the frequency of altruists in each group is expected to equal the frequency in the population, $q_s = q$, and there can be no selection between groups. If, however, during group formation like types are positively assorted, then groups with a high frequency of altruists grow relatively fast in comparison to those with a high frequency of free riders. However, *within* each group, free riders are relatively fit so that the frequency of altruists falls within a given group. If the *between* group effect is larger than the *within* group effect, then the population wide frequency of altruists will increase for $0<q<1$ and eventually go to fixation.

Hamilton was interested in Haldane's suggestion that in a group with a high proportion of altruists, if group fission occurs before free riders overrun the group, then offspring groups could inherit the high proportion of altruists. Since groups with high proportions of altruists grow faster, altruists could potentially increase in frequency at the level of the population.[33]

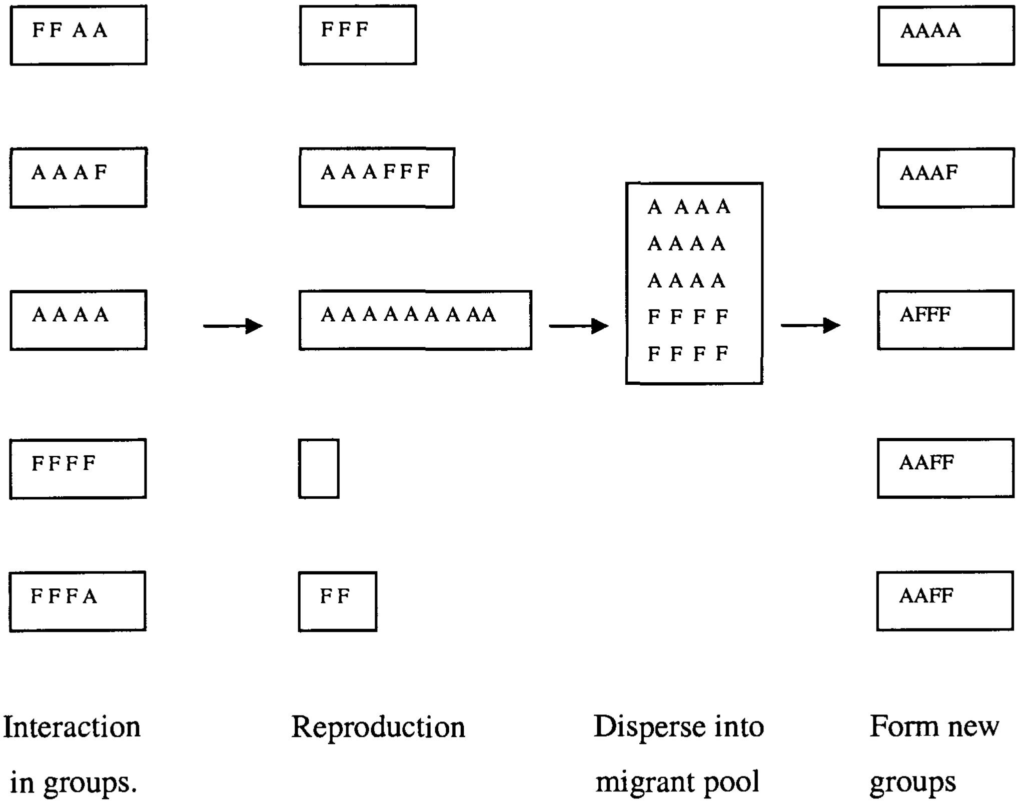
Since Haldane's suggestion is difficult to analyze using the Price equation, Hamilton took an extreme case in which groups completely disperse after an interaction and reproduction phase within groups. When young, asexual organisms interact in groups of size *n* after which they reproduce, the number of offspring depending upon their type (A or F) and the within group frequency of altruists. After reproducing, parents die after which groups completely fission and the offspring join a migrant pool. New groups of

---

[33] See footnote 16 for a discussion of evolutionarily stable equilibria.

size n are formed by randomly drawing members (with replacement) from the migrant pool. The process is repeated. The lifecycle of a given group is shorter than the lifecycle of an individual. A hypothetical cycle is illustrated in figure 4.

Figure 4: Intrademic Group Selection



Although the expected frequency of altruists in each new group is equal to that in the population at large, if groups are small, the frequency of altruists can vary substantially between groups due to random sampling. Can random sampling produce sufficient *between* group variation and positive assortment to allow for the selection of altruism at the level of the population?

To illustrate that altruism cannot evolve to fixation under conditions of random grouping, Hamilton uses a linear approximation of the Price equation. Assume that group fitness increases linearly with the frequency of altruists in the group. Let β_0 be the coefficient obtained by regressing w_S on q_S. The covariance term can be expanded and approximated to give

$$Cov(q_s, w_s) = \sum \frac{n_s}{N}(q_s - q)w_s = \sum \frac{n_s}{N}\frac{(w_s - 1)}{(q_s - q)}(q_s - q)^2 = \beta_0 Varq_s. \qquad [3\text{-}2]$$

By recursive use of equation 1, the expectation term can be expanded as

$$E(w_s \Delta q_s) = E[Cov_s(q_{si}, w_{si})] + E(w_{si}\Delta q_{si}). \qquad [3\text{-}3]$$

The first covariance term represents *between*-individual selection, while the second expectation term represents *within*-individual selection. The fitness of individual i in group s is represented by w_{si}. The frequency of trait A within an individual is represented by q_{si}, which takes the value of 0 for a free rider or 1 for an altruist. Assuming no mutation, an individual cannot change traits within a period and so

$$\Delta q_{si} = 0 \qquad [3\text{-}4]$$

and

$$E(w_{si}\ \Delta q_{si}) = 0. \qquad [3\text{-}5]$$

Again using an approximation in which an individual's within group fitness varies positively and linearly with the within-individual frequency of trait A, and letting β_1 be the coefficient obtained by regressing w_{si} on q_{si} gives

$$E(Cov\ (q_{si}, w_{si})) \cong \beta_1 E[\ Var_s q_{si}]. \qquad [3\text{-}6].$$

Substituting [3-3], [3-4] [3-5]and [3-6] into [3-1] gives the linear approximation for the Price equation;

$$w\Delta q = \beta_0 Varq_s + \beta_1 E[Var_s q_{si}]\ . \qquad [3\text{-}7]$$

Groups with high proportions of altruists have higher average fitness so that $\beta_0 > 0$. Within groups, free riders are fitter so that $\beta_1 < 0$. Hence if the magnitude of β_0 is large relative to β_1 and if the between group variance is high relative to the within group variance with respect to the frequency of the trait, then the frequency of altruists could potentially increase at the level of the population.

The group sampling procedure generates a binomial distribution, which allows for the calculation of the expected frequency of altruists within a given group and also the variance. The expected number of altruists, x_s, in a given group is nq, while the variance of x_s is $nq(1-q)$. The expected frequency of altruists in a group with x_S altruists is $E(q_S)=x_s/n$ and the variance is

$$\mathrm{Var}\, q_s = q(1-q)/n. \qquad [3\text{-}8]$$

Similarly to calculate the variance of the trait among individuals in a given group s with a frequency of q_s altruists, draw sub-groups of size 1 (corresponding to individuals) from a particular group. Let x_{si} represent the number of altruists in a group of size 1 so that $q_{si}=x_{si}/1$. Again using the binomial distribution, the expected value of q_{si} is q_s. The variance of x_{si} in group s is $q_s(1-q_s)$ giving $\mathrm{Var}_s q_{si} = q_s - q_s^2$. Take the expected value to get

$$E[Var_s q_{si}] = Eq_s - E(q_s^2) = q - \frac{q(1+(n-1)q)}{n} = \frac{(n-1)q(1-q)}{n}. \qquad [3\text{-}9]$$

The next task is to find β_0 and β_1 given the fitness parameters. Let 1 be the baseline level of fitness, independent of type. Each altruist incurs a cost of c to donate a gift to each other group member and receives a gift from each of the other x_s-1 altruistic group members. Each gift is shared among n-1 group members. An altruist does not receive a share of its own gift. The average fitness of an individual altruist in a group of size n with x_s altruists is

$$w_{SA} = 1 - c + \frac{x_s - 1}{n-1} b. \qquad [3\text{-}10]$$

A free rider incurs no costs and receives a gift from x_s altruists, each of which is shared with n-1 other group members. The average fitness of a free rider in this group is

$$w_{SF} = 1 + \frac{x_s}{n-1} b. \qquad [3\text{-}11]$$

Combining [3-10] and [3-11] and weighting by the within group frequency of each type gives the expected fitness of an individual in group s,

$$w_s = \frac{x_s}{n}\left[1 - c + \frac{b(x_s - 1)}{n-1}\right] + \frac{n - x_s}{n}\left[1 + \frac{x_s b}{n-1}\right] = 1 + q_s(b - c). \qquad [3\text{-}12]$$

Differentiating [3-12] with respect to q_s gives

$$\beta_0 = b - c. \qquad [3\text{-}13]$$

Adding the individual fitness payoffs [3-10] and [3-11], weighted by the probability that an individual plays a given strategy conditional upon being a member of group *s*, gives the expected fitness of an individual in group *s*,

$$w_{si} = q_{si}\left[1 - c + \frac{b(x_s - 1)}{n-1}\right] + (1 - q_{si})\left[1 + \frac{bx_s}{n-1}\right]. \qquad [3\text{-}14]$$

Differentiating [3-13] with respect to q_{si} gives

$$\beta_1 = -c - \frac{b}{n-1}. \qquad [3\text{-}15]$$

Substituting [3-8], [3-9], [3-13] and [3-15] into the linear version of the Price equation [3-7] gives

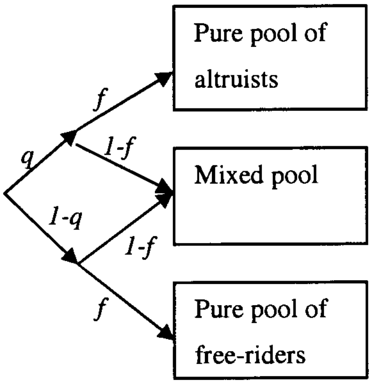
$$w\Delta q = -(1 - q)qc. \qquad [3\text{-}16]$$

Given that [3-15] is less than 0 for $0 < q < 1$, under conditions of random grouping, even if the population-wide frequency of altruists rises in a given period due to chance, over a longer time frame given the law of large numbers, the frequency of altruism will fall until free riding goes to fixation. *Within*-group selection against altruism outweighs *between*-group selection for altruism.

If instead of random sampling, grouping is assortative, such that on average a given altruist associates with a higher proportion of altruists than the population wide frequency, then the population wide frequency of altruists can increase. Hamilton derived a condition illustrating the required degree of positive assortment. Assume that the correlation between any two members of a group is *f*. Such a correlation can be achieved by allowing a fraction *f* of groups be made pure for each type, while a fraction *(1-f)* have the population wide frequency of altruists, *q*.

A given individual is an altruist with a probability *q* and a free-rider with a probability *1-q*. Conditional upon being an altruist, the probability of being in a group consisting purely of altruists is *f*, while the probability of being in a group with a frequency of *q* of altruists is *1-f*. Therefore conditional upon being an altruist, the expected frequency of altruists in the group is $q(1-f) + f$. Similarly conditional upon being a free-rider the expected frequency of altruists in the group is $q(1-f)$. This gives an unconditional group frequency of altruists of *q*. Alternately an individual can be viewed as drawing strategies as follows. In stage 1, the individual draws a strategy for A or F with probability *q* and *1-q* respectively. In stage 2, conditional upon the strategy drawn in stage 1, with probability *f* the individual draws group members from a pure pool of its own type and with probability *1-f* from a mixed pool in which the frequency of altruists is *q*. The three types of groups are: 1) all members are drawn from a pool containing only altruists; 2) members are randomly drawn with replacement from a large group in which the frequency of altruists is *q*; and 3) members are drawn from a pool containing only free riders. The statistics for groups of size *n* for three types of groups are illustrated in figure 5.

Figure 5: Statistics for Assortative Sampling



		$E(q_s)$	$E(x_s)$	Variance within group
Pure pool of altruists	qf	1	n	0
Mixed pool	$(1-f)$	q	qn	$(n-1)(1-q)q/n$
Pure pool of free-riders	$(1-q)f$	0	0	0

To calculate the expected value of within group variance for trait A as indicated in figure 5, add up the within group variance for each type of group and weight by the probability of each type of group to get

$$E[Var_s q_{si}] = (1-f)(n-1)(1-q)q/n. \qquad [3\text{-}17]$$

The variance of the average *within*-group frequency of trait A is

$$Varq_s = q(1-q)[nf + (1-f)]/n. \qquad [3\text{-}18]$$

As in the case of random group formation, substitution of [3-13], [3-15] , [3-17] and [3-18] into the linear version of the Price equation [2] gives $w\Delta q = n(bf - c)$. Altruism is expected to increase in frequency at the level of the population if $w\Delta q > 0$, which yields the condition

$$bf > c. \qquad [3\text{-}19]$$

Hence, parameters that tend to favour selection of the altruistic trait are a high coefficient of assortment and high benefits to recipients relative to costs to donors.

Hamilton's (1975) model illustrates that altruism cannot experience a sustained increase in frequency under conditions of random group formation in which $f=0$. The reason can easily be understood without reference to the formal mathematical model. If groups are formed by randomly drawing n members with replacement from a population with a frequency q of altruists, then a given altruist has the same probability of finding itself in a group with a particular composition of neighbors as a given free rider. For any particular composition of neighbors, by definition, a free rider is always fitter than an altruist within the group. Under random sampling, by chance, there may be sufficient positive assortment of altruists so that the population wide frequency of altruists increases for a particular generation. However, over many generations the population wide frequency of altruists is expected to decrease. The population wide frequency of altruists can only increase on a sustained basis if altruists positively assort such that on average a given altruist groups with a higher proportion of altruists than a given free rider, so that condition [3-19] is satisfied.

2.3.3 Kin Selection/ Inclusive Fitness Revisited

Hamilton (1975) and Wilson (1975) both noted that kin group selection is a special case of intrademic group selection in which kin are positively assorted into kin groups by descent from common parents. The coefficient of assortment f corresponds to the coefficient of relatedness and so condition [3-19] corresponds to Hamilton's rule in the context of kin group selection. For example, in a sibling group, half of an altruist's siblings are expected to be altruists by descent so that $f = 1/2$. Assume that each altruist in the group donates a gift of b, which is shared among the $n-1$ other group members. A given altruist receives a gift of $b/(n-1)$ from each of the other altruists. There are expected to be $(n-1)/2$ other altruists in the group. Hence an altruist is expected to receive a total gift of $b/2$ while incurring a cost of c due to donating. The expected fitness of an altruist is then $x+ b/2 - c$, given the assumption that q, the population wide frequency of altruists, is small where x is the baseline fitness. On the other hand, a free rider's expected fitness is only x because he does not expect to have altruistic siblings and hence expects to receive no gifts, even though he saves c by not donating. Comparing the expected fitness of each type illustrates that an altruist is expected to be fitter than a free rider if $b/2>c$, which is Hamilton's rule for siblings.

In section 2.2.3, Hamilton's rule was calculated using the inclusive fitness approach. Hamilton's (1975) approach clearly illustrates that the process by which altruism can increase in frequency by kin selection is due to Darwinian selection acting upon particular genes. The maximization of inclusive fitness is a byproduct of selection acting at the level of individual genes. If $b/2>c$, altruistic genes and individuals are expected to be fitter than free riding genes because they have a greater chance of inheriting an environment with a higher proportion of altruistic siblings. Hence, the "altruistic" strategy, is actually the "selfish" strategy. Even though free riders would prefer to free ride in a group of altruistic siblings, given that free riders are constrained to inherit free riding siblings, they are unlikely to get the chance to play this preferred strategy.

Wilson and Sober (1994, 1998) point out that it is ironic that inclusive fitness/ kin group selection was seen as an alternative to group selection, given that Hamilton, the inventor of inclusive fitness was later to characterize kin group selection as a special case of group selection. Nevertheless, although it is *not* controversial to characterize kin group selection as a special case of intrademic group selection, it *is* controversial to categorize intrademic selection as "group selection" if this term is reserved for "Darwinian" selection processes that generate group-level adaptations. As kin groups do not have a direct line of descent to a single parent, then kin groups appear to lack sufficient group level heredity.

2.3.4 Evolutionary Game Theory and Intrademic Group Selection

Wilson and Sober (1994, 1997) point out that evolutionary game theory has the same population structure as intrademic group selection. Individuals are matched into groups of size n, in which they interact and then reproduce, along with their strategies, according to group dependent fitness payoffs. Offspring disperse into the migrant pool and the process repeats. In section 2.2.4, grouping was modeled as random. More generally matching can be non-random or assortative. The example is modified to allow for assortative grouping as follows.

Assume that f represents the coefficient of positive assortment defined in section 2.3.2. Using the assortative sampling process of Hamilton (1975) (section 2.3.2), the probabilities that an altruist pairs with an altruist or a free rider are respectively represented as $p_{AA} = f + (1-f)q$ and $p_{AF} = (1-f)(1-q)$. Similarly, the probabilities that a free rider pairs respectively with an altruist or a free rider are $p_{FA} = (1-f)q$ and $p_{FF} = f + (1-f)(1-q)$. Given the payoffs illustrated in figure 3, an altruist's expected fitness is $x - c + b[f + (1-f)q]$, while that of a free rider is $x + (1-f)qb$. By comparing these expected fitness values, it is easy to see that the population wide frequency of altruists will increase if $fb>c$, which is again Hamilton's rule (see Wilson and Sober, 1998, p. 80-82).

Wilson's term "trait groups" refer to groups in which entities interact to affect each other's payoffs at some phase in the lifecycle. Trait groups, hence, correspond to the groups in evolutionary game theory. Wilson (1975) argues that the intrademic population structure is common in nature and is not just limited to the special case of kin selection. For example, each pair of genes at a locus on a diploid genome is an example of a trait group because the gene pair jointly determines the phenotype. For example, if there are two blue-eye alleles at a diploid locus, the phenotype is blue-eyes, while if a blue-eye allele is paired with a brown-eye allele at a locus, the phenotype is brown eyes (each gene at a particular diploid locus is referred to as an "allele"). During meiosis these trait groups "dissociate", after which the independent alleles may find new partners during sex. Another example is groups of parasites or viruses that share a common host. *Within*-group selection favours higher levels of virulence, while *between*-group selection, can favour the evolution of lower levels of virulence because the host survives for longer, helping out the whole group of viruses (Wilson and Sober, 1998, p. 43-50).

Reciprocal altruism can also be viewed as an assortative process (Wilson and Sober, 1998, p. 81-86). Under conditions in which players are randomly matched into groups, the probability of being matched with a free rider does not depend upon a player's strategy and so in this sense matching is not assortative. However, even if matching is random, given that a player's strategy is conditioned upon the types of players with whom he is matched, "actions" end out being played assortatively; a conditional altruist cooperates if matched with another altruist and free rides if matched with a free rider.

2.3.5 Wilson and Sober and Multilevel Selection

The various works of Wilson and also Wilson and Sober (1994, 1998) have played a major role in bringing group selection back into the limelight. By the 1990s, a number of biologists had produced empirical and theoretical support for "group selection". Empirical support includes evidence of female biased sex ratios (see p. 17, footnote 16), selection for low levels of virulence in viruses, and low fecundity rates in parasites. For example, considering group selection and virulence. Wilson and Sober (1998, p. 45) wrote "Richard Lewontin (1970) ...was the first to realize that the evolution of virulence provides a natural test for group selection theory." In order to control the Australian rabbit population, the government introduced a type of virus called Myxoma. Although, initially, the virus rapidly reduced the rabbit population, overtime, the virus became less effective. The most obvious explanation is the selection of wild rabbits, which were most resistant to the virus. Indeed, these wild rabbits were found to be resistant to laboratory strains of the virus. However, not only were the rabbits more resistant, but the viruses had also become less virulent. When laboratory rabbits were inoculated with the virus, it was found that that wild strain was less virulent than the laboratory strain, which is to be expected if selection is acting upon groups, but not if selection is acting upon individuals. For reviews of empirical work see Wade, 1978; Wilson and Sober, 1994, 1998; and Bradley, 1999. In addition to the empirical studies, recent theoretical work suggests that the parameter restrictions used in the earlier group selection models were unnecessarily restrictive (Wilson, 1987). Furthermore, enhanced computer power has reduced the need to make assumptions to facilitate mathematical tractability. In spite of Wilson and Sober's careful documentation of a growing literature in support of "group selection", it has been argued that they overstate the case for group selection based upon their inclusion of intrademic group selection (Maynard Smith, 1987a).

Wilson (1975) categorized intrademic group selection as "group selection" and argued it is an important and widespread process in nature. In many instances, at some stage in the lifecycle, sub-units come together to interact in groups, with fitness consequences, after which they disperse into the population at large before regrouping. Wilson argues that although group selection was initially proposed as a solution to the puzzle of altruism, group selection can occur in a wide variety of situations, in which fitness is affected by grouping. For example two different types of organisms often interact at some stage in their lifecycle, each mutually gaining from the interaction. In this view, communities of organisms consisting of different interacting species are potentially subject to group selection.

Wilson and Sober (1994, 1998) argue that selection of a particular trait can occur at a variety of levels among a hierarchy of units (for example, genes, individual organisms, and groups of organisms) and can generate adaptations at a variety of levels, referred to as multilevel selection. In order to identify the level at which selection acts, they outline a three-step procedure as follows (Wilson and Sober, 1998, p. 102-118). For simplicity, assume that there are two levels of units, individual organisms and groups of organisms. Step 1 is to determine what traits would evolve if selection acted purely at the level of groups. Step 2 is to determine what traits would evolve if selection acted purely at the level of individuals within groups. Step 3 is to determine whether Lewontin's three criteria for Darwinian selection are present at each level.

Darwinian selection of units at a particular level requires; 1) phenotypic variation of the trait among units at that level; 2) a correlation between the unit-level trait and reproductive success of those units; and 3) the unit-level trait must be inheritable. Wilson and Sober allow that resemblance between parent units and offspring units is sufficient evidence that the trait is inheritable.

Wilson and Sober use examples to illustrate that intrademic selection qualifies as "group selection" on the basis of heredity. Consider a locus on a diploid genome in an individual for the trait of interest. The population initially consists of two types of genes (A and a), each at a frequency of 50%. The phenotypes are expressed as follows: AA (height = 0.7), Aa (height = 0.5) and aa (height = 0.3). The pairs break up during meiosis after which they may be re-paired with another gene during sex. For each type of diploid parent unit, the expected frequency of each type of possible diploid offspring and therefore the average height of offspring can be calculated. For example, an AA parent produces an AA offspring with probability 0.5 and an Aa offspring with probability 0.5 for an average offspring height of 0.6. Similarly, the average height of the offspring of Aa and aa parents are respectively 0.5 and 0.4. Therefore, there is a correlation between the parent's height and the average height of the offspring. Since offspring groups resemble parent groups with respect to the trait of interest, Wilson and Sober conclude that the trait is inheritable at the level of the group.

As a second example, they consider groups of organisms. Organisms can be either tall (A) or short (a). Organisms interact in groups of four at some stage in their life cycle with the interaction affecting individual fitness. The groups of offspring then completely dissociate after which new groups are formed by random sampling from the population with replacement. The process is repeated. An AAAA parent group contributes one member to "offspring" groups of possible types AAAA, AAAa, AAaa, and Aaaa. Three members descend from other parent groups and one member descends from the initial parent group and must therefore be an A. Similarly, the possible "offspring" groups that "descend" from an Aaaa parent group are of types AAAA, AAAa, AAaa, Aaaa and aaaa. As in example 1, it can be illustrated that there is a correlation between the average height in parent groups and offspring groups. For example, to calculate the expected height of the "offspring" groups of an AAAA parent find the expected frequency of each type of offspring group conditional upon one member being an A-type and use these to weight the average height of each of the possible types of "offspring" groups.

Maynard Smith (1987a) has allowed that organisms and also the groups, in interdemic group selection, are units of Darwinian selection because all three of Lewontin's criteria are met, including resemblance as a basis for heredity. He did not include intrademic selection on the basis of a lack of group level heredity. In intrademic group selection, "offspring" groups do not descend from a common parent. However, organism "groups" do not descend from a common parent group either, as illustrated by example 1 above. Given that example 1 and 2 both illustrate the intrademic population structure, then if it is conceded that traits can evolve for the good of the diploid pairs in example 1 (or more broadly organisms), then by analogy, traits can evolve for the good of groups (example 2). By allowing organisms to qualify as Darwinian units of selection based upon resemblance, then Maynard Smith appears to open the door to including the groups of organisms in intrademic selection.

To guard against this, he notes that the diploid pairs (example 1) and intrademic groups consisting of groups of organisms (example 2) can be distinguished because there is not normally expected to be *within*-group selection in the former case. In diploid pairs, each gene normally has a 50% chance of being inherited by offspring and so genes in the diploid pair have an equal expected fitness and so are "in the same boat". In groups of organisms, in contrast, in the absence of perfect positive assortment, organisms often have different expected fitness and hence are subject to *within*-group selection. In spite of Maynard Smith's objections, example 2 of Wilson and Sober convincingly illustrates that the groups in intrademic selection have group-level heredity in the sense that offspring groups resemble the parent groups from which they were derived. Note, however, that if groups are formed randomly, as they increase in size, the degree of correlation between parent and offspring groups falls. If, however, groups are formed assortatively such that like types tend to group together, then resemblance could be high or even perfect.

According to Wilson and Sober, the groups in the above examples illustrate criteria 1 and 3 of Lewontin's characteristics for the Darwinian Selection of groups. What about the second criterion, that of correlation between a unit-level trait and the reproductive success of the unit of concern? In example 1, tall individuals might be fitter than medium and short individuals. Hence there is a correlation between the phenotypes (tall, medium, and short) and reproductive success and the diploid units AA, Aa and aa qualify as Darwinian units of selection. Note that *within* groups, the individual genes have equal expected fitness because meiosis is usually fair. Hence there is no within group selection and Wilson and Sober conclude that selection acts only upon diploid groups. The selection process is captured entirely in the first covariance term of the Price equation, while the second expectation term is 0 (equation 2-1, section 2.3.2).

In example 2, suppose that a higher average group height is advantageous from the perspective of groups, but that within groups, short individuals are fitter because they need to eat less. Using the three-step method of Wilson and Sober, tall genes should be selected if groups are the units of selection (step 1), while short genes should be selected if individuals are the unit of selection (step 2). The units at both levels satisfy all three of

Lewontin's criteria for Darwinian selection of a unit (step 3). The level that dominates would then need to be determined empirically.[34,35]

Example 1 differs from example 2 with respect to Lewontin's second criterion. In example 1, within groups, sub-units have identical fitness, hence, they are "in the same boat" because meiosis is usually fair. In example 2, sub-unit organisms have different fitness depending upon their height, hence, sub-units are not "in the same boat". Yet the difference between the examples is only a matter of degree when viewed from the perspective of the Price equation. If the between group covariance term outweighs the within group expectation term, then, between group selection for tall groups outweighs within group selection for short individuals. If resemblance, verses perfect copying fidelity, is accepted as a sufficient criterion for the hereditary basis of a Darwinian selection process, then, the groups in both of Wilson and Sober's examples do qualify as Darwinian units of selection.

It can now be understood why Wilson and Sober have argued that replicators are irrelevant to the question of whether groups qualify as units of selection. They did not need to refer to the underlying gene replicators, given that they rely upon the weak criterion of resemblance between parents and offspring as a sufficient hereditary basis for a Darwinian selection process.

In contrast, the approach of Williams (1966) and Dawkins (1976) requires a much stronger basis of heredity in that in order for a unit to qualify as a Darwinian unit of

---

[34] Consider Williams' example of fast deer (section 1). In this case, if selection acted purely at the level of groups, fast deer would be selected (step 1). Similarly if selection acted purely at the level of organisms, fast organisms would be selected. However, Wilson and Sober agree that the selection of the speed trait can be attributed to competition between and selection of individuals and not to selection acting at the higher level of groups. Genes for grouping are neutral with respect to the speed trait (by assumption). The method of contextual analysis allows for further decomposition of the Price equation and will not be considered in this essay (for a review of this topic, see Goodnight, Schwartz, and Stevens, 1992).

[35] If it is empirically determined that group height suggests selection at the level of groups, then based upon Hamilton (1975), group formation cannot be random but must be assortative.

selection, the unit must be a replicator. Units larger than "genes" (in Williams' sense) do not qualify due to insufficient copying fidelity. Example 1 can be modeled using the genic approach as follows. Given the assumption that the initial population wide frequency of each type of gene is 50%, the A-gene has an equal probability of being in either an AA or an Aa group, while the a-gene has an equal probability of being in either an aa or an Aa group. A-genes find themselves in tall and medium organisms, while a-genes find themselves in medium and short organisms. If height increases the reproductive success of organisms, then A-genes will be selected to reproduce more often than a-genes and therefore A-genes should increase in frequency in the gene pool. In this view, the group level traits in examples 1 and 2 are both examples of emergent phenomena.

This gene-centered approach has the virtue of being simple, but more importantly, the modern understanding of the method of transmission of hereditary information has allowed biologists to be precise about exactly what units are transmitted. While Darwin was obliged to rely upon resemblance, modern biologists are not. Wilson and Sober (and also Maynard Smith) can hence be criticized due to depending upon a weak indicator of heredity (that of resemblance), when more precise methods of describing the process of heredity are available.

Wilson and Sober also argue that this game theoretic/ gene centered approach is subject to what they refer to as the "averaging fallacy". They argue that gene A is only increasing in frequency as the outcome (or product) of a process that is occurring at the higher diploid-locus-unit level. For example, with respect to example 1 above, they argue that AA, Aa, and aa respectively *cause* the phenotypes of "tall", "medium", and "short", where these phenotypes are the relevant traits upon which selection acts because they are not "context dependent". To contrast, they argue that it is incorrect to attribute causation to a single gene because the phenotypic expression of that gene is context dependent. For example, an A-gene in the group AA codes for "tall" while in the group Aa codes for "medium". In agreement with Maynard Smith's (1987b) criticism of Sober (1987a,b), I do not see what is wrong with saying that an A-gene *causes* "tall" *if* it is grouped with another A-gene, while it *causes* "medium" *if* grouped with an a-gene. Furthermore, if single genes are to be excluded from qualifying as Darwinian units on the basis of context dependence, then AA pairs should also be excluded. The phenotypic expression of the AA pair depends upon a host of other environmental factors, including other genes on the genome and the food resources available (Maynard Smith; 1987b). This causation argument of Wilson and Sober is difficult to understand and will not be pursued further here.[36]

It is clear from the discussion of Wilson and Sober that there is still a great deal of controversy about which units can evolve by a Darwinian process. The goal of the remainder of this essay is to suggest a means for resolving this controversy. In spite of criticizing Wilson and Sober for depending upon the weak criteria of resemblance to provide a basis of heredity, there is often marked group level functionality at levels higher than genes. Moreover, given that offspring organisms, and sometimes, groups resemble each other, then surely there is a hereditary basis for such resemblance and therefore for this group level functionality. Is it possible to be more precise about this hereditary basis? In the remainder of this essay, it is argued that the resemblance between offspring and parent units is due to underlying unit-level replicators, which exist at levels higher than genes. In short, Dawkins' replicator approach is used to provide a basis for Wilson and Sober's theory of multilevel selection.

3 A Hierarchy of Replicators: A Proposed Solution to the Group Selection Debate

It is broadly accepted among biologists that the largest organic replicators correspond to Williams' genes.[37] There have then been two extreme ways to explain functionality at the level of organisms and groups. Firstly, as an emergent phenomenon due to selection acting upon genes (Williams, 1966; Dawkins, 1976, 1982) and secondly as an adaptation

---

[36] The "context dependence" argument is used by Lewontin and Sober (1982), Sober (1984, 1987a,b), Gould (2002), and Wilson and Sober (1994, 1998).

[37] Some biologists such as Griffiths and Grey (1994) have argued that vehicles qualify as replicators. Dawkins (1982, p. 97) rejects this argument by arguing that mutations in the vehicle need not be passed on to offspring.

due to selection acting upon vehicles, whereby resemblance provides evidence of a basis for heredity (Wilson and Sober, 1994, 1998).

Similar to Dawkins, I argue that functionality at the level of a unit is not an adaptation that evolved by a Darwinian process unless there is hereditary information underlying that unit. While Wilson and Sober allow that resemblance between offspring and parent units provides evidence of heredity at various levels, including organisms and groups of organisms, I argue that it is possible to be more precise by making use of the replicator criteria. In the following section, it is argued that there are replicators at the level of organisms, and groups in both intrademic and interdemic group selection. Higher level replicators are constructed from lower level replicators by means of inheritable assortative strategies. Higher level replicators are then both emergent phenomena *and* replicators in their own right. Furthermore, for each replicator, an expected vehicle can be identified.

3.1 Inherited Assortative Strategies and Pleiotropy

In the cases of intrademic selection (including kin selection) and reciprocal altruism, the evolution of altruism depends upon some form of positive assortment. Furthermore, as illustrated by Hamilton (1975), in order for altruism to evolve, such positive assortment must be non-random. Under the game theoretic approach, genes for phenotypic traits such as altruism, free riding and eye-colour are referred to as "strategies for actions". Similarly, it can be useful to view the actions of grouping and assortment as being due to strategies for expected group size and composition. If the strategies for grouping are not played at random *with respect to the hereditary information* that specifies the strategies for actions such as altruism, it follows that they are inherited.[38]

---

[38] If strategies for grouping are not played at random with respect to the heredity information, then they are correlated with the hereditary information. In this case, the hereditary information contains the information for grouping, i.e. the grouping strategies. Hence the grouping strategies are inheritable.

To illustrate, consider the case of siblings in kin selection, a special case of intrademic selection. Recall that sibling interact in sibling-groups, after which they disperse into a migrant pool, mate and reproduce to produce a new sibling group. Siblings can play either the altruistic or free riding strategy as coded in their genes. Altruists positively assort with altruists due to descent from common parents. For example, if a sibling is an altruist, there is a 50% chance that each sibling is also an altruist (see section 2.2.3). The action of assorting with siblings can be viewed as an assortative strategy that is *linked* with the gene for altruism. Genes that code for more than one action are said to illustrate pleiotropy. The assortative strategy for an altruist might read "locate in a group of siblings in which each member is expected to be an altruist with a probability ½". Similarly the assortative strategy for a free rider might read "locate in a group of free riding siblings".

The strategy for the expected sibling group size is encoded in the parents' genes and is passed on to offspring. With respect to the assortative strategy, the strategy could be very simple such as "help all small birds in the same nest". The location of the nest is inherited by all siblings due to genes that direct parents to lay all of their eggs in the same nest. The fact that siblings remain in the nest is due to genes for immobility.

Genes allowing for positive assortment can also be due to recognition labels. Dawkins (1982, p. 143-155) considers means by which genes might "recognize" identical genes in other organisms and distinguishes between two types of labels, "armpit effects" and "green-beard effects". In the case of "armpit effects" both the recognition label and the trait for altruism are inherited. The gene for the recognition label and the altruistic action may be located in close proximity on a chromosome so that they tend to be inherited together. Suppose that all siblings inherit a similar smell from parents. A strategy such as "help all individuals which smell like yourself" could serve to direct help to identical genes in siblings. Although a free rider that sports the appropriate smell may receive help without helping in return, giving him the highest possible fitness, he is less likely than an altruist to have altruistic siblings that help like smelling individuals.

Dawkins' "green-beard" effect is more general. Unlike the above "armpit effect", a green-beard gene does not instruct you to inspect yourself for a recognition label and help others with the same label but merely to help others with a particular label, such as a green beard. If all altruists have green beards but free riders do not and if altruists only help individuals with green beards, then altruists can avoid helping free riders allowing for the selection of altruism genes. However, if the genes in an altruistic green-bearded individual mutate so that he becomes a green-bearded free rider, or if the genes in a beardless free rider mutate so that the free rider comes to sport a green beard, he would receive help without incurring the costs of helping. Green-bearded free riders would therefore take over the population and the signal would loose its value. Unlike the armpit effect, linkage between the altruism and green beard gene is not maintained by descent. Dawkins (1982, p. 145) writes:

> Genes are not conscious little devils, able to recognize copies of themselves in other individuals and to act accordingly. The only way for the green beard effect to arise is by incidental pleiotropy. A mutation must arise which just happens to have two complementary effects: the label or the 'green beard', and the tendency to behave altruistically towards like labeled individuals. I have always thought such a fortuitous conjunction of pleiotropic effects too good to be true.

It is interesting that pleiotropic effects are only made explicit in examples involving recognition labels such as green beards or armpits. Such recognition labels are a subclass of assortative strategies. "Location strategies" are also ultimately based upon genes and can also be inherited and linked to specific traits, thereby also qualifying as pleiotropic effects. A group at a particular location can be viewed as consisting of members who have strategies for grouping at that location. If a group by chance has a high proportion of altruists, analogous to the green beard case, it can be invaded from within, due to the mutation of member's strategies from altruism to free riding or from without, due to mutation of outsiders' "location strategies". The possibility for assortment of altruists by spatial means has been the subject of much work by biologists. Similar to the green beard case, getting genes for altruism to become linked to location strategies is not necessarily easy. For example, Wilson, Dugatin and Pollock (1997) find that while high population viscosity serves to facilitate the evolution of altruism due to the possibility of concentrating altruists by descent, it also serves to prevent the "export of altruists" by group fission.

In intrademic and interdemic models, strategies for group formation and group fission may be thought of as the result of assortative strategies, which are often based upon location. These grouping and assortative strategies may have co-evolved with altruism. For example, location serves as the mode of positive assortment in Maynard Smith's haystack model, which is a variant of the standard kin selection model. Females found colonies of offspring in isolated haystacks after which offspring reproduce for several generations before dispersing. Within a given haystack free riders are relatively fit, but colonies with more altruists grow faster and are able to export more progeny to the migrant pool during the dispersal phase. The process is repeated. If the period spent within haystacks is too long then free rider mutants take over within each haystack. However, as in Haldane's model, altruistic genes can increase at the level of the population if groups disperse before free rider mutants become established.

In the Haystack model, the assortative strategy might read: "Mate in the same location in which you were born unless the density of the haystack reaches a critical level. In this latter case, leave the haystack, and if you are female, mate and then find a new isolated haystack." Although, the haystack population structure is exogenous to the model, natural selection could potentially lead to the linkage of assortative and altruistic strategies. In spite of Maynard Smith's early pessimism about the likelihood of altruism evolving under the haystack model, other researchers have illustrated that his initial parameter specifications were too stringent (Wilson, 1987; Wynne-Edwards, 1964).[39]

Although assortative grouping is an important component of the population structures that facilitate the selection of altruism, such grouping is not usually thought of as being

---

[39] Note that it is standard for economists to view location as a strategy. In fact the concept of a spatial address is often used in a very abstract sense to denote some frequency of a particular characteristic which may not literally refer to spatial location (Hotelling, 1929).

the result of hereditary strategies for assortment. To contrast, it is standard to regard phenotypes, such as altruism, as "actions" coded for by hereditary strategies. Hamilton's (1975) model drew attention to the requirement of non-random assortment for the evolution of altruism. If such assortment is non-random with respect to the hereditary strategy for altruism, it must be due to inherited assortative strategies (see footnote 38). In order for altruism to evolve, these assortative strategies must be linked to the gene for altruism such that there is pleiotropy. In the following sections, it is argued that inheritable assortative strategies allow lower-level replicators to construct higher level replicators such that units at levels higher than genes can evolve by Darwinian selection. As a starting point, replicators are identified at the level of organisms.

3.2 Organism Level Replicators

The hereditary information for an individual organism is the genome. For sexually reproducing organisms, it is generally accepted that genomes do not qualify as replicators due to fragmenting by meiosis. Hence, the genic camp, represented by Dawkins and Williams, has argued that organisms are not subject to Darwinian selection, but are emergent phenomena. At the other extreme, Wilson and Sober have argued that adaptations can evolve for the good of organism vehicles, and that organism level heredity is illustrated by the resemblance of parents and offspring. Maynard Smith (1987a) also allows that organisms are subject to Darwinian selection based upon resemblance. Both Dawkins (1982) and Wilson and Sober (1994) have argued that replicators are irrelevant to the units of selection debate. In contrast, I argue that replicators are relevant; I start out by making the case for organism level replicators.

3.2.1 The Ascension of the Genic Approach

For the most part, Darwin wrote of organism level adaptations, based upon resemblance between offspring and parents (Hamilton, 1975). His vagueness regarding heredity was prudent due to a lack of knowledge regarding the nature of the hereditary information. Nevertheless, in the 19th century, particulate factors of inheritance had already been hypothesized.

> Although given a variety of names (physiological units, pangens, biophors, character units), the idea that inherited phenotypic characteristics are determined in some way by distinct elements, now equated with "genes", was well established by the end of the nineteenth century (Forsdyke, D.; 2001).

Given that atoms were providing an explanation for chemical phenomena, it is not surprising that biologists should have hypothesized a discrete particulate basis for genetic traits. Hence, there was much excitement in 1900 when the rediscovery of Mendel's work provided evidence that discrete genetic elements were inherited intact by the next generation. There were plenty of obvious experiments to do and geneticists set out to determine the rules of transmission.

Going against this "genic" mainstream, Bateson (1914), proposed that it was unlikely that the units of inheritance were simply particulate, but must be arranged in some manner. Such an arrangement would qualify as hereditary information, which would be inherited. Bateson postulated that the hereditary information consisted of a "base" to which the independently assorting "transferables" (genes) were "attached" (Forsdyke, D., 2001, p. 67-68).

Bateson's ideas were largely ignored and the genic approach gained momentum. Fisher's (1930/1958) *Genetical Theory of Natural Selection*, integrated the particulate genetic elements into Darwinian theory.[40] In 1953, the "beads on a string" chemical structure of DNA was discovered by X-ray crystallography. This structure supported Bateson's earlier proposal, with the "base" corresponding to the "string" and the "beads" to the

---

[40] In addition, the discovery of the particulate hereditary elements provided a solution to the problem of the elimination of variation due to blending of parental traits in offspring (Fisher, 1930). For example, suppose that at a diploid locus on the offspring's DNA is a short gene and a tall gene, each of which is inherited from a different parent. Although, the phenotypic expression may blend to produce an offspring of medium height, the genes do not blend. The offspring is expected to pass tall and short genes with equal probability to each of its own offspring, not medium genes. If instead, the genes blended in offspring, impossibly high and disadvantageous mutation rates would be required in order to maintain enough variety for natural selection.

"transferables" (genes). However, despite this discovery, geneticists continued to focus upon the "transferables", which conveniently segregated independently making them amenable to statistical analysis and mathematical modeling. Meanwhile, the field of molecular biology developed at a rapid pace providing further support for the "beads on a string" primary structure of the DNA. Furthermore, it was discovered that DNA has complex higher levels of structure. Molecular biologists discovered that only a fraction of the DNA in most species coded for known phenotypic traits, the exons, while the rest of the DNA, the introns and intergenic DNA, had an unknown purpose.

It was within this context that Williams (1966) proposed that the unit of selection was the "gene", the largest fragment of DNA likely to survive meiosis intact. Due to the splintering effects of meiosis, the genetic information for the organism as a whole did not appear to qualify as a unit of Darwinian selection. In the following section, it is argued that the understandable and probably prudent focus on genes, (given, for example, that increased understanding of genes had led to huge advances in medical science), led evolutionary biologists to pay comparatively little attention to the broader structure of the DNA. An alternate trajectory of evolutionary thought flowing from the insights of Bateson might, however, have been pursued. The pursuit of this trajectory leads to the identification of organism level replicators allowing for organism level adaptations.

3.2.2 Meiosis as a Preserver of Similarity

Although genes are shuffled during meiosis, much of the overall informational structure of the genome is conserved. For example, the order and spacing of genes is most usually conserved allowing for an orderly process of shuffling between specific similar genes competing for the same locus. Without such orderly shuffling, some individuals would have excessive numbers of particular genes and not enough of others. Due to interspecies barriers, only genes from within a species can compete for loci on genomes within a species. Mutation aside, competition for a specific locus is limited to those genes that have proven to work well at this locus in the past. Hence, not only is meiosis highly ordered, but the pool of available "transferables" (genes) is limited. Meiosis can hence be regarded as a very orderly process of exchanging genes that have a proven track record in genomes of a particular type of species. Other characteristics of the DNA, such as the overall frequency of the base-pair composition, are also conserved within a species (Forsdyke, D., 2001). Forsdyke argues that the species specific base pair frequency of the DNA, which he refers to as secondary information, may serve to prevent recombination between members of different species or limit recombination between incipient species.

Recombination during meiosis allows for the orderly shuffling of genes providing a basis of variation beyond that available purely due to mutation. However, this may not have been the original role of meiosis (Bernstein et al., 1985; Forsdyke, D., 2001). Meiosis or sex may have evolved very early on in the primordial soup as a means of replicator repair.

Recall that the three properties that determine the success of a replicator are longevity (stability between rounds of replication), copying fidelity and rate of replication. High copying fidelity allows for successful variation to be passed on to offspring (section 2.2.1). Yet, if copying fidelity is too perfect, evolution will eventually come to a halt because of a lack of new sources of variability. Hence, high copying fidelity is required but not too high. Over time, there should be selection pressure for optimal rates of copying fidelity. The early replicators described in chapter 2 of Dawkins (1976) likely had poor copying fidelity, only needing to do slightly better than average in order to increase in frequency in the replicator pool. Hence, for the early replicators, it is likely that there was selective pressure for adaptations for improvements in copying fidelity.

As replicators increased in size via such processes as chance duplication of existing segments, given that the new duplicated sections would inherit the base-pair composition of the older segments, then it follows that the genome as a whole may have had a similar base pair composition. If base pair composition was a globally conserved feature of the genome, this may have allowed for detection and repair of damages. Damaged regions identified by different base pair frequencies could be spliced out and repaired using another part of the genome as a template. Later, the replicators may have further evolved to be able to compare themselves with other replicators in the species in order to splice out and repair damaged regions. The base pair composition could serve as a means by which species members recognized each other in order to prevent repair attempts with members of the wrong species, which would be unlikely to result in improvement. Hence, sex (meiosis) may have evolved early on as a means of preserving the similarity of genomes, (i.e. the base-pair composition, order and location of specific genes, and type of genes "allowed" at a given locus), within a species.

To sum up, meiosis allows for the orderly shuffling of combinations of genes, which have already proven to be successful within a species. Organisms that do not shuffle genes in an orderly manner are less fit. Similarly, interspecies barriers are adaptive, the children of monkey-human hybrids are likely to be less fit. The information regarding the orderly process of meiosis is an inherited global feature of genomes that is adaptive. Part of this hereditary information is the order and spacing of specific loci, and a species-specific pool of "transferables" (genes) due to between species mating and breeding barriers. New theories into the role of meiosis as a means of DNA repair suggest that sex serves to preserve global features of the genome in addition to allowing for recombination. Given this perspective, meiosis appears to act more to preserve similarity of genomes than to fragment it. However the orthodox emphasis on the "fragmenting" effects of meiosis has placed emphasis on the *differences* produced in genomes during meiosis, not the *similarities*. In turn, genomes, the genetic information of the organism in its entirety, have been disqualified as units of selection. The alternative perspective stemming from the insights of Bateson, suggests that perhaps genomes have sufficiently high copying fidelity to qualify as replicators, units for which adaptations can evolve for the good of.

3.2.3 Meiosis as a Preserver of Similarity and Dawkins' Rowing Boat Analogy

Dawkins' rowing boat analogy can be revisited in the context of meiosis as a preserver of similarity. As in section 2.2.6.3, coaches race alternate combinations of rowers to find the fastest combination. At the end of the race, the best boats are lined up in twos and rowers are shuffled between boats. Suppose that rowers tend to row best in a particular seat. Stroke will not do well in the bow and vice versa. Starboard rowers can only row well on starboard and the same with port. If coaches are smart, they will shuffle seats in an orderly manner, not at random. Such coaches will avoid putting a good port rower on starboard or a good stroke in bow. Crews with smart coaches will be selected and so will these coaches. Over time the pool of rowers, corresponding to genes or "transferables", should increase in quality and the method of shuffling should become very orderly.[41]

Analogous to species, there might be two sorts of boats, skulls (where each rower has two oars) and sweep (where each rower has one oar). If rowers are only good at skull or sweep but not both, then its should be advantageous if sweep and skull seat races are held independently. Skull rowers could then avoid "recombining" with sweep rowers, allowing for the selection of better crews than if shuffling occurs independently of preferred oar type. Furthermore, the pool of rowers are selected out of groups of people who have previously illustrated ability as rowers, not from overweight spectators standing on the shore. Information regarding the types of rowers that are allowed in particular seats, the type of rower that can row in a sweep or skull, whether a starboard rower can sit in a port seat, the boat's physical structure, and the set of people from which the rowers are drawn, is inherited by subsequent crews. These structural components of boats will tend to be selected in addition to the selection of individual rowers. The structural information is inherited as a unit that is not fragmented during seat swapping. The selection of individual rowers is only one component of the selection process.

By analogy, meiosis can be viewed as a process that preserves similarity and also as a method that generates variability due to shuffling specific genes. As global features of the genome are conserved during meiosis, and as these are inheritable information, then the structure of the genome qualifies as a replicator. Each locus (seat) can be thought of as a placeholder for the independently assorting genes, which are also replicators.[42]

---

[41] Note that in Dawkins' original rowing boat discussion, he allows for seat specific shuffling, however he does not discuss the structural hereditary information required for such seat specific shuffling.

[42] Note that in humans, less than 2% of the genome consists of exons (sections of DNA that code for phenotypic traits.

3.2.4 Entity Level Adaptations Require Entity Level Hereditary Information

Before more precisely developing an argument for genome level replicators, I want to stress that when looking for adaptations at levels higher than the gene, in my view, the correct approach is to look for more inclusive levels in the domain of hereditary information. Dawkins (1976, 1982) and Williams (1966, 1994) argue that the incorrect approach is to argue that the organism vehicles or putative group vehicles are adaptive units. Hull (1980) considers the possibility that genomes could serve as replicators and concludes that it is possible if populations are sufficiently homogeneous so that crossover does not change the genome too much.

> Recombination has no effect on the structure of genomes in genetically homogeneous populations. However, the more heterogeneous a population is, the more likely that the structure of genomes will be altered by crossover during meiosis. How much alteration can occur before a genome must be considered a new replicator depends on the effect the changes have on the organism as an interactor.[43] How similar is similar enough? Similar enough to respond similarly to similar selection pressures.

Hull focuses upon the hereditary domain but only considers genes while ignoring broader structural aspects of the genome. His concept of a replicator differs from that of Dawkins in the requirement of functional similarity between copies. In contrast to Hull, others who have argued that organisms (in the sense of vehicles) can be replicators have opened themselves up to the charges of violating the "central dogma" by invoking the inheritance of acquired characteristics. Dawkins (1982: 97) writes:

> I have made much of the fragmenting effects of meiosis as a reason for not regarding sexually reproduced organisms as replicators, that it is tempting to see this as the only reason. If this were true, it should follow that asexually reproduced organisms are true replicators, and that where reproduction is asexual we could legitimately speak of adaptations as "for the good of the organism". But the fragmenting effect of meiosis is not the only reason for denying that organisms are true replicators. There is a more fundamental reason, and this applies to asexual organisms as much a sexual ones. To regard an organism as a replicator, even an asexual organism, like a female stick insect, is tantamount to a violation of the 'central dogma' of the non-inheritance of acquired characteristics.

---

[43] Note that Hull's "interactors" are analogous to Dawkins' "vehicles".

In this view, organisms are the aggregate outcome of the phenotypic actions coded for by the individual passenger gene replicators. Organisms (in the sense of vehicles) are clearly not replicators as they do not make copies of themselves, "blemishes and all". If a stick insect looses a leg, this acquired characteristic is not passed on to descendents. This argument does not however disqualify the *hereditary information* of organisms when taken as a whole, i.e. genomes, as replicators. In contrast to what Dawkins appears to say here, the splintering by meiosis argument *is* fundamental to both his and Williams' argument to disqualify genomes as replicators.

3.2.5 An Argument for Organism Level Replicators

The following section utilizes Bateson's concept of a "base" to which are attached independently assorting "transferables" (genes), in order to develop a model of organism level replicators, which in turn provide a basis for organism level adaptations (see section 3.2.1). As the structure of the base is conserved though meiosis, it qualifies as a replicator as do the genes. Imagine a hypothetical simplistic diploid species. Each member of the species has two haploid genomes, each to be referred to as G, upon each of which there are only two loci labeled A and B. Each of these loci is a "placeholder" for two possible genes, where gene-A1 or gene-A2 may locate at locus A and gene-B1 or gene-B2 may locate at locus B.

There are four possible combinations for each haploid genome, which are represented as A1-B1, A1-B2, A2-B1, and A2-B2. Four of the possible diploid combinations are represented below.

A1-B1	A1-B2	A2-B2	A1-B1
A2-B1	A1-B1	A2-B2	A2-B2

The gene replicator A1 has information for two strategies. Firstly, A1 codes for the standard phenotype, such as altruism or eye-colour and secondly for an assortative strategy. The assortative strategy might read "locate at locus A on a genome of type G". This second strategy allows the gene replicator A1 to inherit a copying apparatus on a genome and thus the ability to make copies of itself, a requirement of Dawkins' original replicators (1976, chapter 2). Recall that a strategy for self-copying was not explicitly noted as a requirement of the "gene" replicators defined in chapter 3 (Dawkins, 1976). With regards to the chapter 3 definition of a replicator; "...anything in the universe of which copies are made", it is apparent that unless individual genes have assortative strategies to associate with genomes, then copies will not be made and these genes would fail to qualify as replicators. It may be argued that the fact that the genes inherit a copying apparatus is obvious—it is clear that copies are made of genes and so we don't need to specify a special assortative strategy. Nevertheless, there is an assortative strategy and it is inherited; being explicit about the assortative strategy restores Dawkins' initial chapter 2 sense of the word replicator. The gene replicators A1, A2, B1, and B2 illustrate pleiotropy in that they each code for both the standard phenotype and the assortative strategy.

A fifth replicator, G, to be referred to as the "haploid genome-base replicator" also illustrates pleiotropy. G represents the haploid genome, with "place-holders" locus A and locus B, which specify that locus A can be occupied by either gene A1 or gene A2, and similarly for locus B. G contains information for inheritable strategies for assortment at each respective locus. For example, one strategy might read "at locus A assort with a strategy drawn randomly from the population of available gene replicators A1 or A2". There is a similar assortative strategy for locus B. In the case of diploid cells, there is also an assortative strategy that "instructs" each haploid genome-base replicator to pair with another haploid genome base replicator of the same type in one cell, allowing for the formation of a diploid genome-base replicator of form G-G.

Lastly, assortative strategies for spatial location may be inherited due to genes for limited mobility. Such location strategies are adaptive as they help ensure that the replicator inherits a favorable environment required for successful copying. If a location strategy for a tropical plant seed "mutates" due to unusual winds that blow it to the arctic, these replicators will not be copied. For simplicity, location strategies are ignored in the following section.

When assortative strategies are explicitly accounted for, each of the above-described replicator types satisfies the criterion of a Darwinian unit of selection according to Lewontin's criterion. For example, in the case of the genome-base replicator, there may be structural variants both within and between species and such variants may have differential rates of survival and reproduction (criteria 1 and 2). Both the haploid and diploid forms have high copying fidelity (criterion 3). Genome-base replicators most usually produce perfect copies with respect to structural functional features such as the order of specific loci, which serve as placeholders to be occupied only by members of a limited set of genes.[44] Similarly, the lower-level gene replicators also have high copying fidelity with respect to both the standard phenotypic strategy and the assortative strategy. Due to high copying fidelity, the genome-base and gene replicators satisfy not only Lewontin's hereditary resemblance criterion for Darwinian units of selection but also the more stringent replicator criterion (section 2.2.1).[45]

Hence, the hereditary information of organisms consists of both gene level replicators and organism level replicators each of which can be precisely related to an expected vehicle at the level of the organism. Consider the gene replicator A1, which codes for both a standard phenotype and the assortative strategy to pair with a genome-base replicator at the A-locus. This assortative strategy allows A1 to utilize a specific organism level vehicle in order to get itself copied.[46] This vehicle includes the phenotypic expression of the A1 gene, the expected phenotypic value for the A gene at the other locus ($E[A]$) and the expected phenotypic value for each of the genes at the B locus ($E[B]$). The expected vehicle for the A1 replicator might be represented as:

---

[44] Note that the functional structural basis of the genome base replicator most usually remains intact throughout meiosis, even though there will be changes in the nucleotide sequence. In this functional-structural sense, the genome base replicator is copied perfectly.

[45] It may be argued that the diploid replicators (G-G) fail to qualify as Darwinian units of selection because they lack descent from a common parent. However, the G—G replicator is able to construct a copy of itself in the next generation by means of the inherited assortative strategies; hence it has heredity (Lewontin's third criteria).

[46] Higher level vehicles will be considered in later sections. Note, also, that location strategies can allow the organism level and gene replicators to use the environment as a vehicle.

$A1 \quad - E[B]$

$E[A] - E[B]$

The expected vehicle for a haploid genome-base replicator and a diploid genome-base replicator can similarly be represented as:

$E[A] - E[B]$

$E[A] - E[B]$

The expected vehicle for the haploid and diploid genome-base replicators corresponds to the aggregate phenotype for a typical organism in the species. Note that the haploid and diploid genome-base replicators have the same expected vehicle because the haploid replicators have assortative strategies to pair.

Hence, there is a one-to-one relationship between each replicator at both the gene and organism level and an expected vehicle. Recall that in chapter 2 of Dawkins (1976), there is also a one-to-one relationship between replicators and vehicles. Replicators were envisioned as long strands of DNA (or RNA) that coded for their vehicles. Given that the fragmenting process of meiosis is not discussed until chapter 3, each replicator genome could pass on the information to offspring genome replicators allowing them to reconstruct the same expected vehicle. In chapter 3, this precise one-to-one relationship between vehicles and replicators is lost when replicators are redefined as Williams' genes. These genome fragments are no longer able to replicate themselves unless they assort with a genome. I have argued that it is useful to be explicit about the inheritable assortative strategies that allow genes to assort with genomes, which allows gene-replicators to code for a precise expected vehicle and also to pass on to descendents the information to re-construct the same expected vehicle.

In contrast to chapter 2 usage, in chapter 3 of Dawkins the term "vehicle" becomes somewhat vague. For example, the definition of a vehicle in the glossary of *The Extended Phenotype* (1982) is: "...any relatively discrete entity, such as an individual organism, which houses replicators (q.v.), and which can be regarded as a machine programmed to preserve and propagate the replicators that ride inside it." Making assortative strategies explicit reintroduces the chapter 2 implicit precision in the relationship between replicators and vehicles. Through inheritable assortative strategies, both gene-replicators and genome-base replicators code for expected vehicles, tools by which they copy themselves.

Note also that lower level replicators can construct higher level replicators by way of the assortative strategies. For example, the A-gene replicator has an inheritable strategy to assort at an A-locus on a diploid genome-base replicator allowing for the construction of a larger replicator that consists of an A-gene and the diploid-base replicator. "Offspring" of the A-gene replicator inherit this assortative ability and therefore are able to reconstruct this larger replicator, which has therefore been copied. The information to copy this larger replicator is contained in the two smaller replicators and therefore it is an emergent phenomena. However, it is also a replicator in its own right because it is perfectly copied.

3.2.6 Summary

Maynard Smith (1987a) and Wilson and Sober (1994, 1998) argue that organisms are subject to Darwinian selection because they satisfy Lewontin's three criteria, with resemblance between parents and offspring illustrating the third criteria of heredity. In contrast, Williams (1966) and Dawkins (1976, 1982) have suggested that genes are the largest units of Darwinian selection because genomes lack sufficiently high copying fidelity due to splintering by meiosis and therefore do not qualify as replicators. I have argued that Lewontin's hereditary criteria, in this stricter replicator form, can be applied at the level of organisms. Bateson's perspective drew attention to the fact that the process of meiosis leaves the overall structural features of the genome intact. A genome level replicator, consisting of structural features such as the order of genes, was identified and referred to as the "genome-base replicator".

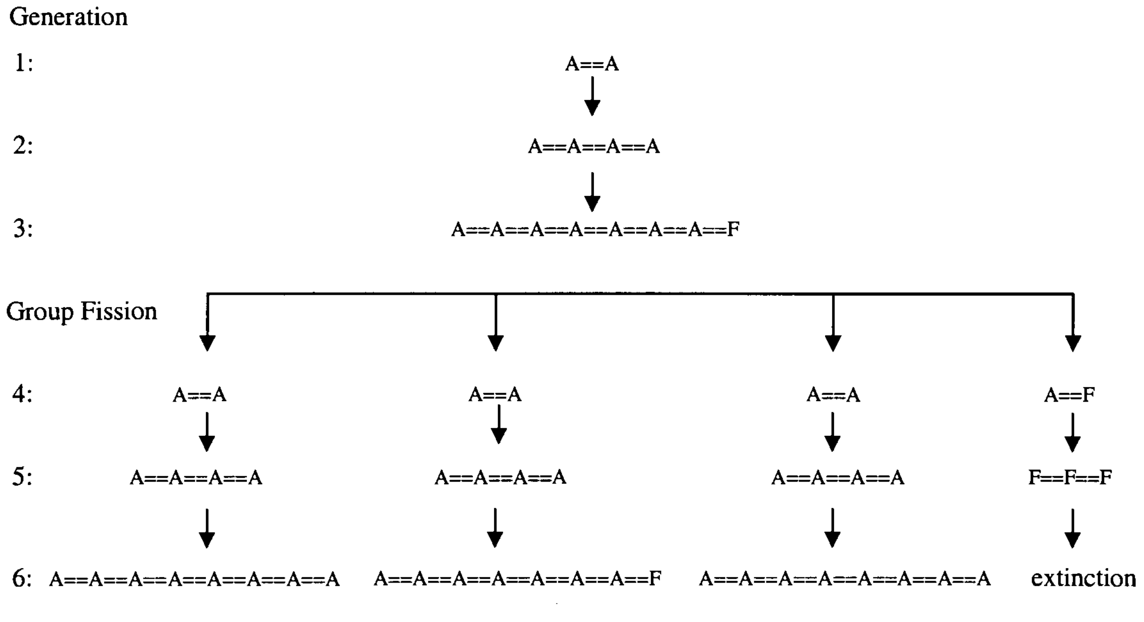
Secondly, it was argued that "genes" in the sense of "genome fragments" are not replicators unless located on a genome. Genes on a genome illustrate pleiotropy in that they code for both the standard phenotypic actions and a strategy to locate at a specific position on a particular sort of genome. These latter strategies were referred to as "assortative strategies". The assortative strategies allow gene replicators to assort with genome-base replicators and to pass this information on to offspring. The corresponding vehicle for a particular gene-replicator is the standard phenotype plus the expected aggregate phenotype of the organism. Offspring gene replicators contain the ability to construct the same expected vehicle as their parents. The genome-base replicator similarly has assortative strategies to group with particular types of genes at particular locations. The corresponding vehicle is the expected aggregate phenotype of the species. Wilson and Sober and Maynard Smith have allowed that due to resemblance between parent and offspring vehicles, organisms satisfy the hereditary criteria to qualify as units of selection. The above section has attempted to illustrate that there is a precise replicator basis for this resemblance. Lastly, it was argued that higher level replicators could be "constructed" from lower level replicators via inheritable assortative strategies and are therefore are emergent phenomena, however this does not disqualify them from being replicators in their own right.

3.3 Group Level Replicators in Interdemic Models of Group Selection

Traditional models of group selection, such as Haldane's model in which groups fission to produce offspring groups, were later categorized as interdemic group selection in order to distinguish them from the intrademic models. Interdemic selection can be categorized as "group selection" of the Darwinian sort according to Lewontin's three criteria. Firstly, groups vary according to some group-level characteristic. Secondly, this group level characteristic determines how fast groups grow and produce propagules, allowing for differential selection between groups. Thirdly, groups have heredity in the sense that offspring groups inherit the group level characteristic of the parent groups. For example if propagule groups are produced by randomly selecting a sub-set of individuals from the parent group, then this offspring group is expected to have the same frequency of altruists as the parent group.

In the following section a simple model is developed in order to illustrate that in the case of interdemic group selection, inherited assortative strategies enable organism level replicators to aggregate to produce group level replicators, which code for group level vehicles. It will be illustrated that the lower-level gene and organism-level replicators also use the group as a vehicle. For simplicity it is assumed that organisms reproduce asexually although the model can easily be adapted to accommodate sexual organisms. Let A= and F= represent organism-level replicators for altruistic and free-riding types. In the example, individual organism altruists and altruistic genes incur fitness costs at a benefit to the group. For example, individuals might incur costs due to secreting a group beneficial chemical. Let the symbol '=' represent an assortative strategy for an organism to remain with the other organisms that were present at birth. The assortative strategy also instructs individuals to form propagules (offspring groups) of size two if the parent group grows to reach a size of eight members. Figure 6 illustrates the process.

Figure 6: A Hypothetical Interdemic Selection Process



In generation 1, there is one group with two altruists. Each altruist reproduces to produce two offspring so that by generation 2 there are four offspring, which are assumed to be altruists. After reproducing each parent organism dies while the offspring remain together. By generation 3, each parent altruist has produced two offspring for a total of eight in the offspring group. Due to one chance mutation, this group contains one free rider. Given that there are now eight members in the group, group fission occurs to create four propagules, each with two members. Three of the four new groups contain two altruists and reproduce so that each group contains four altruists by the 5th generation. These groups of four reproduce and it is assumed that a mutation to free riding occurs in one of the groups. By generation 3, a fourth group contains one free rider and one altruist. The free rider receives help from the altruist without incurring the costs of helping and hence has three offspring while the altruist has none. By generation 5, this group contains only free riders who do not help each other and so go extinct by generation 6. With these hypothetical payoffs and population structure, altruism can evolve and be maintained.

Can group level replicators be identified? Recall that a group level replicator must contain the instructions to make copies of itself. Consider the initial group A==A in generation 1. By generation 4, there are three copies of this A==A group. Hence the A==A replicator has made 3 perfect copies of itself. The information required to reconstruct the group-level replicator, A==A, is contained in the sub-unit organism-level (A=) replicators due to assortative strategies to group. The group level replicator is both an emergent phenomenon and a replicator in its own right. Replicators at multiple levels of the gene, the organism and the group are able to use the group level vehicle to differentially survive and reproduce. Similarly, a nucleotide on a genome can be thought of as a replicator; it possesses all of the information to make copies of itself and due to inheritable assortative strategies these copies locate at the exact same location on an identical gene in the next generation. That a gene can be viewed as an emergent phenomenon due to selection acting at the lower level of a nucleotide does not disqualify a gene as a replicator.

By the same reasoning, the group A==A==A==A is a group level replicator, which also produces three perfect copies. A==A==A==A==A==A==A==F also qualifies in this specific hypothetical example, given that one copy happened to be produced by generation 6. To conclude, group level replicators can be identified in the case of interdemic group selection. These group level replicators provide a precise hereditary basis for the resemblance of offspring groups to parent groups serving to integrate the theory of multilevel selection with the replicator perspective.

3.4 Group Level Replicators in Evolutionary Game Theory/ Intrademic Selection

Maynard Smith (1987a) has allowed that group selection and group level adaptations can occur in interdemic models because groups satisfy Lewontin's criterion for Darwinian units of selection with resemblance between offspring propagules and parent groups providing evidence of a basis for heredity. In contrast, he argues that intrademic models (including kin group selection) do not qualify as "group selection" because of a lack of group-level heredity. Firstly, the "offspring" groups do not directly descend from one parent group, (except in the special case in which new groups consist of only one member), and secondly, there may be within-group differences in fitness. Recall that new groups are formed afresh each generation by sampling from a migrant pool. In the following section, group-level replicators and hence heredity are identified in the case of intrademic group selection. It is argued that descent from a common parent is not a necessary condition for group level heredity—all that is required is that group level replicators possess the hereditary information to construct copies of themselves, albeit somewhat indirectly. The analysis also applies to evolutionary game theory given that intrademic models have the same population structure.

The two cases of random and assortative matching are examined below. Individual organisms possess hereditary strategies to join groups. There are two dimensions to such grouping strategies; firstly are strategies regarding group size and secondly are the "assortative strategies" regarding group composition. The payoff structure is the same as that of Hamilton (1975), section 2.3.2, corresponding to a prisoner's dilemma with respect to individual organism and gene payoffs.

3.4.1 Case 1: Random Matching

Firstly, consider the case in which groups of a particular size (or expected size) are formed by randomly sampling (with replacement) from the population at large so that the only relevant grouping strategies concern group size. For simplicity assume that newly formed groups are initially of size two. From an individual organism's perspective, there are two relevant strategy loci, the locus for either altruism or free riding (A or F) and the locus for the strategy to interact in a group of size two represented by "-". "A-" represents the strategy "play altruist and pair with a member drawn at random from the population" while "F-" is the analogous strategy for a free rider. There are three possible putative group-level replicators represented diagrammatically as group 1 (A--A), group 2 (A--F) and group 3 (F--F).

Individual organisms reproduce according to their fitness payoffs given their own strategy and that of their partners. For example, the free rider in group 2 has the highest possible individual fitness because it is paired with an altruist. It will be demonstrated that these three putative group-level replicators possess the hereditary information to make copies of themselves. It is assumed that there are no mutations.

Consider the first putative group level replicator A--A. After interacting in the pair, each sub-unit organism-level replicator A- is perfectly copied according to its fitness coefficient given that it is paired with an A-. After reproduction, the individual offspring disperse into the migrant pool after which they are randomly paired with a partner that is randomly drawn from the population at large. Let q represent the population wide frequency of A- replicators. A given parent group replicator A--A produces two types of offspring, A--A, which are perfect copies and A--F which are not. A fraction, q, of these offspring are perfect copies of the original parent replicator which only has perfect copying fidelity if q=1. Otherwise, the copying fidelity of A--A is less than perfect. Each "offspring" group replicator consists of one sub-unit replicator that has a direct line of decent to its parent, while the other is drawn from the population at large. Due to decent from two parent replicators, A--A has potentially low copying fidelity which decreases as q falls. Nevertheless, in spite of this low copying fidelity, if q exceeds zero then A--A possesses the hereditary information to potentially produce some perfect copies of itself and hence qualifies as a replicator. The offspring of type A--F increase in frequency as q decreases due to the higher fitness of the F- sub-unit replicators. As q approaches zero, the copying fidelity of A--A approaches zero also. To conclude, A--A qualifies as a replicator but is selected against due to increasingly lower copying fidelity.

Let X- represent an organism level replicator which codes for a random pairing strategy as above, such that X represents a "placeholder" that can be filled with either an A or an F. The expected value of X depends upon the population wide frequency of altruists. For example, if $q = 0.5$, X is expected to be an A- with probability 50% and F- with probability 50%. Consider a group-level parent replicator of the form A--A. A--A has the underlying form A--X, as do offspring, which can be of three possible types; A--A, A--F or F--F, where A--A and A--F have the underlying form A--X. Hence, if $q > 0$, A--X, is potentially able to construct some perfect copies of itself and has a higher expected copying fidelity than A--A .

A--A has a more general form still as represented by X--X, which is perfectly copied to children and grandchildren, and so on. More simply put, groups with two members will produce groups of two members and so on. Hence, ignoring mutations for group size, a perfect copying fidelity group level replicator has been identified for intrademic group selection under random matching. The group level replicator X--X is analogous to but significantly less sophisticated than the genome base replicator in the case of organisms.

Similarly, A--F can produce copies of itself indirectly where one member of each group level replicator's offspring is a direct descendant while the other is drawn randomly from the population at large. There are three types of offspring; A--F, A--A and F--F, with the frequencies of each depending upon the fitness coefficients for the sub-unit organism level replicators. Copying fidelity is less than perfect. For example, if $q = 0.5$, then the copying fidelity of A--F is 50%. As above, the more general underlying group level replicator X--X is copied perfectly. Similarly, for parents of type F--F, there are two types of offspring F--A and F--F, with underlying group level replicators of the form

F--X and X--X, the latter of which is copied perfectly. As q increases due to selection of F- sub-units, the copying fidelity of F--F approaches 100%. Copying fidelity of group level replicators falls as groups get larger because of the dilution of the direct line of descent from a given parent group (except in the case of the most general group level replicators which indicate group size but not group composition). In a large population, for groups of size n, the direct line of descent from a given parent group to a particular offspring copy is now only a proportion $1/n$ of the total group, while a proportion $(n-1)/n$ sub-units descend from the population at large.

To conclude, the group level replicators A--A and F--F only have perfect copying fidelity if their sub-component replicators are at fixation. Replicators of a more general form F--X and A--X have perfect copying fidelity for at least one generation. If there are no mutations for group size, replicators of the more general form X--X have perfect copying fidelity. This most general form of the group level replicator is analogous to the genome-base replicator (section 3.2) with the sub-unit organism-level replicators possessing strategies to assort with this group-base replicator. Hence, under random matching, group level replicators are constructed due to inheritable assortative strategies of the sub-unit organism-level replicators. Eventually, the free riding gene goes to fixation, and the F--F replicator is perfectly copied even though grouping has no adaptive benefit with respect to this trait.

3.4.2 Case 2: Assortative Matching

Consider the case in which grouping strategies specify both group size and composition and groups are of size two. For simplicity assume that there is perfect positive assortment such that Hamilton's coefficient of assortment, $f = 1$. Let '=' represent the strategy to "get into a group of size two and to perfectly positively assort". For example, A= represents an organism level replicator with the strategy to 1) play altruist, 2) match in a pair and 3) pair with another A. Due to the assumption of perfect positive assortment, there are two types of organism level replicators A= and F= and hence only two sorts of group level replicators, A==A and F==F. A= has a higher fitness than F= due to the assumption of perfect positive assortment. After interacting, these organism-level replicators disperse into the migrant pool after which they are re-matched with their own type via the inheritable assortative strategies. Hence, A==A replicators produces only A= offspring organism level replicators, each of which possesses the information to pair with another A= allowing for the construction of another A==A group-level replicator. Hence ignoring mutation, the group level replicator A==A has all of the information to construct perfect copies of itself. Similarly, F==F possesses all the information to make copies of itself.

As in the case of random matching, such copying is indirect because only one sub-unit in each offspring group-level replicator has a direct line of descent to a given parent. Nevertheless, these group level replicators do possess hereditary strategies allowing them to make perfect copies of themselves. They not only satisfy Lewontin's criteria for Darwinian units of selection, but also the stronger Dawkins' replicator requirement. Firstly, there are two group level variants A==A and F==F. These variants vary in fitness; A==A is fitter than F==F. Lastly, offspring A==A and F==F, not only resemble parents but are perfect copies. While it is true that the sub-components A= and F= possess all the information required to reconstruct the group level replicators, which are indeed emergent phenomenon, as in the case of interdemic group selection, this does not disqualify them as group-level replicators.

Consider the case of a mutation, in which an A= mutates to F* where F* represents the strategies to free ride, group in a pair, and pair with an A-type. Additionally, suppose that organism replicators of type A= cannot distinguish between A= and F* types. Now group level replicators of type A==A are no longer perfectly copied due to the presence of F* in the population at large. Offspring group level replicators include both A==A and A=*F. F* will increase in frequency so that eventually the copying fidelity of A==A approaches zero. In the case of kin selection, altruists are positively assorted by descent protecting against the "invasion" of F* types.

Consider the case of imperfect positive assortment, for example a fraction, *f*, of partners are drawn from a pure pool of the same type, while a fraction *1-f* are drawn from a mixed pool as in Hamilton (1975). In this case, the group level replicators, A==A and A==F, make *some* perfect copies of themselves and still qualify as replicators. The level of copying fidelity is increasing in f, the coefficient of assortment. Again the group level replicators A==X and F==X have even higher copying fidelity than A==A and F==F. In the case of A==X, X represents the "expected value" of the second group member, where the probability that this member is an altruist is $f + (1-f)\ q$ and the probability that this member is free rider is $(1-f)\ (1-q)$. Similarly, for the replicator F==X, X represents the "expected value" of the second group member. As f approaches 1, the copying fidelity of A==A increases until it is perfect. As long as Hamilton's (1975) rule (equation 3-19) is satisfied, the copying fidelity of A==A is sufficient to ensure that altruism increases in frequency at the level of the population.

3.4.3 Summary

Based upon satisfying Lewontin's three criteria for a Darwinian selection process, Wilson and Sober (1994, 1998) have argued that intrademic group selection qualifies as "group selection", a process that can generate group-level adaptations. Although it has been disputed that such groups have the third hereditary criteria, they illustrate that there is a correlation between group level characteristics of parent and offspring groups (section 2.3.5) and argue that a direct line of descent to a single parent group is not necessary for group-level heredity. In this section, it was argued that group-level replicators provide the basis for this correlation. Sub-unit replicators (for example at the level of organisms) are able to construct group-level replicators by way of inheritable assortative strategies. Hence group level replicators sometimes possess the inheritable information to instruct offspring sub-units to re-construct copies of the parent group-level replicator. The degree of copying fidelity is determined by the extent to which grouping is assortative and the benefits of altruism relative to the individual fitness costs as summarized by Hamilton's (1975) rule (equation 3-19). Although group-level replicators are emergent phenomena, they are also replicators in their own right. Replicators at multiple levels, including nucleotides, genes, organisms, and groups, are able to use the group-level vehicle to differentially copy themselves. In the case of random matching, there are also perfect copying fidelity group-level replicators, but these take a very general form in which only expected group size is specified. Under random matching, the copying fidelity of more specific group-level replicators such as A--A is too low to allow altruism to evolve.

Overall, section 3 has identified replicators at multiple levels including organisms, and groups for both the cases of interdemic and intrademic group selection. Adaptations can evolve for the good of replicator units at various levels in the biological hierarchy. While these units of selection are in the domain of hereditary information, each replicator has a corresponding expected vehicle in the domain of actions. The theory of a hierarchy of replicators, provides a reconciliation of the theory of multilevel selection with the views of the genic camp, the latter of which requires that the units "for which adaptations evolve for the good of" are replicators in the domain of hereditary information.

4 Inherited Assortative Strategies, Natural Selection, and Adaptation

Hamilton (1975) drew attention to the requirement of non-random grouping in order for the population wide frequency of altruistic genes to increase on a sustained basis. This required $bf > c$, where f is the coefficient of positive assortment and each altruist incurs a cost of c in donating b units of fitness to the group. Hamilton assumed that the coefficient of assortment was exogenous to the model allowing for a focus upon the single locus for altruism or free riding. The coefficient of assortment indicates the degree of linkage of the altruism and free rider strategies to the strategies for positive assortment. This coefficient can be regarded to be a measure of the outcome of a selection process acting at a second locus for positive assortment. If the strategies for positive assortment have reached equilibrium, then f can be held exogenous. If $f = 0$, there is no linkage and grouping is random with respect to group composition, while if $f = 1$, there is perfect positive assortment. Alternately, the selection of strategies for each locus could each be explicitly modeled, however at the expense of tractability.

Recall that in section 3 it was argued that gene and organism-level replicators construct group-level replicators by way of inheritable assortative strategies. In order for such groups to qualify as replicators, the objects of a Darwinian selection process, it must be the case that the assortative strategies are inheritable. Otherwise, any increase in frequency in the altruistic gene at the level of the population is merely due to chance positive assortment into groups, cannot be sustained over time, and cannot be viewed as an adaptation from the perspective of genes, organisms, or groups.

In the next section, a broad view is taken in order to illustrate that standard natural selection can also be viewed as a non-random grouping process, the result of inheritable assortative strategies. This follows from the fact that the word "selection" means "putting entities into groups", while "Darwinian" refers to a specific non-random means of doing so. As in the case of Hamilton's 1975 model, for purposes of mathematical tractability, it is easiest not to make these assortative grouping strategies explicit, although it is possible to do so. It will be illustrated that inheritable assortative strategies underlie the fitness coefficients, which can be related to the coefficient of assortment. Hamilton's approach is used to model standard natural or Darwinian selection in two ways. In model 1, the strategies for positive assortment are held exogenous but underlie the fitness coefficients. In model 2, the strategies for positive assortment are made explicit.

4.1 Model 1: The Price Equation and Standard Natural Selection

There are two types of players, organism players (O-players) and resource players (R-players). O-players play either strategy A or B, which correspond to genes competing for a locus. In each generation, each O-player is randomly matched with an R-player. The R-player gives resources to the O-player conditional upon its type, with the fitness coefficient providing a measure of the number of donated resources. A-types receive more resources than B types and are hence fitter. The expected number of offspring for A and B-types are respectively represented by w^A and w^B with corresponding relative fitness values of w^A/w and w^B/w, where w is the population wide average number of offspring. Under this standard approach, the fitness coefficients convey information about the relative success of each match.

Alternately, standard natural selection can be modeled using Hamilton's framework. Assume that there are groups indexed by s, where q_S represents the within-group frequency of gene A. Initially each group contains one O-player and one R-player. Let $n_S = 1$ represent the number of O-players in group s. If the O-player plays strategy A or B then q_S is respectively 1 or 0. N represents the initial size of the organism-player population, which is also equal to the number of groups. The number of A-genes in group s is $n_S q_S$. Aggregating across groups and dividing by the number of groups gives the population wide frequency of the A-gene,

$$q = \frac{\Sigma n_S q_S}{N}. \qquad [4\text{-}1]$$

Each s-group also contains an R-player that donates resources to O-players dependent upon the recipient's type. This O-player then reproduces to produce $w_S n_S$ offspring, after which the parent dies. Let q_S' represent the frequency of gene A among group s offspring. The change in frequency of gene A at the level of the population is

$$\Delta q = \frac{\Sigma w_S n_S q_S'}{wN} - \frac{\Sigma n_S q_S}{N}. \qquad [4\text{-}2]$$

For simplicity let the population wide average fitness equal 1. The change in frequency of gene A is described by the Price equation (equation 3-1; section 3.2.2) as

$$\Delta q = \operatorname{cov}(w_S, q_S) + Exp(w_S \Delta q_S). \qquad [4\text{-}3]$$

If it is assumed that there is no mutation, then $\Delta q_S = 1$ and the expectation term in equation 4-3 equals 0. Using the linear approximation of section 3.2.2, equation [4-3] simplifies to $\Delta q = \beta_0 Var q_S$, where β_0 is the coefficient obtained due to regressing w_S on q_S. Differentiating the average fitness of a member of the population, $q_S w_A + (1 - q_S) w_B$, by q_S gives $\beta_0 = w_A - w_B$. Again assuming binomial sampling to generate groups, the variance of q_S is represented as $Var q_S = q(1-q)$. Substituting into equation [4-3] gives the change in the population wide frequency of gene A over one generation,

$$\Delta q = (w_A - w_B) q (1-q). \qquad [4\text{-}4]$$

In the case in which the A or B gene is at fixation, respectively corresponding to $q=1$ or 0, then given the assumption of no mutation, $\Delta q=0$. If $0 < q < 1$, then given that $w_A > w_B$, the A-gene will increase in frequency and eventually go to fixation. Utilization of the Price equation illustrates that natural selection can be modeled as a process in which organism players are grouped with resource players. The fitness coefficients represent the payoffs due to the match. Any assortative grouping strategies have not been made explicit and matching was modeled as random. In model 2 an alternate approach is utilized in order to illustrate that natural selection can be modeled as a process of assortative grouping.

4.2 Model 2: Making Assortative Grouping Strategies Explicit

In a second model, O-players are matched into groups within each of which there are n_R R-players. R-players can play either strategy G, in which the R-player gives a particular quantity of resources to the O-player in its group or strategy Z in which zero resources are handed over. Let μ represent the population wide proportion of resource players that play strategy G while a fraction $1-\mu$ playing strategy Z. The notation for O-players is identical to that of model 1. Let μ_S represent the proportion of resource players that play strategy G in a given *s*-group, while $1-\mu_S$ play strategy Z. Again, let q_S represent the initial frequency of A-genes in group *s*. As in model 1, it is assumed that there is only one O-player in each group so that q_S takes on the value of either 0 or 1.

Initially assume random matching such that each O-player randomly groups with n_R resource players in each group *s*. The O-player in each *s*-group is assumed to reproduce proportionate to μ_S, the fraction of R-players who play strategy G, such that $w_S = k\mu_s$, where *k* is a proportionality constant. For simplicity let $k=1$, so that $w_S = \mu_S$. As in model 1, the population wide frequency of A-genes in the first generation of offspring can be represented as $q' = \sum \frac{q_S' n_S \mu_S}{N\mu}$. [4-5]

Again taking the difference, $q' - q$ and rearranging gives the Price equation, which can be written as $\Delta q = Cov(q_S, \mu_S)$ under the assumption of no mutation with $\mu=1$. Given the assumption of random group formation with respect to resource player composition, there is no correlation between q_S and μ_S so that the covariance term and Δq are both equal to 0. Any change in the frequency of q is not due to Darwinian selection, but to random drift.

Suppose instead that groups are formed assortatively. Using Hamilton's 1975 method, assume that R-players are drawn with replacement from three separate pools. In pool 1, all R-players use strategy G, in pool 2, all R-players use strategy Z and in pool 3, the frequency of the G-strategy is equal to its frequency in the population at large. The within pool frequencies of G-strategy players are respectively $\mu^1 = 1$, $\mu^2 = 0$ and $\mu^3 = \mu$. Assume that A-strategy O-players draw a fraction f of group resource members from the pool 1 and a fraction $1-f$ of group resource members from pool 3. An A-type then expects to be matched into a group in which the frequency of R-players using the G-strategy is $f+(1-f)\mu$. Similarly, B-types draw a fraction f from pool 2 and a fraction $(1-f)$ from pool 3 for an expected frequency of $(1-f)\mu$ R-players using the G-strategy. As in model 1, the covariance term simplifies to $\Delta q = \beta_0 Var(q_S)$, where β_0 is obtained by regressing the expected value of μ_S (the within group fraction of G-strategy players) on q_S to give $\beta_0 = f$. Given binomial sampling, the within group variance of q_S is represented as $Var(q_S) = q(1-q)$, so that

$$\Delta q = f(1-q)q . \qquad [4\text{-}6]$$

If $f=0$, grouping is random such that on average the within group frequency of A-types does not change. If $f>0$, grouping is assortative and A-types increase in frequency if $0<q<1$. Comparing equations [4-4] and [4-6] illustrates that $f = w_A - w_B$. Hence, the difference between the fitness coefficient corresponds to the coefficient of positive assortment where this coefficient is due to O-player strategies to assort into groups with differential amounts of resources.

The second model illustrates that standard Darwinian selection can be interpreted as a process of assortative grouping, whereby organisms play inheritable strategies to positively assort with resources. If grouping is random, A-types might be selected due to chance assortment with resources. However, such selection is not Darwinian but corresponds to drift. The exercise illustrates that fitness coefficients can be interpreted as summary statistics representing the outcome of an underlying assortative grouping process.

4.3 Summary

Selection is another word for sorting entities from a larger group into smaller groups, a process that can be either random or non-random with Darwinian selection falling into the latter category. This non-random grouping can be viewed as being due to inheritable assortative strategies, which allow for differential survival and reproduction. In this section, it was argued that "Darwinian" selection processes can be viewed as being the result of selection upon inheritable strategies for assortative grouping.

Hamilton's framework was used to illustrate that standard natural selection can be modeled as the outcome of inheritable assortative strategies for grouping with 'resource players'. In model 1, organism players were randomly matched with resource players who donated resources according to organism specific fitness coefficients. In model 2, organism players were assortatively grouped with two types of resource players, which differed according to the number of resources that they donated to organism players.

Comparison of the two models illustrated that the difference in fitness coefficients for each of the two organism types (model 1) was equivalent to the coefficient of assortment of organism types with resource types (model 2). This coefficient was analogous to the coefficient of assortment in Hamilton's model.

In cases in which organism types do not vary with respect to inheritable strategies for assortment with resources, but differential assortment occurs by chance, then any selection is not "Darwinian" but is merely the result of random drift. Inheritable assortative strategies allow gene-level and organism-level replicators to use the resource base as a vehicle by which to differentially make copies of themselves. Similarly, in the case of standard group selection, it is by way of inheritable assortative strategies that individual gene and organism level replicators are able to use group level vehicles to differentially replicate.

Can group level replicators be identified in the case of standard natural selection? Given that the resource base is composed of living organisms, it literally replicates itself allowing both the organism and gene-level replicators to duplicate themselves. These organism replicators can be viewed as positively assorting with the resource base replicators. As both O-types and R-types replicate themselves, it can be argued that there is a group level replicator although it is somewhat one-sided because R-types do not depend upon O-types to reproduce by assumption. Lewontin's three criteria for Darwinian selection are satisfied. Firstly, the group level replicator for A-types differs from that of B-types and so there is variation between groups. Secondly, there is differential fitness based upon these variants and thirdly, the group level replicator is able to make perfect copies of itself conditional upon a non-changing resource environment. Suppose that the R-organisms that constitute the resource base go extinct. In this case, the O-organism and O-gene level replicators are unable to make copies of themselves; indeed they no longer satisfy the criterion of replicators, i.e. an entity's status as a replicator depends upon the environment.

Modeling standard natural selection as an assortative grouping process is complex, while the concept of a group level replicator in which groups are composed of both the hereditary information of organism types and resource types is confusing and probably not of much practical use. Indeed, the exercise illustrates that the production of a consistent theory of multilevel selection is difficult and in need of more work. More importantly from the perspective of this essay, the exercise serves as an interesting thought experiment by placing group selection within a broader context. In section 3, the group level replicators identified for both interdemic and intrademic group selection were more concrete.

Overall, the exercise drew attention to the two general types of grouping (or selection), the type that is based upon heritable assortative strategies and the type that are not. In purported cases of group selection, unless group formation is due to inherited assortative strategies, the advantages of group structure cannot be passed on to offspring. The offspring cannot then (except by chance) form new groups and pass on group advantages to their offspring. In cases in which group strategies are played by chance and are beneficial any "group selection" is not Darwinian and there cannot be group level adaptations. The term group selection was traditionally reserved for cases in which there are group level adaptations. Researchers must be careful that they do not attribute apparent group level harmony to "group selection", when all that is happening is what might be referred to as fortuitous neighborhood effects due to random drift.

5 Conclusions to Part I

The group selection debate has persisted with two extreme views, the genic approach (Dawkins, 1976, 1982; Williams, 1966) and the multilevel selection view (Wilson and Sober, 1994; Gould, 1998, 2002). In the genic view, any functionality of units at levels larger than genes including organisms and groups of organisms is an emergent phenomenon due to selection acting upon genes, the largest units of DNA likely to survive the process of meiosis intact. Higher level units are not the subject of Darwinian selection processes and adaptations do not evolve for the good of units at these levels. In this view, the units of selection are replicators, which in the biological case correspond to genes.

To contrast, Wilson and Sober have argued that adaptations can evolve for the good of units at a variety of levels and that adaptations evolve for the good of vehicles (the aggregate actions of coalitions of genes). Resemblance between offspring and parents of organisms and sometimes groups of organisms provides evidence of heredity at that level, a requirement for Darwinian selection of units at that level. In this view replicators are irrelevant to the units of selection debate. As illustrated by the Price equation, if there is sufficient positive assortment of sub-units within a group, then between group variation may outweigh within group variation allowing for the evolution of traits which benefit the group, but would reduce the fitness of sub-units in the absence of positive assortment.

The debate has been remarkable in its persistence and the topic has attracted attention among biologists, social scientists and also lay audiences. Perhaps the widespread interest is due to the possible implications of group selection for human nature. How can we reconcile the better side of human nature with individualistic Darwinism? On both sides of the debate, biologists have received support from philosophers, often in the form of full-length books, which at times may have done more to confuse than enlighten. At times semantic impasses have been reached, with a lack of ability to communicate by cold hard scientific reasoning. For example, Maynard Smith (1987b) argues that a gene for trait X *causes* trait X while Sober (1987 a,b) insists that as long as action X is context dependent, then gene X cannot *cause* it. The debate at times has become quite passionate. For example, Dawkins in the commentary to Wilson and Sober (1994) wrote; "They are zealots, baffled by the failure of the rest of us to agree with them. I can sympathize: I remain reciprocally baffled by what I still see as the sheer, wanton, head-in-bag perversity of the position that they champion."

Hull (1980) with reference to the difficulty of resolving outlying questions involving group selection wrote; " Comparable stalemates in the history of science have tended to result from everyone concerned taking for granted something so fundamental that no one in their right mind would question it." In this light, this essay suggests that the fundamental point which has been taken for granted by almost everybody is Williams' (1966) argument that the largest units of selection are "genes". These "genes" correspond to fragments of DNA, which are small enough to survive fragmenting by meiosis sufficiently often to have the copying fidelity required to potentially leave a lineage of descendants. Both Dawkins and Wilson and Sober agree that replicators are irrelevant for the group selection debate, which is about vehicles, but for different reasons. For Dawkins, vehicles are simply emergent phenomenon, not to be glorified. Dawkins (1994) writes; "*The Extended Phenotype* (Dawkins; 1982...) is best seen as an attack on the organism and this should be music to W&S's ears. I coined the term "vehicle" not to praise it but to bury it."

To reconcile the views that the largest organic replicators are genes with the theory of multilevel selection, Wilson and Sober are left with the option of elevating vehicles to the status of the fundamental units of selection for which adaptations evolve for the good of. Among other arguments, in order to justifying moving the units of selection from the domain of hereditary information to vehicles, they utilize somewhat confusing philosophical arguments regarding the nature of causation. Trivers (1998) in a harsh review of Wilson and Sober (1998) refers to "verbal tricks", which he does not identify, but seem to correspond to these philosophical arguments. Wilson and Sober depend upon resemblance between offspring and parent units at various levels in the biological hierarchy in order to establish that there can be heredity at the level of organisms and groups of organisms. They also provide many empirical examples of traits that benefit groups but which would not benefit individual organisms or genes in the absence of assortative grouping. Importantly, they draw attention to the Price equation, which provides a theoretical basis for separating the within and between group components of selection.

For a unit to be the object of a Darwinian selection process there are three requirements (Lewontin, 1970). It is required that there is variation in a trait among units at some level, that this trait is correlated with differential survival and reproduction of the unit, and that the unit level characteristic can be inherited by offspring units. Wilson and Sober argue that resemblance provides evidence of heredity in the case of organisms and also the groups in both interdemic and intrademic selection. I have argued that given our current understanding of the chemical basis of heredity, that it is possible to do better than mere resemblance and that the replicator concept can be applied both at the level of organisms and sometimes groups of organisms.[47] I have argued that resemblance between

---

[47] Maynard Smith (1987a) also allows that resemblance provides evidence of a basis of heredity, and, in contrast to Dawkins, concedes that adaptations may evolve for the good of organisms and, sometimes, higher level groups in the case of interdemic group selection.

offspring and parent units at a variety of levels, which is not simply due to chance but persists throughout the generations, has a basis in the hereditary domain in the form of a replicator at that level.

It was argued that there is a hierarchy of replicators in the domain of hereditary information. Lower level replicators such as genes can be viewed as coding for two traits, the standard phenotype such as altruism or eye-colour and the assortative trait to assort with other replicators, which allow for the construction of larger replicators at the level of organisms and groups of organisms. As the assortative strategies are inheritable, offspring gene-level replicators are able to re-construct the same higher level replicators. Hence, the higher level replicators possess the information to construct perfect copies of themselves.

Firstly, higher-level replicators were identified at the level of genomes. It was argued that the basic structure of the genome is conserved during meiosis. The genome-base replicator (in haploid and diploid forms) was defined as the basic structure of a genome with "placeholders" possessing assortative strategies to associate with a limited set of possible gene-replicators. The lower-level gene-replicators can be viewed as possessing inheritable assortative strategies to assort with a genome-base replicator at specific "place-holder" locations on the genome. These assortative strategies are passed on to descendent gene-replicators, which inherit the ability to assort with an identical genome-base replicator allowing them to reconstruct the same expected vehicle as parents. The vehicle for the genome-base replicator is just the average species aggregate phenotype. For each replicator in the hierarchy an expected vehicle can be identified.

Next, it was argued that for the cases of interdemic and intrademic group selection that gene and organism level replicators aggregate to form group level replicators by means of inheritable assortative strategies for grouping. For example, in the case of sibling groups, gene-replicators for altruism code for both the altruistic trait and for the ability to assort in a sibling group in which at least 50% of the members are also expected to be altruists. Individual gene and organism level replicators use the kin group as a vehicle. As in the case of organism level replicators, group level replicators are both emergent phenomena and replicators in their own right.

Intrademic group selection does not appear to qualify as a process that generates group level adaptations because offspring groups do not descend from a single parent group and hence appear to lack group level heredity. The theory of a hierarchy of replicators suggests that descent from a common parent group is not necessary. Offspring organism level replicators possess assortative strategies, which allow them to reconstruct the same group level replicators of their parents. Group level replicators contain the hereditary information for their own self-copying even though the component sub-units need not descend from a single parent group. If Hamilton's rule is satisfied, selection between groups outweighs selection within groups and adaptations can be said to evolve "for the good of" both group-level and lower level replicators; The assortative strategies serve to synchronize the interests of lower and higher-level units.

Hamilton (1975) illustrated that assortative grouping is a necessary condition for the evolution of altruism. Although he did not explicitly model positive assortment as being due to strategies, using this game theoretic tool provides a framework in which it is possible to construct the group-level replicators. Also, Hamilton's paper makes it clear that altruism cannot evolve if group formation is random. The assortative strategies must be inheritable, otherwise, any group level functionality is merely due to chance and is not the outcome of a Darwinian selection process. Section 4, in which Hamilton's framework was used to model standard natural selection, illustrated that in essence natural selection is a process of non-random grouping, with variation and differential fitness between units being due to inheritable assortative strategies for grouping.[48]

The theory of a hierarchy of replicators serves to reconcile the views of Dawkins (1976, 1982) / Williams (1966) with those of Wilson and Sober and Gould by employing the

---

[48] The work of Romanes (1897) provides a lucid description of the two types of isolation (i.e. selection). Selection is either non-discriminate (random) or discriminate (non-random). Only in the latter case can selection generate adaptations.

powerful replicator concept (based upon Williams' genes), while also providing a basis for multilevel selection.[49] Although resemblance provides an indicator of heredity, given current knowledge it is possible to use the more precise replicator concept. It was argued that there is a one-to-one correspondence between replicators and expected vehicles. Accordingly, higher-level replicators imply higher-level expected vehicles and hence resemblance between parents and offspring units at these higher levels. This is consistent with the arguments of Wilson and Sober and Maynard Smith, that resemblance provides evidence of higher level heredity.

Given that the proposed theory of a hierarchy of replicators is simple, then why wasn't it proposed earlier? In other words, why is there almost ubiquitous agreement among biologists that the largest organic replicators correspond to Williams' genes? The historical context presented in this essay suggests various reasons. Firstly, given that the particulate approach had proven so successful in physics and chemistry, there was a predisposition to look for a particulate basis for heredity in biology. The rediscovery of the independently assorting Mendelian factors of inheritance in 1900 provided support. Although the discovery of chromosomes and later the helical chemical structure provided evidence that the genes were arranged on a base, the particulate focus upon genes remained. This focus is understandable because it was correctly believed that characterizing the genes would reap substantial benefits in such areas as medicine and agriculture.

Secondly, for population geneticists and evolutionary game theorists, focussing upon one gene at a time was both useful and also mathematically tractable. Evolutionary game theory is easier to model under the assumption of random assortment. Simultaneously modeling the standard phenotype and strategies for assortment can generate messy mathematics and is not as easily amenable to elegant closed form solutions. To contrast, in object-oriented computer programming, in which organisms have locations on a spatial

---

[49] Note that although Williams' genes refer to pieces of DNA large enough to most likely survive meiosis intact, given that genome-base replicators and group level replicators may make perfect copies of themselves, in "spirit" they qualify as "genes" in Williams' sense.

lattice, strategies for assortment must be explicitly specified. It is not however obvious that changes in location can be viewed as mutations in hereditary information, analogous to mutations of the standard gene strategies. The closed form mathematical approaches are often difficult for non-mathematical biologists to understand, which makes it difficult for these biologists to criticize without a considerable time investment. Today computer simulations, which model the co-evolution of multiple traits are relatively easy to understand so that hopefully there will be less of these communication problems. Communication problems due to the mathematical nature of the debate may have slowed its resolution.

Assumptions made for tractability, which have later been judged to have been "too restrictive to get group selection to work" originally suggested that group selection was unlikely to be important in practice. Wilson has pointed out that another reason that group selection was originally considered to be unlikely was because it was originally proposed as a solution to the puzzle of altruism. However, group selection is not merely limited to situations in which there is conflict between units at different levels (Wilson, 1975, 1977).

Thirdly, the inheritable assortative strategies, which are fundamental to understanding group selection and also the concept of group level replicators did not "jump out" at biologists as obvious strategies as did the strategies for standard phenotypes such as eye-colour or altruism. For example, an assortative strategy for location, might partially be determined by genes for immobility and also the actual physical location of the parent, which is literally inherited by offspring. In standard game theory, under the assumption of random matching, although it is possible to explicitly note the strategies for grouping, this is not standard practice. When biologists adopted game theory to the evolutionary context, it was not obvious that the actual grouping process could be modeled as the result of strategies. Moreover, for the type of questions that evolutionary game theory was attempting to answer, there was no need to explicitly model the grouping strategies.

Fourthly, Williams focused upon meiosis as a process that fragments the DNA as a basis for defining "genes", which Dawkins later argued were the largest organic replicators. The traditional focus upon single genes provided a predisposition to see meiosis in this light. However, a shift of the lens illustrates that meiosis serves to preserve the overall structure of the genome between generations. Indeed, recent theories and empirical work suggest that the original purpose of meiosis may have been a mechanism of DNA repair, a means of increasing the copying fidelity of the genome (Bernstein et. al, 1985; Forsdyke, D., 2001). Bateson's (1914) had originally proposed that the hereditary information consisted of a "base" to which independently assorting transferable particles (genes) were attached. If this line of thought had been seriously considered, more attention might have been given to the overall structural features of the genome.

Fifthly, Dawkins' book, *The Selfish Gene* has had a large impact upon biological thought. Chapter 2 develops the replicator concept prior to developing the concept of meiosis such that the replicator is originally envisioned as an entire genome. In chapter 3, the organic replicators are equated with Williams' genes. Given that a replicator is defined as "anything in the universe of which copies are made", it is obvious that such "genes" are copied and are hence replicators. However, they are only copied because they are located on genomes. I have argued that the location of genes at specific loci on genomes can be viewed as being due to inheritable strategies to assort with genomes. These strategies are inherited by descendent gene-replicators, which in turn inherit the copying apparatus of the genome and also the genome vehicle. Although it is obvious that genes are copied, I have suggested that it is useful to be explicit about this process by clearly identifying the inheritable assortative strategies. If a "replicator" does not possess all of the information for its own copying, then perhaps it should not be considered to be a replicator. For example the bible is a replicator as it has been copied many times by humans. However, if a particular bible is lost in the dessert and is never copied, it is not a replicator. In this case the assortative strategy of an environment of humans is part of this bible replicator's strategy to be copied.

In this essay, it has been argued that organism and group level replicators are both replicators in their own right and also emergent phenomena due to selection acting at lower levels. Then why bother with multiple levels of selection? Why not simply focus upon the gene replicators while being explicit about the assortative strategies? As argued by Wilson and Sober, the standard explanations for altruism, reciprocal altruism and kin selection can be categorized as forms of group selection. However, these processes are also due to the selection of individual genes, including the assortative strategies. Does employing the concept of group selection facilitate modeling and understanding? Firstly, and perhaps most importantly, the theory of a hierarchy of replicators due to inheritable assortative strategies has served to integrate the theory of multilevel selection with the genic view. Secondly, it may at times to useful to focus on higher levels. For example, even though gene replicators may be viewed as emergent phenomena due to selection acting upon nucleotides, which have inheritable assortative strategies to locate at specific positions on genes, focussing on individual nucleotides is not very useful from the perspective of population genetics. Similarly, in order to understand altruism among humans or social insects, a group perspective may provide insights. Lastly, the group perspective may predispose us to take a different view of human nature. For example the term reciprocal altruism predisposes us to consider humans in the light of cold rational self-interest. While it is true that reciprocal altruism benefits individual humans, it is also of benefit to the group. Utilizing both perspectives may help social scientists to develop a more complete picture of human nature. Social scientists have been interested in and contributed to the biological debate, particularly with regards to the extent that humans are altruistic or selfish. It is hoped that this essay has provided a biological basis that will contribute towards further understanding.

Part II

Altruism, Information and Behaviour

1 Introduction

1.1 Altruistic Psychology and Economic Theory

Humans have long debated the relative importance of self-interested verses social motivations. In *The Wealth of Nations*, Adam Smith noted the potential for some alignment of selfish and social interests in market economies. However, despite this apparent moral validation for selfishness, his other book *The Theory of Moral Sentiments* investigates emotions such as sympathy and their role in human affairs. Economic theory has for the most part been constructed upon the foundations of materialistic self-interest and policy recommendations have been made in accordance.[1] The human sympathies, to contrast have largely been neglected although recently economists have increasingly begun to incorporate social emotions into utility functions (Hirsheleifer, 1987; Andreoni, 1990; Frank, 1994). Findings by experimental economists are being interpreted as consistent with an important role for human social emotions in situations in which individual decisions are believed to affect the welfare of others (Fehr and Gachtër, 2000, 2002; Gintis, Bowles, Boyd, and Fehr, 2003; Andreoni, 1995). Additionally, although traditional Darwinian theory provides a theoretical basis for human selfishness, over the past 40 years biologists have produced a large body of theoretical and empirical work that challenges this individualistic focus (Wade, 1978; Wilson and Sober, 1994, 1998; Gould, 2002). Social emotions such as guilt, shame, empathy, self-righteousness, and revenge, affect the utility derived from actions with perceived social consequences. Individuals with preferences such that they derive utility directly from the welfare of others are described as altruistic.[2]

---

[1] The term "private material" is used to denote the benefits and costs that fall upon an individual due to the consequences of his own decisions. These material costs and benefits do not include the psychological effects of the social emotions, experienced due to beliefs that decisions have affected others.

[2] Agents are modeled as utilitarian in the sense that Mother Theresa acquires psychological emotional benefits if she believes that she is helping others. In this sense, she can be regarded as selfish, however, she will still be referred to as a genuine altruist as she incurs 'material' costs from helping others. There are further distinctions regarding the types of altruism (Wilson and Sober, 1998, p. 17, 199-222).

In economics, it is standard to use benchmark cases such as perfect competition and monopoly. With regards to the sociality of preferences, there are two benchmarks: Pure free riders are motivated entirely by material self-interest, while pure altruists psychologically internalize the net external benefits that they believe they have inflicted upon others.[3] For example if person A believes he has imposed $100 of damages upon person B, A experiences $100 worth of psychological costs. The assumption of pure altruism is extreme, but given the evidence as well as theoretical rationale for at least somewhat altruistic preferences, it is perhaps no more extreme than the other commonly used benchmark of *Homo economicus*. Flipping the assumptions can motivate different questions. Instead of finding altruistic behaviour puzzling, it is now the observance of many incidences of apparent free-riding that is in need of explanation. For example, instead of finding voting and donating to charities paradoxical, the puzzle is a public lack of engagement.

An instructive case is the theory of public goods provision. Under the assumption of self-interest, the under-provision of public goods is easily explained as the result of free-riding. This theory has recently been challenged both theoretically (Gintis, 2000; Gintis and Bowles, 2003, 2004) and empirically (Fehr and Gachtër, 2000; Gintis, Bowles, Boyd, and Fehr, 2003). In experimental multi-period public goods games, agents are anonymously given opportunities to contribute to a common pool. Donations are multiplied and then shared equally among all players. The game is repeated a number of times. Fehr and Schmidt (1999) report moderately high contributions in early rounds (40% to 60%), however such games rapidly unravel such that by the 10th round, individuals contribute nothing or close to nothing. In debriefings, experimental subjects explain their behaviour as retaliation based upon anger as provoked by the free riding of others (Andreoni; 1995).

---

[3] The more extreme benchmark of spiteful preferences is not considered.

As a modification, Fehr and Gachtër (2000) allow players to punish non-contributors.[4] As punishers incur a monetary loss when inflicting punishments, punishing is referred to as "costly". Punishers are referred to as "second order altruists", who contribute to a public good, the deterrence of free riding. Punishing can be regarded as "irrational" because higher private payoffs can be attained by free riding in the sense of waiting for others to incur the costs of punishing. In contrast to games with no opportunities to make costly punishments, in partner treatments in which group membership did not change between rounds, the level of contributions increased to almost full cooperation by the 10th round.[5] Early contributors are observed to impose costly punishments upon non-contributors. As players learn that others both contribute and "irrationally" punish, they have rational incentives to cooperate.[6] Punishers appear to be driven by retributive emotions and possibly by an altruistic desire to contribute to the public good. These results suggest a new interpretation of standard public goods games in which there are no opportunities for costly punishments. In such games, the only way to punish is to free ride in the sense of not contributing. Hence, much apparent "free riding" may be driven by emotions that serve a broader purpose of promoting altruistic behaviour. Paradoxically, it can then be argued that such "free riding" is altruistically "motivated".

The consideration of altruistic preferences is providing alternate explanations for phenomena as well as motivating researchers to ask different questions. In this vein, I consider some possible implications of altruistic preferences for behaviour and information about the social consequences of behaviour. The social consequences of

---

[4] Other experiments that find that subjects are eager to engage in costly punishment of non-contributors include Dawes, Orbell and Van de Kragt, 1986; Sato, 1987; Yamagishi, 1988, 1988, 1992; and Ostrom, Walker and Gardner, 1992, as referred to in Gintis (2000).

[5] The importance of retaliatory emotions is supported by comparison of three treatments. In personal treatments, in which group members are the same for all ten rounds, contributions increase over time to a situation of almost 100% cooperation. In stranger and perfect stranger treatments, for each round group members are randomly selected anew. In the latter case, players are given the guarantee that they will never play the same player twice. In these latter two treatments, contributions are initially lower than in personal treatment and remain at a constant level.

[6] Here, "irrational" refers to decisions that result in private material net costs.

one's actions are often far from obvious. Both altruists and self-interested types have incentives to search for information about the social consequences of their actions. The motivations for each type however differ. Self-interested types are motivated purely by the *private material* net benefits of their decisions. For example, they may be concerned about retaliation by others. Altruists, on the other hand, are motivated by a direct concern for the social consequences of their actions.[7]

1.2 Sources of Information and Supply Side Bias

Sources of information about the social appropriateness of behaviour include the media (both non-fiction and fiction), education systems, publications by experts and government, conversations, and direct observation. Assuming that there exists some best objective estimate of the social consequences of activity choice, the attainment of an objective view may be difficult and costly due to factors such as the abundance of information, the existence of multiple and biased perspectives, and the cognitive difficulty of the task.[8,9] The media have been the topic of many studies due to their importance as sources of information. Media bias has been attributed to multiple factors including the concentration of media ownership, the difficulty of accessing the credibility of information sources including those provided by vested interests, the dependence upon corporate advertising for finance, and the subconscious imposition of the agendas and biases of media professionals (Chomsky, 1989; Anderson, 1997; Harris,

---

[7] Andreoni (1990) distinguishes between "altruistic" and "warm glow" giving. Under altruistic giving, the utility of others is directly incorporated into the utility function while under warm glow giving, the altruistic action provides utility directly. It is difficult to see why an individual would receive a warm glow unless he believes that his giving will increase the utility of others. Surely he has already formed prior beliefs about the consequences to others. Hence, warm glow giving must ultimately be based upon altruistic preferences. As the distinction between warm glow giving and altruism appears to be a matter of modeling convenience, I make no distinction.

[8] It is commonly argued that the media is biased. Some communication theorists argue that this implies that it is possible in theory to provide unbiased views. In contrast, constructionists argue that there is no way to form an objective interpretation of reality--all interpretations are subjective (Anderson, 1997, p. 47).

[9] Note that such an objective view may include acknowledgement that expected social costs and benefits cannot be calculated.

1998; Beder, 2002). Particular media lend themselves to particular types of bias. For example, time limits in television broadcasts, can lead to simplistic and sensationalistic coverage making it difficult for viewers to attain more than a superficial understanding of complex world issues (Anderson, 1997; Harris, 1998; Wicks, 2001).

1.3 Demand for Information and Demand Side Bias

Although much of the communications literature focuses upon supply side aspects of information provision, particularly by the media, there is a growing literature on the effects of demand by audiences (Wicks, 2001). At the most basic level, material that does not appeal to the demands of audience members generates low ratings, sales and availability. Media suppliers attempt to gage the demands of audiences when deciding upon content. For example, the Nielsen ratings provide feedback on the size and composition of the audience for television shows.[10] Libraries and bookstores stock books largely in response to demand. Although the demand for information may reflect the audience's desire to obtain an "objective truth", psychological factors provide incentives for individuals to "selectively expose" themselves to information allowing for "demand side bias". Such selective exposure includes selective attention or selective avoidance of information that respectively supports or challenges one's existing attitudes (Baron and Byrne, 1994, p. 155).[11] In the current information age, information abundance and issue complexity make it costly and difficult both for suppliers to provide and demanders to expose themselves to a balanced spectrum of information. This increases the ease with which audiences can selectively expose themselves to information. Overall, the current view is that suppliers and demanders of information and the external objective reality interact in a complex way to determine the content of media and other information sources.

---

[10] The demands of television audiences may be having an increasingly large effect upon content due to the internet. For example, at fan web sites for popular television shows, viewers express opinions about the direction of plots, sometimes expressing intent to discontinue viewing if particular plot directions continue.

[11] Note that in subsequent sections, for purposes of conciseness, the term "selective attention" sometimes refers to both the phenomena of selective attention and avoidance, as will be apparent from context.

1.4 Information Demand and Cognitive Dissonance

Private material interests often conflict with social interests. A situation of psychological conflict can arise due to exposure to information suggesting that an individual's behaviour damages others. In such a psychological state, an individual is said to experience cognitive dissonance due to the perception that his behaviour conflicts with his attitudes regarding the social appropriateness of the behaviour.[12] An altruist experiences dissonance if he chooses to engage in activities that he believes to be asocial. Like hunger, dissonance is uncomfortable and individuals seek to reduce it by changing behavior, beliefs or attitudes.

Selective attention to information can allow people to change their beliefs and attitudes providing a potential means to reduce dissonance. Characteristics of the available information such as the existence of multiple viewpoints, information abundance, and complexity, potentially allow for the formation of a wide variety of beliefs regarding a behaviour's social effects and facilitate selective attention. Psychologists find that individuals selectively attend to or avoid information that respectively supports or challenges their beliefs about the social appropriateness of their behaviour (Oskamp, 1991; Wicks, 2001). Selective filters can then bias views even if in principle the supply side characteristics of information would allow for the construction of an objective view.

1.5 Economic Work on Cognitive Dissonance

Despite widespread acceptance of the importance of the cognitive dissonance in the field of social psychology (Oskamp, 1991, p. 238-250; Baron and Byrne, 1994, p.157-166), there is a limited amount of work within the field of economics. Hirshman (1965) considers some consequences of cognitive dissonance for development theory. Development researchers, circa 1965, focused upon attitude changes as a means of inducing behavioural changes. For example, attitudes towards community sharing may

---

[12] Cognitive dissonance is potentially experienced when agents experience inconsistencies between cognitive elements, where cognitive elements refer to items of knowledge, information, attitudes and beliefs that a person holds about himself and his surroundings. Among competing psychological theories for attitude formation, dissonance theory is broad—containing some of the other theories as sub-components (Oskamp, 1991, p. 239-240).

conflict with a potential entrepreneur's motivation to accumulate capital. However, according to dissonance theory, attitudinal changes are likely to succeed behavioural changes. For example, if opportunities to accumulate capital arise and are taken, people may experience dissonance if their actions conflict with their beliefs regarding appropriate social behaviour. To reduce such dissonance, they may then change their attitudes regarding capital accumulation. Hirshman writes; "The art of promoting development may therefore consist primarily in multiplying the opportunities to engage in these dissonance-arousing actions and in inducing an initial commitment to them". Later, economists used dissonance theory to provide explanations for a diverse range of topics including the irrational behaviour of workers regarding safety equipment (Akerlof and Dickens, 1982), irrational investment behaviour (Gilad, Kaish and Loeb, 1987) and risk adverse behaviour of government workers (Brady, 1993). For general discussions by economists see Akelof and Dickens (1982), Gilad et. al. (1987), and Hosseini (1997).

1.6 Economic Work on Dissonance Due to Asocial Behaviour

Psychological research suggests that dissonance is particularly likely to occur if one believes that one's actions have caused harm to others (Oskamp, 1991, p. 242). Social consensus attitudes regarding appropriate social conduct provide an important standard by which to evaluate the social consequences of one's actions. It follows that the violation of accepted moral codes of conduct is likely to induce dissonance. Published work by academic economists on dissonance due to the violation of socially accepted codes of conduct is limited to Hirshman (1965) and Rabin (1994). Hirshman does not explicitly relate the underlying basis of attitudes regarding appropriate social behaviour to altruistic preferences, however such a link should strengthen his argument. In Rabin's model, agents consider themselves to be moral people who face a tradeoff because they incur psychological costs due to increasing the level of some activity that is privately materially desirable but imposes costs upon others. Dissonance is modeled as varying directly with the absolute value of the difference between the actual level and beliefs about the appropriate social level of the activity. Agents can reduce dissonance by selectively avoiding events that would expose them to information telling them that their activity choices are respectively prosocial or asocial. The avoidance of such events is assumed to be costly. Rabin uses the example of fur coat owners who may forgo attending rock concerts in order to avoid exposure to information about animal cruelty. Rabin assumes that it is possible to attain objective truth and that there is a cost to maintaining false beliefs regarding the social consequences of one's activity choices. An individual's beliefs regarding the socially appropriate level of the activity is determined by the social consensus attitude that in turn depends upon the average level of the activity.

Rabin finds paradoxically that higher levels of dissonance (corresponding to higher levels of society wide altruism) can lead to increases in the level of the asocial activity. Although more altruistic preferences should increase dissonance leading to less of the asocial activity, in order to reduce dissonance, agents selectively avoid information. The aggregate effects of such selective attention can generate a change in social consensus attitudes allowing for an *increase* in the socially acceptable level of the activity, in turn leading to more asocial behaviour. Rabin focuses upon parameter specifications that can generate this counterintuitive result.

1.7 Overview

As many if not most decisions have social consequences, it is surprising that Rabin's paper has not generated more follow-up. Perhaps this is because he illustrates his argument with the not particularly controversial or costly decision to not wear fur. Some more costly and controversial examples include changing consumption habits in developed countries in the face of both a deteriorating environment and high levels of poverty in the third world. I will argue that altruistic preferences, as motivated by psychological emotional effects, in conjunction with selective attention to information have important and widespread consequences for both activity choices and social consensus attitudes regarding the degree to which activity choices are prosocial or asocial.

Activity choices in conjunction with beliefs regarding the social appropriateness of these activity choices can induce psychological emotional effects leading to selective attention and avoidance of information, which in aggregate may affect social consensus attitudes regarding appropriate social behaviour. Likewise, social consensus attitudes can affect individual beliefs regarding the social consequences of activity choices affecting activity choices. Hence, there is two-way feedback. In Rabin's model there is a unique equilibrium dependent upon parameter specifications, however, in many situations in which behaviour affects others, there may be multiple equilibria. For example, Hirshman informally discussed the problem of transition between equilibria in the context of development. As both activity choices and social attitudes may depend upon starting conditions, there is potentially an important role for time. The mutual reinforcement of activity choices and information regarding the social consequences of behaviour may make it difficult to move between equilibria.

Although, altruistic preferences can motivate individuals to behave in ways that promote the social good, paradoxically such preferences may lead to socially undesirable behaviour. If altruists have correct information about the social consequences of their behaviour, they will attempt to behave in the public interest. However, due to dissonance induced selective attention to information, information about the social consequences of behaviour may be biased, allowing individuals to behave in a privately materially self-interested manner while erroneously believing that they are behaving in the public good. Hence, analogous to the findings in public good games, as discussed above, while it is common to attribute asocial behaviour to *Homo economicus*, ironically altruistic preferences may take at least part of the blame.

The structure of the essay is outlined below.

Section 2 provides an overview of the relevant literature on the social emotions and the evolution of altruistic preferences including work by evolutionary biologists, psychologists, anthropologists, and experimental economists. The review suggests that there is sufficient theoretical and empirical support for economists to give more serious consideration to the effects of altruism and the social emotions upon behaviour.

Section 3 provides an evolutionary explanation for the phenomenon of selective attention to information in order to reduce dissonance. It is suggested that the phenomenon of dissonance-induced selective attention to information about the social consequences of behaviour is an adaptation that evolved because it facilitated the downward adjustment of group size and hence the sphere of altruism during times of resource scarcity, while allowing the social emotions to facilitate broader-based altruism during good times. Understanding how psychological traits were adaptive in ancestral environments can serve as basis for understanding behaviour in post forager societies in which the scale, complexity and mechanisms of information transmission have undergone major changes.

Section 4 develops an analytical framework which is then used to investigate a role for cognitive dissonance and selective attention to information during transition from traditional to market type economies. The discussion focuses upon the interaction between changes in moral attitudes in the form of religious doctrine and the adoption of market-type activities. The focus upon cognitive dissonance and selective attention provides a basis by which to compare the arguments of Tawney (1926) and Weber (1904-5). The timing of attitudinal and behavioural changes is broadly consistent with the views of Hirshman and Tawney, who stress a direction of causation in which behavioural changes precede attitudinal changes.

Section 5 provides an application to the current environmental problem. Due to high private material costs of changing behaviour and high external costs of continuing current behaviour, the situation is ripe for problems of dissonance induced selective attention to information. The framework prediction that social consensus information should come to understate the magnitude of environmental damages is consistent with empirical findings based upon media content analysis and case studies (Anderson, 1997, Beder, 2002). While the literature attributes such bias to supply side factors including vested interests and media characteristics, the dissonance/ selective attention hypothesis provides a complementary demand side explanation, particularly given that supply side bias and also issue complexity may facilitate selective attention to information.

In section 6, a model is used to investigate some possible effects of altruistic preferences upon information provision and behaviour in the context of environmentally damaging transport choices. The model illustrates possibilities for the mutual lock-in of activity choices and environmental information. A dynamic setting allows for the examination of both the effects of starting conditions upon behaviour and information, and of the relative efficacy of various policies and their timing. The model suggests that the provision of accurate information may have important implications for environmental policy.

Section 7 provides an overview, and makes suggestions for future work.

2 Emotions and Social Behaviour

2.1 The Puzzle of Altruism

The debate regarding the degree to which humans are selfish or altruistic has persisted at least since the Ancient Greeks. In 1859, Darwin provided a theoretical basis for a more definitive answer to the question. According to traditional Darwinism, human nature should be expected to display characteristics that are the outcome of ruthless competition between and selection of individuals that are relatively well adapted to their environment. The theory of natural selection was to strengthen the theoretical basis for employing material self-interest as a causal basis for explaining human behaviour. As a consequence of the individualistic focus of Darwinism, the many observed incidences of "apparent" altruism among humans and also other species became puzzles in need of explanation (Gould, 1990).

The assumption of material self-interest not only has theoretical backing but also explanatory power. Models based upon the assumption are often consistent with empirical observations. However, assumptions and models that generate predictions consistent with the data need not be correct. For example, the under-provision of public goods can either be explained as the outcome of free riding, as predicted under the assumption of material self–interest, or due to the existence of second order altruists who withhold contributions in order to punish low contributors (see section 1.1).

In the latter part of the 20th century, biologists and some social scientists were to increasingly challenge traditional Darwinism and consequently the assumption of individual self-interest employed by social scientists. A new emerging view of human nature suggests that people can be psychologically motivated to behave altruistically, with an important role for the human social emotions. This section provides an overview of the arguments.

2.2 Traditional Explanations of Altruism

20th century biologists proposed three explanations for the many observed incidences of "apparent" altruism; group selection (Haldane, 1932; Wright, 1945; Wynne-Edwards, 1962), reciprocal altruism (Darwin, 1871; Trivers, 1971) and kin selection (Hamilton, 1963, 1964). By the early 1970s, a theoretical and empirical basis for kin selection and reciprocal altruism had been provided while group selection, although deemed theoretically possible, was considered to be unlikely (Maynard Smith, 1964; Williams, 1966; Eshel, 1972). By the 1980s the former two theories had received popular acceptance, while group selection was popularly discredited. Biologists shifted their focus from organisms to genes, where the latter were accorded the status of the fundamental units of selection (Fisher, 1956; Hamilton, 1963, 1964; Williams, 1966; Dawkins, 1976, 1982). According to this "gene's eye view", biological altruism at the level of the individual is explainable as the outcome of selfishness at the level of gene. Similarly, in the case of reciprocal altruism, altruism is only "apparent" from the individual's perspective as gains are expected to exceed losses in the context of repeated interactions.

Economists concerned with a biological basis for human preferences have for the most part accepted reciprocal altruism and kin selection as a theoretical basis for human altruism.[13] As these theories limit the scope of altruism to near kin and to situations of reciprocal interactions, they could readily be accommodated into mainstream economic theory. Examples include the inclusion of the utility of kin in individual utility functions and the substitution of households for individuals as maximizing units. Following Axelrod and Hamilton (1981), reciprocal altruism was incorporated into economic

---

[13] Examples of economic papers that consider group selection include Samuelson (1993), Hirshleifer (1999), Gintis (2000), and Gintis and Bowles (2003a).

rational choice models, whereby individuals take into account the costs of punishment by reciprocators.

2.3 On the Importance of the Social Emotions

Yet, reciprocal altruism and kin selection do not easily provide explanations for the many observed incidences of "apparently" altruistic behaviour directed towards non-kin with no apparent deliberate intent to secure future reciprocity. For example, individuals donate money anonymously to charities and tip waiters whom they do not expect to see again. One explanation is that such "genuine" altruism is an "evolutionary relic" of traits that were adaptive in ancestral environments. In ancestral environments, it may often pay an individual to behave altruistically, given the high likelihood of reciprocation due to repeated interactions (Earle and Johnston, 2000). To contrast, in post forager societies, individuals are often exposed to people with whom they do not expect to interact regularly and so such altruistic behaviour may be maladaptive. Given that the post forager time frame is short in evolutionary terms, there has not then been sufficient time for selection against these putatively maladaptive altruistic traits. This evolutionary relic explanation is reasonable unless it is maintained that reciprocation is a product of rational deliberation; If materially self-interested individuals wholly depended upon reason to decide when to reciprocate in the ancestral environments, they should not be expected to tip waiters, etc. in the current environment.

Recent economic experiments using ultimatum and public goods games (section 1.1) suggest that individual's derive utility from fair outcomes and are often willing to incur high costs to "irrationally" punish non-cooperators. These results suggest that individual behaviour is influenced by more than rational consideration of personal gain. Such empirical results need to be examined with care and have been challenged. Donations to non-kin in situations in which there is no expectation of direct reciprocation by recipients can be explained as an outcome of rational choice with the goal of establishing an altruistic reputation. For example, large tips may signal generosity to friends. Charities often publicize donations. In experimental contexts, hidden incentives such as trying to please the person running the experiment may come into play (Ridley, 1996, p. 136-141).

Despite these criticisms, evidence suggests that emotions, not reason, often provide the driving force for altruistic acts. For example in debriefings following public good experiments, subjects who initially made high contributions explain their subsequent lower contributions as motivated by anger followed by the desire to extract revenge upon low contributors. In ultimatum games recipients irrationally turn down stingy offers apparently to punish stingy donors (Sigmund, Fehr, and Nowak, 2002). These results suggest the need for a biological basis for human social behaviour, beyond that explainable by deliberative self-interested reciprocal altruism and kin selection. However, reciprocal altruism can still provide a basis for altruistic behaviour if agents are modeled as "boundedly" rational.

Simon (1993) argued that individuals who are emotionally wired to behave as "genuine" altruists might have higher fitness than materially self-interested types, where the latter make decisions according to the expected present value of net private material benefits. Due to both the cognitive difficulty of calculating the expected net benefits of altruistic acts based upon predictions of future reciprocation, Simon argued that self-interested types are likely to make mistakes. Similarly, "boundedly" rational types who experience emotional costs (benefits) when behaving in discordance (accordance) with rules of thumb may make mistakes by helping non-reciprocators. However, these emotionally driven altruists may make fewer mistakes than rational self-interested types, raising the formers' biological fitness due to the receipt of more reciprocal lifetime benefits. Hence, prosocial emotions as a motivating force for altruism may be individually adaptive.

Trivers (1971, 1981) argued that emotions have an important role in the facilitation of reciprocal altruism. As free riding can be a matter of degree, it may be difficult for altruists to detect subtle cheating. Hence, subtle cheats may be selected, in turn providing selection pressure for altruists to detect subtle cheats. Altruists may guard against subtle cheats by limiting reciprocation to those who display genuine emotional responses such as gratitude, empathy, and guilt. In turn, the ability of subtle cheats to mimic emotional responses can then be selected, and so on. However, subtle cheats need not give themselves away due to an inappropriate lack of emotional responses, if they actually experience the social emotions. Self-deception, whereby an emotional cheat convinces himself that he is not free riding but is behaving in the public interest, can reduce dissonance and the associated emotions allowing a cheat to avoid detection. Accordingly, Trivers argues that self-deception may be evolutionarily advantageous.

Selective attention to information, as discussed in section 1.3 and 1.4, may facilitate such self-deception. For example, consider food-sharing in hunter-gatherer groups. Suppose that group member X makes a low contribution to the group's pooled food supply. Member Y does not know whether the low contribution is due to free riding (hoarding or shirking) or poor luck. Member Y may also want to free ride by shirking and hoarding. If Y selectively interprets X's low contribution as due to free riding, and his own free riding as punishing X for free riding, he has provided himself with a justification for not sharing.

Although, Simon's altruism serves to increase biological fitness and is hence selfish from an individual's reproductive perspective, it is still "motivationally genuine". In place of rational deliberate planning in expectation of future reciprocation, altruistic emotions serve as a guide. Prediction of emotional responses can serve as a deterrent to free riding and so there is an interplay between emotions and reason.

Adding to Simon's argument regarding the cognitive difficulty of rational reciprocal altruism are reputation effects. A major impetus for the evolution of sophisticated human cognitive capabilities may have been the need to remember and keep score of altruistic behaviour in reciprocal interactions (Trivers, 1971; Nowak and Sigmund, 1998). Further complicating the reciprocal calculus is the possibility of indirect reciprocity, whereby donors direct help towards individuals with altruistic reputations even though they have not directly received help from recipients. Individuals also seek to establish altruistic reputations.

To conclude this section, there is theoretical and empirical support for an important role of the social emotions in facilitating reciprocal altruism and kin selection. Moreover, selective attention to information may serve to reduce dissonance helping to mask emotions allowing cheats to avoid detection.

2.4 Challenges to the Individualistic/ Genic Focus

Based upon kin relations in forager groups and also plenty of possibilities for reciprocation, evolutionary psychologists and social scientists have almost exclusively adopted reciprocal altruism and kin selection as an evolutionary basis for human altruism (Badcock, 2000). This individualistic (or genic) focus has however been increasingly challenged within the field of biology, while group selection, the differential selection of groups, is receiving renewed attention. For example, the very biologists who were responsible for the initial discrediting of group selection have all used group selection to explain a variety of biological phenomena (Hamilton, 1975; Williams, 1992; Maynard Smith, 1995). Furthermore, a mathematical theoretical basis as well as empirical support has been compiled over the past 35 years (for polar reviews, see Wilson and Sober, 1994, 1998; Ridley, 1996).

The social emotions promote cooperation within groups potentially increasing the fitness of groups of altruists relative to groups of free riders. The problem for group selection is to explain how to prevent free riders, who are fitter *within* the group, from taking over the group. Models of group selection illustrate that if competition *between* groups is high relative to competition *within* groups, then cooperative strategies can increase in frequency (Maynard Smith, 1964; Eshel, 1972; Hamilton, 1975; Wilson, 1975). If altruists positively assort with each other, thereby avoiding interactions with free riders, both individual altruists and groups of altruists have higher fitness than individual free riders and groups of free riders. Kin selection (Hamilton, 1975; Wilson, 1975) and reciprocal altruism (Wilson and Sober, 1998) have both been characterized as special cases of group selection.[14] In kin groups, altruists are positively assorted by decent. As

---

[14] Note that this categorization is controversial. Kin selection is a special case of intrademic group selection. Maynard Smith (1976, 1987) and Williams (1992) accept the theoretical possibility of intrademic group selection but do not categorize this as "group selection" based on an apparent lack of group level heredity.

kin share genes, altruists are more likely to inherit altruistic siblings by descent and free riders are more likely to inherit free rider siblings. Groups of altruistic kin can hence be fitter than groups of selfish kin according to Hamilton's rule (1964). Similarly, reciprocal altruism allows for the positive assortment of altruists but depends upon relatively complex strategies whereby altruists are able to recognize free riders and withhold cooperation, for example by ostracism (Gintis and Bowles, 2004). In the presence of second order altruists, free riders can be rationally motivated to cooperate in order to avoid punishment.

Group selection may have particular applicability to humans due to large returns to cooperating in groups for purposes of food procurement, defense, information sharing, and risk sharing. The extension of helping relationships to non-kin can be adaptive from both the individual and group perspective.[15] Anthropologists find that hunter-forager groups did not consist entirely of kin (Wilson and Sober, 1998, p. 134).

Anthropologists initially discounted group selection as likely to be unimportant for humans due to data suggesting that warring in ancestral environmental was too infrequent to allow for strong between group competition. Recently however, Boehm (1997) has argued that low levels of within group competition may have been sufficient to compensate (also see Soltis, Boyd and Richardson, 1995). Ethnographic data suggests that for *Homo sapiens*, within group variability in fitness was reduced by leveling devices both due to innate factors and intentional equality based social traditions. Boehm (1993) describes human ancestral groups as reverse dominance hierarchies in contrast to other primates.[16] Moreover, recent economic experiments (Ozwald and Zizzo, 2001) indicate that the social emotions may serve as "within group leveling devices". Participants were given money and allowed to make bets resulting in unequal wealth distributions. In a burning phase, subjects were allowed to pay money in order to confiscate other people's winnings. 62% of subjects paid money to burn away

---

[15] Associating with non-kin also decreases the chances of inbreeding.

[16] In a dominance hierarchy, there is a "pecking order" with dominant individuals receiving priority access to resources.

the winnings of others with the least wealthy focusing upon destroying the winnings of the wealthiest individuals.

The theory of group selection suggests a rosy view of human nature due to social emotions that promote within group cooperation (Wilson and Sober, 1998, p. 132-194). However, there is a nasty flip side; Although more intense between-group competition selects for altruism towards in-group members, such between-group competition corresponds to "war" and behaviour towards out-group members may be far from altruistic (Hamilton, 1975).

2.5 Summary

The research of theoretical biologists, anthropologists, and psychologists suggest that social emotions provide an important impetus to prosocial behaviour in humans. Anthropologists and biologists have argued that the dynamics of primitive groups may have served to positively assort altruists, allowing for the selection of genes for altruistic behaviour. Selected characteristics include the social emotions and the ability to identify free riders and to selectively exclude them from groups. Such selective exclusion can be broadly categorized as a form of reciprocal altruism even when emotionally driven.

Although the orthodox theories of kin selection and reciprocal altruism, have been categorized as special cases of group selection, this is not a consensus view. Due to the semantic confusion, a case can be made for dropping the term "group selection" entirely. However, viewing human psychological characteristics as at least partially having evolved to promote group welfare provides a different perspective on human nature. To contrast, the term "reciprocal altruism" paints a picture of self-interested rational deliberation in order to secure reciprocal benefits.

3 An Evolutionary Theory of Cognitive Dissonance and Selective Attention

3.1 Some Benefits and Costs of Selective Attention

The social emotions provide an impetus for prosocial behaviour and were selected in ancestral environments because they increased individual and possibly group fitness. Accordingly, psychological costs and benefits are experienced if an altruist believes that his behaviour has social consequences. In ancestral environments, as today, the social consequences of one's actions need not be obvious, providing altruists with incentives to search for information. Selective attention to and avoidance of information may allow individuals to convince themselves that privately materially beneficial, yet asocial behaviour, is prosocial, thus potentially undermining the adaptive benefits of the social emotions. Hence, the evolution of selective attention to and avoidance of information poses a puzzle.

Trivers (1971) proposed that selective attention to information could facilitate subtle cheating by emotional cheats (section 2.3). If a cheat believes that his actions have harmed others, he may experience dissonance and the associated emotional display may reveal his cheating. If however, he selectively attends to information, thereby convincing himself that his privately materially preferred activity choices are prosocial, he may avoid an emotional display.[17] Although self-delusional emotional cheats could potentially coexist with altruists as an evolutionarily stable equilibrium, if such selective attention is too widespread, the cooperative benefits of the social emotions are undermined, so serving to lower group fitness. Furthermore, as selective attention distorts the perception of objective reality, if invoked too readily individuals may make poor decisions. An alternate theory is proposed below.

---

[17] Trivers (1971) discusses developmental plasticity allowing individuals to adapt their level of altruism to environmental and cultural circumstances. However, he also discusses evidence (Krebs, 1970) that altruistic tendencies have a genetic basis.

3.2 An Evolutionary Theory of Context Dependent Selective Attention

During the late Pleistocene, during which *Homo sapiens* evolved, humans were subject to large resource fluctuations. There is disagreement regarding the extent to which these fluctuations were due to open access problems and/or to external factors such as weather and climate. The analysis of skeletal remains provides evidence of severe population contractions occurring at a frequency of approximately every 30 years (Gintis and Bowles, 2004). The picture painted is one in which groups are frequently forced to re-locate, coming into competition with each other for resources. (Gintis and Bowles, 2004). Due to such fluctuating environments, both private incentives and behaviour that is desirable from the group perspective may change over time, complicating the task of deciding which behaviours best promote the public good.

To illustrate how the social desirability of particular behaviours may be context dependent, consider food-sharing. Food sharing is beneficial from both the perspectives of individuals and of groups as it allows for risk-pooling. Groups that share food should be fitter, but only as long as there is sufficient food for the whole group. To give a stark example, assume that a group has enough resources to remain comfortably above subsistence. Although on average individuals bring home enough food, on some occasions they may go hungry. The group can benefit by pooling food, allowing individuals to inter-temporally smooth consumption. For simplicity, assume that an equal sharing rule maximizes both individual and group fitness.[18] Now suppose that there is a severe resource shortage, such that under the equal food-sharing rule, all individuals receive less than the subsistence food level, resulting in widespread starvation. To contrast, if sharing is *unequal*, *some* of the group can survive raising the fitness of both the *group* and those who receive larger shares than under the equal sharing rule.

---

[18] Trivers (1971) states that in forager societies there is no evidence that altruism diminishes if groups are above optimal size. However, warring and ostracism both illustrate a reduction in the sphere of altruism.

One solution to the problem of food scarcity is to reduce group size, so that the group is split into "in-group" members, with whom sharing is continued, and "out-group" members, with whom sharing is reduced or halted. The process can be thought of as one in which group members default from altruism to free-riding with respect to particular group members—now outsiders, while maintaining or even increasing their level of altruism towards remaining in-group members.[19] The sphere of altruism is hence reduced.

It may be objected that reducing group size reduces the number of contributors to the shared food supply and hence the amount of shared food. However, given that in the ancestral evolutionary environment, food was either hunted or gathered with little production, then under conditions of resource scarcity, there was diminishing marginal returns to labour. Adding an additional person to a group need not add much to the food supply but may add substantially to the group's food requirements.

From the group's perspective, it should also be adaptive to reduce the sphere of altruism in a coordinated manner. For example, suppose that there are three group members, A, B and C, with only enough food to maintain two people at subsistence, given pooling. Assume that sharing refers to an equal split of pooled resources. If A decides to share only with B and not with C, while B decides to share only with C and not with A, while C decides to share only with A and not with B, the group will go extinct. If however, A and B both share with each other but exclude C, A and B will survive, while C will not.

The above example illustrates that group fitness enhancing behaviour may vary according to resource scarcity. In good times, food sharing, as promoted by pro-social emotions, benefits the group. In hard times, the private incentives to hoard increase changing the margin between the private material benefits of not sharing and the psychological costs of "free riding". With very real prospects of starvation, individuals

---

[19] Boehm (1996) has argued that the need to pool risk is higher in crisis situations. This might provide selective pressure for increased food-sharing (altruism) towards remaining in-group members.

may default to the privately efficient behaviour—to not share with respect to some group members with whom they previously shared. The cost of such free riding is cognitive dissonance, as behaviour now conflicts with beliefs regarding appropriate social behaviour. Three ways to reduce the sphere of altruism are voluntary fission, ostracism and the breaking of peaceful relations between bands, i.e. war. The latter two methods may induce large amounts of dissonance as individuals are behaving less altruistically towards persons with whom they previously had cooperative or neutral relations. As ostracism of individuals or family groups was often a death sentence, it follows that dissonance could be non-trivial. As these actions, albeit distasteful, are adaptive from the perspective of the group and also of remaining in group members, or of winners in the case of war, we should expect the evolution of psychological traits that promote them. However, the social emotions and induced dissonance may serve to prevent the downward adjustment of group size during hard times. Hence, the social emotions that serve to maintain group cooperation in good times may prove to be a liability in hard times. A solution to this dilemma is that dissonance can be adjusted in a context dependent manner so that it is reduced during times of scarcity but maintained at high levels during good times. Some mechanisms of reducing dissonance in hard times are listed below.

1) Psychological experiments suggest that under situations in which individuals feel that they have no choice, dissonance is reduced (Oskamp, 1991). Such reduced dissonance could be evolutionary adaptive serving to reduce the psychological costs of antisocial behaviour during times of crisis. Meanwhile, higher levels of dissonance may serve to maintain broader based altruism during good times.

2) Individuals that remain within the in-group could increase the level of altruistic acts directed towards each other, inducing positive social emotions that help to counteract the dissonance caused by the asocial treatment of excluded members. A possible example of such compensatory altruism in modern societies is the cinematic depiction of the Mafia boss as a great family man. Individuals that inflict harm upon others are found to

increase altruism towards third parties. Trivers (1971) suggests that such reparations are made in order to reduce guilt and also to establish oneself as a genuine altruist.

3) Making decisions as a group may serve to reduce individual responsibility for harm inflicted by ostracism or war. On the basis of ethnographic data, Boehm (1997) argues that emergency decisions, such as ostracism or war declarations, were often made by group consensus. If a consensus could not be reached, groups would often disband. Context-dependent social norms, understood to be adjustable during times of crisis, may also serve as a guide to group decision making. For example, in modern times killing persons from other nations is socially acceptable in times of war, while not in times of peace. Groups with adjustable cultural social norms might be fitter, helping the groups to survive along with the adjustable social norms. For a discussion of the selection of cultural social norms, see Dawkins (1976, p. 189-201) or Boyd and Richardson (1990).

4) Individuals may selectively attend to or selectively avoid information in order to alleviate dissonance.

5) During times of resource shortages, group identity may contract. Psychologists find that humans form strong group identities and are found to act more altruistically towards in-group members (Akerlof and Kranton, 2000). If the scope of group identity is narrowed, individual's who were previously in-group members become out-group members. Asocial behaviour towards out-group members may now result in less dissonance. Modern history has illustrated that "us verses them" distinctions are exacerbated in times of scarcity. For example, after World War I, Germans were forced to pay large war reparations, which contributed to widespread poverty. Jews, an identifiable sub-group, became scapegoats while the German identity was strengthened. For a more recent example, the attack on the World Trade Centre (Sept, 2001) was followed by a wave of positive American sentiment within America.[20]

---

[20] For a discussion of the relevance of group identity for economic theory, see Akerlof and Kranton (2000). It remains to integrate this work with the literature on altruism.

Points 4 and 5 are expanded upon below. There are two basic ways to free ride in forager societies—hoarding and shirking. Reasons for not sharing include hoarding, laziness, low ability or bad luck. During times of food shortage, incentives to hoard increase, while private food yields fall, reducing the ability to share. In hard times, it may hence be difficult to distinguish between stinginess and low abilities to share. A justifiable reason for not sharing is to punish others for free riding. These considerations give rise to possibilities for selective attention to information whereby X interprets Y's low contributions to risk pooling as free-riding, and X interprets his own hoarding as second order altruism (see section 1.1).

Indirect reciprocity allows for further considerations. Suppose that X reports Y's stingy behaviour to Z. Due to resource scarcity, Z may also be contributing less than normal and would like a reason to not share with Y. Hence, Z also has incentives to selectively attend to X's information and so on. Such "gossip" may allow a core sub-group to selectively attend to information so that Y is excluded from shared food. In such a manner, a group could coordinate with respect to particular individuals or sub-groups, thus reducing the size of the core group and so promoting the group's survival.

Similarly, a group may selectively attend to information in order to justify attacking another group. In both cases of ostracism or warring, small normally tolerated misdemeanors may now become justifications for retributive actions. Furthermore retributive actions may be construed as second order altruism.

Additionally, selective attention to information can potentially facilitate group consensus decision making, allowing the group to coordinate with respect to the ostracism of a sub-group or attack of another group.

3.3 Summary

Cognitive dissonance and ensuing social emotions may serve to promote altruism during good times, allowing groups of altruists to grow relative to groups of free-riders. During times of crisis, it may be optimal to contract group size or to attack another group with whom peaceful relations were previously maintained. Both actions constitute a reduction in the sphere of altruism, potentially causing dissonance. In order for dissonance to serve as a deterrent to asocial behaviour in good times, it follows that dissonance due to "free-riding" in hard times may be non-trivial, potentially thwarting the group's ability to make group-beneficial size adjustments. As a solution, I proposed that dissonance reducing selective attention/ avoidance of information is an adaptation with an increased probability of being invoked in hard times due to an increase in the private material benefits of free riding, relative to the perceived social costs. Selective attention/avoidance of information may hence allow for context dependent adjustment in the level of dissonance, raising the fitness of both the group and remaining in-group members. Other methods for reducing dissonance were also discussed. Some empirical data was found to be consistent with the theory, while formal modeling remains for future work.

4 Dynamics of Behaviour and Information about the Social Consequences of Behaviour

4.1 Information About the Social Consquences of Behaviour in Hunter-Gatherer Societies

Since leaving the evolutionary ancestral environments, human societies have increased in complexity and scale, with a wide array of institutional arrangements. Persons may expect to interact with many persons within their lifetimes, often in a highly impersonal manner. These societies are often subject to changes due to factors such as technological innovations, population fluctuations, resource depletion, wars, and the development of new institutions and social arrangements. The social implications of behaviour may be difficult to determine, particularly during periods of rapid change. Although, as in forager societies, information about social consensus attitudes is still attained by the direct observation of others and interpersonal communication, there is an increasing dependency upon specialized sources of information, including the media, government, clergy and secular experts. Furthermore, improved information storage technologies allows for better communication over space and time. As time progresses, there is an increasing abundance and variety of information sources to which individuals can selectively attend to or avoid. Dissonance and selective attention/ avoidance can arise if individuals adopt activity choices that they *believe* are asocial. What are the implications of selective attention/avoidance of information in order to reduce dissonance in the more complex societies that succeeded the agricultural revolution? Should these psychological traits facilitate or hinder the alignment of private behaviour with the public interest? What are the dynamics of information about the social consequences of behaviour and behaviour and how do they interact? A framework is developed in order to investigate these issues.

4.2 Framework: The Selection of Activity Choices with Externalities

In the next section a framework is developed which will be used to examine some possible implications of altruistic preferences for activity choices, information about the social consequences of activity choices and how these affect each other.

Altruists would prefer to choose activities from which they acquire private material net benefits and also provide net benefits to society. Private material and social interests can be in conflict, creating a tradeoff for altruists. Although an altruist may be willing to incur private material costs to confer a net social benefit to society, if the anticipated dissonance due to negative externalities is lower than the anticipated private net benefits of an activity choice, even an altruist will "rationally" choose the asocial activity choice. A "pure altruist" internalizes all perceived external costs that he believes himself to have inflicted upon others as discussed in section 1.1. A "pure free rider", in contrast, is unconcerned about the welfare of others and is only concerned about private-material net benefits. Intermediate types, between these two benchmarks, internalize some of the externalities that they believe they have imposed upon others.

In spite of good intentions, pure altruists may behave in a socially efficient manner only if they hold correct beliefs about externalities; information may be biased so that it does not necessarily reflect the true or "society's best expert consensus estimate" of the costs or benefits of the externality.[21] After choosing what he believes to be an asocial activity choice, an altruist will experience dissonance because his actions are discordant with his beliefs that he is an altruist. On the other hand, if he believes that he has behaved in a prosocial manner, he will experience a warm glow because his actions are consistent with his beliefs that he is an altruist. Ex post, dissonance or warm glows may induce individuals to selectively avoid or attend to information in order to change their beliefs about external cost or benefits in order to reduce dissonance as discussed in section 1. The aggregate effects of such selective attention affect the demand for information, which can in turn affect the supply of information received at the beginning of the next period. Hence, activity choices and information about the social consequences of these activity choices can potentially co-evolve with two-way feedback.

For simplicity, assume that there are two activity choices, A1 and A2. Let $p_t \in (0,1)$ represent the proportion of individuals in the society who choose A1 at time *t*, where $p_t=1$ is preferable to all other allocations of a social efficiency basis. The social and private net benefits due to an individual *i* adopting activity *j* for a given level of p_t are respectively represented by $\pi_{ij}^S(p_t)$ and $\pi_{ij}^P(p_t)$, where $j \in (1,2)$. Let $D_{ij}(p_t)$ represent the net external costs due to *i*'s decision to choose activity *j*. Hence, the social net

---

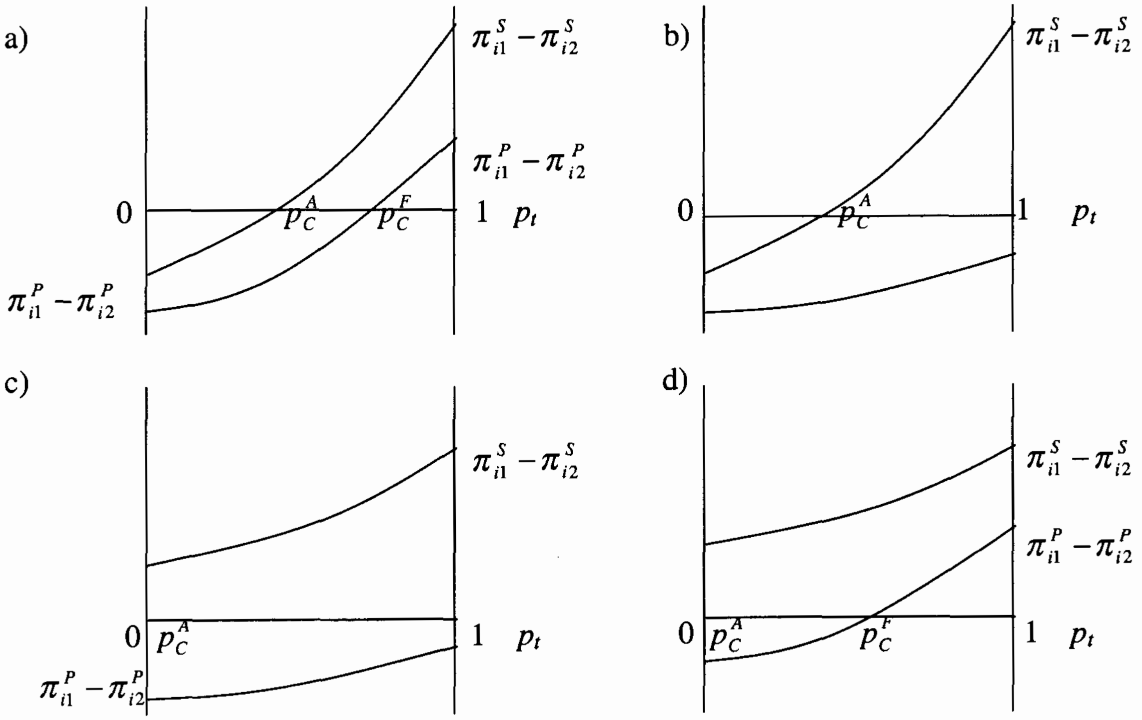
[21] Society may not know the truth about costs or benefits. The "best expert consensus view" might for example refer to the estimates compiled using the scientific methods that are generally accepted by experts. To take a simple example, assume that 20% of the scientists believe that a cost is \$100, while 80% believe that a cost is \$200. A best consensus estimate might be $0.2 \times \$100 + 0.8 \times \$200 = \$180$, assuming that all scientists are of equal caliber. If the *true* value of damages is \$190, the best social consensus estimate would be biased relative to the truth. The social consensus view, that is the average view of a member of society, might differ from the best expert consensus estimate. For example, the media might be biased so that the average person in society (including non-experts) believes that costs are \$120. The social consensus view is then biased relative to both the expert consensus view and the truth. To focus upon bias that is generated from the demand side due to dissonance created due to "consumption" activity choices, the framework focuses upon bias of the "social consensus view" relative to the "best expert view". Accordingly, the "socially efficient allocation" refers to that which would be chosen by unbiased experts.

benefit of *i*'s activity choice is $\pi_{ij}^S(p_t)=\pi_{ij}^P(p_t)+D_{ij}(p_t)$. Let $D_i(p_t)=D_2(p_t)-D_1(p_t)$ represent the net external cost due to *i*'s choice of A2 relative to A1, given p_t. Given p_t, the net social benefit due to *i*'s choice of A1 relative to A2 is the net private material benefit of A1 plus the net external cost of A2 relative to A1.

$$\pi_{i1}^S(p_t)-\pi_{i2}^S(p_t)=\pi_{i1}^P(p_t)-\pi_{i2}^P(p_t)-D_i(p_t) \qquad [4\text{-}1]$$

Four possible situations are represented in figure 4-1. The net social and private benefits of A1 verses A2, respectively represented by $\pi_{i1}^S(p_t)-\pi_{i1}^S(p_t)$ and $\pi_{i1}^P(p_t)-\pi_{i1}^P(p_t)$, are assumed to be increasing in p_t. The latter private effect reflects a positive "user" externality, whereby the net private benefits of A1 verses A2 increase in the number of individuals who use A1. For example, if more people use public transport (the case to be examined in section 6) the private material net benefit to other users increases due to greater efficiency of service, for example more frequent service, more routes and lower gasoline costs per mile.

Figure 4-1: Net Private and Social Benefits of A1 verses A2



For simplicity assume that individuals are identical with respect to all characteristics except for the level of altruism. Figures 4-1 a and b represent situations in which it is *socially* desirable for a given individual to choose A1 if $p_t > p_C^A$ or A2 if $p_t < p_C^A$. The superscript 'A' represents "altruist" indicating that a pure altruist with perfectly unbiased information about damages chooses A1 if $p_t > p_C^A$. p_C^A represents the critical mass proportion of A1 players, above which, given the behaviour of others, it is socially desirable for a given individual to choose A1. In cases c and d, it is always socially desirable to play A1, independent of the choices of others, so that $p_C^A = 0$; If their information about social payoffs is unbiased, pure altruists choose A1, independent of p_t.

Let p_C^F represent the critical mass proportion of A1 players necessary to induce pure free riders to choose A1. In cases a and d, $0 < p_C^F < 1$, so that if $p_t > p_C^F$, given the behaviour of others, each player privately materially prefers A1 (or A2 if the equality is reversed). In cases b and c, each player privately materially prefers A2, independent of p_t so that if the society consists entirely of pure free riders, all choose A2. For any p_t, where $p_C^A < p_t < p_C^F$ (a and d) or $p_C^A < p_t \leq 1$ (b and c), there is a prisoner's dilemma.

For intermediate types, the point of indifference between the two activity choices lies between p_C^A and p_C^F.

Actual behaviour depends upon information about the social consequences of activity choices. For example, if individuals receive information that the net external costs of A1 do not exceed those of A2, then they will believe that the social net benefits are identical to the private material net benefits. In this case, altruistic types will behave like free riders and $p_C^A = p_C^F$. Information could either overstate or understate net damages, $D_i(p_t)$. If information understates (overstates) damages, then p_C^A will fall below or exceed its level in the case of unbiased information.

4.3 Initial Conditions and Activity/ Information Lock-in

Figure 4-1 illustrates the potential for multiple equilibria with sensitivity to initial conditions. Consider an initial situation (not represented in figure 1) in which both the socially efficient and privately efficient allocation is *p=0*, where all players choose A2. Now suppose that changed circumstances alter social payoffs, changing the socially efficient allocation from $p_t = 0$ to $p_t = 1$. In cases a and b, even in a society consisting entirely of pure altruists, given the initial condition, $p_t = 0$, the society "locks-in" to the socially inefficient equilibrium ($p_t = 0$). In cases c and d, depending upon the number of altruists or extent of altruism and beliefs about social payoffs, the society may remain at the $p_t = 0$ equilibrium. In case d, sufficient conditions for social efficiency are that the proportion of pure altruists exceeds p_C^F and there is perfectly unbiased information about social payoffs. In case c, given perfectly unbiased information, pure altruists and free riders respectively choose A1 and A2 so that social efficiency is only attainable if all players are pure altruists.

4.4 Application: Differential Growth Rates

The framework can be applied to the problem of differential growth rates between societies, both throughout history, and today. For example, why did the market-based growth-takeoff of the past millennium occur in Northern Europe? Why do many countries in the 3rd world remain poor?

Let A2 and A1 represent activities respectively found in traditional and market based economies.[22] Figure 4-1 illustrates some potential difficulties of making the transition to market based economies given initial conditions and the fact that private payoffs depend upon the behaviour of others. A further complication arises due to beliefs about the

---

[22] Note that the terms "market activities" and "capitalistic activities" are used interchangeably, where "capitalistic activity" refers to activities found in capitalist economies. These include not only the practices required to support large scale capital projects, such as legal markets for borrowing an lending at market-determined interest rates, but also practices such as market-determined prices for consumption goods, and increased trade in contrast to barter and community planning.

social consequences of behaviour. In traditional economies, beliefs about the social merits of community sharing, fair trades, price ceilings, etc. may help to reinforce A2 behaviour. Such beliefs may have become established over time after people observe the adverse short-term social consequences of deviating form traditional practices. For example, in the medieval times, people were well aware that market determined prices could have extremely adverse consequences on the poor in times of scarcity. Starting in a traditional economy, in which all players choose A2 so that $p_t = 0$, assume that there is an exogenous change such that the new private and social payoffs are illustrated as in figure 4-1 c or d. In a society consisting entirely of pure altruists with perfect information about externalities, as the critical mass threshold for pure altruists is $p_C^A = 0$, all individuals switch to A1 because it is both socially and privately preferred. However, even if people are pure altruists, but continue to incorrectly believe that A1 is asocial, they will continue to play A2, so that the system remains at the $p_t = 0$ equilibrium. People may continue to incorrectly believe that new practices are asocial due to previous observations in which experiments with such practices led to adverse social consequences. Social theorists (such as, the clerical scholastics in the Middle Ages) may also have developed theoretical support for past practices and it may take time to develop social theory appropriate for the new circumstances.

In the case of transition to market economies, various factors may contribute to inertia in changing beliefs about the social appropriateness of market activities. Research suggests that in both ancestral evolutionary environments and post-ancestral traditional economies, free riding, in both the senses of not sharing and shirking, is considered to be a social vice. Ethnographic data on modern hunter-gatherer societies is used to make inferences about behaviour in ancestral environments (for example, see Boehm, 1993, 1997). For example, in hunter-gatherer societies, stinginess is often subject to social ostracism. In post-ancestral traditional economies, such as feudal Europe or in today's less developed countries, examples of not sharing include activities such as capital accumulation, charging high interest on loans or high prices to those in extreme need, low wages, and the privatization of property. These activities are often considered to be social vices, for example, in Feudal Europe there were prohibitions against usury and schedules that set out rules for just prices. In the Middle Ages, the Catholic Church characterized such usurious activities as avarice, a deadly sin, punishable by fines, jail sentences, penance, physical punishment, and excommunication. If, as suggested by Boehm (1993, 1997), (see section 2.4), there is an innate basis for an "egalitarian ethic", then there may be an innate basis for finding such capitalistic activities morally distasteful.

Once social norms, such as prohibitions against usury, are established, they can be reinforced by punishments, either by an authority such as the Catholic Church, (as in the above example), or less centrally by the community at large, (as in the case of social ostracism). Such punishments then affect the private material benefits of capitalistic activity choices and are an example of a "user" externality, whereby the severity of the punishment depends upon the number of users. Additionally private material benefits are affected by the development of institutional structures, such as systems of property rights, which co-evolve with capitalistic activities. These institutional structures also provide an example of "user" externalities—the more capitalistic activities, the more supportive institutional structures and the greater the benefit of capitalistic activities. Moreover, as social welfare systems develop in response to negative social effects of increasingly capitalistic economies, (partially due to a break down of traditional social welfare institutions), adverse consequences due to markets may fall also. In short, both net private benefits and net external benefits of A1 verses A2 may depend upon the proportion of the society involved in capitalistic behaviour, p_t.

Hirshman (1965) used the theory of cognitive dissonance and selective attention to argue that attitudinal changes are likely to succeed behavioral changes. Hence attempting to change attitudes in order to promote "development" may be ineffective or costly. He argues that market-type behaviour and institutions will be adopted, given enough time.

> The latecomer will stumble more or less absent-mindedly into such behaviour as pursuit of individual profit, entrepreneurial risk-taking, ...and will then gradually lead to those changes in attitude and basic beliefs which were thought to be prerequisites to the just-mentioned modes of behaviour. The art of development may therefore consist

> primarily in multiplying these opportunities to engage in these dissonance-arousing actions and in inducing an initial commitment to them (Hirshman, 1965).

Transition to market economies need not be inevitable because, firstly, private material payoffs depend upon the behaviour of others, and secondly, due social consensus information, altruists may erroneously continue to believe that current activity choices serve the public good, so thwarting behavioural changes. Moreover, individuals may selectively attend to information that serves to morally validate current activity choices, so affecting social consensus attitudes. Hence information and behavioural choices may mutually reinforce each other, so facilitating the lock-in of each.

4.5 Framework to Investigate the Dynamics of Activities and Information

Given that social consensus information about the social consequences of activity choices may serve to reinforce current behaviour and vice versa, there is an interesting problem of the dynamics of information and behaviour. Figure 4-2 provides a framework with which to examine these dynamics. Note that time subscripts are referred to in the discussion but not in the text in order to illustrate the circular flow nature of the diagram. The boxes in the diagram are numerically labeled to facilitate comprehension. Arrows are referred to according to the number on the source and terminal boxes. For example, arrow 1-3 refers to the arrow that starts at box 1 and ends at box 3.

Figure 4-2 is described as follows. A given individual (*i*) chooses an activity, based upon beliefs about the private material and external consequences of each activity out of the set of possible activities (box 1). Individual *i*'s activity choices have private material consequences for himself (arrow 1-4) and external consequences to others (arrow 1-3). Similar to individual *i*, the activity choices of others (box 2) affect *i*'s private material benefits (box 4) through the user externality (arrow 2-4). For example if many people are lending at market interest rates, there is less chance that *i* will be punished for doing so. Secondly, the proportion that chooses each activity may affect the magnitude of *i*'s user externality if its size is sensitive to the number of users (arrow 2-3). The upper box (I) is referred to as "reality" because it indicates "actual" or "true" payoffs.

Box II, indicated by "Beliefs of *i*", represents individual *i's* beliefs about the private material net benefits of his own period *t* activity choices given the behaviour of others (box 5) and his beliefs about the external net benefits of these choices (box 6). It is assumed that the number of individuals who will choose each activity is accurately predicted by all individuals. In order to focus upon the formation of beliefs about external consequences, assume that beliefs about private material benefits are correct. At the end of period *t*-1, prior to choosing activities in period *t*, agent *i* forms beliefs about the social consequences of each activity choice. To do so, he makes use of social consensus information about the external consequences of activity choices (box 7). Such social consensus information may be provided by an external authority such as the church and also by talking to others to find out the average attitude of neighbors in the community. This social consensus information is represented as I_{t-1}^{P}, where the subscript "*P*" represents "pooled".

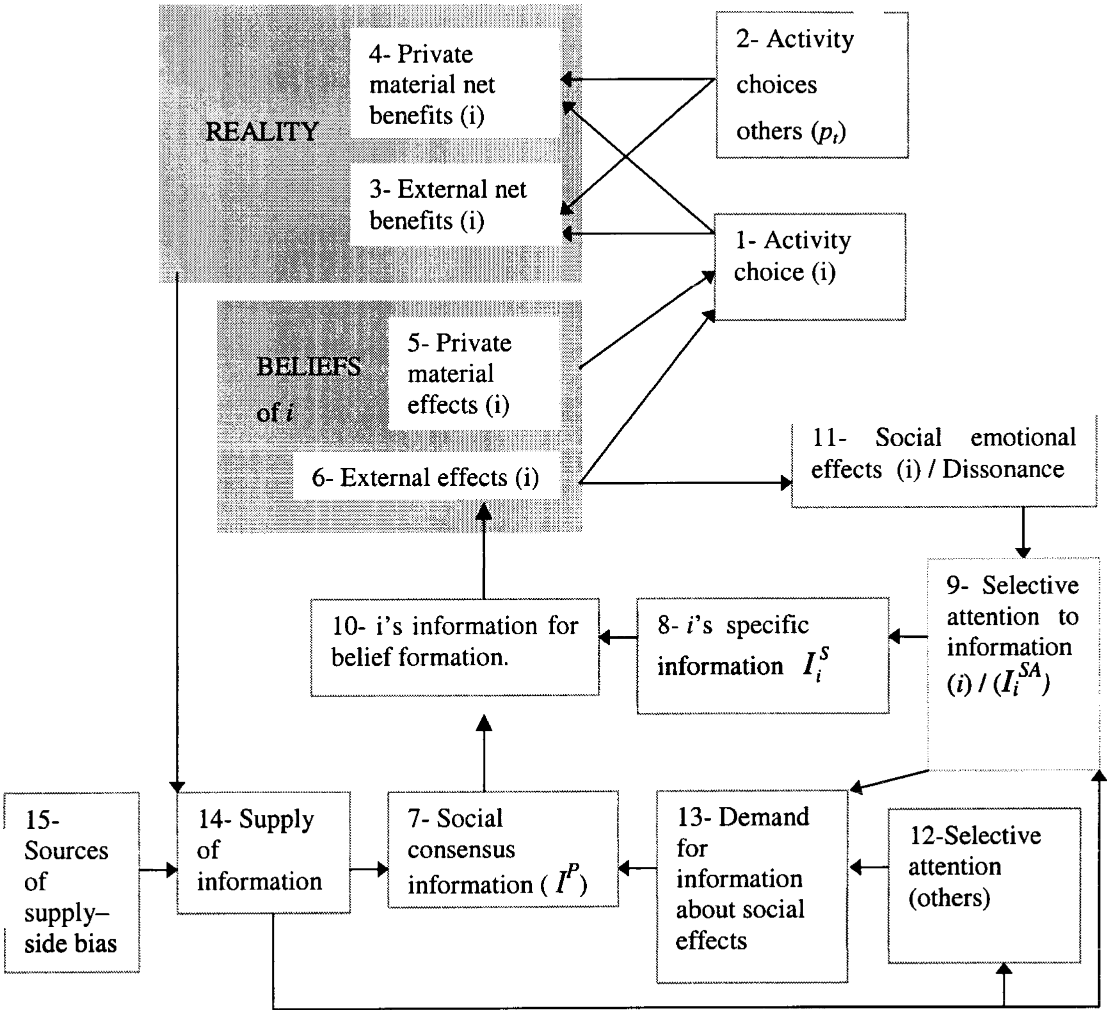
There may also be an individual specific component to such information, referred to as $I_{i,t-1}^{S}$ due to factors such as exposure to individual specific information sources including personal prior observations (box 8). For example, the individual may be a member of subsets within the society, which are exposed to different information sources. For example, literate people may have access to different information sources than illiterate people. Moreover, the individual may have selectively attended to information after choosing period *t*-1's activity (box 9) and this information may be retained in memory to affect the specific component of information utilized in period *t* (arrow 9-8). The individual specific and pooled information jointly determine the information that *i* uses to form period *t*'s beliefs, referred to as $I_{i,t-1}$ (box 9), as indicated by arrows 9-8 and 7-8. $I_{i,t-1}$ is then used to determine *i*'s period *t* beliefs about external consequences (box 10, arrow 8-10).

Individual i's beliefs about external net benefits are assumed to be formed at the beginning of period t based upon information received by the end of period t-1 (arrow 10-6). These beliefs are used to estimate the dissonance or warm glows that will be received from each of the period t activity choices. After belief formation, individual i maximizes private utility by choosing the activity which yields the highest possible net private benefits, which include both private material net benefits and also the expected dissonance costs or benefits from warm glows. These beliefs affect period t activity choice as indicated by arrows 5-1 and 6-1. Aggregating such decisions across all individuals determines p_t, the proportion of individuals who choose A1 in period t. In aggregate such decisions determine p_t, the proportion of individual's who choose A1 in period t.

Given beliefs about external effects (box 6), after choosing period t's activity, individual i may experience psychological costs and benefits due to the social emotions (box 11), potentially inducing selective attention to information (box 9) in order to reduce dissonance or receive warm glows (arrow 11-9). The degree of selective attention depends upon the degree of dissonance or warm glows based upon the magnitude of the believed external effect, the individual's level of altruism, propensity to evaluate information truthfully, and the nature of the information available.

The aggregate effects of selective attention by individual i (box 9, arrow 9-13) and of others (box 12, arrow 12-13) may contribute to the aggregate demand for information about the social consequences of activity choices at the end of period t (box 13). The aggregate demand for information interacts with the supply of information (Box 14) to determine the social consensus information (I_P^t) (box 7) that will be used to form beliefs at the start of period $t+1$ (arrows 13-7 and 14-7). The information supply is determined by information supplier's interpretation of the true consequences of activity choice (reality) (arrow I-14), and the extent of supply side bias. Bias may occur due to factors such as vested interests, the medium of the information source, and selective attention on the part of suppliers, indicated as "sources of supply side bias" (box 15, arrow 15-14). The social consensus information at the end of period t is used to determine beliefs about external consequences in the next period *t*+1 (through arrows 7-10 and 10-6), in turn, affecting period t+1 activity choices, as before. The process is repeated. Note that an individual's activity choice affects others *directly* through the "activity" externality and potentially *indirectly* through the effect upon social consensus information.

Figure 4-2: Framework of Dynamics of Activities and Information



4.6 The Effects of Heterogeneous Agents

Given the payoff structure in figure 4-1, with homogeneous agents and the starting condition, $p_t=0$, society may lock in to the traditional economy, where all choose A2. However, in reality individuals are heterogeneous due to factors such as variation in altruistic tendencies, private material net benefits, propensities to selectively attend to information, and exposure to information regarding socially acceptable behaviour. Membership to various groups, specialized skills, and innate aptitudes contributes to such variation.

Given the initial conditions, $p_t = 0$, and social consensus information that provides moral validation for A2, consider a change that increases the private material net benefits of A1 for a subsection of the population, that is, market-activities become more desirable. Although information may remain locked in, serving to socially validate traditional A2 activities, particular individuals or groups may choose A1 based upon factors such as individual specific high material private benefits, low levels of altruism, and isolation from social consensus moral information. If some of these A1 players have at least somewhat altruistic preferences, they may experience dissonance creating a demand for information regarding the social merits of A1. If a supply is induced, the increased availability of A1 validating information may increase the chance that others adopt A1. Activity choices and socially validating information can hence co-evolve, reinforcing each other in a positive feedback loop.

4.7 Cultural Group Selection

Recent efforts to explain long run growth emphasize multiple causation, and sensitivity to starting conditions based upon non-linearities (Rosenberg and Birdzell, 1986; Mokyr, 1990; Landes, 1998). Explanations regarding why one country grew faster than another may be attributable to a peculiar coincidence of circumstances. On the other hand, broad features of the environment may serve to increase the probability of such a chance coincidence (Hirshman, 1965; Lipsey and Carlaw, 2003). Hayek (1979) notes that market type activities may emerge by chance through "a process of winnowing and sifting, directed by the differential advantages gained by groups from practices adopted for some unknown and perhaps purely accidental reasons" (sited by Caldwell, 2000). Hayek (1988) commenting upon the adoption of practices favorable to the emergence of market-based economies, writes:

> ...the extended order resulted not from human design or intention but spontaneously: it arose from unintentionally conforming to certain traditional and largely *moral* practices, many of which men tend to

> dislike, whose significance they usually fail to understand, whose validity they cannot prove, and which have nonetheless fairly rapidly spread by means of an evolutionary selection—the comparative increase of population and wealth --of those groups that happened to follow them (emphasis in the original, sited by Caldwell, 2000).

In terms of the framework, in a very populated society, in which individuals are fairly homogeneous with respect to tastes and environment, including exposure to information sources, it may be difficult to achieve the critical mass in order to enter the basin of attraction of the market economy. Hayek draws attention to the possibility that it may be easier for subsections of the population to coordinate upon A1 activities. If figure 4-1 is interpreted as applying to groups instead of the population at large, then given the smaller population of groups it follows that a smaller number of individuals need adopt A1 in order to achieve critical mass. Furthermore, within groups, it may be easier to coordinate upon a new within group social consensus regarding the social consequences of behaviour. Isolation from society wide information sources may shield the within group social consensus.

Groups that adopt market-based activities and supporting moral attitudes can grow faster in terms of population and wealth, potentially allowing for greater political power. As these groups grow, the group consensus information that provides moral validation of A1 may then spread to other groups, increasing the probability that these groups also adopt A1.[23]

The following section provides an historical case study that examines changes in information about the social appropriateness of behaviour and activity choices during transition from traditional to market based economies in Europe. The analysis attempts to determine whether the timing of informational and activity changes are consistent with the framework developed above. Two sources of information—religious and secular, are discussed.

---

[23] Zywicki (2000) notes that Hayek's ideas about cultural group selection were discredited along with biological group selection. He suggests that Hayek's ideas be revisited given the current reconsideration of group selection.

4.8 Application: Religion and Market Activities in the Late Middle Ages

4.8.1 Religion as a Source of Social Information

Religious services provide information about the morality of behaviour. Expectations about private material benefits and costs of behavioural choices, such as afterlife punishments and rewards, and social inclusion, serve as incentives to comply with religious behavioural prescriptions. It then *appears* to follow that if people need to be told how to behave in a pro-social manner, that this may not be their natural inclination. However, given at least somewhat altruistic preferences, religious services also serve to satisfy a demand for information regarding appropriate moral conduct as motivated both by social emotions and the private material benefits of complying with consensus social norms. Accordingly, religions emphasize love of others, use guilt to induce moral action, and emphasize "warm glows" from good works.

Religion was an important source of moral information during the Middle Ages, the period immediately preceding the European growth takeoff. To investigate the role of moral information during the transition from predominantly feudal to market based economies, religious attitudes should then be a key point of focus. Rosenberg and Birdzell (1986, p. 128) write:

> At least in the sixteenth and seventeenth centuries, the source of morality had to be religion...The social teachings of the Catholic Church came from the Middle ages. During the Middle ages, the custom of the manor rigidly prescribed the terms of manorial economic relationships, and in the towns the rules of the guilds were nearly as comprehensive. A morality inherited from a medieval economy based on faithful compliance with customary relationships could not have been expected to fit a commercial economy in which individual choice and bargaining had superceded custom as the basis for exchange, and which was eventually to displace most of the customary relationships. The prohibition of the charging of interest is the most often cited example of Church doctrine running head-on into the needs of a rising merchant class. But something more important was missing; a moral outlook that would facilitate, encourage, and legitimize the rising world of market relationships.

The quote draws attention to the potential difficulty of changing moral attitudes to accommodate new and potentially welfare enhancing behaviours.

In terms of the framework developed in sections 4.2 to 4.7, religious services provide information regarding the social appropriateness of behaviour. For simplicity, assume that there are two choices of religious services, R1 and R2, which respectively provide information morally validating either of the two activity choices, A1 and A2. A1 activities include practices such as the market determination of prices, the ownership of private property, and markets for borrowing and lending for the purposes of consumption smoothing and/ or investment. A2 activities include practices such as Church determined price schedules, common land holdings, informal consumption smoothing through both informal reciprocal relationships with neighbors and more formal pooling with the Church serving as an intermediary, and usury prohibitions.[24] Starting out in a traditional economy, the majority of people adopt A2 activities and attend R2. Due to changed circumstances, the net social benefits of A1 increase relative to A2. However, given starting conditions, it remains privately desirable for most individuals to continue to choose A2, creating a high demand for R2. For *some* individuals however, the high private material net benefits of A1 verses A2 may induce a switch to A1, but at a cost of dissonance because R2 labels these activities as amoral. Dissonance may induce a demand for R1 information, potentially inducing a supply. The increased availability of R1 may induce others to adopt A1. A1 and R1 can hence potentially co-evolve.

According to Hirshman's argument that attitudes are difficult to change in the absence of behavioural changes, some degree of adoption of market-type activities should predate changes in attitudes regarding their morality. If, also, moral validation of behaviour is an important factor influencing human decision making, then changes in private material incentives may be necessary but not sufficient to affect transition to a market-

---

[24] For a discussion of the consumption smoothing devices available in the Middle Ages, see Reed and Bekar, 2003.

oriented economy. The framework of sections 4.2 to 4.7 suggests that due to positive feedback loops between behaviour and morally validating information, for economies that do manage to make the transition, changes in information regarding the morality of market activities may be large and rapid.

4.8.2 Changes in Behaviour and Religious Doctrine in the Late Middle Ages

That a moral basis for market-type activities is a necessary condition for transition to market-type economies is consistent with the broad historical facts. By the late Middle Ages, many economies, including the ancient Greeks and Romans, the Islamic World, China, and Western Europe had experienced substantial capitalistic activities, including a merchant class, international trade, banking, and money loaning. However, these activities were limited in scope, in comparison to the growth takeoff that was to emerge out of the 16th century in Britain and Holland. Consistent with the model, throughout the latter half of the 15th century and 16th century, information regarding alternatives to orthodox Catholicism was on the increase in Europe (Johnson, 1976, p. 269-279). In 1517, the German Monk, Martin Luther addressed his 95 theses outlining corruption in the Church to the Archbishop of Mainz.[25] The critique protested a new issue of indulgences that were to finance both the rebuilding of the St Peters Basilica in Rome and the Archbishop's debts to the Fugger bankers (Green, 1998, p. 16). The new source of religious information, Lutheranism, broadly appealed not only to the urban bourgeoisie, but also to peasants, nobles, and kings and was to spread rapidly throughout Germany and into other parts of Europe (Green, 1998, p. 24). Other variants of Protestantism, such as Calvinism, developed elsewhere. The rapid adoption of this new

---

[25] Some late medieval Church "corruption" may have been altruistically motivated. An altruistic Pope should aim to maximize the number of souls saved, for example by taking actions to maintain or increase the power of the Church. Increasing the moral requirements for attaining salvation makes it easier to sell indulgences and relics. Large expenditures upon ornate church property and clerical costumes can be justified as a means to convince the laity of the power of God, thereby increasing compliance to church doctrine, so in turn saving souls. Given that such practices are morally questionable, but in line with the private material interests of the clergy, there may have been some clerical selective attention to information. Another explanation of large church expenditures is the "hostage capital" argument of Klein and Leffler (1981).

religious information, particularly in urban centres, suggests that there was already a bottled up demand. I will attempt to evaluate whether and to what extent this revolution in information was a response to a bottled up demand for information that would provide moral validation for market-type activities.

It is historically well established that there was a large change in private material incentives inducing an increased amount of capitalistic behaviour both prior to and coincident with the spread of Protestantism (Tawney, 1926, p. 66-79). This included an earlier commercial revolution in the 12th and 13th centuries, followed by population decline due to plagues in the 14th. The 15th century was a period of recovery, with a gradually increasing population, and a commercial recovery, which was eventually to flow into industrial capitalism. Rosenberg and Birdzell (1986, p. 37) write "...the latest period in which we can view the institutions of medieval Western society as functioning in an approximately normal way is the thirteenth century, and even that may be too late a date for Italy." After the 14th century disasters, plagues and population reductions, "...the fifteenth century was a period of recovery, but recovery did not take the form of a return to medieval institutions."

Factors affecting private material incentives included technological innovations in agriculture, metallurgy, the military, mechanization due to the waterwheel, and transport technologies such as three-masted ships (Mokyr, 1990, ch. 3). Secular power was becoming increasingly centralized, largely due to military innovations, which served to increase the optimal scale of armies. Given frequent wars, nobles in need of finance had incentives to accommodate business activities as a source of tax revenue (Rosenberg and Birdzell, 1986, p. 63). The Church competed with the crown for the tax base, using some of the revenue to provide social services. In times of war, it is prudent to shift resources from social services towards military expenditure. A leaner reformed church was in the interest of kings who engaged in activities such as appropriating Church lands, in turn decreasing the ability of the Church to provide social services. Reed and Clegg (1994), discussing the English case, argue that people predicting the appropriations correctly deduced the increased risk of loaning to the church, for example by buying annuities in order to smooth consumption, and decreased their loans accordingly. Lacking finance, the quality of church social services further deteriorated and the demand for private alternatives increased. Other factors included the increasing use of money to purchase land, the break down of the liege system, population expansion with a consequent surplus of free wage labour, the growth of inter- European and global trade, and the continued expansion of urban centres. Although the above listed factors are incomplete, they do help to paint a picture of the Late Middle Ages. This was a period of unprecedented change with the private sector expanding in size and wealth while the Church's ability to provide social services and to accommodate doctrine to the new conditions was under strain.

4.8.3 The Demand for New Moral Codes in the Late Middle Ages

Tawney (1926, 97) argued that there existed a strong need for new moral attitudes regarding capitalistic activities in the late Middle Ages. These activities had created a severe strain on the moral teaching of the Catholic Church, which had supplied a morality that had evolved to accommodate a predominantly agricultural feudal economy. Although, Catholic doctrine had by the late Middle Ages deviated substantially from the original written interpretations of the teaching of Christ, these were also suited to traditional primarily agricultural economies.

A common feature of ethical codes, independent of the time period is that they promote contribution to the public good by discouraging the two basic forms of free riding, shirking and not sharing. Catholicism discourages shirking by both appealing to people's sense of ethical duty, both rationally and emotionally, and by including good works as an essential determinant of "justification". Medieval Catholicism categorized non-sharing activities as avarice, a deadly sin.[26] Avaricious activities include hoarding, the charging of interest, and extortionary pricing. Accordingly, the Church officially

---

[26] "Justification" refers to the process of becoming acceptable for salvation, which is a Christian requirement for entry into heaven.

discouraged and sometimes prohibited the charging of interest and laid out rules regarding the setting of fair prices and wages (Dempsey, 1935; De Roover, 1958).[27]

Small-scale agricultural economies facilitate the workings of reciprocal altruism. Charging interest during times of hardship is viewed as exploitation, working against the spirit of community sharing. Loans made in kind reduce the need to worry about nominal interest. Urbanization strains the system of reciprocal altruism. The use of money and existence of banking services, that developed during the commercial revolution of the 12th and 13th centuries, made it easier to hoard and thus avoid detection for stinginess. Money, is an excellent store of value, in comparison to perishable food, although it does deflate. Furthermore, with a reduced guarantee that the community, especially, in later times, given the weakening of church social services, will bail out individuals in hard times, the incentives to take personal responsibility for one's own consumption smoothing increase. Not only is it easier to hoard, but there are also greater incentives to hoard. Furthermore, the late medieval devaluation of currency made it reasonable to charge nominal interest.

With the increasing commercialism of the Late Middle Ages, the line between usurious and reasonable business practice became difficult to determine. Businessmen and also clerical scholars, such as the Schoolmen attempted with difficulty to classify clear-cut cases of usury (Tawney, 1926, p. 42-55). In earlier times, cases of usurious behaviour had been easier to identify as communities were smaller, and loans had primarily been made in times of need in order to smooth consumption, not to finance investments. Church arguments against usury focused upon fairness, the protection of the poor, and appealed to the scriptures.

The charging of interest on lending and borrowing was advantageous from the perspective of the emerging business classes. For example, merchants need to pay suppliers on delivery, making it advantageous to borrow until merchandise is sold. It is

---

[27] For a chronology of usury restrictions, see Reed and Bekar, 2003.

easier to attain financing for large-scale capital projects, such as shipbuilding, mining, and land purchase if the borrower agrees to pay interest. It is hence not difficult to create arguments regarding the social merits of charging interest. For example, it was acceptable that a partner in a business venture took his share of profits, provided he shared the risks (Tawney, 1926, p. 42).

However, charging interest can have adverse consequences for the poor; a starving person with no option other than to borrow at a high interest rate, effectively pays extortionate prices for current consumption. Clyde and Bekar (2003) argue that the existence of formal markets for lending and borrowing at interest attracted money from wealthier individuals. This, in turn, reduced their probability of participating in informal pooling, through networks of relatives and neighbors, and formal pooling, by contributing to the Church in good years with the assurance of receiving payments in bad years. Reed and Bekar provide evidence that the development of such "capital markets" coincided with the implementation of usury restrictions in the 13th and 14th century.

Similarly, arguments can be provided regarding the merits of market determined prices. For example, during a shortage, high prices provide a signal for increased production. However, negative effects upon the poor are also readily discernable. During a shortage the poor may not be able to afford basic commodities while the rich overindulge. Tawney (1926, p. 41) discussing late medieval attitudes about prices writes: " 'To leave prices of goods at the discretion of the sellers is to give rein to the cupidity which goads almost all of them to seek excessive gain.' Prices must be such, and no more than such, as will enable each man to 'have the necessaries of life suitable for his station.' ".[28] The late medieval debates regarding the morality of market activities can be readily

---

[28] De Roover (1958) argues that this view of just price is inaccurate writing; "According to the majority of the doctors, the just price did not correspond to the cost of production as determined by the producer's social status, but was simply the current market price, with this important reservation: in cases of collusion or emergency, the public authorities retained the right to interfere and to impose a fair price." The "doctors" refers to the clerical scholars, such as Thomas Aquinus.

appreciated today, for example there is still heated debate over rent control, privatization of basic services, and minimum and union wages.[29] In the 3rd world, there is often popular support for leftist governments.

In spite of late medieval official prohibitions on usury, in practice there was often a double standard with relaxed application to nobles and upper level clergy relative to lesser merchants and free artisans. "No reasonable judgment of the medieval denunciation of usury is possible, unless it is remembered that whole ranges of financial business escaped from it almost altogether. It was rarely applied to the large-scale transactions of kings, feudal magnates, bishops and abbots...Popes regularly employed the international banking-houses of the day...and sometimes forced the repayment of their debts by threat of excommunication (Tawney, 1926, p. 44-45)." Tawney argues that in spite of this double standard, as leniency towards usury was not systematically applied, the Church was only partially able to accommodate the needs of the business class. Furthermore, those in violation of Church rules, not only had to contend with the psychological consequences of social emotions, but also with fear regarding the status of their salvation, potential excommunication, and social ostracism.

Tawney's arguments that church policy created a frustrating environment for the emerging business classes may be disputed. This is particularly so, given that usury prohibitions were relaxed during the 16th century, a period in which commercialization was proceeding at a rapid pace. For example, Reed and Bekar (2003) argue that usury restrictions were primarily enforced upon consumption loans and that "church rhetoric" regarding usury on business loans did not hold in practice. They write:

---

[29] In the Late Middle Ages, fair wages were not a key point of focus for social theorists concerned about fair prices (Tawney, pg. 268). Slow to appreciate the role of a new growing wage earning proletariat, they failed to accommodate it into social theory. Moreover, he provides evidence that the bourgeoisie held an attitude of contempt towards wage earners writing that the bourgeoisie "appeared at times to consign to collective perdition almost the whole of the wage-earning population".

> The Church was constrained to argue from the basis of natural law, which made it difficult to draw a credible distinction between consumption and investment lending. A further rationale follows from our general contention that the primary motivation for the Church, with respect to usury laws, was concern for the poor. Church doctrine simultaneously offered salvation in exchange for charity and damnation in return for usury. Yet loopholes were offered for investment lending. The result was that the merchants and bankers who engaged in investment financing (at interest) were strongly motivated to offset the damnation effect of this activity with compensating contributions to charity. Allowing lenders a clear conscience with respect to usury would have negated this result.

Tawney also makes extensive note of this double standard, however he argues that the relaxed standards were limited mostly to the worlds of high finance (upper level clergy, nobles, and larger-scale business capitalists). The frustrated business classes to which Tawney refers include small business people, primarily in urban centres, including craftsmen, merchants, and traders. Furthermore, Tawney notes that usury prohibitions were unevenly applied. For example, small business people and also local clergy at distances far removed from Rome were less likely to be aware of the usury loopholes for investment loans. Similarly, today, small business owners are less likely to participate in offshore tax shelters as large-scale multinational companies. As a further point, the existence of charitable donations that were required to offset the damnation effect of engaging in usury, (as mentioned by Reed and Bekar), suggest, in support of Tawney's claim that charging interest, even on business loans, was widely held to be a morally dubious activity. Indeed, the requirement of having to make such charitable donations was privately materially costly.

4.8.4 Catholicism and the Moral Validation of Unequal Wealth Distributions

In spite of the restrictions upon business, the Catholic Church had prospered over a 1500 year history based upon its ability to accommodate both the needs of the poor and the wealthy.[30] Since, the 3rd century, the Church had accepted various forms and substantial amounts of money payments, in exchange for salvation (Johnson, 1976, p. 80). For

---

[30] In the Middle Ages, inequality was accepted as part of God's plan, each person playing a role in the Medieval social organism (Tawney, 1926, p. 14-16).

example, penance allows the Church to reinstall good standing with respect to salvation by forgiving sins. Of particular concern to 15th and 16th century reformers were the sacraments such as penance and confession and also the sale of relics and indulgences.[31] These practices generated revenue allowing for the provision of social services and also the maintenance of Church power.

The opportunity to use money payments as a means of penance implicitly assigns social value to money, and the underlying means to generate it. Money can in effect be used to cancel out sins.[32] By allowing for these exchanges, the Church implicitly provided a degree of moral validation for surplus generating business activities. There were mutual benefits from such trades. Individuals were able to accumulate wealth and purchase some degree of assurance that such accumulation would not harm their prospects for salvation. Meanwhile the church gains by "skimming off" a share of the surplus, allowing it to provide social services and also to maintain its power (Johnson, 1976, p. 314). Other factors that may have contributed to a moral validation of unequal wealth include the double standard on usury, and the status of envy as a deadly sin. Envy is a social emotion that serves as a motivation to level wealth. The existence of significant levels of capitalistic activities in the late Middle Ages attests to the fact that the Catholic Church provided substantial accommodation for the needs of an emerging merchant class.

4.8.5 Information Demand and Supply: The Morality of Market Activities

The Catholic Church was able to accommodate a significant amount of market based activities, which increased somewhat unevenly from the 11th century onward, with high points in the 12th and 13th centuries, plagues in the 14th century and recovery in the 15th century. According to framework 4.2, there should have been an increase in demand for R1 information in order to provide moral validation for capitalistic activities (A1). This postulated demand would have served to alleviate dissonance due to activity choices that

---

[31] The Church practiced price discrimination in the sale of indulgences.

[32] "Officially" indulgences do not "work" unless the sinner is truly sorry for his sins.

conflicted with Church doctrine and commonly accepted attitudes about morally appropriate behaviour. An additional motivation is the desire to believe that one's behaviour will allow one to achieve salvation.

A supply of new religious behavioural prescriptions was to emerge in the latter half of the 15th century for a number of reasons. High education and literacy rates among urban bourgeoisie allowed non-clerical people access to the scriptures and also to the works of Antiquity. The dissemination of the latter created a revival of the classical rationalist methods of inquiry. The humanist clerical scholars came to believe that God wanted the people to use reason to understand the principles of the good Christian society (Johnson, 1976, p.331). This led to an increasingly rationalist method of analysis of the Christian scriptures by both clerics and the educated laity. By the mid 15th century, the printing press was in widespread use leading to the dissemination of the Classical literature, vernacular editions of the scriptures, older Christian texts such as the original Greek version of the New Testaments, and contemporary religious writings. Such works were translated into the vernacular, further increasing accessibility. Literate people were both able and eager to read the scriptures for themselves (Green, 1998, p. 25). Clerical scholars discovered that the Latin versions of the New Testaments deviated considerably from earlier versions. This, in conjunction with the rationalist methods of inquiry, led to increased questioning of the contemporary practices of the church (Johnson, 1976, part 5).

Throughout the Late Middle Ages, according to Tawney's thesis, due to the increasingly commercial environment, there existed conflict between private material incentives and behaviour and orthodox religious doctrine regarding the moral appropriateness of such behaviour creating possibilities for dissonance. In the latter part of the 15th century, not only had commercial activity made a comeback after the previous century of plagues, but there was also an increased supply of alternate religious literature and high literacy rates. This new supply of information afforded an unprecedented opportunity for selective attention to information in order to reduce dissonance and also anxiety regarding salvation. An important contribution to the supply of "alternative literature" was the writing of Erasmus, who believed that the gospel of St. Paul was the most accurate interpretation of Christ's teachings. This teaching focused upon basic moral principles in contrast with the specific behavioural rules laid forth in the tomes of cannons in use by the Church. Erasmus regarded practices such as penance, indulgences, and relics as being in conflict with Christ's teachings and encouraged the laity to read the bible in order to develop a personal relationship with God. Additionally, he argued that justification could be attained by faith in conjunction with good works as laid out by St. Paul. The works of Erasmus were popular. Johnson (1976, p. 271) writes:

> His *In Praise of Folly* (1511) went into thirty-nine editions before 1536- some of these were very substantial...There were some years, it has been calculated, when between one-fifth and one-tenth of all books sold in Oxford, London and Paris were by Erasmus. In the 1530s, 300,000 copies of his *Greek New Testament* were circulating, and over 750,000 of his other works. He was a new phenomena, a living world best-seller.

The success of Erasmus indicates a high demand for new religious information by the increasingly literate growing urban bourgeoisie

The works of Erasmus conflicted with orthodox Catholic doctrine in a variety of ways including his attacks upon the sacraments and the selling of indulgences and relics. The Church taught that salvation could be attained by good works, providing incentives to comply with Church behavioural rules. Individuals were encouraged to take confession, which often generated much anxiety thereby motivating them to undertake penance as a means to cancel out sins. Encouraging individuals to develop a personal relationship with God by reading The Bible threatened the role of the Church as an intermediary between God and the people; the Pope's role as the sole authority in interpretation of the scriptures and God's word was challenged. Allowing for justification by faith would lessen Church power to use the desire to achieve salvation as a tool to induce compliance with church rules, opening the door for the self-justification of activity choices. The problem of the self-justification of privately preferred activity choices is potentially exacerbated by the fact that the Christian scriptures are subject to multiple interpretations due to ambiguities and inconsistencies. [33] From the perspective of an altruistic cleric, who believes in the Orthodox teachings of the church, the very role of the Church as a saver of souls was threatened. From a more pragmatic point of view, the reduced ability of the Church to influence behaviour threatened revenue and power.

Hence, prior to 1517, when Luther sparked off the Reformation, there already existed a demand and supply of information challenging the Church's behavioural rules and clerical structure[34]. Although influenced by Erasmus, Luther, in contrast, argued that justification could be attained by faith alone. Like Erasmus, Luther attacked practices such as indulgences and many of the sacraments. The ideas of the Reformation spread rapidly throughout Germany, Switzerland, Britain, the Lowlands and France, making some headway into Spain and Italy. The pattern roughly fits the model. A change in private material incentives in the late Middle Ages led to an increase in capitalistic behaviour creating dissonance due to a discordance between behaviour and beliefs regarding the its morality and also due to the possibility of reduced prospects for salvation. A large supply of variant sources of religious information in conjunction with high literacy rates facilitated selective attention to information in order to reduce dissonance and anxiety regarding salvation. This selective attention created yet more variants, which could then be copied rapidly via the printing press. Importantly this new information included the discrediting of various sacraments and indulgences and also the doctrine of justification by faith, facilitating selective attention to religious doctrine. As the new religious doctrine may facilitate the ease with which new behaviours are adopted (Weber, 1904-5), and vice versa, informational and behavioural changes can be rapid.

---

[33] The existence of multiple sects in the early Church (Johnson, 1976, part 2) and also the proliferation of variants of Protestantism after the Reformation illustrate the ease with which variant forms of Christianity are created. Theological debates over the Trinity were often a source of schism.

[34] Earlier heretical movements include the Bohemian Hussites and English Lollards (Thomson; 1998; 211-219).

4.8.6 Effects of New Religious Variants on the Moral Validation of Market Activities

Justification by faith along with encouragement to read religious texts led to an increased number of religious variants, which provided a variable amount of accommodation for capitalistic activity. Tawney argues that in some cases the reformers moved backwards against the interests of the business class. For example, Luther formally increased restrictions upon usurious behaviour, such as charging interest (Tawney; 1926, p. 94).[35] Luther did however break the Catholic Church's monopoly to dictate the rules regarding the morality of behaviour. Effectively, justification by good works, with the Church as an intermediary conveying God's word regarding what constitutes good works, had allowed the church to maintain a barrier to entry to new religious variants. The 16th century variants of Protestantism came to be characterized by a less hierarchical structure than Catholicism, ranging from fully democratic decision making to cases such as Lutheranism, in which founders laid out basic church doctrine and limited individual freedom to interpret the scriptures (Allen, 1995). It was perhaps inevitable that some of these Protestant variants would come to provide a high level of moral validation for the activities of the rising merchant class. In some locales, capitalist behaviour and beliefs regarding the morality of this behaviour would come to mutually support each other. Such regions might experience higher wealth allowing for population growth and territorial expansion. The religions that accommodated

---

[35] The Catholic Church relaxed usury restrictions in the 14th and 15th century and tightened them in the 16th. This earlier relaxation may have been in response to a reduced need for consumption smoothing by the poor, (and therefore less need to for capital market regulation), due to rising real wages, and relatively equal land distribution (Reed and Bekar, 2003). It also may have served to provide some degree of accommodation of business interests. In the latter stages of the 15th century and in the 16th century, wages fell and land holdings became less equal providing the Church with an increased incentive to prohibit usury due an increased need for consumption smoothing (Reed and Bekar, 2003). However, in addition to Reed and Bekar's arguments, the more stringent usury prohibitions can be viewed as a response by the Catholic Church to the Protestant Reformation which provided a strong impetus for within-Church reform (Jones, 1995). Another interpretation is that the Catholic Church, now no longer a monopoly provider of moral information, and unable to accommodate both the needs of the poor and the business classes opted to specialize in religious behavioural prescriptions that facilitated the short-term interests of the poor.

capitalism best and capitalism itself could then evolve by a process of cultural group selection (see section 4.7).

One particularly successful variant was Calvinism, which from its conception took a pragmatic approach to the interests of business, for example by officially allowing the charging of interest (Tawney, 1926, p.102-132). Weber argued that the Protestant work ethic, especially under Calvinism, was fundamental to the expansion of capitalism.

> The religious valuation of restless, continuous, systematic work in a worldly calling, as the highest means of asceticism, and at the same time the surest and most evident proof of rebirth and genuine faith, must have been the most powerful conceivable lever for the expansion of . . . the spirit of capitalism (Weber, 1958, p.172).

Calvin extended justification by faith to allow for double predestination according to which God chooses to elect particular individuals for salvation while the non-elect are destined for damnation. As all is predetermined, individuals cannot use good works to attain salvation. Success in this worldly life, for example, birth into the capitalist class, provides evidence of one's election. According to Weber, Calvinists have incentives to work hard in order to convince themselves and others that they were members of the elect. Furthermore, as Calvinism (and also Lutheranism) shunned displays of wealth, money was reinvested in new projects. To support this thesis, Weber notes that the Protestant countries experienced the growth take off in the industrial revolution, and since then have continued to maintain higher levels of wealth (Weber, 1956, 35-46). Even in England, where Calvinism was a minority religion, there is some evidence, albeit disputed, that successful capitalists were often disproportionately Calvinist (Rosenberg and Birdzell, 1986, p. 131). While Weber by no means takes a uni-causal view in explaining differential growth rates, he lays stress on the direction of causation running from the Protestant work ethic to capitalistic behaviour. Tawney, on the other hand, stresses the incentives of capitalists to adopt Protestantism. Rosenberg and Birdzell (1986, p. 131), summarizing the view of some of Weber's critics write.

> One might argue, not that Protestantism created capitalism but that capitalism created Protestantism. By this Weber's critics have essentially meant that Protestantism offered a set of beliefs which were highly congenial and flattering to the successful capitalists, who therefore embraced it. Or, less invidiously, one might argue that the new merchant capitalist class felt religious and moral needs not satisfied by the religious institutions of feudalism, thereby creating a vacuum which Protestantism filled.

Johnson (1976, p. 314-318) criticizes Weber arguing that the primary outcome of the Reformation was a division of nations into Protestant and Catholic nations and that the Protestant countries industrialized first, largely due to a greater degree of toleration for business and alternate religions. In Catholic countries during the Counter Reformation, tolerance towards alternatives to orthodox Catholicism was dramatically reduced. During the Inquisitions, the persecution of Protestants, Jews, and Catholics critical of the orthodox Church, forced many to seek exile in Protestant countries. This meant an increase in the entrepreneurial class in Protestant countries and a drain from Catholic countries, given that exiles were often members of the urban business class.[36] Furthermore, immigrants excluded from social networks, may adopt entrepreneurial activities out of necessity. Johnson (1976, p. 314) criticizes Weber further by arguing that Catholicism itself produces high levels of anxiety due to insecurity regarding salvation and hence promotes a strong work ethic. After the Reformation in Protestant countries the elimination of indulgences and other revenue generating practices prevented the church from "creaming off" such a large share of the surplus of work.[37] This surplus could instead be directed towards reinvestment in the private sector.

During the 15th century, the need for reform was acknowledged at various levels within the Church (Jones, 1995, p. 28-48). However, the Pope would not agree to a council of upper level clergy, which might challenge his position as authority as sole interpreter of

---

[36] Crouzet (1991) provides empirical evidence that Protestant exiles from the Spanish Lowlands, originally the banking and trading capital of Northern Europe, formed a significant portion of the business elite in London in the 17th and 18th centuries.

[37] Note that in terms of putting this surplus to productive use, the Catholic system improved upon Pagan systems of burying considerable wealth, including scarce tools, with the dead.

the scriptures with consequences for Church power and control over doctrine. When a council finally met at Trent in 1545, reformist elements of the church were absent, largely due to the intolerance of the Inquisitions. The council enhanced the Pope's position as sole authority on scriptural interpretation (Jones, 1995, p. 65-78) strengthening the top-heavy hierarchical structure regarding the setting of doctrine, reducing the Church's ability to reform from within. Unlike Protestantism, there was then a limited possibility for the creation of new religious behavioural codes to satisfy the needs of business. Moreover, during the Counter Reformation, the Church tightened up the rules regarding capitalistic behaviour.

Ultimately, which countries remained Catholic and which adopted Protestantism was partly a matter of chance but also included systematic factors such as geographical distance from Rome, the relative autonomy of urban bourgeoisie, and political structure. For example, Calvinism developed in the Swiss Cantons, where the urban bourgeoisie had a high level of autonomy. The private material incentives and personal traits of nobles offer an example of both systematic and chance factors. In England, Henry the Eighth in need of war finance had incentives to appropriate property from monasteries, weakening the power of the Church. Chance factors regarding England's adoption of Protestantism include the Pope's refusal to annul Henry's marriage to his first wife Catherine of Aragon, who had failed to produce a son. In order to marry Ann Boleyn, he declared himself supreme authority of the new Church of England. After Henry's death, Protestantism was intensified under his son Edward VI. Edward happened to die young after which Catholicism returned under the next monarch in line, Mary Tutor. Mary, by chance, also had a short life leaving the throne to Ann Boleyn's daughter Elizabeth. Elizabeth reinstalled Protestantism due to her illegitimate status under Catholic law (Green, 1998, p. 46-69). The success of Luther had apparent random elements such as his decision to become a monk, purportedly, after a near death experience with a lightening bolt. He was lucky that the local noble, Frederick the Wise the Elector of Saxony had fallen into conflict with the Church over the 1517 issue of indulgences that created a tax drain in his jurisdiction (Green, 1998, p. 14-18).

Protestantism may have served to provide moral validation for the preferred private material activity choices of both Frederick and Henry.

The example of the Reformation illustrates the potential complexity of the dynamics of moral information in conjunction with changes in behaviour. Indeed, I have painted broadly —nevertheless, the timing of historical events is consistent with the model in that changes in behaviour appear to predate major changes in religious doctrine. In order to provide stronger empirical validation for the model, it would be necessary to separate out the effects of motivation to achieve salvation from the effects of altruistic social emotions. That non- religious people today still attempt to validate their actions, in terms of social merit suggests that at least some of the informational change may have been induced by dissonance. Ozment (1975) provides evidence of substantial variation in the degree of attention to orthodox Church practices in the Late Middle Ages. For example, many people avoided confession for years on end, sometimes waiting until they were on their deathbed. Hence, the fear of damnation may have had a variable effect. Furthermore, religions appeal to guilt as well as to "warm glows" attainable through love for others and good works. Another factor inhibiting empirical testing of the model is the myriad of positive feedback effects due to informational cascades and dialectical interaction between the Weber hypothesis (see section 4.8.6 above) and the cognitive dissonance hypothesis.

4.9 Economic Ideology as Information about the Social Consequences of Market Activities

With the collapse of the feudal system in the 15th century, the social welfare system provided by the Church had largely broken down and the void was not to be filled until much later (Tawney, 1926, p. 253-273). For Calvinists, although the plight of the poor could be justified as a clear indicator of non-election, this could only be pushed so far as Calvinists also emphasized charity. Moreover, in England, which was eventually to provide the birthplace for the industrial revolution, Calvinists were a minority religion. Given the growing inequality in the 16th and 17th centuries, the new religious doctrine may have provided some moral validation for inequality. Social theorizing was, however, to move increasingly out of the religious realm into the secular. An increasingly literate and educated elite would extend the methods of classical rational inquiry to develop secular theoretical validation of capitalistic activities.

200 years after the century of the Reformation, Mandeville (1714) and Adam Smith (1778) were to turn medieval ethics on its head by supplying a social justification for markets as an allocation mechanism.[38] Although inequality was an undesirable bi-product of markets, the poor, although relatively worse off may be absolutely better off than under non-market systems. The concept that individual vice can lead to aggregate virtue was at odds with community based social attitudes which had existed throughout the Middle Ages and throughout most of human evolution. This secular moral justification for markets may have served to decrease the dissonance of the well-off classes, who had high standards of living, yet were surrounded by the squalor of the working classes. Smith's ideas may not have been so popular if the working proletariat had driven the demand for reading material, yet these classes were illiterate. Under both the feudal and market systems, distinct class identities, a form of group identity (section 3.2), may have provided a further means to alleviate the dissonance of the well-off classes.

Today the ideological debates regarding the mix of markets and central planning as a means of allocating resources continue. Along the political spectrum, the preferred policy mix is often rationalized for the social good. For example, advocates for increasingly liberalized trade use economic theory to argue that increased foreign investment will allow the third world to develop. On the other side, socialists argue that foreign investment in the third world is a form of imperialistic exploitation. In both cases, the social justification promoted by the ideological viewpoint is often aligned with the private interests of its advocates. 1st world labour, employing the arguments of Marx, aligns with the leftist anti-globalization movement. Capital advocates free trade

---

[38] Due to space considerations, I refer only to Smith and Mandeville. These systems of thought trace a causal path through the work of many social theorists from Aristotle to the Schoolmen of the Late Middle ages, followed by the Merchantilists of the 16th and 17th centuries.

by employing the arguments of Smith, Hayek, and others. The rich can argue that their situation is based upon merit, the opposite of sloth, while the poor can argue that they lack opportunity. The lack of agreement regarding the social merits of alternative economic systems, yet often fervently held beliefs in support of particular ideological viewpoints is suggestive of a high degree of selective attention to information.

4.10 Summary

The framework developed in section 4 allows for the examination of the dynamics of activities and information about the social consequences of these activities and how they interact. Many activities are characterized by positive user externalities, allowing for multiple equilibrium with sensitivity to starting conditions and the possibility of lock-in. Information and attitudes regarding the morality of such activities may serve to reinforce lock-in or slow transition between equilibrium for two reasons. Firstly, individuals may choose activities based upon their perceptions of the social consequences and secondly, when activity choices and moral beliefs are in conflict, dissonance-induced selective attention may lead to changes in moral information.

Hirshman (1965) used the theory of dissonance to argue that the emphasis upon changing attitudes to induce behavioural changes was misplaced; in the absence of activity changes, individuals are unlikely to adopt new attitudes. Accordingly, to affect transition to market economies it is necessary to change private material incentives such that agents adopt capitalistic behaviour. Attitudes should then change in response to induced dissonance. Similarly, Tawney (1926), looking at the European transition, argued that the substantial increase in the amount of market activity during the Late Middle Ages and the consequent strain upon the moral teachings of the Catholic Church was an important causal element for the Reformation. In contrast, Weber (1904-5) controversially laid stress on the direction of causation running from a change in religious attitudes to the adoption of capitalistic activities.

The emerging literature in support of the importance of altruistic preferences suggests that the theories of both the Weber and Tawney/ Hirshman have important implications for transition to market economies. Altruists have incentives to search for and selectively attend to information in order to justify currently privately desirable activity choices in so doing amplifying inertia in attitudinal changes. Altruists should also be hesitant to adopt activities that violate social consensus attitudes, which amplifies inertia in the adoption of new activity choices. With respect to transition to market economies, altruistic preferences should be predicted to further slow transition. However, if transition does finally occur, due to two-way feedback we might also expect to observe large changes in moral information.

The timing of the Reformation supports that Hirshman/ Tawney hypothesis in that substantial capitalistic activity predated the Reformation. Protestantism allowed for justification by faith and increased freedom in scriptural interpretation leading to the creation of religious variants. Regions with a relatively autonomous business sector could then selectively interpret the scriptures to support the activities of business. The fact that Calvinism, a religion that provides substantial support for Capitalistic activities, evolved in the industrial Swiss Cantons while the agricultural Cantons remained Catholic (Landes, 1998, p. 178-179) is consistent with the thesis of both Tawney/ Hirshman and Weber. The Protestant Cantons could then differentially grow due to higher wealth, allowing for the spread of both the capitalistic activities and the supporting moral validation. Hayek (1988) suggested that lock-in may be easier to overcome if such changes are initiated within groups.

As altruistic preferences may amplify that importance of changes in moral information and attitudes as causal elements for differential growth rates, economic historians may find it fruitful to pay more attention to moral attitudes. Rosenberg and Birdzell (1986) write that the focus of many "modern moralists" on concerns about economic inequality, due to market behaviour, as providing an impediment to the adoption of capitalistic activities, is misguided. "The perception of poverty as morally intolerable in a rich society had to await the emergence of a rich society, considerably later than the period with which we are now concerned." Yet as I have argued above, inequality has been a persistent feature of human post hunter-gatherer environments. According to the dissonance thesis, if inequality is present and a stable feature of the system, moral systems may evolve to provide social validation. In the Late Middle Ages, moral angst may have been an important causal factor leading to the change in moral information in the form of new religion. After the new information is adopted, the moral angst is then contained, i.e. dissonance is alleviated, and becomes indiscernible as the well-off classes have successfully selectively attended to information.

According to the dissonance thesis, it is in times of rapid behavioural changes, as induced by structural changes in society, during which there may exist a mismatch between information about appropriate moral behaviour and actual behaviour. During these times we may expect to see large changes in moral information that may be amplified by two-way feedback as the information comes to provide moral support for activity changes. The rapid and widespread changes in religious information both during and before the Reformation provide broad empirical support for the altruism/ dissonance thesis as do the timing of later secular justifications. More empirical work is needed.

5 Application: Environmental Damages

5.1 Explaining Environmental Damages Given the Assumption of Altruism

Moving to current times, the environmental problem provides a clear example in which private material interests often conflict with the social good. Environmental economists explain the environmental problem as the outcome of selfish behaviour on the part of individual economic agents attempting to maximize net private benefits while ignoring external costs to third parties. The associated policy tools focus upon bringing material private incentives into line with society's interests with comparatively little attention being given to moral suasion and the provision of accurate information regarding environmental damages.

Although the pursuit of material self interest, in spite of what are often substantial external environmental costs, surely explains much environmental damage, much observed behaviour suggests a brighter side to human nature. Governments and environmental groups utilize moral suasion and provide environmental information through advertising and education programs. Consumers donate money to environmental organizations and are willing to pay a premium for "green" goods. The creation of a green image has become an increasingly important part of corporate public relations (PR). In 1990 and 1995, US firms respectively spent $500 million and $1 billion on environmental PR (Beder, 2002, p. 108). These activities suggest that policymakers, environmentalists, and businesspeople believe that individuals can to some extent be motivated to make private sacrifices for the public good. Green consumption choices may be motivated by genuine concern about the state of the environment, whereby green consumers hope to decrease their own "biological footprint" and also to set an example for others. Green consumers may also be motivated by the desire to establish socially conscientious reputations.

While costly green behaviour is puzzling given the extreme assumption of *Homo economicus,* under the polar assumption of pure altruism (section 1.1), it is the prevalence of environmentally damaging behaviour on the part of consumers that needs explanation. Under the latter assumption, the most obvious explanation is that consumers possess biased information leading them to underestimate the social impacts of their activity choices upon the environment. A growing literature on the media and the environment provides both theoretical and empirical support for substantial bias in the environmental information provided by the media and also of misconceptions held by the public. This literature is briefly reviewed, followed by a discussion of how altruism, cognitive dissonance, and selective attention may contribute to media bias in the direction of under-reporting the environmental damages due to conventional activity choices.

5.2 Media bias: A Demand and Supply Approach

There is both empirical and theoretical support for the existence of bias in the environmental information provided by the media, with bias largely although not exclusively on the side of understating the extent and consequences of environmental damages (Anderson, 1997; Allen, Adam and Carter, 1999; Smith, 2000; Beder, 2002). In order to utilize the framework developed in section 4, sources of information bias are categorized below according to demand or supply side (Section 1.2-1.4). Most of the literature focuses upon supply-side bias, while the effects of the interaction between audiences and content are less well understood (Anderson, 1997, ch. 6). Demand-side factors include audience preferences for environmental verses other types of news and more generally entertainment, given an abundance of information and time constraints, and also psychological characteristics and prior knowledge base.

Altruistic social emotions may contribute to demand side bias due to selective attention to information in order to reduce the dissonance that might result if individuals were to accurately acknowledge the links between private behaviour and environmentally induced social damages. Two factors potentially contribute to high dissonance in the environmental arena. Firstly, it is often materially privately costly to adopt environmental alternatives. For example, there are often large time costs to using public transport verses automobiles. Secondly, privately materially preferable activities often generate (or are predicted to generate) high environmental damages and associated social costs.[39] For example, climate change models predict that the global mean surface temperature will increase by 1.4 to 5.8^0 C by the year 2100. Predicted effects include increasingly variable weather including droughts, grave damages to ecosystems, species extinction, rising sea levels and flooding, and serious socio-economic consequences for many regions (IPPC, 2001, p. 72). The high private material costs of adopting environmental activities in conjunction with estimated high social costs of not doing so create a situation that is ripe for dissonance and selective attention to information. The factors of issue complexity, uncertainty in costs and benefits of damage reduction, and supply-side bias, contribute to the ease with which environmental information can be selectively attended to.

5.3 Uncertainty and Issue Complexity

Two characteristics of the environmental problem lend themselves to the creation of biased views regarding the socioeconomic consequences of damages. Firstly, there is often a high degree of uncertainty regarding the relationship between human behaviour

---

[39] Many environmental damages are predicted to fall upon future generations. The high degree of altruism of parents towards children increases the possibility of dissonance.

and environmental and associated social damages. For example, in the case of global warming, although there is a scientific consensus regarding the range of expected temperature changes, the probability distribution over outcomes is uncertain. For given temperature change estimates, there is then uncertainty regarding the socio-economic consequences due to factors such as the resiliency of ecosystems and climate, and abatement and mitigation costs. Such uncertainty is partly the result of non-linearities due to positive and negative feedback effects. For example, an intensified hydrological cycle is predicted to lead to increased cloud cover, either amplifying or dampening warming, depending upon the nature of the cloud cover. Many cases of environmental damages are characterized by uncertainty in the sense that probabilities cannot be precisely assigned to the set of possible outcomes.

Secondly, environmental issues are often complex. Basic comprehension requires some knowledge of the sciences of climate, ecology, chemistry, health, and also social aspects such as politics and economics. The attainment of a broad understanding is difficult even for experts, who need to acquire cross-disciplinary expertise, and more so for information consumers, including the media, policy makers, and the public. Experts face a difficult task of conveying complex ideas to the media who, in turn, face challenges such as determining the quality of expert opinion and conveying complex information to the public in a comprehensible format (Anderson, 1997, p. 128, Stocking, 1999; Zehr, 1999). For example, conveying the basic principles of probability is difficult. Some possible implications of uncertainty and issue complexity for the demand and supply of environmental information are outlined below.

5.4 Uncertainty, Information Complexity and Environmental Information

Issue complexity and uncertainty facilitates the ease of supplying biased perspectives, whether with intent or not. Facts and statistics can be selectively chosen to support particular viewpoints. For example, Christy (2002, p. 11), in a book entitled *Global Warming and other Eco-Myths* (Bailey, 2002), writes "our planet's life system has experienced multimillennial periods of both warmer and colder weather, successful adaptation is a defining characteristic of every living plant and creature in our world today." Although the statement is true, it is misleading in that it suggests that the climate change that is expected to occur over the next hundred years is not unusual. For example, the predicted *rate* and *magnitude* of warming, to be expected over the next 100 years, exceeds estimates of temperature changes that have occurred over the past millennium (IPCC, 2002, p. 49). Over a multi-millennial verses centennial timeframe, species and ecosystems have orders of magnitude more time to "migrate" to different latitudes (IPCC, 2002, p. 94). Even if similarly rapid periods of climate change have occurred, in the past, although it is true that some species have adaptive features that allow them to survive, others go extinct.

Christie (2002, p. 11-15) then goes on to argue that uncertainty, due to non-linearities in climate change models, casts serious doubt on their predictions. For example, he sites the uncertainty of the effects of increased cloud cover upon warming. However, such uncertainty is acknowledged and discussed at length by the ICPP (2002, p. 31-32). Modelers conduct sensitivity analysis allowing for error estimates due to a myriad of uncertain factors, including the effect of warming upon cloud cover.

Of particular importance is the difficulty of conveying the scientific concepts of risk and uncertainty to both the media and the public. In the case of global warming, there is uncertainty regarding issues such as the predicted rate of warming and the consequent magnitude and timing of effects upon ecosystems and economies, specific regional impacts, and the dollar costs of mitigation verses adaptation options. However, there is also a broad consensus among climate experts regarding many aspects of the basic science. For example, the theory that increased concentrations of greenhouse gases enhance the greenhouse effect is well established. Empirical studies illustrate a rapid increase in the atmospheric concentration of greenhouse gases during the 20th century (IPCC, 2002, p. 33). In concordance with theory, mean surface temperatures have trended upward over the past 50 years; the majority of experts agree that there is an anthropogenic basis for such warming (IPCC, 2002, p. 51). The predictions of the many climate change models used as a basis for the IPCC reports are accepted by the vast majority of experts as providing a best available range of estimates regarding the timing of temperature changes. Furthermore, ecologists predict grave and in some cases irreversible negative impacts upon ecosystems and biodiversity (IPCC, 2002, p. 94). The net economic costs of climate change are also projected to be serious and increasing with the rate of warming, with costs falling disproportionately upon less developed countries (IPCC, 2002, p. 67-78). Such costs include agricultural loss due to droughts and flooding; fishery's loss due to destruction of marine ecosystems; health effects due to migration of disease vectors, increased water scarcity and degradation in some regions; and higher temperatures. Uncertainty regarding the predictions of theories is a normal feature of science. As new data are collected and theories revised, predictions are updated. Scientific debate and the questioning of consensus views is accepted to be an important feature of a healthy scientific environment. The problem arises in conveying the concepts of uncertainty to the public.

Proponents for taking limited action on climate change have laid stress upon uncertainty as a reason for delaying action to mitigate climate change. Survey studies (Wilson, 1997) suggest that the broad public has interpreted uncertainty to mean that there is a great deal of disagreement among experts regarding the science of global warming and the need to take substantive action to mitigate greenhouse gases. However, the IPCC reports, representing the research of approximately 3000 experts across various disciplines from 150 countries, categorize predictions according to whether they are generally considered to be robust verses those for which there is a great deal of uncertainty. Hence, the IPCC reports can reasonably be said to represent the expert consensus view.

Media researchers acknowledge that uncertainty and issue complexity constitute a broad challenge to media, which in combination with conversation constitutes the primary source of information for the public (For example, see Allan et. al., 1997; Anderson, 1997; Chapman et. al, 1997; Smith, 2000; p. 15-63). Lack of scientific training by the media as well as limited expertise in conveying the ideas of science creates a barrier to the provision of accurate information (Anderson, 1997, p. 128). Scientists face a difficult task in conveying the essence of scientific ideas to the media and the public. Furthermore, it is costly for the media, who are often under considerable time pressure, to evaluate the authority of experts. As a result, media often cite information from a variety of sources, leaving interpretation up to audiences who often lack the relevant knowledge required to form an objective view.

Overall, issue complexity and uncertainty facilitate the creation of a large, various, and potentially biased supply of information contributing to the ease with which individuals may selectively attend to information. In the next section, supply-side bias is examined under two categories, the structure of the media and the effects of vested interests upon information provision.

5.5 Supply Side Media Bias: Media Structure and Organization

The media environment is highly competitive. Different media agencies compete with each other for audience attention. Within agencies, news competes with other forms of information, while environmental news competes with other news. Within the environmental news, a variety of viewpoints, issues, and sources vie for attention (Chapman et. al., 1997, p. 36-42). What features of the media, including its structure and medium, determine the quantities and qualities of environmental coverage?

Coverage is partly determined by the beliefs of media providers regarding which issues will attract audience attention. Editors choose stories that they believe will generate interest, while journalists attempt to gage which stories will be chosen by editors (Chapman et. al., 1997, p. 36-42). Audience attention generates revenue both from direct sales and advertising revenue. Content that does not appeal to the interests of audiences will be selected against, either by putting suppliers out of business or by driving rearrangements in media organization and content. Although it is obvious that complex environmental issues might have difficulty competing with less complex high entertainment value stories, such as the personal lives of movie stars, within the environmental news itself, coverage can be biased towards particular attention-grabbers. For example, seal deaths, due to pollution or clubbing, generate widespread attention, partly due to the use of empathy generating visuals (Anderson, 1997, p. 121). Knowledge of the media process, and which types of stories are likely to receive airtime, in turn guide the action agenda of environmental groups. Greenpeace uses satellites to beam news-ready video and print coverage of its actions to major news networks (Dale, 1996, ch. 7).

The media uses framing techniques to generate audience attention (Anderson, 1997, p. 121-128; Miller and Riechert, 1999). For example, stories that can be framed within standard "story scripts" may be considered to be more news worthy. An example of a well-used story script contains conflict with clear-cut villains, victims and heroes. The case of environmentalists battling on the high seas to save the whales from "greedy" whalers easily fits into this story script while more complex issues, such as global warming, can be accommodated less easily without distortion. Media may shy away from story lines that cast the audience in the role of villain.

Environmental issues have an increased chance of becoming news items if they are "event centered" (Anderson, 1997, p. 121). Environmental groups accordingly make use of publicity stunts. For example in the Brent Spar controversy, Greenpeace generated a large amount of media attention when it occupied a redundant oil-rig in an attempt to prevent its sinking and release of toxic waste into the marine environment (Hansen, 1997). In cases such as global warming, in which environmental damages occur gradually, discrete events upon which to centre story lines are less readily identifiable.

Limits on time frames in television news broadcasts or print space in newspapers lend themselves to overly simplified depictions of the environmental news (Anderson, 1997, p. 67). Although complex topics may be avoided altogether, when covered, insufficient information may be presented so that audiences gain a poor understanding of the issues. As many environmental stories have a long time frame, they may not fit easily into the 24-hour time frame utilized by the mainstream media (Anderson, 1997, p. 123).

It is traditional media practice to report conflicting viewpoints, both to create interest and also to provide a balanced view. The provision of opinions and absence of an overall synthesis is consistent with the press's objective of providing balance without advocating a particular viewpoint. Indeed, an important role of the press is to provide a voice for those who question consensus views. Unfortunately, this approach is proving problematic when it comes to environmental reporting (and science reporting in general) because audiences often lack the science background to produce their own synthesis. As the media is a major source of science information for the public, providing a synthesis is arguably an important role of the media (Anderson, 1997, p. 49; Wilson, 1997). Exacerbating the problem is that due to time and space limitations, it is often the case that only two opposing sources are cited. In the case of global warming, Wilson argues that such binary opinion reporting is problematic because the consensus viewpoints held by the majority of climate change experts are given comparable weight to those put forth by a minority. Wilson argues that this has contributed to the widespread misunderstanding that there is little agreement among climate experts.

The media has a large dependence upon government sources as news claims-makers. Anderson (1997, p. 70) provides mixed evidence regarding the success of government departments to directly influence the environmental news. A less direct effect, given the media's dependency upon government sources, is that, given the low priority of the environment on government agendas, the proportion of government generated stories is then biased towards too little environmental coverage.

Given an increasing variety of information sources such as the internet, specialized television channels, and magazines, audiences are increasingly able to seek alternatives to main-stream media. For example, environmental groups, government, conservative think tanks, and corporate front groups provide environmental web-sites. Whether the increased variety of viewpoints contributes to more balanced views is questionable as media consumers may have difficulty accessing the legitimacy of such information. The increased diversity of sources may, however, contribute to a greater heterogeneity of viewpoints regarding the severity of environmental problems.

5.6 Supply-Side Bias: Vested Interests

Evidence and theory suggest that private material interests, both within the media and on the part of news claims-makers, contribute to bias in the environmental news. The following section investigates the vested interests of corporations, environmental groups, government and media providers. Although vested interests have incentives to choose privately materially desirable activity choices, if altruistic, they have incentives to selectively attend to information that justifies their actions. Consequently, vested interests may believe that the information that they provide is accurate.

5.6.1 Corporate Influence on the Media

Corporate private material incentives to provide biased environmental information can be readily identified. Producers of environmentally damaging commodities have large financial stakes in maintaining sales, providing incentives to overstate the social benefits of their commodities while understating environmental damages. Beder (2002) provides empirical support for the existence of substantial corporate bias in the provision of environmental information. These are summarized below.

Corporations employ a large variety of tools to get their messages across. According to Beder (2002), these include the financing of conservative think tanks, setting up corporate environmental front groups, hiring public relations firms to push messages to policy makers and media, product advertising, censoring programming content that conflicts with products advertised, and providing free "environmental" educational materials for schools. Furthermore, corporations may indirectly affect the media through their influence upon government agendas, affecting news through the media's dependence upon government sources.

Corporate advertising promotes consumption of products at levels and with qualities that often cause high environmental damages. Advertisements focus upon the benefits of consumer commodities, while drawing attention away from the social costs. From the late 1980s into the 1990s, corporations embarked upon a phase of green marketing. Based upon content studies, there was a decline in green advertising in the late 1990s, when advertisements began to focus more upon environmental backdrops (Corbett, 2002). Although, there has been substitution into the production of less environmentally damaging goods, in many cases, products were simply reinvented as green goods (Beder, 2002, ch. 11). Many such "green" products cause high damages relative to available alternatives. Furthermore, given the promotion of many only marginally useful but environmentally damaging products and that the ultimate objective of advertising is to increase sales, advertising is overwhelmingly "brown", not green (Corbett, 2002). It follows that advertisements provide biased information regarding the environmental damages due to many consumer activities.

Advertisements communicate information regarding acceptable social behaviour. Audiences not only receive information about product attributes, but also about the beliefs of marketers—namely that others will purchase the products, thereby indirectly conveying information that the product is or will be found to be socially acceptable. The individuals depicted in advertisements often possess stereotyped characteristics of the target market, aiding in identification with the advertisement. For example a sports utility vehicle (SUV) advertisement depicting "outdoor-adventurer types" indirectly conveys information that these outdoor types drive SUVs and that therefore SUVs are socially acceptable to this group. But as environmental concern is a correlated attribute of outdoorsy types, information that environmental types believe that SUVs cause a socially acceptable amount of environmental damages may be indirectly conveyed.

Most media providers, including public television, depend heavily upon corporate advertising or sponsorship for finance, providing incentives to please corporate sponsors. Beder (2002, p. 182) documents media admissions to corporations directly manipulating programming content. For example, corporations have threatened to withdraw sponsorship for programs in which content is perceived to be antagonistic to their products. Although evidence of such manipulation is provided in a variety of cases, making a broad case is difficult; Media outlets and advertisers would obviously be hesitant to admit to such manipulation.

Advertising affects media content indirectly because willingness to pay for advertising spaces depends upon client perceptions of the size and characteristics of the audience (Beder, 2002, p. 181; Corbett, 2002). Corbett (2002) writes:

> Advertisers and media programmers engage in a dance to fill each other's needs, each having a vested interest in constructing certain versions of the world and not others. According to Turow (1999). 'the ad industry

> affects not just the content of its own campaigns but the very structure and content of the rest of the media system' (p. 194). At the same time, media develops formats and tones for their outlets and programming deemed to be most acceptable to audiences that they hope marketers find most attractive. What this means for programming is that the upscale twenty-something audience—the most appealing segment to advertisers—will find itself represented in more media outlets than older men and women to whom only a small number of highly targeted formats are aimed. According to the researchers of the green marketing boom, the segments of the population most committed to the environment do not belong in this twenty-something group (Ottman, 1993).

As a consequence, programs depicting environmentally friendly lifestyles or documentaries about the environment may receive less exposure as they conflict with the lifestyle choices of the primary target audience.

Another issue is the increasing concentration of the media including the corporate conglomerates of consumer commodity and media/ entertainment firms. This creates conflicts of interest between objective information provision and the desire to promote their own products. For example, General Electric, named as one of the worst eight toxic polluters in America in 1992 and 1993 owns NBC, one of America's largest media conglomerates (Beder, 2002, p. 223-231).

Corporations provide finance to conservative think tanks such as The Competitive Enterprise Institute, and The Heritage Foundation. In turn these think tanks provide information to the media, government, and schools in formats such as interviews, pamphlets, reports, videos, and books such as Bailey (2002). These groups tend to understate the severity of the environmental problem, the need for more substantive environmental regulation and changes to consumer life styles. Think tank researchers are hired because they espouse a particular ideological perspective and understand that employment is contingent upon continuing to do so. Although the same claim can be made for hiring by environmental groups, the key difference is that due to corporate financing, conservative think tanks have deeper pockets, helping them to influence information flows. The media utilize corporate think tanks as sources for environmental views (Beder, 2002, ch. 6). Although sources are often cited, readers need not be aware that the perspective is provided from a particular ideological bent, nor that the source receives corporate finance.

Corporations set up environmental front groups that provide information to media, politicians, policy makers, and schools. The titles of such groups make them difficult to distinguish from legitimate grassroots environmental groups or even government institutions (Beder, 2002, ch. 2). According to Beder these front groups employ public relations firms to create the image of a high level of grass route support, helping them to influence policy makers and also the media. Similar to conservative think tanks, they provide information to governments, media, and schools. For example, Beder writes:

> The Global Climate Coalition, a coalition of fifty US trade associations and private companies representing oil, gas, coal, automobile and chemical interests, uses these sorts of arguments to fight restrictions on greenhouse gas emissions. Its tactics have included distributing to hundreds of journalists a video which claims that increased levels of carbon dioxide will increase crop production and help to feed the hungry people of the world.[40]

While it is true that some regions may experience higher agricultural yields (mid and high latitude regions), in other areas including most subtropical and tropical regions, yields are predicted to fall with net negative effects increasing with the severity of warming. Furthermore, the adverse effects of warming are predicted to fall disproportionately upon less developed countries (IPCC, 2002, p. 71).

Another concern is corporate influence upon environmental information provided in schools. For example, 40% of U.S. schools have agreed to screen 12 minutes of Channel 1 per day in return for loans of televisions, VCRs, and satellite dishes.[41] The content of Channel 1, owned by corporations such as Phillips and Time-Warner, includes 2 minutes of targeted advertisements, three minutes of world and national news

---

[40] "These sorts of arguments" include the claim that global warming will not happen and the emphasis upon uncertainty as a reason for not taking action.

[41] 90% of the students are expected to watch Channel 1 90% of the time according to the deal.

and seven minutes of features of interest to adolescents (Beder, 2002, p. 164). Beder (2002, p. 166) writes:

> Teachers are being overwhelmed by free and unsolicited classroom curriculum materials provided by public relations firms, corporations and their industry associations. For example, some environmental programs for children include Project Learning Tree, and Caretakers All, sponsored by the beef and forestry industry respectively.

According to Beder (2002, p. 169), "The Consumers Union describes Caretakers All as 'pure one-sided image-building for farmers and ranchers, disguised as lessons on land conservation'." Some corporate information providers and also parents working in environmentally damaging industries argue that the environmental information currently provided in schools is biased towards overstating environmental damages, while ignoring product benefits (Satchell, 1996). As teachers often lack expertise in environmental issues, they may lack the ability to critically examine environmental information. Furthermore, students may have a low propensity to question corporate learning materials because they are provided within the educational context.

Corporations hire public relations (PR) firms, which often work with corporate front groups. The activities of PR firms include creating "green" images and mounting political lobbies. According to Beder (2002, ch. 2), PR firms prepare persuasive information packages for politicians and attempt to create the illusion of grass root support for environmental front groups. Government policy makers have incentives to support corporate interests due to "revolving doors" conflicts of interest, whereby they increase their chances of finding well paid employment with corporate PR firms and lobby groups upon leaving government (Beder, 2002, p. 118-121).

Beder (2002, ch. 4) documents cases in which corporations have mounted libelous suits against environmentalists critical of corporate activities. She argues that although these suits are rarely upheld in court, they can be used to generate fear in order to silence anti-corporate opinions.

To conclude, there is both theoretical and empirical support for substantial corporate bias in the provision of environmental information. Due to the interest of maintaining sales of products at levels and with characteristics that damage the environment, the direction of bias is an overstatement of the benefits of environmentally damaging products and understatement of associated damages. Moreover, if corporate information providers are altruists, they have incentives to engage in dissonance-induced selective attention in order to convince themselves that they are providing accurate information about the "green-ness" of their products.

5.6.2 Environmental Groups and the Media

Individuals working for environmental groups hope to protect the environment, providing incentives to emphasize environmental damages, while laying less emphasis upon the economic costs of mitigation. For example, in 1988, Greenpeace provided speculations that pollution was contributing to the deaths of large numbers of seals in the North Sea. Scientists attributed the deaths to a virus and although admitting the possibility that pollution had contributed by lowering the immunity of the seals, they were unable to scientifically verify a linkage (Anderson, 1997, ch. 5). In the Brent Spar oilrig occupation case, after achieving substantial news coverage, Greenpeace admitted that they had overestimated the amount of oil remaining on the platform. The seal deaths and oilrig occupation stunt offered opportunities for event-centered coverage and also the use of empathy generating visuals. Although the message was biased with respect to environmental particulars, the essence of the message regarding the level of pollution in the oceans requires more media attention (Hansen, 1999). Yet a more reasoned less propaganda-based approach is less likely to have delivered the message or to have generated the media's interest.

Due to the widespread perception among environmentalists of a corporate bias in the press, environmental groups may justify information bias as a means to compensate for bias presented elsewhere. For example, corporate advertisements are a form of propaganda, depending upon visual images, music and other psychological manipulation techniques to convey messages. Daly (1999) commenting on the non-profit organization, Ad Busters, which makes use of anti-corporate advertisements in an attempt to provide some balance to the "consume more" messages of corporate advertising, states; "One has a right to fight fire with fire, but my personal hope is that reasoned argument still has a place, and that right and wrong will not ultimately be decided only be warring advertising images." Interestingly, advocates for both environmental and producer groups claim that the media is biased towards providing information antagonistic to their own viewpoint. These are respectively a "corporate" and a "liberal" bias. On both sides, there is then some perceived justification to provide propaganda in order to correct the perceived bias of the other side.

Although, on the surface, it would appear that environmental groups should if anything overstate environmental damages, incentives also work the other way. Environmental groups wishing to maintain hope (and also charitable donations) may have incentives to understate environmental damages. Dale (1996, p. 125-130) argues that Greenpeace, the most highly financed environmental group in the world, may have avoided campaigns critical of the lifestyle's of the Northern middle class due to fear of the erosion of its primary base donor base. Dale suggests that this may explain the decision not to pursue an anti-car campaign. Greenpeace representatives, however, cite a host of other reasons, including their belief that such a campaign would prove unsuccessful. Beder (2002, p. 131) cites cases in which corporations provide finance to environmental groups, also creating possible conflicts of interest.

Lastly, members of environmental groups having adopted environmentally friendly lifestyles, often at considerable private material cost, may have incentives to selectively attend to information that understates the economic costs of mitigation relative to environmental damages. Similarly, educated low-income people should experience relatively low dissonance due to adopting a pro-environmental stance. For example, a university student who cannot afford to drive a car may experience psychological benefits by believing that cars cause high environmental damages.

5.6.3 Government and the Media

As briefly discussed above, close links between industry and government may bias government agendas toward corporate interests to the neglect of environmental concerns affecting both the quality and quantity of media coverage. In "revolving doors" conflicts of interest, government workers acquire expertise regarding the workings of government. These people may have incentives to cater to the needs of industry in order to increase the chances of being hired as private sector lobbyists upon leaving government. Beder (2002, p. 118) writes:

> When the Republicans lost office in 1992 there was a mass movement of government officials to the lobbying and PR firms. Ralph Nadar's group, Congress Watch, tracked 300 of them: over half moved to Washington DC lobbying and PR firms. The door swings both ways, and former lobbyists often become part of the government, there they have a unique opportunity to help their former clients.

Furthermore, in the United States, there are revolving doors between corporate financed conservative think tanks such as The American Enterprise Institute and the administration. According to *The Economist* (1991, vol. 319 (7708), p. 23-26),

> The revolving door between governments and think tanks is well established...When Ronald Reagan took office, no fewer than twenty of the research fellows at the neo-conservative American Enterprise Institute (AEI) joined his administration. Now that Mr Reagan has left power, many of his appointees, such as Jeane Kirkpatrick and Richard Perle, are working at AEI. Every American think–tank director has a dream and a nightmare. The dream is to house the next administration; the nightmare is to house the last one. AEI seems to have managed both in the course of a decade (sited by Beder, 2002, p. 84).

Other factors that may influence government agendas and hence media coverage include corporate election finance and electoral time-frames. The short time-frame of electoral terms elevates the political importance of the short-term economic costs of environmental reforms, for example unemployment, relative to the more distant environmental benefits, biasing policy agenda towards the former.

For an opposing perspective, Brady (1993) argues that risk averse policy makers have incentives to put into place unwarranted and excessively stringent greenhouse gas abatement policies due to concern that if the worst case scenarios become a reality, policy makers will suffer personal consequences, such as discontinued employment. He then argues that policy makers have incentives to selectively attend to information that overstates the costs of global warming in order to reduce the dissonance that would result from more balanced attention to the broader spectrum of information. Brady uses arguments similar to those presented by Christy above to argue that the consensus view overstates the need to abate greenhouse gases. However, the dissonance theory can be used to make the opposite argument. As policy makers are members of the class of people most responsible for greenhouse gas emissions, acknowledgment of the link between personal consumption behaviour and environmental damages may provide incentives to selectively attend to information that understates these costs.

5.6.4 Vested Interests Within the Media

Due to corporate finance, private media editors and producers have incentives to provide content that they believe will appeal to the target markets to which the advertised commodities most appeal. In turn, writers of both news and non-news have private incentives to provide content that they believe editors will choose out of an oversupply of possible programs or stories. In the case of news, if the type of stories provided in order to satisfy (advertiser) editor requirements conflict with those that editors (reporters) believe that they should provide, then a situation of dissonance is created. For example, a journalist might believe that an article on logging should include the concerns of both loggers and environmentalists, but that editors face pressures to avoid an article that stresses environmental costs. The reporter might then lay less stress on environmental damages. Brown, a reporter for *The Guardian* writes:

> The word 'environment' also has high boredom value; in fact, if I do not want to do a story, the useful key phrase is 'of course, this is an important environment story'... On the other hand, if I do want to do a story, it is important not to mention the environment at all, and make it sound as little to do with our specialism as possible (Smith, 2000, p. 67).

Media providers have incentives to selectively attend to their own coverage as being balanced or truthful. Furthermore, as media professionals are drawn from the middle and upper income groups, the largest per capita contributors to environmental damages, there are incentives to selectively attend to information that understates damages.

5.6.5 Summary: Supply-Side bias

Supply-side bias in the provision of environmental information occurs due to characteristics of the information itself, including issue complexity and uncertainty. Firstly, inaccurate and biased coverage may occur due to characteristics of the media, including its structure and organization, and limitations due to standard accepted formats for the particular medium. Misunderstandings may occur due to a lack of scientific expertise within the media as well as the lack of expertise by scientists in communicating information to lay audiences. Secondly, vested interests including corporations, environmental groups, government and media professionals may intentionally or subconsciously bias environmental information. The evidence suggested that corporate vested interests have contributed to substantial bias in the direction of underestimation of environmental damages and, importantly, the link between consumer behaviour and environmental damages. Altruistic preferences may contribute to supply-side bias if vested interests selectively attend to information, allowing them to believe that the information that they provide is accurate.

5.7 Demand-Side Bias and Cognitive Dissonance

Due to the magnitude, scope and nature of supply-side bias in environmental information provision, it is perhaps not surprising that media researchers have focused upon supply with little emphasis on demand. Although, some demand-side factors have been discussed above, for example the selection of media coverage that it is believed will appeal to audience interests, it has received little systematic consideration. The focus upon supply-side bias may be partly due to its ease of explanation based upon material self-interest. Considering altruistic preferences draws attention to demand-side factors.

As many information sources are at variance with consensus expert views on the state of the environment and associated social consequences, it seems reasonable to inquire as to whether selective attention to information in order to reduce dissonance is helping to sustain bias in the provision of environmental information. Suzuki (2002) draws attention to this "ostrich phenomena" in his environmental science column.[42]

> ....while 99 percent of climate scientists think global warming is a serious problem ..., the media love to focus on the few who disagree...Earlier this spring a European statistician wrote a book saying that the global environment is healthy and actually improving. The many scientists, environmentalists, and citizens who worry about most environmental problems are misguided, he argued. They just don't understand nature like he does. In spite of the facts that this man is not an expert in the fields he is critiquing and has been widely discredited by his peers, he has become a media darling. He has had speaking engagements across North America, front-page stories, editorials, and more. Why is he so popular? Simple: He assuages our guilt about ecological problems. Like a travelling tonic salesman, he tours the land telling us what we would prefer to hear, making us feel better about ourselves and the world. Contrast that with the United Nations Environment Program's report, "The State of the Environment: Past, Present, Future?" released last week. It's pretty depressing stuff.

Due to dissonance, we might expect the under-reporting of information that links consumer behaviour to environmental damages and associated social consequences. Consistent with the cognitive dissonance thesis, as evidence of damages due to global warming increased, empirical studies suggest that media coverage proportionately decreased (Williams, 2001, p. ix). Supply-side bias facilitates the ease with which individuals can selectively attend to information in order to reduce dissonance. Due to the "mass" component of the media, people then receive biased news helping to reinforce the moral attributes of their current behaviour. The consensus current behaviour is itself an important source of information regarding society consensus beliefs regarding the social appropriateness of behaviour. For example, the large-scale use of sports utility vehicles in conjunction with the assumption that people are somewhat altruistic provides information that others believe the behaviour is socially acceptable. Behaviour and self-justifying environmental information can then reinforce each other. As much consumer behaviour became established before the magnitude of

---

[42] Suzuki is referring to the book *The Skeptical Environmentalist* by Bjorn Lomborg. After publication, the book was discredited in general science journals such as *Scientific American* (Rennie, 2002).

environmental damages became known, starting conditions may have contributed to the lock-in of many asocial consumption choices.

5.8 User Externalities and the Environmental Problem

There are two types of externalities that are important for the environmental problem. These are environmental damages and the less emphasized "user externalities", in which the number of individuals who choose an activity affects the private material net benefits that others derive from the same activity. These externalities can result in the payoffs illustrated in figure 4.1.

An example, to be modeled in section 6, is the choice of private transport methods. As more people use public transport, the environmentally less damaging alternative to automobiles, the private net material benefits of public transport increase, due to more routes and an increased frequency of service. These user externalities potentially allow for "critical mass" levels of public transit use, above which it becomes privately materially beneficial for a given individual to switch from cars to pubic transport.

The concept of private material critical masses has more general application to environmental problems, due to the high cost of time in the "consumer society". For example, consider two persons (X and Y) competing for promotions in the work place. X uses public transport, composts, spends time shopping for less packaged goods, gardens without pesticides or power tools, avoids over-packaged fast food, and does not annually dispose of clothing in order to keep up with office fashion. Y, on the other hand, drives a car, creates large amounts of garbage, and keeps up with the latest fashions. As X has less time to devote to the workplace, he may forgo promotion.[43] Y makes it more difficult for X to achieve promotion because he has raised the bar at work. The non-environmental activities of Y have imposed a negative user externality upon X. The more people that choose lifestyle Y, the higher the net private material benefits of lifestyle Y and the lower the benefits of lifestyle X.

---

[43] Children and double income families add to substantially to these time costs.

Another factor, partly eluded to above under fashion, is prestige gained from conspicuous consumption. As prestige is a relative phenomenon, it also represents an example of a positive user externality. For example, if most people on a street have large houses, the person with a small house may feel socially inferior. Each large house owner then imposes a negative "prestige" externality upon persons with small houses, potentially creating an arms race in house size. There are both environmental and prestige externality benefits to coordinating on smaller houses.

To conclude, positive user externalities are associated with many environmentally damaging activities. As illustrated in section 4, these externalities have theoretical implications for the dynamics of activities and moral information. Namely, positive user externalities create the possibility for multiple equilibria with sensitivity to starting conditions. As many social norms were established before the seriousness of associated environmental damages was appreciated, activities may have locked-in to low welfare equilibrium. Furthermore, dissonance induced selective attention may reinforce lock-in to low welfare equilibrium because information may come to provide moral justification for currently privately materially preferred activity choices.

5.9 Summary

Evidence and theory suggest that there is substantial under-coverage and understatement of environmental damages by the media. Such bias is facilitated by characteristics of environmental issues including complexity and uncertainty, characteristics of media structure, media format, and the influence of vested interests. Demand side factors affecting environmental coverage have received less consideration. A potentially large source of demand side bias is dissonance-induced selective attention both by information consumers and suppliers. As many accepted activity choices provide high private material benefits, but only at a cost to the environment, acknowledgement of best expert estimates of damages may cause high levels of dissonance leading to selective attention. Activity choices and information about environmental damages may come to reinforce each other as discussed in section 4. Furthermore, positive user externalities related to activity choices may make it difficult to switch towards more environmentally friendly alternatives.

In line with Hirshman (1965), due to dissonance and selective attention, behavioural changes might need to precede changes in information and beliefs regarding environmental damages. Accordingly, an effective approach to the environmental problem appears to be the utilization of standard policies that target material private incentives. In the absence of switching to more environmentally friendly activity choices, policies that aim to increase the accuracy of environmental information might be largely ineffective due to dissonance induced selective attention. Yet the fact that a proportion of the population does make efforts to acquire accurate information about environmental damages and is also motivated to undertake privately materially costly actions for the public good suggests that there may be a role for policies that target information accuracy.

In the next section, a model is used to investigate some possible dynamics of activity choices and the accuracy of information regarding environmental damages under the assumption of somewhat altruistic preferences. The effects of factors such as increases in expert estimates of damages and policies that target private material incentives and also those that target the accuracy of information are considered.

6 A Model

A model, based upon the framework developed in section 4, investigates some possibilities for the dynamics of activity choices and information and beliefs regarding the social appropriateness of activity choices. In each of the two case studies of sections 4 and 5, both private material and social net benefits often depend upon the number of users. The model below is discussed with reference to the environmental problem and more specifically the problem of personal transport choices.

6.1 Model Overview and Activity Choices

6.1.1 Private Net Benefits, Externalities and Social Net Benefits

As in section 4.2, there are two possible activity choices. A1 and A2 respectively represent public transport and automobiles and p_t represents the proportion of individuals choosing A1 in period t. The allocation $p_t = 1$, in which all individuals choose A1, is assumed to be socially efficient. $\pi_j^P(p_t)$ and $\pi_j^S(p_t)$ respectively represent the net private material benefits and net social benefits due to an individual's choice of activity j.

For the purposes of tractability and to focus upon the effects of altruism, assume that individuals are identical with respect to all parameters affecting activity choice with the exception of their level of altruism. Alternatively, each individual (i) can be viewed as possessing the average level of any attribute other than the level of altruism, where none of these other characteristics are correlated with this level. To contrast, the framework developed in section 4.2 allowed for heterogeneity along multiple dimensions. Note that for notational simplicity, in contrast to section 4, the individual subscripts, represented by i, are dropped for variables that do not differ between individuals. In reality, private material benefits vary due to factors such as income levels, residential location, family size, and time costs. For example, the time costs of high-income people may increase the desirability of cars while low-income people may experience low opportunity costs from using public transport. Also, it is assumed that individuals possess identical information about the social effects of their activity choices. In reality individuals differ due to factors such as the degree of exposure to mass media verses individual specific information sources, prior knowledge and cognitive abilities, and individual propensities to selectively attend to information in order to reduce dissonance. Differences in these propensities are due to factors such as the individual's level of altruism and his ability and desire to conduct truthful information searches.

There are two types of externalities, a user externality and a non-user externality. The user externality, represented as $\zeta_j(p_t)$, corresponds to the net external benefits due to an individual's use of activity j that is received by others when they undertake activity j. For example, if more people use public transport, the net benefits of public transport increase relative to cars due to more frequent services and more routes. Furthermore, increased use of public transport increases the chance that car drivers will experience social criticism. In the case of cars, more users may lower automobile prices due to increasing returns to scale in production. Additionally, dependency upon automobiles affects the spatial density of cities, increasing travel time and increasing the relative timesaving benefits of cars. The user externality is reflected in the fact that $\pi_1^P(p_t) - \pi_2^P(p_t)$ is assumed to be increasing in p_t.

The second externality $D_j(p_t)$ represents non-user net external costs (or damages) due to an individual's use of activity j that do not occur through the user effect. The example to be considered is air pollution, with cars causing more pollution per passenger mile on average than public transport so that $D_2(p_t) > D_1(p_t)$.

The social net benefits due to an individual's use of activity j, which include both his private material benefits and the two externalities, are represented as $\pi_j^S(p_t) = \pi_j^P(p_t) - D_j(p_t) + \zeta_j(p_t)$. Taking the difference, gives the net social benefits that society would gain if the individual switches from A2 to A1, represented as

$$\pi_1^S(p_t) - \pi_2^S(p_t) = \pi_1^P(p_t) - \pi_2^P(p_t) + D(p_t) + \zeta(p_t). \qquad [6\text{-}1]$$

$D(p_t) = D_2(p_t) - D_1(p_t)$ represents the addition to non-user net external costs due to an individual switching from A1 to A2, i.e. pollution damages saved due to switching from A2 to A1. $\zeta(p_t) = \zeta_1(p_t) - \zeta_2(p_t)$ represents the gain in external user net benefits if an individual switches from A2 to A1. Note that in section 4.2, no notational distinction was made between the user and non-user net damages while individual damages were accounted for such that $D_i(p_t)$ (section 4.2) corresponds to $D(p_t) + \zeta(p_t)$ (section 6). As illustrated in figure 4-1 (section 4.2), A2 is assumed to be privately preferred if $p_t = 0$. For cases 4a and c, A2 is privately materially preferred for all p_t, while in cases 4b and d, there exists a critical mass level of A1 players, p_C^F, above which A1 is privately materially preferred. The net social benefit of A1 verses A2, $\pi_1^S(p_t) - \pi_2^S(p_t)$, is increasing in p_t such that at some threshold level, p_C^A, where $p_C^A < p_C^F$, pure altruists with unbiased information about social damages choose A1.

6.1.2 Beliefs About Damages

Although information about damages may be individual specific as discussed in section 4.5, in the current model it is assumed that all individuals receive identical information at the start of each period. In period *t* individuals contribute to the determination of the pooled information to be received in the start of period *t+1*. After choosing period *t* activities, they may selectively attend to information in order to reduce dissonance. This selective attention contributes to the demand for information, which affects the supply next period. The process of information determination will be modeled shortly. The information index, which measures the degree of information bias is assumed to be a real number such that $I_t \in \Re \,|\, 0 \leq I_t \leq 1\}$. Each individual receives information of quality I_{t-1} at the start of period *t*, where I_{t-1} was endogenously determined last period. Note that although I_{t-1} depends upon p_{t-1}, this dependency will be ignored currently for the purposes of notational simplicity.

Individuals use the pooled information to estimate non-user externalities (pollution damages) prior to choosing activities. A lower limit on bias corresponds to $I_{t-1} = 0$, corresponding to a situation in which no information is received about damages, in which case no dissonance is experienced. At the opposite extreme, if $I_{t-1}=1$, the individual receives the truth about damages and dissonance is maximized. If no information is received about damages, no dissonance is experienced. For intermediate values of I_{t-1}, information is partially biased and some dissonance may be experienced.

Believed damages are assumed to be the product of the information index and the true damages from each activity choice. $D_j^e(p_t, I_{t-1}) = I_{t-1} D_j(p_t)$ represents an individual's beliefs about his external damages due to choosing activity *j* and $D^e(p_t, I_{t-1}) = I_{t-1} D(p_t)$ represents his beliefs about the net damages due to switching from A1 to A2. For simplicity, it is assumed that per capita damages are independent of the proportion of people who choose each activity. The damages due to activity *j* are represented as D_j, while the net damages due to switching from A1 to A2 are represented as *D*. This assumption is reasonable for environmental damages from cumulative pollutants such as greenhouse gases but not for non-cumulative local air pollutants, in which case damages per person increase with p_t.

6.1.3 The Level of Altruism

Altruists are assumed to incur cognitive dissonance if they believe that they have inflicted harm upon others. Dissonance costs depend upon two factors 1) beliefs about external damages and 2) the individual's level of altruism. Let $\theta_i \in \{\Re \mid 0 \leq \theta_i \leq 1)$ serve as an index for an individual's level of altruism where the level of altruism is increasing in θ_i. $\theta_i = 0$ and $\theta_i = 1$ respectively represent the benchmark indices for pure free riders and pure altruists. Individual levels of altruism vary due to factors such as the genetic variation in the level of social emotions, the plasticity to be socialized and socialization from environmental factors such as parents, other role models, and life events (Krebs, 1970). Due to choosing activity j, an individual experiences dissonance costs that are assumed to be the product of beliefs about damages and the altruism index. The period t dissonance costs that an individual receives due to choosing activity j are represented as $C_{ij}(\theta_i, I_{t-1}) = \theta_i I_{t-1} D_j$. His dissonance costs due to switching from A1 to A2 are represented as $C_i(\theta_i, I_{t-1}) = \theta_i D^e = \theta_i I_{t-1} D$. [6-2]

If there is no information bias, a pure altruist with $\theta_i = 1$ internalizes all believed external damages to receive a dissonance cost of D_j from each activity choice or D if he switches from A1 to A2. A pure altruist receives no dissonance costs if $I_{t-1} = 0$, which corresponds to the maximum possible level of information bias. If information is intermediate between these values, an altruist partially internalizes damages. A pure free rider with $\theta_i = 0$ experiences no dissonance, independent of the degree of information bias. Persons who are partially altruistic with $0 < \theta_i < 1$ internalize some damages if $I_{t-1} > 0$.

6.1.4 Individual Utility

An individual's utility from activity j is the sum of the private benefits of activity j minus his beliefs about the damages caused by his activity [6-2] plus a last term,

$p_t D_1 + (1 - p_t) D_2$, which represents the non-user externalities (damages) that befall *i* due to other people's choices of activities.

$$U_{ij}(p_t) = \pi_j^P(p_t) - \theta_i I_{t-1} D_j + p_t D_1 + (1 - p_t) D_2 \qquad [6\text{-}3]$$

Taking the difference between the utility due to A1 and A2 gives the net private benefit to *i* due to his choice of A1 relative to A2 [6-4]. This difference represents the net private benefit that *i* receives due to switching from A2 to A1, which can be either positive or negative.

$$NB_i(\theta_i, I_{t-1}, p_t) = U_{i1}(\theta_i, I_{t-1}, p_t) - U_{i2}(\theta_i, I_{t-1}, p_t) = \pi_1^P(p_t) - \pi_2^P(p_t) + \theta_i D I_{t-1}. \qquad [6\text{-}4]$$

The private net benefit of A1 verses A2 for individual *i* is equal to the difference between the net private material benefits plus the net dissonance savings from A1.

6.1.5 The Timing of the Model

The timing of the model is as follows:

1) At the start of period *t*, each individual receives information about the magnitude of damages, I_{t-1} (determined last period).
2) Based upon this information and the individual's degree of altruism, each individual chooses the activity that maximizes his utility, which is equivalent to choosing the activity which yeilds the highest private net benefits [6-4]. The individual is able to anticipate dissonance effects when making his decision.
3) Each individual recieves the private net benefits based upon the private material and dissonance costs [6-4].
4) After choosing the activity, the individual may experience dissonance, which can now be reduced by selectively attending to information regarding non-user damages. Such selective attention contributes to an individual's demand for particular sources of information. The process of information selection is modelled in section 6.5.

5) These individual demands for information are aggregated across society to determine the aggregate demand for information which in turn contributes to the determination of the pooled information received at the beginning of the next period (section 6.5).
6) The sequence is repeated. The system is in equilibrium if no individuals have incentives to change activity choices and if the information index does not change.

The model is developed in two stages. Firstly the dynamics of activity choice are determined for a given level of pooled information. Secondly, the dynamics of the pooled information are determined for a given level of activity choices. Lastly, the two systems are combined in order to examine the co-evolution of activity choices and information bias. Comparative statics are used to investigate the effects of changes in environmental damages, policies, and other exogenous factors that affect private material incentives and also the accuracy of information.

6.2 Assumptions Regarding the Net Benefits of Activity Choices

For tractability, specific functional forms are employed. An individual's private material benefit of A1 is represented as $\pi_1^P(p_t) = \bar{\pi}_1 + \beta p_t^2$, where $\beta > 0$ determines the size of the positive user externality due to others' use of A1.[44] An individual's private material benefit from A2 is represented as $\bar{\pi}_2 > 0$. The net private material non-user benefits of A1 verses A2 can be written as $\bar{\pi} = \bar{\pi}_1 - \bar{\pi}_2 < 0$. The net private benefits of A1 verses A2 simplify to

$$NB_i(\theta_i, I_{t-1}, p_t) = \bar{\pi} + \beta p_t^{\,2} + \theta_i I_{t-1} D . \qquad [6\text{-}5]$$

To ensure that $p_t = 1$ is socially efficient requires that $\pi_1'(p_t) > 0$ and $\bar{\pi} + \beta + D > 0$ so that at $p_t = 1$ the net damages outweigh the net private material benefits of A2 relative to A1.

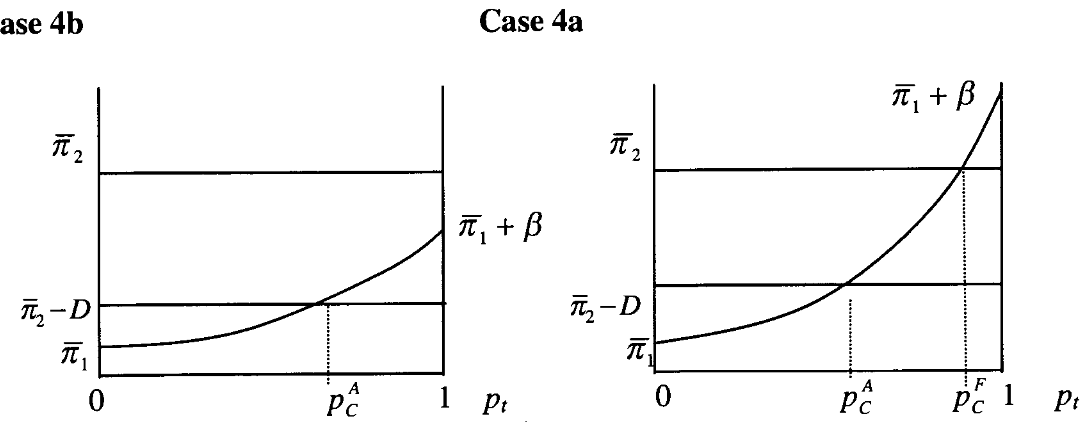
---

[44] Private material net benefits of the social activity (A1) are represented by $\pi_1(p_t) > 0$ with $\pi_1'(p_t) > 0$ and $\pi_1''(p_t) > 0$, indicating that the private material benefits of A1 increase in the number of users at an increasing rate. This assumption is reasonable given the network externalities present in public transport systems and also allows for the consideration of multiple equilibria.

The four cases illustrated in figure 4.1 (section 4.2) are examined. Cases 4.1 a and b, depicted in figure 6.1, represent situations in which it is socially preferable for a given individual to play A2 if $p_t=0$, but not above some critical level of p_t, referred to as p_C^A where $0 < p_C^A < 1$. To ensure that A2 is privately preferable at low levels of p_t, it is assumed that $\bar{\pi} < D$.[45] In case 4b, $\bar{\pi} < \beta$, so that A2 is privately preferable for all levels of p_t. In case 4a, $\bar{\pi} > \beta$, there is a critical mass p_C^F, above which pure free riders choose A1, where $p_C^A < p_C^F < 1$. The critical mass threshold for a pure altruist with perfectly unbiased information about damages is p_C^A. More generally, the critical mass threshold for a given individual depends upon his specific beliefs about damages and his level of altruism.

In cases 4.1 c and d, A1 is socially preferable for all levels of p_t, ensured by $\bar{\pi} + D > 0$. In case 4.1c it is privately materially desirable to play A2 for all levels of p_t, ensured by $\bar{\pi} + \beta < 0$. In case 4.1 d there is a private material critical mass in which $0 < p_C^F < 1$, ensured by $\bar{\pi} + \beta > 0$.

Figure 6.1: Model Parameters and Net Private and Social Benefits of A1 verses A2



---

[45] Although low income types who privately prefer A1, even if $p_t = 0$, are not directly included in the model, they can be indirectly considered by increasing $\bar{\pi}$, the baseline net private benefits of A1 at p_t=0, due to their effect on the user externality.

6.3 The Dynamics of Activity Choice Given Exogenous Information

If the net benefits of A1 verses A2, $NB_i(\theta_i, I_{t-1}, p_t) > 0$ (<0), [6.5], then the individual chooses A1 (A2). Assign $\theta_i = \hat{\theta}(p_t, I_{t-1})$ so that $NB_i(\theta_i, I_{t-1}, p_t) = 0$, such that i is indifferent between the two activities. To find $\hat{\theta}(p_t, I_{t-1})$, the altruism index for the indifferent individual, set equation [6-5] equal to 0 yielding

$$\hat{\theta}(p_t, I_{t-1}) = -(\bar{\pi} + \beta p_t^{\,2}) / I_{t-1} D \,. \qquad [6\text{-}6]$$

Individuals with $\theta_i > \hat{\theta}(p_t, I_{t-1})$ prefer A1, while those with $\theta_i < \hat{\theta}(p_t, I_{t-1})$ prefer A2. If all individuals choose their preferred activity, then the proportion of the population choosing A1 is $1 - \hat{\theta}(p_t, I_{t-1})$. This proportion is stationary for a given level of I_{t-1} if

$$p_t = 1 - \hat{\theta}(p_t, I_{t-1}) \,. \qquad [6\text{-}7]$$

Substituting [6-6] in [6-7] and solving yields the conditions for interior stationary values of p, given I_{t-1}.

$$p^*(I_{t-1}) = \frac{I_{t-1} D \pm \sqrt{(I_{t-1} D)^2 - 4\beta(\bar{\pi} + I_{t-1} D)}}{2\beta} \qquad [6\text{-}8]$$

The two values of $p^*(I_{t-1})$ are subsequently referred to as $p_L^*(I_{t-1})$ and $p_H^*(I_{t-1})$, respectively representing the lower and higher root. Substituting $\theta_i = 1 - p_t$ into [6-5] gives the net private benefit of A1 verses A2 for the "marginal" individual with altruism index $\theta_i = 1 - p_t$.

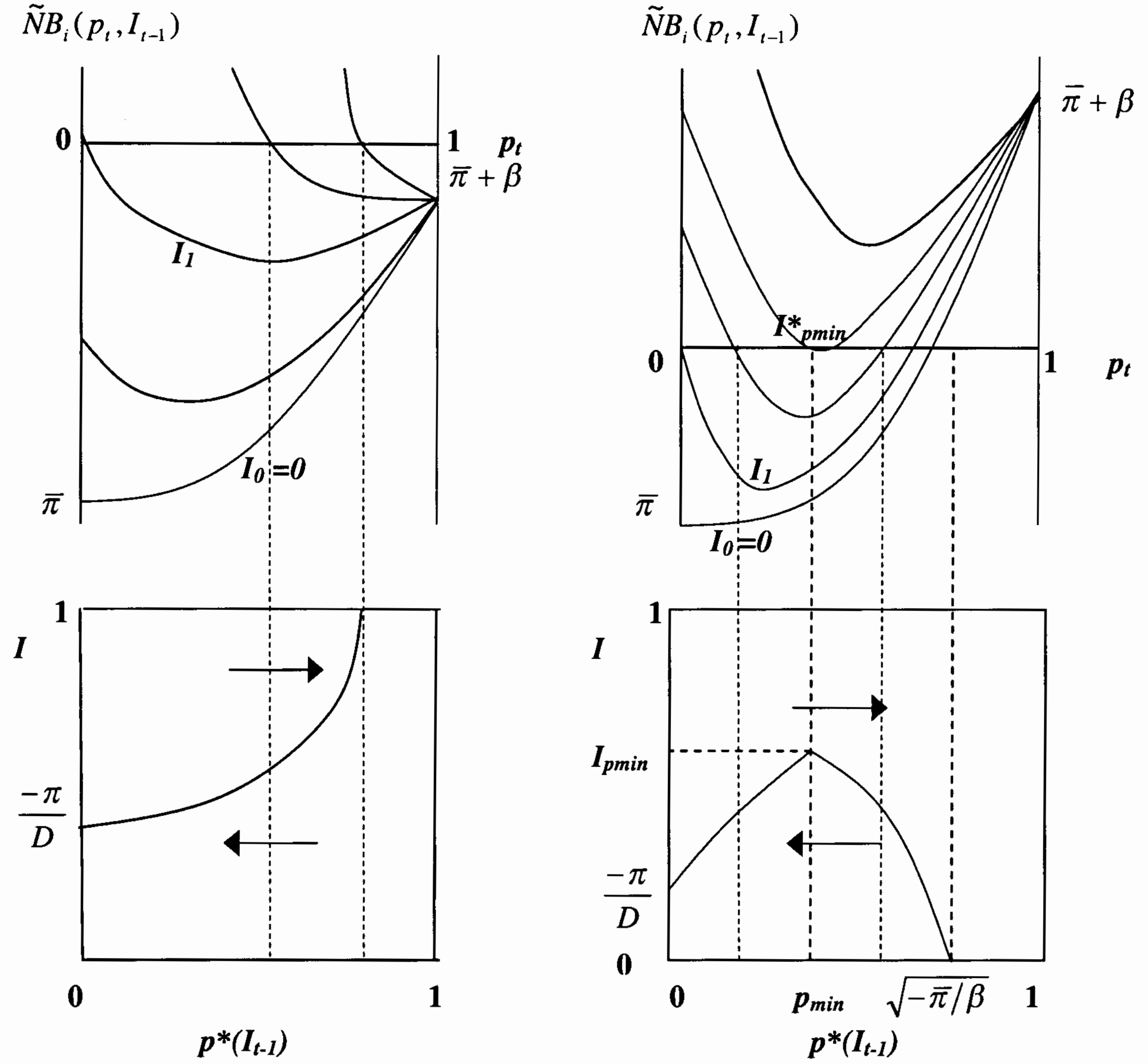
$$\tilde{N}B_i(p_t, I_{t-1}) = \bar{\pi} + \beta p_t^2 + I_{t-1} D - p_t I_{t-1} D \,. \qquad [6\text{-}9]$$

Equation [6-10] represents a parabola with a slope of $2\beta - I_{t-1} D$ and a minimum value at $p_t = p_{min} = I_{t-1} D / 2\beta$. Increasing p_t has two opposing effects on $\tilde{N}B_i(p_t, I_{t-1})$. Firstly, $2\beta p_t$, the marginal effect of p_t on the positive user externality increases, in turn raising the chance that the marginal individual will choose A1. Secondly, $I_{t-1} D$, the marginal effect of increasing p_t on the psychological emotional cost, falls because the marginal individual is less altruistic, reducing the chance that he will choose A1.

Figure 6.2 illustrates how to construct phase diagrams, which indicate the direction of motion of p_t given I_{t-1}. Case 1 corresponds to figure 4a and d, while case 2 corresponds to figure 4b and c. In the top panel, $\tilde{N}B_i(p_t, I_{t-1})$ is plotted for various levels of I_{t-1}; The parabola shifts upwards as I_{t-1} increases. Four critical levels of I_{t-1} are required to construct the phase diagrams; $I_0 = 0$, $I_1 = -\bar{\pi}/D$, $I_2 = I^*_{pmin}$, and $I_3 = 1$. To find the level of I_{t-1} for which $p^*(I_{t-1})=0$ is a stationary point, set $\tilde{N}B_i(0, I_{t-1}) = \bar{\pi} + I_{t-1}D = 0$. Hence ($p^*=0$, $I_1^* = -\bar{\pi}/D$) corresponds to a stationary point. For $0 \le I < -\bar{\pi}/D$, $p^*_L(I_{t-1})<0$, giving a corner stationary point at $p^*(I_{t-1})=0$. For each level of I_{t-1}, the minimum of the parabola occurs at $p_{\min} = I_{t-1}D/2\beta$. p_{min} is stationary if $\tilde{N}B_i(p_{\min}, I_{p\min}) = 0$, giving $p^*_{\min} = 1 - \sqrt{(\beta + \bar{\pi})/\beta}$ and the corresponding I_{t-1} for which p^*_{min} is stationary, $I^*_{p\min} = \left(1 - \sqrt{\beta + \bar{\pi}/\beta}\right) \times 2\beta/D$.

The phase diagrams (bottom section of figure 6.2) are constructed from the top graphs by plotting $p^*(I_{t-1})$, the stationary values of p_t for each level of I_{t-1}. $p^*(I_{t-1})$ is indicated on the top graphs by the levels of p_t for which the parabolas intersect the $\tilde{N}B(p_t, I_{t-1}) = 0$ axis. The arrows on the phase diagrams indicate the direction of motion for p_t given I_{t-1}.

Figure 6.2: Construction of Phase Diagrams for p_t Given I_{t-1}

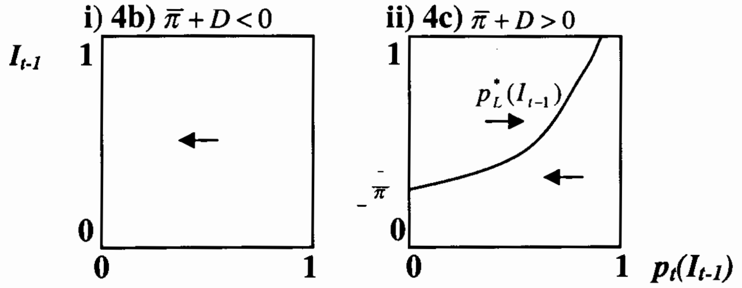

Case 1: $\bar{\pi} + \beta < 0$

Case 2: $\bar{\pi} + \beta > 0$

The set of possible phase diagrams for all four cases is illustrated in figure 6.3.

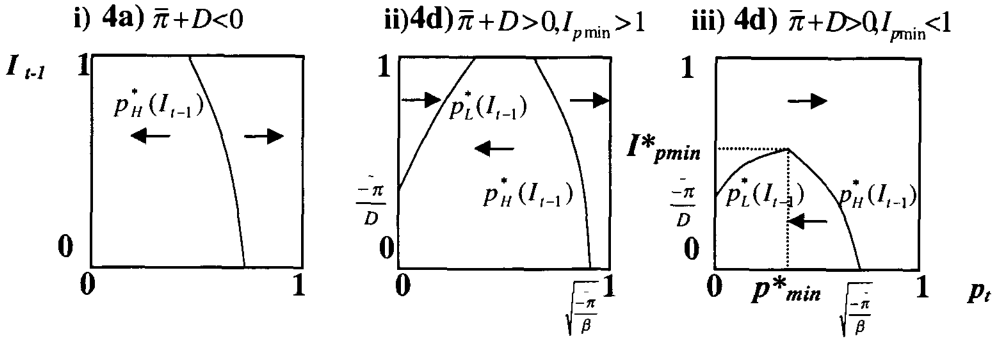
Figure 6.3: Direction of Motion of p_t given I_{t-1}

Case 1: $\bar{\pi} + \beta < 0$


i) 4b) $\bar{\pi} + D < 0$

ii) 4c) $\bar{\pi} + D > 0$

Case 2: $\bar{\pi} + \beta > 0$



i) 4a) $\bar{\pi}+D<0$

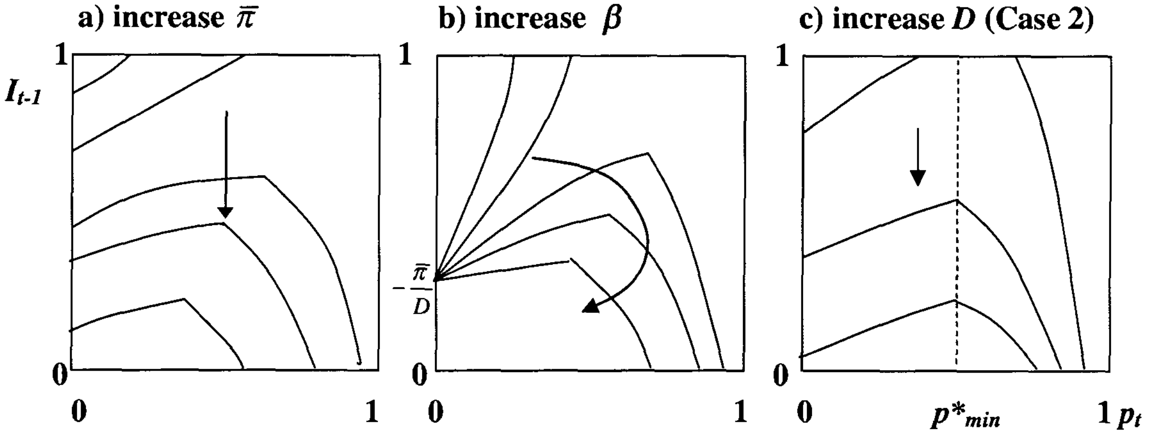
ii) 4d) $\bar{\pi}+D>0, I_{p\min}>1$

iii) 4d) $\bar{\pi}+D>0, I_{p\min}<1$

6.4 Effects of Policies and Exogenous Changes in Damages Upon p^* (I_{t-1})

Maintaining the assumption that I_{t-1} is exogenous, phase diagrams are used to examine the effects of changes in the parameters β, $\bar{\pi}$, and D upon the $p^*(I_{t-1})$ loci, as illustrated in figure 6.4. Secondly, policies that target information accuracy or limit the level of A2 may serve to constrain the set of possibilities.

Figure 6.4: Shifts of the $p^*(I_{t-1})$ Locus



a) increase $\bar{\pi}$

b) increase β

c) increase D (Case 2)

An increase in the net private non-user benefits of A1 verse A2, $\bar{\pi}$, can be attained through policy instruments including automobile taxes, carbon taxes, subsidies to public transport, technologies that raise the relative cost of automobiles, bus lanes, oil shortages, or an increase in the number of low income types. Figure 6.4a illustrates that increasing $\bar{\pi}$ shifts the $p^*(I_{t-1})$ locus downward as indicated by the arrow. A sufficiently large increase in $\bar{\pi}$ shifts the $p^*(I_{t-1})$ locus downward so that $I_{t-1} > I^*_{pmin}$, allowing the system to converge to the $p^*(I_{t-1}) =1$ equilibrium. At high levels of information bias, the magnitude of a price policy such as a tax on A2 needs to be larger, increasing costs.

An increase in β reflects an increase in the magnitude of the user externality associated with A1. As β increases, the $p^*(I_{t-1})$ locus shifts as indicated by the direction of the arrow. For low levels of β, only the most altruistic types choose A1 and only if $I_{t-1} > -\bar{\pi}/D$. As β increases, A1 becomes materially more desirable so that for $I_{t-1} > I^*_{pmin}$, the proportion of A1 users increases until $p^*(I_{t-1}) = 1$.

Figure 6.4c illustrates an increase in the level of damages for case 2, due to a shift in the $p^*(I_{t-1})$ loci as indicated by the arrow. As damages rise, the number of A1 users may increase as the $p^*_L(I_{t-1})$ increases for a given level of I_{t-1}. If the locus shifts such that $I_{t-1} > I_{pmin}$, the whole population may come to adopt A1.

Quantity policy instruments can potentially be used to constrain activity choice or information bias. For example the number of A2 users can be limited by placing a ceiling on the number of automobile licenses. Information bias can be targeted by means of education programs, government advertising campaigns, or the regulation of corporate advertising. In figure 6.4a, if information bias is high, a large tax on cars is required to shift the $p^*(I_{t-1})$ locus down far enough such that $I_{t-1} > I_{pmin}$. An alternate approach is to target the accuracy of information causing I_{t-1} to increase, allowing for a lower tax. Targeting information accuracy in the absence of price instruments, such as a tax on A2, is insufficient to bring about social efficiency for starting conditions with $p_t < p^*(1)$ and $I^*_{pmin} > 1$.

In the sections 6-1 to 6.4 the information bias index was held exogenous. The effects of activity choices upon the information index due to induced dissonance and selective attention are modeled in the next section.

6.5 The Demand for Information to Alleviate Cognitive Dissonance

In the above section, utility maximizing individuals chose activities based upon private material and psychological payoffs given the "pooled" information received at the start of each period. An individual may then experience dissonance, the magnitude of which depends upon his beliefs regarding external damages and his level of altruism. Agents may *ex poste* selectively attend to information in order to alleviate dissonance, affecting the demand for information. The individual information demands are aggregated to get the aggregate demend for information, which contributes to the supply of information received at the beginning of the next period, in turn affecting activity choices and so on.

The assumption that agents receive identical information prior to activity choice may be regarded as a benchmark case in which the dependence on shared information sources is maximized. At the opposite benchmark, at the start of the period, each individual receives the individual specific information demanded last period. As this individual specific information validates last period's actions, there is no incentive to change actions. Individuals "hear what they want to hear" and information has no effect on behaviour. The pooled information benchmark is examined below.[46]

To model the information generating process, assume that there are two information sources, which for heuristic purposes are referred to as newspapers. If society has a correct estimate of damages, then A1 is socially efficient. Alternately, *D* can be regarded as the expert consensus best estimate of damages. Newspaper 1 (N1) provides truthful information regarding damages with an information index of 1. Newspaper 2 (N2) provides mixed information such that the on net the information index is 0 for maximum possible bias.[47] Altruistic A2 players that read N1 experience dissonance due to receiving information that their behaviour is asocial. N2, on the other hand, serves to mitigate these concerns. After choosing activities, individuals can selectively attend to the news source that provides the greatest dissonance relief. A2 players can attempt to

---

[46] Due to the complexity of the model, intermediate cases in which there is a pooled and individual specific component to information received are not analyzed.

[47] Newspapers play the analogous role to religions (R1 and R2) in section 4.

reduce dissonance by increasing their demand for N2. Similarly, for A1 players, reading N1 may mitigate concerns that their net private material losses are socially justified; A moral boost from behaving socially is only sustainable if agents believe that that A1 is socially preferred.[48]

For a given level of information accuracy and damages, dissonance is increasing in the level of altruism. Hence, more altruistic types are likely to have a greater impact upon the aggregate demand for information about damages. The intensity of demand for such information by an individual of type θ_i is described as $M(\theta_i) = \gamma\theta_i^{\rho}$, where $\gamma > 0$ and $\rho > 0$.

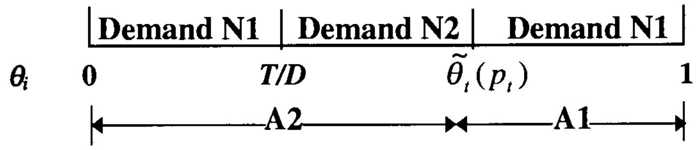
Altruistic A2 players experience dissonance if $I_{t-1}D > 0$ and should prefer N2, while altruistic A1 players should prefer N1. However, information that is perceived to be false should not affect beliefs about damages. Assume that N1 readers recieve a reward T (in utility metric), where $T \in \Re$ is the average net reward for searching truthfully. N2

---

[48] The way in which individuals form beliefs under situations of uncertainty has long been a concern of economists. Under the traditional approach, individuals form subjective beliefs that assign probabilities to each possible state of the world and then use Bayes' Rule to update these beliefs. In this model, due to issue complexity, information consumers delegate the task of information search and belief formation to the media, here represented by two newspapers, thereby simplifying their own belief formation process to a binary task of choosing between two sources of information. By assumption, N1 and N2 respectively provide unbiased and biased information. Although the model does not specify the process by which these newspapers estimate damages, they each could conceivably use Bayesian updating. If so, N1 has engaged in "truthful" methods of forming subjective beliefs over probability distributions while N2 has been influenced by supply-side factors, which have biased the process of forming subjective beliefs. Information consumers (the newspaper readers) are influenced by the psychological factor of dissonance and hence their behaviour departs from standard assumptions of Bayesian updating. Presumably a Bayesian newspaper reader would form subjective beliefs about the probability that each newspaper provides accurate content and then employ Baye's rule to update beliefs. For discussions of departures form Bayesian learning in a psychological context, see Rabin (1998) and Rabin and Schrag (1999). Note that it is not necessary to assume that information consumers are myopic, updating is carried out by information specialists, the media, and past social consensus beliefs do affect this updating.

readers receive no reward. T should increase if the population has universal access to high quality education because people will have the factual and knowledge base to understand and to receive the benefits of comprehension. In a participatory society in which individuals believe that they can affect policy outcomes, there will be greater individual rewards from making the effort to understand social issues. Indeed, as more persons engage in understanding these issues, individuals can experience benefits due to shared conversations. T decreases (and can even become negative) as the costs of attaining unbiased information increase. For example if the population is poorly trained in critical thinking and lacks a factual and knowledge base to understand the relevant information, it will be costly to acquire unbiased information. Supply side bias (section 5) and information abundance may also increase the costs of truthful search. For simplicity, no relationship between the desire to search truthfully and the altruism index is assumed. A1 players recieve $\pi_1(p_t)+T$, for reading N1 but only $\pi_1(p_t)$ for reading N2 and so choose N1 if $T>0$. If $T<0$, A1 players will choose N2. A2 players receive $\overline{\pi}_2-\theta_i D+T$ for reading N1 and only $\overline{\pi}_2$ for reading N2. Hence, A2 players choose N1 if $\theta_i < T/D$, or N2 if the equality is reversed.

Figure 6.5: Demand for Information about Damages



Let $\tilde{\theta}_t(p_t)=1-p_t$, where $\tilde{\theta}_t(p_t)$ is distinguished from $\hat{\theta}$ in that equilibrium with respect to activity choice is not assumed. Individuals with $\theta_i > \tilde{\theta}_t(p_t)$, play (A1, N1) while those with $\theta_i < \tilde{\theta}_t(p_t)$ choose A2. If truthfulness of search is low such that $T/D < \tilde{\theta}_t(p_t)$, a fraction of the population, $\tilde{\theta}_t(p_t)-T/D$, chooses N2. The most altruistic A2 players have the greatest incentives to search less truthfully due to higher dissonance. These altruistic A2 players have a large effect upon the aggregate demand for information due to a high intensity of demand. Hence, altruism combined with a low truthfulness can generate a high demand for misinformation. Under the standard pure free-rider assumption, with $\theta_i = 0$ for all agents, there are no dissonance effects to generate this demand for misinformation.

The individual demands are aggregated in order to generate the aggregate demand for each newspaper. An individual's demand for information is assumed to be proportional to their intensity of demand $M(\theta_i)$, so that for individuals on the interval $d\theta_i$, the share of the total demand for information is $M(\theta_i)d\theta_i$. Integrating over θ_i, given the bounds illustrated in figure 6.5, gives the aggregate demand for each newspaper. The aggregate demand for N1 is

$$AD_t^{N1}(p_t) = \int_{\tilde{\theta}_t(p_t)}^{1} M(\theta_i)d\theta_i + \int_0^{T/D} M(\theta_i)d\theta_i = \frac{\gamma}{\rho+1}\left(1-\tilde{\theta}_t(p_t)^{\rho+1} + (T/D)^{\rho+1}\right). \qquad [6\text{-}10]$$

The aggregate demand for N2 is

$$AD_t^{N2} = \int_{T/D}^{\tilde{\theta}_t(p_t)} \gamma M(\theta_i)d\theta_i = \frac{\gamma}{\rho+1}\left(\tilde{\theta}_t(p_t)^{\rho+1} - (T/D)^{\rho+1}\right). \qquad [6\text{-}11]$$

Similarly, the aggregate demand for both newspapers is

$$AD_{Total} = \frac{\gamma}{\rho+1}. \qquad [6\text{-}12]$$

Dividing [6-10] and [6-11] respectively through by [6-12] gives the relative demand for each newspaper;

$$RD_t^{N1}(p_t) = 1-\tilde{\theta}_t(p_t)^{\rho+1} + (T/D)^{\rho+1} \quad \text{and} \quad RD_t^{N2}(p_t) = \tilde{\theta}_t(p_t)^{\rho+1} - (T/D)^{\rho+1}. \qquad [6\text{-}13 \text{ a \& b}]$$

Since N1 reports unbiased information about damages, its information content is $I_{N1}=1$, while N2's information content is $I_{N2}=0$. The aggregate demand for information content for the newspaper industry as a whole is a weighted average of the information contents of the two newspapers, where the weights are the respective relative aggregate demands. If $p_t < 1-T/D$ and $T/D \geq 0$, then the aggregate information demand is

$$I_t^D(p_t) = RD_t^{N1}(p_t)I_{N1} + RD_t^{N2}(p_t)I_{N2} = 1-\tilde{\theta}_t(p_t)^{\rho+1} + (T/D)^{\rho+1}. \qquad [6\text{-}14a]$$

If $p_t \geq 1-T/D$ and $T/D \geq 0$, then $I_t^D(p_t) = 1$. [6-14b]

If $T/D < 0$, then $I_t^D(p_t) = 0$. [6-14c]

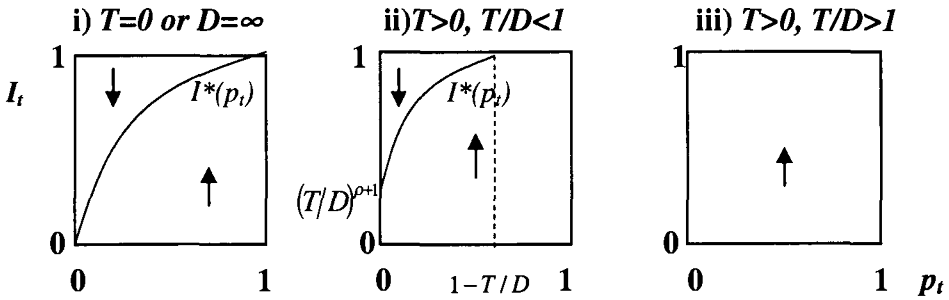
Given the proportion of A1 players, $p_t = 1 - \tilde{\theta}_t(p_t)$, the steady state condition for the pooled information content can be derived. If $I_t^D(p_t) < I_{t-1}$, where I_{t-1} is the supply of pooled information received at the start of period t, then the supply of accurate information exceeds the demand and the information index falls to become more biased next period. Similarly, information becomes less biased if the inequality is reversed. If $I_t^D(p_t) = I_{t-1}$ the information index has reached an equilibrium for a given level of p_t. Noting that $\tilde{\theta}_t(p_t) = 1 - p_t$, and substituting in [6-14a], for this case, the stationarity condition for the information index given p_t is

$$I^*(p_t) = 1 - (1 - p_t)^{\rho+1} + (T/D)^{\rho+1}. \qquad [6\text{-}15]$$

$I^*(p_t)$ is increasing in p_t at a decreasing rate indicating that the steady state level of information bias falls as more people choose A1. The slope is decreasing in p_t because as the number of A2 players fall, there is less demand for N2 in order to reduce dissonance. Furthermore, as p_t increases, the altruism index of the marginal individual falls providing less incentive to demand N2. Substituting the boundary conditions for p_t into [6-16] gives $I^*(0) = (T/D)^{\rho+1} \geq 0$ and $I^*(1) = 1 + (T/D)^{\rho+1} \geq 1$.

Three possible phase diagrams are illustrated in figure 6.6. Arrows indicate the direction of motion of I_t for a given level of p_t. In case i, with low positive truthfulness of search or alternately infinitely high damages, $I^*(p_t)$ is increasing at a decreasing rate in p_t and $I^*(0) = 0$ and $I^*(1) = 1$. In case ii, with low positive truthfulness relative to damages such that $T/D < 1$, $I^*(p_t)$ increases at a decreasing rate if p_t increases and $1 > I^*(0) > 0$ and $I^*(1) > 1$. In case iii with $I^*(0) \geq 1$, I_t is increasing for all p_t if $I < 1$ with a stationary state at $I^*(p_t) = 1$. In a fourth case, not illustrated, if $T<0$, I_t is falling for all levels of p_t.

Figure 6.6: Phase Diagrams for I_t given p_t



i) $T=0$ or $D=\infty$ ii) $T>0$, $T/D<1$ iii) $T>0$, $T/D>1$

6.6 Effects of Parameter Changes upon the $I^*(p_t)$ Locus

The $I^*(p_t)$ locus can be shifted by changing T, ρ *or* D. Increases in damages increase dissonance *ceteris paribus* and may hence induce agents to demand more biased information. In case 6.6i, there are no benefits from truthful search and hence damages do not affect the $I^*(p_t)$ locus. In case 6.6ii, with $T<D$, an increase in damages shifts the $I^*(p_t)$ locus shifts downward to the right. In case 6.6iii, an increase in damages has no effect as long as $T>D$. If D surpasses T, the phase diagram comes to resemble case ii.

An increase in the truthfulness of search, T, has the opposite effect of an increase in damages. In case 8.6-ii the $I^*(p_t)$ locus shifts upward to the left, decreasing the equilibrium level of information bias given p_t. Policies to increase T reduce the effects of selective attention, helping to lower the level of information bias.

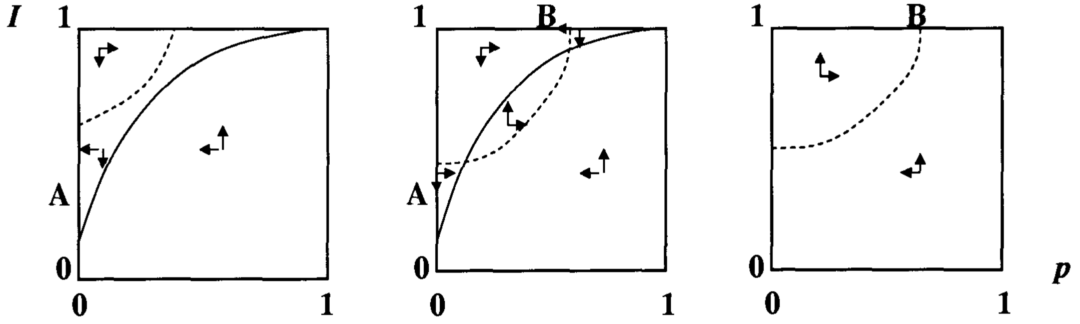
Lastly, an increase in ρ raises the relative intensity of demand for information of the most altruistic types, causing the $I^*(p_t)$ locus to shift downwards in case 8.6-ii. This reflects the fact that altruistic A2 players have a relatively large influence upon the demand for N2.

6.7 Dynamics for the Complete System

The dynamics of activity choices, $p_t(I_{t-1})$, and the information index, $I_t(p_{t-1})$, have been examined above while taking the respective arguments as given. In the following sections, these arguments are endogenized and phase diagrams are used to jointly determine the dynamics of activity choices and informaiton bias. The examples are labelled according to their correspondence with figures 6.3 and 6.6. The solid and dashed lines respectively represent the $I^*(p_t)$ and $p^*(I_{t-1})$ locus.

The phase diagram (figure 6.7) corresponds to figure 6.3 (Case 1ii) with $\bar{\pi}_2 + \beta < 0$ indicating that all individuals privately materially prefer A2 for all levels of p_t. Damages are relatively high with $\bar{\pi}_2 + D > 0$; it is socially efficient for a given agent to choose A1 for all p_t. In case i, if information bias is low, the most altruistic types choose A1. However low information bias cannot be maintained due to a low levels of truthful search relative to damages. There is a stable equibrium at point A($p^*=0$, $I^* = (T/D)^{\rho+1}$). In case ii, the private material benefits of A1 are higher than in case i inducing more agents to choose A1 for a given level of I_{t-1}. If information bias is initially high, the system may again converge to the equilibrium at point A. However, for a large variety of starting conditions, the system converges to point B, where information bias is low and the most altruistic types choose A1. In case iii, $T/D > 1$ such that I increases for all $I_{t-1}<1$; There is a stable equilibrium at B with $I^*=1$ and the most altruistic types choosing A1.

Figure 6.7: Case 1-ii, $\bar{\pi} + \beta < 0$, $\bar{\pi} + D > 0$



i) 6.6-ii, *T>0, T/D<1*

ii) 6.6-ii, *T>0, T/D<1*

iii) 6.6-iii, *T>0, T/D>1*

Figure 6.8 corresonds to figure 6.3 (case 1ii) with a small user externality ($\bar{\pi} + \beta < 0$) and low damages ($\bar{\pi} + D < 0$), such that if p_t=0 it is socially efficient for a *given* individual to choose A2. As damages are low, so is dissonance, which is then of insufficient magnitude to offset the private net benefits of A2 so that p falls for all $p_t>0$. In case i, $T/D < 1$, leading to dissonance induced information deterioration. As information bias rises, so does dissonance causing p_t to fall. There is a stable equilibrium at point A ($p^*=0,\ I^* = (T/D)^{\rho+1}$). In case ii, truthfulness of search is relatively high allowing for perfectly unbiased information, however as p_t falls for all $p_t>0$, there is a stable equilibrium at A ($p^*=0,\ I^*=1$).

Figure 6.8: Case 1-i, $\bar{\pi} + \beta < 0$, $\bar{\pi} + D < 0$

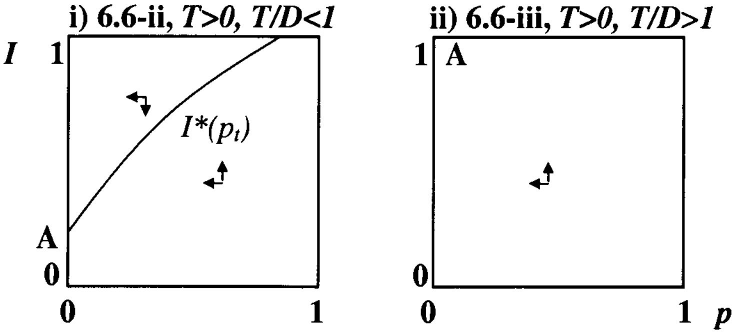
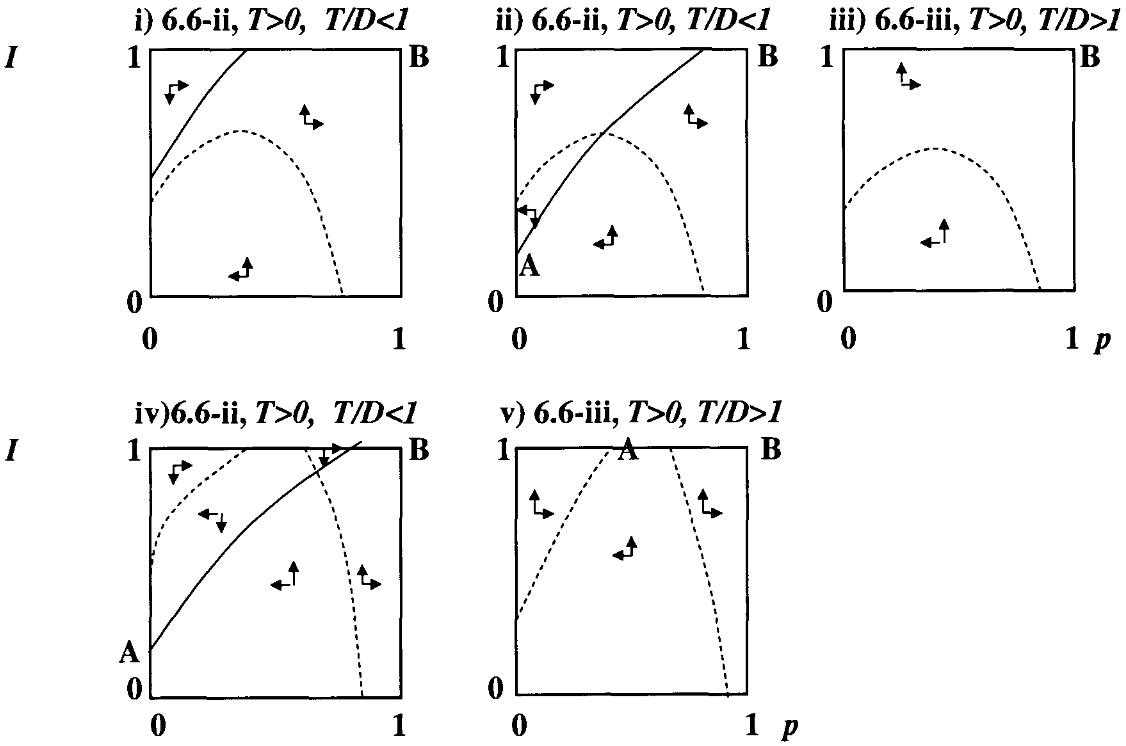


Figure 6.9 illustrates case in which $\bar{\pi}_2 + \beta > 0$ such that all individuals privately prefer A1 for $p_t > p_C^F$. Damages are high relative to private non-user material benefits ($\bar{\pi}_2 + D > 0$), such that for all p_t, it is socially efficient for a given player to choose A1. In cases i-iii, $I^*_{p\min} < 1$, so that if I initially exceeds I^*_{pmin}, the proportion of A1 users tends to rise allowing for convergence to point B($p^*=1,\ I^*=1$). In case ii, $T/D < 1$ so that if initially p_t is low and $I_{t-1} > I^*(p)$, information deteriorates allowing for convergence to point A ($p^*=0,\ I^* = (T/D)^{\rho+1}$). An increase in T shifts the $I^*(p_t)$ locus up moving the system from case ii to case i, allowing for convergence point B.

In cases iv and v, $I < I^*_{pmin}$, so that given initial conditions $p < p_H{}^*(I)$, p and I remain in the basin of attraction of the low welfare equilibrium,A, independent of I. In case iv there are two equilibria; points A ($p^*=0,\ I^* = (T/D)^{\rho+1}$) and B($p^*=1,\ I^*=1$). In case v, there are two equilibrium at points A($p^*=0,\ I^*=1$), and B($p^*=1,\ I^*=1$). A policy such as a tax on A2 shifts the $p^*(I)$ locus downward, lowering I^*_{pmin}. For example, if the system starts out at point A in case v, a tax on A2 could move the system to that depicted in case iii allowing for convergence to the high welfare equilibrium (B). However, if instead the system starts out at point A in case iv, a much higher tax is required to shift the *p*(I)* locus down far enough to achieve the situation in case i, allowing for convergence to point B. A lower tax could however be combined with policy to increase *T*.

Figure 6-9: Case 2-ii, iii, $\bar{\pi}+\beta>0$, $\bar{\pi}+D>0$


i) 6.6-ii, *T>0, T/D<1*

ii) 6.6-ii, *T>0, T/D<1*

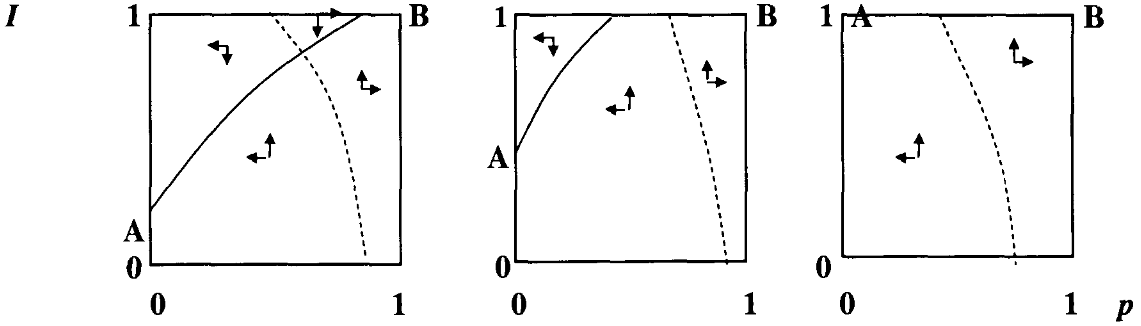
iii) 6.6-iii, *T>0, T/D>1*

iv)6.6-ii, *T>0, T/D<1*

v) 6.6-iii, *T>0, T/D>1*

Figure 6.10 represents a situation with a large user externality ($\bar{\pi}+\beta>0$), such that if p_t exceeds the critical mass, p_C^F, it privately materially pays for all individuals to switch to A2. Damages are low ($\bar{\pi}+D<0$) such that if *p=0*, it is socially efficient for a given individual to play A2. In all three cases, given initial conditions in which $p>p^*_H\,(I_{t-1})$, the system can potentially converge to the high welfare equilibrium at B(*p**=1, *I**=1). If initially $p<p^*_H\,(I_{t-1})$, the system converges to A (*p*=0*, $I^*=(T/D)^{\rho+1}$) in cases i and ii or A(*p**=0, *I**=1) in case iii. If p_t is initially low, policies that target information accuracy or truthfulness are insufficient to attain social efficiency.

Figure 6.10: Case 2, 4a) $\bar{\pi} + \beta > 0$ $\bar{\pi} + D < 0$



i) 8-ii, *T>0*, *T/D<1* ii) 8-ii, *T>0*, *T/D<1* iii) 8-iii, *T>0, T/D>1*

The above is a representative sample of the set of possible phase diagrams for the complete system. In order to precisely determine which equilibrium is reached given initial conditions, the time dynamics would need to be modelled. For example, if activities are costly to switch relative to information choices, then vertical movements should be rapid relative to horizontal movements. This situation seems likely given that switching news sources is inexpensive relative to switching tranport choices. The latter may require sale or purchase of an automobile and possibly residential location. There may however also be inertia in changing information sources. For example, concession that prior information choices were poor may call into question an individual's perceptions regarding his status as a discerning information searcher. Agents organize information in modules referred to as schemas; new information that conflicts with existing schemas may be ignored (Wicks; 2001). Other desirable components of the initial news source, unrelated to environmental damages may also be forgone.

6.8 Shifts of the $I^*(p_t)$ and $p^*(I_{t-1})$ Loci due to Policy and Exogeneous Shocks

The effects of policies and exogenous shocks depend upon starting conditions and parameter specifications. A few representative examples are discussed below.

6.8.1 Changes in the Level of Perceived or Actual Damages

Automobiles became a social norm prior to the widespread scientific appreciation of enviromental damages. The initial situation then corresponds to a situation of low damages both due to low levels of greenhouse gases in the atmosphere and lack of scientific research. With further research and greenhouse gas accumulation, the best scientific estimate of damages, corrsponding to D, increased (IPCC, 2001). The initial situation might correspond to figure 6.11-i, point A, in which $I = 1$ and all but the most altruistic types choose automobiles.[49] An increase in damages has two effects. Firstly the $p^*(I)$ locus shifts downward due to increased dissonance inducing altruistic types to switch to public transport. Secondly A2 players experience increased dissonance inducing selective attention to N2 causing the $I^*(p_t)$ locus to shift down. Recalling that $I^*(0) = (T/D)^{\rho+1}$ and that $p^*(I_{t-1}) = 0$ if $I_{t-1} = -\bar{\pi}/D$, there are then two possible cases.

Firstly, in figure 6.11-ii $T + \bar{\pi} > 0$ so that the $I^*(p_t)$ locus is situated above the $p^*(I_{t-1})$ locus. T is sufficiently large that the system converges to the high welfare equilibrium at B. Secondly, in figure 6.11-iii, $T + \bar{\pi} < 0$ so that the $I^*(p_t)$ locus is situated below the $p^*(I_{t-1})$ locus allowing selective attention to generate large reduction in information accuracy. If news sources are costly to change relative to activity choices, the system can converge to point B. However, in the likely case of high inertia in activity choices, information accuracy deteriorates relatively quickly and the system converges to the low welfare equilibrium at A. Hence, given a combination of low T and high private material benefits of A2, an increase in damages is unlikely to move the system out of the basin of attraction of the low welfare equilibrium. To conclude, an increase in damages and/or experts best estimate of damages has an ambiguous effect upon the amount of damaging behaviour and information accuracy.

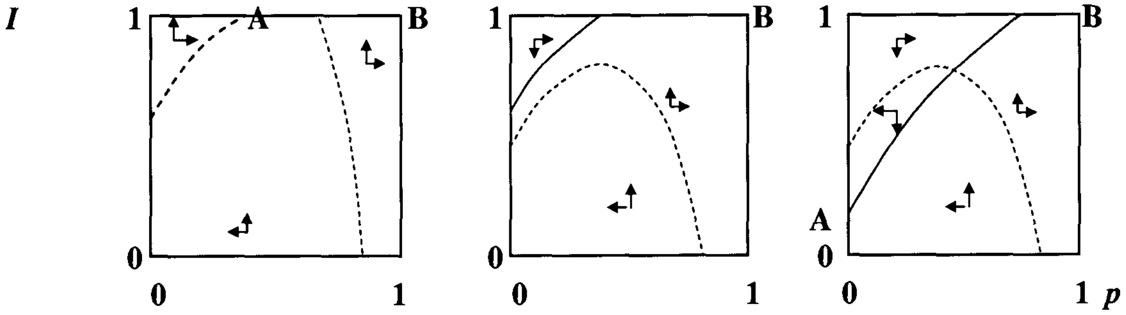
Studies suggest that media coverage of global warming has not increased proportionately as expert estimates of damages have been revised upwards (Williams, 2001). Furthermore, in North America greenhouse gase emissions from private transport vehicles have increased in spite of increased fuel efficiency due to factors such as increases in car size. According to the model, these findings suggest a low truthfulness of search. However, although low information accuracy may be due to

---

[49] Alternately, if there is no critical mass given low initial taxes on A2, the initial situation corresponds to figure 6.7-iii.

dissonance in conjuction with a low T, it may also be explainable due to efforts by vested interests such as oil and automobile manufacturers to provide information that understates damages (section 5). Policies to achieve the situation in case ii include raising the relative private material benefits of A1 and increasing T. The optimal combination would theoretically take into account cost effectiveness and political feasiblity.

Figure 6.11: Case 2, $\bar{\pi} + \beta > 0$, (3-2-ii and iii)



i) 6.6-iii or 6.7-iii low D

ii) 6.6-ii, $T + \bar{\pi} > 1$ increase D

iii) 6.6-iii, $T + \bar{\pi} < 1$ increase D

6.8.2 Policies that Target Private Material Net Benefits

Standard incentive based policies affect the $p^*(I_{t-1})$ locus. Supposing that the system starts out at point A in figure 6.12-i, with a low or no tax on gasoline or cars corresponding to the hypothesized current status quo.[50] Given a high damage to truthfulness level, information accuracy is initially low. Case ii illustrates that given a continued low T/D ratio, a large tax is required to move the system to the equibrium at B. An alternate is to combine a lower tax with a policy to increase T (case iii), with the optimal policy mix depending upon the relative costs of raising T verses implementing the tax. Another potential problem is that as individuals switch to A1, vested interests may step up their efforts to supply information that understates the social costs of A2, increasing the cost of truthful search so that the $I^*(p_t)$ locus shifts down.

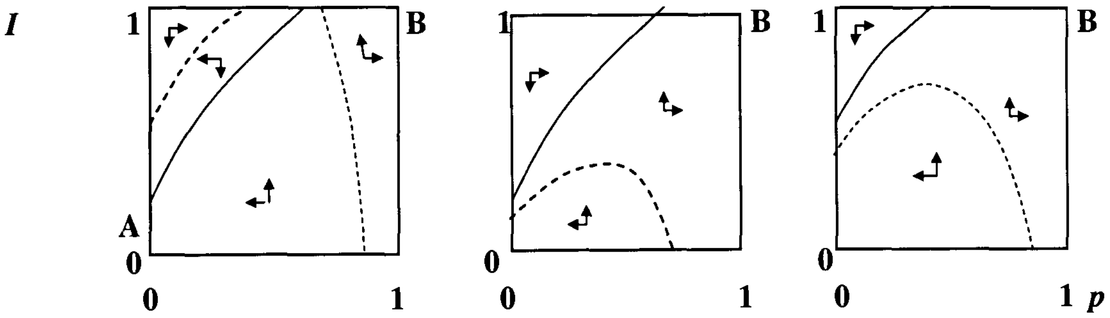
Price policies that target private material incentives are the standard approach advocated by environmental economists. Yet it is currently difficult to implement policies such as

---

[50] Alternately, the initial situation corresponds to figure 6.7-iii (see footnote 39).

carbon or automobile taxes due to an absence of political consensus. The example illustrates that the deterioration of the accurancy of information due to dissonance may contibute to thwarting the attainment of such a political consensus. Policies that aim to increase T or to temporarily raise I above equilibrium levels might facilitate the achievement of such a consensus.

Figure 6.12: Case 2, $\bar{\pi} + \beta > 0$, (3-2-ii and iii)



i) 6.6-ii or 6.7-iii, no tax **ii) 6.6-ii, high tax** **iii) 6.6-ii, low tax, increase T**

6.8.3 Policies Targeting the Level of Information Bias and Truthful Search

Section 5 provided evidence of bias in the supply of environmental information due to factors such as vested interests, characteristics of environmental information and media structure. These supply side factors increase the cost of truthful information search, lowering T. Policies to counteract these effects include increasing T and raising the accuracy of information above its equilibrium level. Policies to increase T include regulating corporate advertising with a possible switch to a higher proportion of public finance, training of scientists to convey scientific ideas in a format that the public can comprehend, and improved scientific training of the media. Governments can attempt to increase the accuracy of information by providing programs and advertisements about the environment, similar to anti-smoking advertisements currently in widespread use and also by increasing enviromental education requirements in schools and teacher training.

In cases in which $I_{t-1} > I_{pmin}$ and $I_{pmin} < 1$, policies to increase the accuracy of information above its equilibrium level can allow for convergence to the high welfare equilibrium, as long as information does not depreciate too rapidly. However given that private transport choices may be costly to change relative to news sources, such a campaign might need to be maintained for a long time with potentially high costs.

Depending upon governments to directly intervene to affect the information provided by media and schools has the drawback that governments may have incentives to provide biased information. A more decentralized but longer term approach for increasing T is to educate audiences to be more discerning information searchers. This includes low-cost universal access to high quality information sources including government and academic publications, public libraries, internet, and higher level education. Education programs should emphasize the development of critical thinking. This might include training students to be more discerning media consumers, for example, by teaching students to be aware of psychological propensities to selectively attend to information and also to be influenced by corporate advertising. Although in a highly technological society, it is important to train specialists, a greater emphasis on general education may provide students with the background to be more discerning media consumers (Saul; 1995).

6.8.4 Moral Suasion

Policies of moral suasion appeal to people's sense of duty to behave in a prosocial manner. In terms of the model, an effective moral suasion policy increases the average level of altruism. The expression for an individual's social emotional costs can be modified to read $C_i(\theta_i, I_{t-1}) = \gamma\theta_i I_{t-1} D$ where $0 \leq \gamma \leq 1$ where γ is a parameter that can be targeted by moral suasion policies. Increasing γ increases the costs of dissonance shifting the $p^*(I_{t-1})$ locus downwards. The increased dissonance may induce individuals to selectively attend to information causing a downward shift in the $I^*(p_t)$ locus. As in the case of increases in damages, moral suasion policies may either increase or decrease the level of asocial activities, depending upon the magnitude of T relative to γD. This result is in line with the findings of Rabin (1994) (section 1.6). A higher sense of societal obligation may also induce more discerning media use, raising T.

6.8.5 Group Identity

In the model, it was assumed that altruism is directed equally towards all members of society. However, as discussed in section 3, the scope of altruism may be limited to persons within groups with which individuals most strongly identify. Unless an individual has a strong sense of humanitarian identity, the majority of damages due to an individual's greenhouse gas emissions fall upon persons who are outsiders to the groups with which he most stongly identifies. A lower degree of altruism directed towards outsiders, should reduce the psychological social costs of damages. Group considerations can partially be accomodated by limiting the scope of altruism such that $G \in (0,1)$ represents the fraction of the population with which the average individual identifies. The benchmarks $G = 1$ and $G = 0$ respectively represent a humanitarian identity and the standard case of *homo economicus*. The psychological costs to an individual due to dissonance can be written as $C_i(\theta_i, I_{t-1}) = G\theta_i I_{t-1} D$. Similar to the cases of increasing damages and the degree of altruism, increasing group identity has an indeterminate effect upon activity choice. If T is large relative to $\bar{\pi}$, then a policy to broaden group identity could decrease A2 use. However if T is relatively low, A1 use could fall as G increases.

Assuming that G is exogenous does not take into account the possiblity that group identity may itself be a choice variable, serving to reduce dissonance as discussed in section 3. By reducing group identity, one reduces the sphere of relevant parties upon which damages fall, allowing for dissonance reduction. For example, an individual may increase altruism towards those with whom he more closely identifies in order to compensate for damages imposed upon larger but less strongly identified groups. An SUV driver that is concerned about contributions to greenhouse gases, could reduce dissonance by lavishing gifts upon close friends and family. Unfortunately, both behaviours contribute to the environmental problem, however gratitude is more likely to be received in the latter case. A more complete model might also take into account group specific information, and include group identity as choice variables.

6.9 Model Summary

It is standard for economists to explain the environmental problem as the result of selfish individuals who lack concern for the external costs of their behaviour. Accordingly, modeling individuals as altruists would intuitively be expected to eliminate the environmental problem. However, in contrast, the model illustrates that due to dissonance-reducing selective attention to information, altruistic preferences can theoretically lead to the same (or lower) levels of environmental activity choices as predicted under the material selfishness assumptions. The altruism model makes the additional prediction that the social consensus information may come to understate the magnitude of environmental damages. Consistent with the model, there is empirical evidence of media bias, whereby environmental issues are undercovered and damages understated. However, further empirical work is needed because due to factors such as vested interests, the assumption of material self-interest would also be expected to generate supply side bias on the side of understatment of environmental damages. Empirical methods to distinguish between the two hypothesis might include surveys and the statitistical analysis of media content and reader attributes to test for a correlation between activity choices and beliefs about related environmental damages.

Many environmentally damaging activities became social norms before the magnitude of the damages was appreciated by experts. The model suggests that if it is costly to acquire accurate information and/or if individuals have poor training in critical thinking (low T), then as damages and/or expert estimates of damages rise, both information accuracy and prosocial behaviour may decrease. Issue complexity, uncertainty in damages, and supply side bias make it costly to acquire accurate environmental information, and so T is likely to be low for many types of damages. Given the assumption of low T, the model is consistent with empirical evidence suggesting that media coverage of global warming has not kept pace with expert estimates of damages. As vested interests may also step up efforts to provide information that under-reports environmental damages, more empirical work is needed to determine to what extent this purported relative decrease in the accuracy of information is due to dissonance induced selective attention.

A model in which the world is inhabited by altruists, intuitively suggests that there should be an important role for policies of moral suasion. However, given a low T, the model illustrated that increases in the level of altruism through moral suasion could paradoxically lead to increases in asocial behaviour. This finding suggests that policies that facilitate truthful information search might not only lead to higher levels of information accuracy and prosocial behaviour but could also enhance the effectiveness of moral suasion policies. Furthermore, as research into environmental damages procedes, uncertainty may be reduced, lowering the cost of information search, potentially decreasing the level of information bias and increasing the level of prosocial behaviour. If critical thresholds are reached, changes in beliefs about environmental damages and activity choices may be dramatic.

Unfortunately, the model predicts a low level of prosocial behavior for a wide variety of particular parameter specifications even if information is perfectly unbiased (examples include figure 6.3; case 1-i and ii, case 2-i and ii). In these cases policies that raise the private material net benefits of A1 verses A2 are required to induce more prosocial behaviour. In case 2-i and ii, although a high welfare equilibirum exists, it is impossible to reach, given asocial starting conditions. Futhermore, dissonance induced selective attention may lead to more information bias, reinforcing lockin to the low welfare equilibrium. In line with Hirshman (1965), changing behaviour through policies that target private material incentives may be a necessary precondition to changing beliefs about appropriate social behaviour. In other words, under a variety of parameter specifications, the standard approaches recommended by environmental economists are still required under the altruism assumption. *Ex post*, information bias may fall on its own as there is no longer a need to reduce dissonance.

Unfortunately, prior to implementing price based policies, asocial activity choices may lead to a dissonance induced deterioration in the accuracy of information, making it difficult to achieve a political consensus for policy. Moreover, government efforts to increase information accuracy in order to attain political support for policy may be costly due to dissonance induced depreciation and also the provision of misinformation by vested interests. The model suggested that a means to overcome this "chicken-egg" type problem is to combine lower relative taxes on A2 with policies that target information, either by temporarily shocking information above equilibrium levels or by increasing the truthfuness of search (see figure 6.9). The advantage of the combined approach is that it may be easier to attain political consensus for a lower tax. Theoretical models could be constructed in order to determine the cost effective combination of policies. Futhermore, political costs to the parties that implement the policy may be lower than anticipated because a switch to the prosocial activities may reduce dissonance and information depreciation helping to generate a political consensus.

The model assumed a binary discrete choice between activities while in reality activities can be consumed in variable quantities. For example, families can own fewer cars and reserve the automobile for specific tasks such as moving furniture or trips to areas with poor public transport. Allowing for non-discreteness has two effects. Firstly, the costs of changing behaviour in degrees may not be high, increasing the probability that partial activity switches will help to achieve critical mass. On the other hand, small insignificant changes may serve to reduce dissonance and hence the psychological incentives to make significant changes.[51]

Although the model focused upon heterogeneity in the level of altruism, other sources of heterogeneity could be examined. For example, in the case of transport choices, low-income types tend to privately prefer public transport, while in other cases high-income types choose the more environmentally friendly products. An example of the latter is the consumption of organic food by high-income types due to high prices. After the intial adoption by high-income types, prices may fall due to innovations, learning by doing, and scale effects. Information about the health and environmental benefits of

---

[51] This may also be an undesirable byproduct of recycling, whereby individuals feel good about recycling products such as recyclable juice and pop bottles, when more environmental alternatives such as eating locally grown fruit and drinking tap water exist.

organic foods may increase through advertising and the media. The price drop could induce middle income types to consume organic foods and so on.[52]

7 Conclusions to Part II

The research of evolutionary biologists, psychologists, and experimental economists suggests that altruistic preferences as motivated by the social emotions may provide an important impetus for prosocial behaviour in humans (see sections 1 and 2). If this is the case, then why is it that much human behaviour seems to be theoretically explainable and empirically consistent with the standard assumption of private material self-interest? A possible answer is that the social emotions provide an impetus for prosocial behaviour but only if such behaviour is not too privately materially costly. If private material interests reach some critical level, individuals may default to free riding but at the cost of dissonance. To reduce dissonance, they may selectively attend to information, the aggregate effects of which can contribute to an incorrect consensus regarding the social appropriateness of behaviour. As a result, in environments in which there are strong incentives to free ride, both the altruism and the standard model of *Homo economicus* can generate similar predictions regarding human activity choices.

The altruism model in conjunction with the dissonance thesis makes the additional prediction that information may come to provide moral validation for standard activity choices. Both of the case studies of section 4 and 5 provided empirical observations consistent with the model. In the case of European transition to market economies, a substantial increase in market activities preceded the changes in information that provided for their moral validation. The timing and rapidity of the Reformation is consistent with a process in which changes in moral information and activity choices reinforce each other in a positive feedback loop. Religious moral information may have contributed to lock in of traditional activities serving to slow transition from feudal to market economies.

---

[52] Note that the environmental merits of organic verses standard pesticide-intensive farming are not well established.

The environmental situation offers possibilities for a high degree of dissonance and selective attention due to what is often a high degree of discordance between private material incentives and activities that promote the social good. As activities often became social norms prior to the availability of expert estimates of substantial environmental damages, consumers were effectively hit with information shocks conveying the news that many of their privately preferred activity choices were asocial. Given what are often substantial private material costs of switching activities, the model predicts that these shocks should induce dissonance and selective attention. Moreover, selective attention is enhanced due to factors such as characteristics of the environmental problem, including issue complexity and uncertainty, media structure, information abundance, and supply side bias due to vested interests. Accordingly, the model predicts that environmental information will be biased in the direction of an understatement of damages. The model is consistent with empirical studies suggesting that the media understates environmental damages and provides insufficient coverage of the environment relative to its importance as a social problem. Moreover, in the case of global warming, as expert estimates of damages increased, there is evidence that the media has provided proportionately less coverage.

Although the altruism model is consistent with these observations, it is inconclusive. In the case of market transition, a model based upon material self-interest, namely the fear of damnation, may explain the switch to religious variants that provided a greater degree of moral validation for market activities. In support of the altruism/dissonance hypothesis, there is evidence of heterogeneity in late medieval piety. In the environmental case, supply side bias based upon vested interests, media structure, and issue complexity provide the standard explanations for biased media coverage. Much media bias is hence potentially explainable based upon the standard assumptions of private material self-interest. Furthermore, as expert estimates of damages increase, firms fearing regulation costs and loss of market share have increased incentives to provide misinformation. Hence, the assumption of material self-interest is consistent with both biased media coverage and evidence suggesting that environmental coverage has declined as expert estimates of damages have increased.

Hence, in spite of consistency of the model with the data, more empirical work is required. In the case of environmental damages, surveys could be used to attempt to find a correlation between activity choices and beliefs regarding the extent of environmental damages. Statistics from media content analysis and target market characteristics may also allow for the testing of an empirical relationship between activity choices and information about the sociality of those choices.

Nevertheless, the increasing empirical support for the altruism hypothesis does call into question the extent to which social scientists depend upon material self-interest as a causal assumption. Hence, this essay can best be viewed as an exploratory study that builds upon the growing literature in altruistic economics. The essay illustrated that altruistic preferences in conjunction with the selective attention/ dissonance thesis may have important implications for behaviour and policy. In particular, in the environmental arena, policies that target information accuracy may complement standard policies that target private material incentives. Furthermore targeting information accuracy in conjunction with moral suasion, a policy that intuitively might be recommended under the altruism assumption, may be ineffective unless preceded by policies that target private material incentives. But there is a "catch-22" as a political consensus for substantive environmental policy is currently absent. The model did however suggest that policies that target information accuracy could reduce the required stringency of standard policies. Furthermore, as individuals switch activity choices, information accuracy might increase on its own due to a reduced need to alleviate dissonance. Future work might also involve an investigation of a cost-effective mix between the two types of policies. Importantly, in the case of global warming, as uncertainty regarding damages and mitigation costs is reduced, it may be easier to search truthfully, raising information levels in turn potentially causing a large change in activity choices with two-way feedback.

Evolutionary theories suggest that the prosocial emotions evolved to promote cooperation in groups, increasing both individual and group welfare (Wilson and Sober, 1994, 1998). Given that selective attention potentially undermines cooperation, it is difficult to explain. I argued that selective attention to information may have evolved as an adaptation that served to reduce dissonance due to asocial activities that were undertaken during resource shortages in ancestral evolutionary environments. During hard times, the marginal private benefit of free riding may increase, inducing agents to behave asocially, but at the cost of dissonance. Free riding can enhance both individual and group fitness. In the latter case the group must also coordinate upon a subgroup with which to free ride with respect to. As information is shared, selective attention to information can potentially be coordinated, allowing for a within group consensus regarding who to ostracize or declare war upon. Formal modeling and empirical support remains for future work. The evolutionary theory, although tentative, suggests that high degrees of selective attention may be expected in situations with high levels of interpersonal competition, for example in market economies.

In the model it was assumed that all agents chose activities based upon a common information pool. However in contemporary society, there is a large and various information supply, potentially allowing for much heterogeneity in the information received. This heterogeneity may allow for group specific activity choices and associated moral information. Similarly in the Late Middle Ages, a town that had adopted significant levels of market activities might develop its own "market-friendly" religion, thereby reinforcing market activities and so on. The new religion and market activities could then spread by a process of group selection. To contrast, consider a specific group that adopts environmentally friendly activities and then selectively attends to information provided by sources such as environmental groups and alternative news sources. Because the benefits due to the group's reduction of damages fall upon both insiders and non-environmental types, there is no between group benefit to environmental behaviour and so this behaviour and information cannot spread by group selection.[53]

---

[53] In ancestral environments, as environmental damages were largely local, environmental social customs and possibly innate propensities to take care of the environment may have evolved by group selection. It is plausible that the "natural ascetic" evolved as a means to promote environmental conservation in ancestral environments.

Much economic literature attempts to provide comparative moral justifications for policies and economic systems. This essay suggests that one approach to studying the history of economic thought is to view it as an evolving system of moral justification. For example, if a large literate proletariat had exited in the 18th century, would a socialist literature have developed earlier? Some other applications to consider include the different attitudes regarding the morality of slavery in the North and South prior to the Civil war in view of different incentives regarding the economic desirability of slave verses wage labour. At a more microeconomic level, academics make decisions regarding how much time to allocate to research verses teaching. An altruistic professor might selectively attend to information that suggests that spending more time on research serves the social interest. Promotion and prestige largely depends upon the quality and quantity of publications, partially because it is easier to place a numerical value on research verses teaching quality. Both the altruism and material self-interest assumptions yield a similar prediction; too much research. However, the former assumption gives the additional possibility of an incorrect social consensus regarding the appropriate mix of teaching and research. Another example is an altruistic student who believes that others cheat and therefore that he will do relatively poorly unless he cheats too. His cheating may be justified by arguing that failing on a relative basis will not allow him to contribute optimally to society. Of course cheating raises the bar, as does more research effort on the part of academics, potentially reinforcing the asocial behaviour. The problem with achieving a socially efficient allocation in these two examples is intense competition at the level of the individual with little if any between group competition. To contrast, in the case of firms, high levels of between-firm competition may select for customs that promote within-firm cooperation.

Sometimes it is claimed that the morality of behaviour can be judged by intent not consequences. However, this may be a is dangerous rule of thumb because altruists may selectively attend to information deluding themselves into believing that their own private materially preferred activity choices also serve the social good. As a check to this problem, individuals look to social consensus views regarding appropriate moral behaviour. Unfortunately, as argued in this essay, such social consensus views may be incorrect particularly in situations in which there are high levels of interpersonal competition. The model drew attention to the importance of policies that promote truthful information search such as education in critical thinking. The essay suggests that an important component to teaching critical thinking is to educate people about the nature of cognitive dissonance and associated selective attention. When making decisions one can attempt to safeguard against selective attention by asking the following question: Does the social validation that I provide accord with my own private material incentives? If so, then the decision needs to be evaluated further, for example by examining the activity choices and moral validation provided by those holding different private material incentives.

Bibliography

Akerlof, G.A., & Dickens, W. (1982). The economic consequences of cognitive dissonance. *American Economic Review*, *72*, 307-319.

Akerlof, G.A., & Kranton R.E. (2000). Economics and identity. *The Quarterly Journal of Economics*, CVX (3), 715- 752.

Allen, D.W. (1995). Order in the church: A property rights approach. *Journal of Economic Behaviour and Organization*, *27*, 97-117.

Allen, S., Adam, B., & Carter, C. (1999). The media politics of environmental risk. In S. Allen, B. Adam & C. Carter (Ed.), *Environmental risks and the media* (pp. 1-26). London: Routledge.

Anderson, A. (1997). *Media, culture and the environment*. London: University College London Press.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm glow giving. *The Economic Journal*, 100 (401), 464-477.

Andreoni, J. (1995). Cooperation in public goods experiments: kindness or confusion. *American Economic Review*, 85 (4), 891-904.

Avilés, L. (2002). Solving the freeloaders paradox: genetic associations and frequency-dependent selection in the evolution of cooperation among non-relatives. *Proceedings of the National Academy of Science*, 99 (22), 14268-14273.

Axelrod, R., & Hamilton, W.D. (1981). The evolution of cooperation. *Science*, 211 (4489), 1390-1396.

Badcock, C. (2000). *Evolutionary psychology : A critical introduction*. Malden, MA: Polity Press.

Bailey, R. (2002). *Global warming and other eco-myths: How the environmental movement uses false science to scare us to death*. Roseville, CA: Prima Publishing.

Baron, R.A., & Byrne, D. (1994). *Social psychology: Understanding human interaction*, 7th ed. Boston: Allyn & Bacon.

Beder, S. (2002). *Global spin: The corporate assault on environmentalism*, 2nd ed. White River Junction, VT: Chelsea Green Publishing Company.

Bergstrom, T.C. (2002). Evolution of social behaviour: Individual and group selection. *Journal of Economic Perspectives*, 16 (2), 67-88.

Boehm, C. (1993). Egalitarian behaviour and reverse dominance hierarchy. *Current Anthropology*, 34 (3), 227-254.

Boehm, C. (1996). Emergency decisions, cultural-selection mechanics, and group selection. *Current Anthropology*, 37 (5), 763-792.

Boehm, C. (1997). Impact of the human egalitarian syndrome on Darwinian selection mechanics. *The American Naturalist, 150*, S100-S121.

Boyd, R., & Richardson, P.J. (1990). Group selection among alternative evolutionary stable strategies. *Journal of Theoretical Biology, 145*, 331-342.

Bradley, B.J. (1999). Levels of selection, altruism, and primate behaviour. *The Quarterly Review of Biology* 74 (2).

Brady, G.L. (1993). Global warming: cognitive dissonance, incipient regimes, and rent-seeking. *Journal of Public Finance and Public Choice*, 11 (1), 41-52.

Caldwell, B. (2000). The emergence of Hayek's ideas on cultural evolution. *The Review of Austrian Economics*, 13 (1), 5-22.

Canals, J., & Vega-Redondo, F. (1998). Multi-level evolution in population games. *International Journal of Game Theory, 27*, 21-35.

Chapman, G., Kumar, K., Fraser, C., & Gaber, I. (1997). *Environmentalism and the mass media: The north south divide*. London: Routledge.

Chomski, N. (1989). *Necessary illusions: Thought control in democratic societies.* Toronto: House of Anansi Press Ltd.

Clegg, N.W. & Reed, C.G. (1994). The economic decline of the church of medieval England. *Explorations in Economic History*, 31 (2), 261-280.

Courbet, J.B. (2002). A faint green sell: Advertising and the natural world. In M. Meister & P. Japp (Ed.), *Enviropop: Studies in environmental rhetoric and popular culture* (pp. 141-160). Westport, Connecticut: Praeger Publishers.

Crouzet, F. (1991). The Huguenots and the English financial revolution. In P. Higonnet, D.S. Landes & Rosovsky H. (Ed.), *Favorites of fortune: technology, growth, and economic development since the Industrial Revolution* (pp. 221-266). Cambridge, MA: Harvard University Press.

Dale, S. (1996). *McLuhan's children: The Greenpeace message and the media.* Toronto: Between the Lines.

Daly, H.E. (1999). *Ecological economics and the ecology of economics: Essays in criticism.* Cheltenham, UK: Edward Elgar Ltd.

Darwin, C. (1859). *The origin of species by means of natural selection or the preservation of favoured races in the struggle for life.* London: John Murray.

Darwin, C. (1871). *The decent of man and selection in relation to sex.* New York: Random House.

Dawkins, R. (1976). *The selfish gene.* Oxford: Oxford University Press.

Dawkins, R. (1982). *The extended phenotype.* Oxford: Oxford University Press.

Dawkins, R. (1989). *The selfish gene,* 2nd revised ed. Oxford: Oxford University Press.

Dawkins, R. (1994). Commentary on Wilson and Sober: Group selection. *Behavioral and Brain Sciences, 17,* 616-617.

Dempsey, B.W. (1935). Just price in a functional economy. *American Economic Review,* XXV, 471-486.

De Roover, R. (1958). The concept of just price: Theory and economic policy. *Journal of Economic History,* 418-434.

Earle, T., & Johnson, A.W. (2000). *The evolution of human societies: From foraging group to agrarian state* (2nd ed.). Stanford, CA: Stanford University Press.

Eaton, B.C., & Morrison, W.G. (2003). Evolutionary stability in the investor_opportunism_retaliation game. *Journal of Bioeconomics, 5,* 27-45.

Eshel, E. (1972). On the neighbor effect and the evolution of altruistic traits. *Theoretical Population Biology, 3,* 258-277.

Fehr, E., & Gachtër, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review,* 90 (4), 980-984.

Fehr, E., & Gachtër, S. (2002). Altruistic punishment in humans. *Nature, 415,* 137-140.

Fehr, E., & Schmidt, M. (1999). A theory of fairness, competition and cooperation. *Quarterly Journal of Economics,* 114, 817-868.

Fisher, R.A. (1930/ 1958). *The genetical theory of natural selection.* New York: Dover Publications Inc.

Forsdyke, D.R. (2001). *The origin of species revisited.* Kingston & Montreal, Canada: McGill-Queens University Press.

Frank, R.H. (1994). *Microeconomics and behaviour,* 2nd ed. Boston: McGraw Hill, Inc.

Gilad, B., Kaish, S., & Loeb, P.D. (1987). Cognitive dissonance and utility maximization: A general framework. *Journal of Economic Behaviour and Organization, 8*, 61-73.

Gintis, H. (2000). Strong reciprocity and human sociality. *Journal of Theoretical Biology, 206*, 169-179.

Gintis, H., & Bowles, S. (2003). The evolution of strong reciprocity: cooperation in heterogeneous populations. *Theoretical Population Biology*, 65, 17-28.

Gintis, H., & Bowles, S. (2004). Prosocial emotions. Santa Fe Institute Working Paper, #02-07-028.

Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behaviour in humans. *Evolution & Human Behavior, 24*, 153-172.

Goodnight, C.J., Schwartz, J.M., & Stevens, L. (1992). Contextual analysis of models of group selection, soft selection, hard selection, and the evolution of altruism. *The American Naturalist, 140*, 743-761.

Gordon, H.S. (1954). The economic theory of a common property resource. *The Journal of Political Economy, 62*, 124-142.

Gould, S.J. (1990). *The individual in Darwin's world: The second Edinburgh medal address.* Edinburg: Edinburg University Press.

Gould, S.J. (1998). Gulliver's further travels; The necessity and difficulty of a hierarchical theory of selection. *Philosophical Transactions of the Royal Society of London, 353*, 307-314.

Gould, S.J. (2002). *The structure of evolutionary theory.* Cambridge, MA: Harvard University Press.

Grafen, A. (1984). Natural selection, kin selection and group selection. In J.R. Krebs & N.B. Davies (Ed.), *Behavioral ecology: An evolutionary approach* (pp. 62-84). Oxford: Blackwell Scientific Publications.

Green, V. (1998). *The European Reformation.* Stroud, UK: Sutton Publishing.

Griffiths, P.E., & Gray, R.D. (1994). Developmental systems and evolutionary explanation. *Journal of Philosophy*, XCI (6), 277-304.

Haldane, J.B.S. (1932/ 1966). *The causes of evolution.* Ithaca, NY: Cornell University Press.

Hamilton, W.D. (1963). The evolution of altruistic behaviour. *The American Naturalist,* XCVII (896), 354-356.

Hamilton, W.D. (1964). Genetical evolution of social behaviour I and II. *Journal of Theoretical Biology, 7,* 1-52.

Hamilton, W.D. (1975). Innate social aptitudes of man: An approach from evolutionary genetics. In R. Fox (Ed.), *Biosocial anthropology* (pp. 133-155). London: Malaby Press.

Hansen, A. (1999). Claims-making and framing in British newspaper coverage of the Brent Spar controversy. In S. Allen, B. Adam, & C. Carter (Ed.), *Environmental risks and the media* (pp. 55-72). London: Routledge.

Hardin, G. (1968). The tragedy of the commons. *Science, 162,* 1243-1248.

Harris, R.J. (1994). *A cognitive psychology of the mass media,* 2nd ed. New Jersey: Lawrence Erlbaum, Associates, Publishers.

Hirsheleifer, J. (1987). On emotions as guarantors of threats and promises. In J. Dupré (Ed.), *The latest on the best: Essays on the evolution of optimality* (pp. 307-326). Cambridge, MA: MIT Press.

Hirsheleifer, J. (1999). There are many evolutionary pathways to cooperation. *Journal of Bioeconomics, 1,* 73-93.

Hirshman, A.O. (1965). Obstacles to development: A classification and a quasi-vanishing act. *Economic Development and Cultural Change,* 13 (4), 383-393.

Hosseini, H. (1997). Cognitive dissonance as a means of explaining economics of irrationality and uncertainty. *Journal of Socio-Economics,* 26 (2), 181-189.

Hotelling, H. (1929). Stability in competition. *The Economic Journal, 39*: 41-57.

Hull, D.L. (1980). Individuality and selection. *Annual Review of Ecology and Systematics. 11,* 311-332.

Johnson, P. (1976). *A history of Christianity.* Massachusetts: Halliday Lithograph Corporation.

Jones, M.D.W. (1995). *The Counter Reformation: Religion and society in early modern Europe.* Cambridge: Cambridge University Press.

Klein, B., & Leffler, K.B. (1981). The role of market forces in assuring contractual performance. *The Journal of Political Economy*, 89 (4), 615-641.

Krebs, D.L. (1970). Altruism—An examination of the concept and a review of the literature. *Psychological Bulletin*, 73 (4), 258-302.

Landes, D.S. (1998). *The wealth and poverty of nations.* New York: W.W. Norton and Company.

Lewontin, R.C. (1970). The units of selection. *Annual Review of Ecology and Systematics. 1*, 1-18.

Lipsey, R.G., Bekar, C., & Carlaw, K. (2004). *Economic Transformations: General Purpose Technologies and Sustained Economic Growth*, forthcoming. Retrieved July 28, 2004, from http:// www.sfu.ca~rlipsey/res.html

Mandeville, B. (1714). *The fable of the bees or private vices, publick benefits.* London: J. Roberts.

Maynard Smith, J. (1964). Group selection and kin selection. *Nature, 201*, 1145-1146.

Maynard Smith, J. (1976). Group selection. *Quarterly Review of Biology*, 51 (2), 277-283.

Maynard Smith, J. (1987a). How to model evolution. In J. Dupré (Ed.), *The latest on the best: Essays on the evolution of optimality* (pp. 119-131). Cambridge, MA: MIT Press.

Maynard Smith, J. (1987b). Reply to Sober. In J. Dupré (Ed.), *The latest on the best: Essays on the evolution of optimality* (pp. 147-149). Cambridge, MA: MIT Press.

Maynard Smith, J. (1995). *The major transitions of life.* New York: W.H. Freeman.

Miller, M.M., & Reichert, B.P. (1999). Interest group strategies and journalistic norms: News media framing of environmental issues. In S. Allen, B. Adam & C. Carter (Ed.), *Environmental Risks and the Media* (pp. 45-55). London: Routledge.

Mokyr, J. (1990). *The lever of riches: Technological creativity and economic progress.* New York: Oxford University Press.

Nowak, M.A., & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature, 393*, 573-577.

Nunney, L. (1985). Group selection, altruism, and structured-deme models. *The American Naturalist.* 126 (2), 212-230.

Oskamp, S. (1991) *Attitudes and opinions*, 2nd ed. Englewood Cliffs, New Jersey: Prentice Hall Inc.

Ozment, S.E. (1975). *The Reformation in the cities: The appeal of Protestantism to sixteenth-century Germany and Switzerland.* New Haven & London: Yale University Press.

Ozwald, A., & Zizzo, D. (2001). Are people willing to pay to reduce other people's incomes? *Annales D'Economie et de Statistique*, 0 (63-64), 39-65.

Price, G.R. (1970). Selection and covariance. *Nature*, *277*, 520-521.

Rabin, M. (1994). Cognitive dissonance and social change. *Journal of Economic Behaviour and Organization*, *23*, 177-194.

Rabin, M. (1998). Psychology and economics. *Journal of Economic Literature*, XXXVI (March), 11-46.

Rabin, M., & Schrag, J.L. (1999). First impressions matter: A model of confirmatory bias. *The Quarterly Journal of Economics*, February, 37- 46.

Reed, C.G., & Bekar, C.T. (2003). Religious prohibitions against usury. *Explorations in Economic History*, 40, 347-368.

Rennie, J. (2002). Misleading math about the earth. *Scientific American*, 286 (1), 61- 71.

Ridley, M. (1996). *The origins of virtue: Human instincts and the evolution of cooperation.* New York: Penguin Books.

Romanes, G.J. (1897). *Darwin and after Darwin: Part III.* Chicago & London: Open Court Publishing Company.

Rosenberg, N. and Birdzell (Jr), L.E. (1986). *How the West grew rich: The economic transformation of the industrial world.* New York: Basic Books Inc.

Samuelson, P.A. (1993). Altruism as a problem involving group verses individual selection in economics and biology. *American Economic Review*, 83 (2), 143-148.

Satchell, M. (1996, June 10). Dangerous waters? Why environmental education is under attack in the nation's schools, *U.S. News and World Report*, p. 63.

Saul, R.J. (1995). *The unconscious civilization.* Canada: House of Anansi Press.

Sigmund, K., Fehr E., & Nowak, M.A. (2002). The economics of fair play. *Scientific American*, 286 (1), 83-87.

Simon, H.A. (1993). Altruism and economics. *The American Economic Review*, 83 (2), 156-161.

Smith, J. (2000). *The daily globe: Environmental change, the public and the media.* London: Earthscan Publications Ltd.

Sober, E. (1984). *The nature of selection—Evolutionary thought in philosophical focus.* Cambridge, MA: MIT Press.

Sober, E. (1987a). What is adaptionism? In J. Dupré (Ed.), *The latest on the best: Essays on the evolution of optimality* (pp. 105-118). Cambridge, MA: MIT Press.

Sober, E. (1987b). Comments on Maynard Smith's 'How to model evolution'. In J. Dupré (Ed.), *The latest on the best: Essays on the evolution of optimality* (pp. 133-145). Cambridge, MA: MIT Press.

Sober, E., & Lewontin, R.C. (1982). Artifact, cause and genic selection. *Philosophy of Science*, 49 (2), 157-180.

Soltis, J., Boyd, R., & Richardson, P.J. (1995). Can group functional behaviors evolve by cultural group selection? *Current Anthropology*, 36 (3), 473-494.

Stocking, H. (1999). How journalists deal with uncertainty. In S.M. Friedman, S. Dunwoody & C.L. Rogers (Ed.), *Communicating uncertainty: media coverage of new and controversial science* (pp. 23-41). London: C.L. Laurence Erlbaum Associates.

Suzuki, D. (2002, May 31). Time to pull our heads out of the sand. *Science Matters.* Retrieved July 28, 2004, from http://www.davidsuzuki.org/About_us/Dr_David_Suzuki/Article_Archives/weekly05310201.asp

Tawney, R.H. (1926/1998). *Religion and the rise of capitalism.* New Brunswick, NJ: Transaction Publishers.

Thomson, J.A.F. (1998). *The Western Church in the Middle Ages.* New York: Oxford University Press.

Trivers, R.L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, *46*, 35-57.

Trivers, R.L. (1981). Sociobiology and politics. In E. White (Ed.), *Sociobiology and human politics* (pp. 1-43). Toronto: Lexington.

Trivers, R.L. (1998). As they would do to you: A review of "Unto others: The evolution and psychology of unselfish behavior" By Elliot Sober and David Sloan Wilson.

Skeptic, 6 (4). Retrieved July 28, 2004, from http://www.skeptic.com/mag64.html

United Nations International Panel on Climate Change. (2001). *IPPC Third Assessment Report-Climate Change 2001.* Retrieved July 28, 2004, from http://www.ippc.ch

Wade, M.J. (1978). A critical review of the models of group selection. *Quarterly Review of Biology,* 53 (2), 101-114.

Weber, M. (1904-1905/1958). *The Protestant ethic and the spirit of capitalism.* New York: Charles Scribner's Sons.

Wicks, R.H. (2001). *Understanding audiences: Learning to use the media constructively.* London: Lawrence Urlbaum Associates Publishers.

Williams, G.C. (1966). *Adaptation and natural selection: A critique of some current evolutionary thought.* Princeton, NJ: Princeton University Press.

Williams, G.C. (1992). *Natural selection: Domains, levels and challenges.* New York: Oxford University Press.

Williams, G.C., & Williams, D.C. (1957). Natural selection of individually harmful social adaptations among sibs with special reference to social insects. *Evolution,* 11 (1), 32-39.

Williams, J.L. (2001). *The rise and decline of public interest in global warming: Toward a pragmatic conception of environmental problems.* Huntington, New York: Nova Science Publishers, Inc.

Wilson, D.S. (1975). A theory of group selection. *Proceedings of the National Academy of Sciences.* 72 (1), 143-146.

Wilson, D.S. (1977). Structured demes and the evolution of group advantageous traits. *The American Naturalist.* 977 (111), 157-185.

Wilson, D.S. (1987). Altruism in Mendelian populations derived from sibling groups: The haystack model revisited. *Evolution,* 41 (5), 1059-1070.

Wilson, D.S. (1989). Levels of selection: An alternative to individualism in biology and the human sciences. *Social Networks, 11,* 257-272.

Wilson, D.S. (1990). Weak altruism and strong group selection. *Oikos,* 59 (1), 135-140.

Wilson, D.S., & Dugatin, L.A. (1997). Group selection and assortative interactions. *The American Naturalist,* 149 (2), 336-351.

Wilson, D.S., Pollock, G.B., & Dugatin, L.A. (1992). Can altruism evolve in purely viscous populations? *Evolutionary Ecology, 6*, 331-341.

Wilson, D.S., & Sober, E. (1989). Revising the superorganism. *Journal of Theoretical Biology, 136*, 337-356.

Wilson, D.S., & Sober, E. (1994). Reintroducing group selection to the human behavioral sciences. *Behavioral and Brain Sciences*, 17, 585-654.

Wilson, D.S., & Sober, E. (1998). *Unto others; The evolution and psychology of unselfish behaviour.* Cambridge, MA: Harvard University Press.

Wilson, K.M. (1999). Communicating climate change through the media: Predictions, politics and perceptions of risk. In S. Allen, B. Adam & C. Carter (Ed.), *Environmental risks and the media* (pp. 201-217). London: Routledge.

Wright, S. (1945). Tempo and mode in evolution: A critical review. *Ecology, 26*, 415-419.

Wynne-Edwards, V.C. (1962). *Animal dispersion in relation to social behaviour.* Edinburg: Oliver and Boyd.

Wynne-Edwards, V.C. (1964). A reply to Maynard Smith's rejoinder. *Nature, 201*, 1147.

Zehr, S.C. (1999). Scientists' representations of uncertainty. In S.M. Friedman, S. Dunwoody & C.L. Rogers (Ed.), *Communicating uncertainty: media coverage of new and controversial science* (pp. 3-21). London: C.L. Laurence Erlbaum Associates.

Zywicki, T.J. (2000). "Was Hayek right about group selection after all?" Review essay "Unto others: The evolution and psychology of unselfish behaviour" by Elliot Sober and David Sloan Wilson. *Review of Austrian Economics*, 13 (1) 81-95.