Psychological Sentiments & Economic Behaviour

Psychological Sentiments and Economic Behaviour

PROEFSCHRIFT

Ter verkrijging van de graad van doctor aan de Universiteit van Tilburg, op gezag van de rector magnificus, prof.dr. F.A. van der Duyn Schouten, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op

vrijdag 10 oktober 2003 om 14.15 uur

door

Jeroen van de Ven,

geboren op 10 april 1975 te Huissen.

PROMOTOR: Prof.Dr. Th.C.M.J. van de Klundert

To my parents,

Contents

Acknowledgements

During the past four years or so, I have been working on this thesis. This has been a very enjoyable time for me, if not only because of the topic. How amusing is it to read about the irrationalities of other people! The negative side of studying some of the psychological literature, however, is that you also learn a bit too much about yourself. I realized only too well how well I fit the evidence of irrational behaviour myself. Thus, I tend to attribute failures to other people or to forces beyond my power, taking credit for successes myself. After a purchase, I tend to experience an unpleasant feeling of doubt, after which I try to rationalize my choice. I certainly don't treat sunk costs as sunk. And worst of all, I have postponed writing this preface until the very last moment, not being able to recall all the brilliant ideas I had for it. All this is, to the very least, a bit disturbing: being irrational is one thing, but being aware of it another. Some people may want to stay ignorant about such matters, and I can only encourage them not to continue reading.

However difficult it is to give credit to others, I will have to admit that I have been surrounded with such nice, skilful, and helpful people that any errors can only be blamed on me. Above all, I am much indebted to my supervisor professor Theo van de Klundert. Starting during my undergraduate studies, he has been a continuous source of inspiration for me. I respect his broad views on economics and his never-ending enthusiasm. As a supervisor, he has shaped many of my current views. I will be very proud if somebody recognizes me as being one of his students. Part of my time as a Ph.D. I spent at GREMAQ in Toulouse. It

was a very stimulating environment there. Special thanks go to professor Jean Tirole for supervising me during that time, and for providing me with ideas for a chapter. Furthermore, I would like to express my gratitude to Eric van Damme for encouraging and helping me to go there.

Thanks also to the committee: Professors A.L. Bovenberg, W. Güth, M.J. James, G.J. Mellenbergh, C. van Ewijk, and Doctor J.A. Smulders. I appreciate their time and efforts a lot and I am very proud to have them in my committee.

I was also extremely fortunate to share the office with Richard Nahuis and Luuk van Kempen. Luuk kindly saved my ideas from the paper shredder and we cried together on more than one occasion (having hot Argentenean soup or vlammetjes at the drunken horse). We even managed to keep a fish in our office, though at the sacrifice of our coffee machine. I suspect that our jogging outfit is legendary. Nobody encouraged me more than Richard. Needless to say that I am very happy to work at the same place as him again.

Next, I would like to thank two co-authors. It is a pleasure and honour to work together with Enrico Diecidue and Anton Souvorov. I am impressed by their brightness and kindness. I am confident in an enjoyable and fruitful future cooperation.

There were many other people who helped me out on more than one occasion. Among others, Jan Boone, Jeffrey James, Karim Sadrieh, and Sjak Smulders always had their door open for questions and chats. I am happy that they were so quick in understanding and fixing my problems, otherwise this preface would probably have been postponed even longer. In addition, I would like to thank the people who read and commented on several chapters. Among others, Martin van Tuijl deserves credit for stimulating my interest in science.

Everyday had three focal points. With Swiss preciseness the coffee group gathered for lunch and coffee breaks (for a composition of the group, see e.g. CentER Ph.D. thesis 111). One person in particular deserves special mention. Theo Leers was one of the finest sorts of people who made life beautiful. He was a great young scientist and very funny person. I just wish he would still be around.

During the past four years many colleagues became dear friends. I would like to thank them all, in particular Riccardo.

Writing a thesis cannot be done without the help of secretaries. I would like to thank the secretaries at the department of economics and CentER. Aude

Schloeßing in particular I wish to thank. Without her help, I would have been lost in France for sure.

My family provided continuous support throughout my life, if only indirect. Heel erg bedankt papa, mama, oma, Margreet, Maurice, Suzanne en Thomas.

Finally, Michèle, thank you for being with me. You unite all of the above mentioned qualities.

1
Introduction

People do not always behave as economists expect them to do. Not so long ago, my girlfriend, who holds a Ph.D. in economics, had to decide which health club to go to. One of these clubs had a higher membership fee than the other, but also provided more facilities. In the end she made the decision to go for the expensive one. Surprisingly, it was not the high price-quality ratio that ultimately mattered most, but, she reasoned, the fact that if she paid a high contribution, she would feel committed to actually go.

1.1 Rational economic man

Economics students are taught early in their study about the rational economic man[1]. The economic man at least knows his preference ordering (which satisfies transitivity and completeness), and given this ordering plus some constraints he attempts to attain his most desired bundle of goods. Moreover, economic man is incredibly good at solving optimization programs to calculate the best, say, consumption to savings ratio. In other words, he is usually depicted as selfish as well as smart.

[1] Rationality is a delicate concept. For current purposes, I need not define it in a precise way. Hereafter I drop the term and just speak of economic man and the economic paradigm (see section 1.3.1).

Being smart means, among other things, that decisions are based on weighing marginal benefits against marginal costs (a consequence of maximizing utility). This means that the optimal frequency of visiting a health club is independent on membership fees, since such fees are fixed costs and do not influence marginal costs. If people fail to disregard sunk costs (or do not feel that they should be treated as such), then there is something wrong with modeling those people as economic men. If even economists cannot be modeled as economic men, there is something seriously wrong. In this thesis, individuals are taken to be less selfish and smart.

1.2 Defending rational economic man

The economic man is quite often practical for reasons of tractability. Clearly, however, it is not a very accurate description of most people in real life. Friedman and Savage [1948] have nevertheless defended the use of economic man on the grounds that it does not matter so much whether the assumptions underlying the model are truly accurate, as long as the predictions are (see also Thaler [1980]). Economic man need not literally and consciously make the necessary calculations. In a well known passage, they compare the assumptions behind the economic man who calculates, say, the optimal lifetime savings plan, to a billiard player who has to predict all the movements of the balls and therefore essentially needs to solve a system of equations. They reason:

> "... it seems not at all unreasonable that excellent predictions would be yielded by the hypothesis that the billiard player made his shots *as if* he knew the formulas. (...) It would in no way disprove or contradict the hypothesis, or weaken our confidence in it, if it should turn out that the billiard player had never studied any branch of mathematics and was utterly incapable of making the necessary calculations..." (Friedman and Savage [1948, 298], italics in original).

Of course, modeling a billiard player as if he solves a system of equations yields more accurate predictions when one considers excellent players than for a notorious beginner. Thus, the model would arguably be more fruitful to describe games at championships than at the average elderly home.

A similar argument goes through for using the standard assumptions behind economic man. Modeling an agent as if he can solve an optimization program is likely to yield better predictions for some agents than for others. One is in particular tempted to think that agents who are actively engaged in market transactions can reasonably be modeled this way. The market would punish those agents who do not behave as if they solved the optimization programs properly. These agents would realize losses, a situation that cannot be sustained for a very long time. Or at least they would on average make less profits and be competed away. According to similar logic, the market would leave no room for other than purely self-interested agents. Setting a price that is perceived as fair but is not competitive, induces losses as well, leaving the opportunistic agents in the market.

We would thus be left with the (as if) maximizing and selfish agents as market participants. Since economics is in particular oriented towards studying markets, the "as if" assumption seems innocent in this field. In fact though, neither one of the above claims is necessarily true. Both non-maximizing and boundedly selfish players can survive market forces. Arbitrage opportunities are limited, apparently even in the realm of financial markets (Mullainathan and Thaler [2000]). Learning by agents may over time lead to the competitive equilibrium, but learning itself is often a costly and slow process.

All in all, the defense of modeling agents as if they are maximizing selfish agents seems unwarranted on many occasions. Moreover, to the extent that the market will surpress non-maximizing or unselfish agents, it is still interesting to study what kind of heuristics (like simple rules) agents would use otherwise, or with what kind of sentiments they are equipped. Only then is it possible to judge market efficiency in comparison to other institutions that bolster these sentiments more than the market does (Rabin [2002]).

1.3 Economics and psychology

Assuming less intelligent and selfish agents than usual is not always straightforward or even useful. Sometimes the agents make mistakes but is it not well understood how they reason or why they reason other than is assumed. At other times, the mistakes are just too small to consider. It also happens that agents make mistakes but over time converge to the predicted equilibrium. To illustrate: in an experiment by Nagel [1995], individuals had to state a number in the in-

terval [0, 100]. The one closest to two thirds of the average received a price. It is easy to see that in equilibrium all individuals choose 0. In the experiment, this was not the case. In the first round, the mean and median were around 33. The mean number did converge to the equilibrium thereafter.

Many times, however, agents make 'mistakes' which are predictable, important, and for which there are good explanations. In those cases, it makes sense to model agents as psychological man instead of economic man.

In this dissertation the consequences of various psychological sentiments are scrutinized. In order to place the chapters in a broader framework, it is useful to consider the categorization by Rabin [1998] and Tirole [2002]. They survey the literature that departures from the economic paradigm. I briefly discuss this categorization and some of the interesting contributions in the literature so far.[2]

1.3.1 The economic paradigm

Tirole [2002, 634] summarizes the economic paradigm as follows. The individual is thought of as "maximizing at each instant t over some action set A_t the expectation of the present discounted flow utility of consumption $u_\tau(c_\tau)$ given the information I_t he has accumulated prior to date t":

$$\max_{A_t} E\left[\sum\nolimits_{\tau \geq t} \delta^{\tau - t} u_\tau(c_\tau) \middle| I_\tau\right]. \tag{1.1}$$

Disentangling the maximization program, the following elements can be distinguished: the utility function, beliefs, discounting, and optimization. For each of these elements, violations of the usual assumptions are identified. In the remainder of the section some of the many contributions are highlighted. Some of them are elaborated upon more in later chapters.

1.3.2 Preferences

The starting point of most economic analysis is a concave utility function that is only a function of individual i's own bundle of consumption goods: $u_i = u(c_i)$ with $u'(c_i), u''(c_i) > 0$. There are many indications that this functional form does not capture many subtleties that enter the agent's decision. Both the functional

[2] The most extensive survey is that by Rabin [1998]. The survey by Tirole [2002] includes some of the most recent contributions to the field.

form of how much utility is derived from consumption as well as the idea that only own consumption matters do not reflect the true complexity of real behavior.

$$u''(c_i) \gtreqless 0$$

To begin with, consider individuals' choices isolated from interaction with other agents. The existing evidence from experiments suggests that the utility function contains a reference point at the status quo. For gains, the utility function is indeed concave as is usually assumed, implying risk aversion. For losses, however, the utility function turns convex, implying risk seeking behavior (see Tversky and Kahneman [1992]). Individuals are loss averse in the sense that a small loss compared to the status quo is not outweighed by an equal gain. These properties together, and some more, are elements of what Kahneman and Tversky [1979] dubbed prospect theory, now one of the most well known theories.

$$u_i = u(c_i, c_j, \cdot)$$

It is often suggested that people derive not only pleasure from own consumption, but also from the happiness of their friends and relatives, from the fact that other people behave nicely towards them, from possessing more wealth than their neighbours, etc. In short, they have social preferences: they care about the payoffs of others (Charness and Rabin [2002]).

Consider for example the following series of experiments. In the dictator game, one person has the power to distribute a sum of money over another person and himself. He can do so in any way he wishes. Often, the other person is anonymous. The observation that the player who divides the money (the proposer) usually does not keep all the money to himself points to a notion of fairness. Apparently, agents derive pleasure from being fair and it is not regarded as a fair distribution to keep all the money.

In an extended version of the game, the ultimatum game, the proposal of the player is not immediately implemented. The second player (the responder) gets a chance to either accept or reject the proposal. If he rejects, nobody gets anything. If he accepts, the sum is divided as proposed. After a rejection there is no second opportunity to make a proposal (at least not with this combination of players). It can be expected that the responder will not always accept positive but small

enough offers. Since he may dislike unequal distributions, he may decline offers where he gets too little and the other too much in his view. This is indeed found.

Interestingly, however, we can learn more from this game. As it turns out, it is not only the (inequality of the) distribution that the players care about, but also the intentions of players, something conjectured by Rabin [1993] in a theoretical paper. To see the relevance, consider the following special case of the ultimatum game taken from Falk et al. [1999]. In this game, the proposer can only choose between two possible offers. Either he chooses the offer (8,2) (keeping 8 for himself and leaving 2 for the responder) or the offer (x, y), where x and y are varied among different treatments. In the first variant (x, y) is set to $(8, 2)$. This means that the proposer has no other choice than to propose $(8, 2)$. It turns out that this offer is rejected in about 20 percent of the cases. In a second variant, (x, y) is set to $(5, 5)$ giving the proposer the opportunity to split the sum of money exactly in two. From the proposers who offered $(8, 2)$ in this treatment, no less than 45 percent is rejected. This, despite the fact that given the proposal, the distribution is identical in both treatments. The natural interpretation is that in the latter case, responders were angry because the proposer did have the option of an equal split, but he decided to go for the unequal distribution anyway.

Collecting data from experiments, a preference for an equal distribution and reciprocal behaviour (that is, rewarding nice behaviour and punish stingy players) now seems a robust finding (Fehr and Schmidt [1999], Bolton and Ockenfels [2000], Charness and Rabin [2002]). It is nonetheless also an established phenomenon that agents try to distinguish themselves from others rather than trying to become equal. Worries about status is one of the most recurring patterns in all cultures (see Wright [1994], Van Kempen [2003], and chapter 3). Inequality aversion and status seeking behaviour need not be mutually exclusive if status is rewarded for making the distribution more equal. Often, however, they are. To my best knowledge, conditions for when status seeking behaviour prevails over inequality aversion are unknown, but would be interesting for future research.

Finally, a recurring theme in the recent literature is that people are intrinsically motivated (Deci and Ryan [1985]). Intrinsic motivation refers to an inner state of satisfaction from being involved in an activity. The activity itself is rewarding. This contrasts with the idea that people only work for monetary (extrinsic) rewards or in a controlling environment. This partly explains why people supply voluntary labour or work harder and longer hours than is expected from them.

This concept is firmly established in psychology, but has not received much attention from economists (Frey [1997b]).

Stability of preferences

A final note on the utility function concerns the stability of preferences. A sizeable body of research shows a picture of a remarkably labile nature of preferences (see Slovic [1991]). Illustrative of this are the following well-known examples.

First, there are endowment effects: once goods are part of one's endowment, the valuation immediately increases sharply. This effect is present even if the subjects are made familiar with the object on beforehand, thereby excluding learning arguments as an explanation (Loewenstein and Adler [1995], Thaler [1980]). In the experiment by Loewenstein and Adler [1995] for instance, subjects indicated their preference for a mug, and based on this the predicted mean selling price was \$3.73. A few minutes later, when the participants were actually given mugs, the mean selling price increased to \$4.89.

Secondly, framing effects take place: the choice of agents is sensitive to the way that a choice problem is formulated. For example, the valuation of a gamble is sensitive to whether the outcomes are framed as gains or losses relative to the status quo (Tversky and Kahneman [1992]).

Third, individuals adjust to the state they are in. Such treadmill effects are discussed for instance by Kahneman [2000, 686]:

> "Anyone who has bathed in a cool pool, or in a warm sea, will recognise the basic phenomenon. As one adapts, the experience of the temperature of the water gradually drifts towards "neither hot nor cold," and the experience of other temperatures changes accordingly. A temperature that would be called warm in one context may feel cool in another."

Another instability is caused by the simple fact that individuals have difficulties in remembering how they felt about something. Retrospective evaluations of past utilities are known not always to be reliable (Kahneman [1994]). This is one reason why individuals have imperfect self-knowledge (see Baumeister [1998]). We tend to forget how difficult it actually was to give up an addiction and therefore try

to do so over and over again[3]. Right after the struggle of the attempt to quit smoking, a person may decide never to try again. Yet after a while he tends to forget and tries to give up his habit again.

1.3.3 Beliefs

I now turn to the second element of the economic paradigm: beliefs. The expected utility functional is the standard framework for decision making under uncertainty. Agents are assumed to maximize the sum of utilities of outcomes linearly weighted by the corresponding probabilities that these outcomes occur. Mathematically, individuals are assumed to maximize:

$$\sum_{i=1}^{n} p_i u(x_i), \tag{1.2}$$

where outcome x_i is received with probability p_i for $i = 1, ..., n$, and with a utility function $u(x)$ over outcomes (see e.g. Varian [1992]).

I already pointed out that the utility function itself does not satisfy the usual assumptions of global concavity (see Kahneman and Tversky [1979]). But even if it would, Rabin [2000] shows that within the framework, risk aversion cannot be sensibly explained. He makes this clear by the following observation. Suppose an individual turns down a bet that gives him a 50 percent chance of winning \$110, and another 50 percent of losing \$100. Then, if he is an expected utility maximizer, this same individual should also turn down a bet which gives him a loss of \$1,000 with a 50 percent chance, no matter what the possible gain would be. That even a bet with a 50 percent chance of winning, say, \$10 billion and a 50 percent chance of losing \$1,000 is turned down by anyone can be said to be counterintuitive, except perhaps for the credit constrained people.

The Allais paradox is an early contribution showing the limitations of the expected utility theorem. The paradox lies in the choices of subjects between two sets of lotteries. Denote by $X = (p_1, x_1; p_2, x_2; ...; p_n, x_n)$ a lottery X which gives a prize x_1 with probability p_1, a prize x_2 with probability p_2, etc. Consider first the following two lotteries, taken from Kahneman and Tversky [1979]:

[3] Likewise, people are not perfectly able to predict their feelings. "Most people are very suprised to learn that paraplegics are not always miserable and that lottery winners are not particularly happy" (Kahneman et al. [1997, 396]).

$A = (0.33, 2500; 0.66, 2400; 0.01, 0)$,
$B = (1, 2400)$.

Thus, lottery B is degenerate and gives a prize of 2400 with certainty. Out of these two lotteries, 82 percent prefers lottery B. Consider next the following set of lotteries:

$C = (0.33, 2500; 0.67, 0)$,
$D = (0.34, 2400; 0.66; 0)$.

Out of these two lotteries, 83 percent prefers lottery C. These choices are, however, inconsistent with each other. This is, then, the paradox. To see this, note that preferring B to A implies (normalizing without loss of generality $u(0) = 0$):

$$0.33u(2500) + 0.66u(2400) < u(2400), \tag{1.3}$$

or equivalently:

$$0.33u(2500) < 0.34u(2400). \tag{1.4}$$

However, preferring lottery C to D implies the opposite of equation (1.4):

$$0.33u(2500) > 0.34u(2400). \tag{1.5}$$

To gain in descriptive power, the expected utility model needs to be refined.

Optimism: $\pi \neq p$

Non-expected utility models have refined the expected utility model in the utility domain but also in the domain of beliefs (see Starmer [1998] for a survey). For example, rank-dependent utility models assume that probabilities, p, are transformed by a weighting function, $\pi(p)$. The intuition behind the weighting function is that individuals do not only pay attention to the specific probability that a particular outcome occurs, but also to the probability of an outcome in comparison to other outcomes (Diecidue and Wakker [2001]). An individual will for example take into account the probability of getting a certain outcome or something better.

Calibrating the weighting function it is found that individuals tend to overweigh both high and low probabilities. This can explain phenomena such as optimism and pessimism. It also explains why people are at the same time risk-seeking and risk-averse, in other words, why they participate in lotteries and insure themselves

against unforeseen events (see Diecidue and Wakker [2001]). This is predicted by the theory because the small probability of ending up with the jackpot of the lottery and the small probability that their house burns down are both overweighted.

Overconfidence

Besides being optimistic, individuals are also often overconfident: 90 percent of adults rate themselves as better than average drivers[4], and 25 percent of high school seminars rate themselves in the top 1 percent on the ability of getting along with others (see Baumeister [1998]).

Part of the overconfidence is the result of ignorance of relevant information. This happens even if the costs of obtaining information is insignificant. Blackwell's theorem, on the other hand, says essentially that individuals should never ignore freely available information. The idea is that actions are based on information. More information enables agents to design better strategies (Grant et al. [1998]).

The first thing to note is that Blackwell's theorem holds for expected utility maximizers. However, I have already argued that expected utility theory does not give an accurate description of behaviour. If, on the other hand, someone is not an expected utility maximizer, he should sometimes prefer less information (Grant et al. [2000], Wakker [1988]).

There are intuitive reasons why individuals may prefer to have less information. First, individuals may have an intrinsic value of ignorance due to psychological sentiments such as anxiety, hope, or fear (Grant et al. [1998], Ahlbrecht and Weber [1997]). Not all people would like the idea to know it when they are terminally ill. Secondly, individuals may attempt to self-commit. This is very much linked to dynamic inconsistency (i.e. overweighing the present), something I turn to now. I come back to the relation between overconfidence and self-commitment in section 1.4.

1.3.4 Discounting

The third element in (1.1) concerns discounting. Suppose you get the choice between \$50 now or \$100 in 2 years. Which one do you prefer? A majority of the adults from a sample report that they prefer to have \$50 now. Now consider the

[4] Although this can be in line with the true distribution, the distribution would have to be very skewed. The second example (ability of getting along with others) is not compatible with any distribution.

choice between \$50 in 4 years or \$100 in 6 years. Almost no adult prefers the \$50 in 4 years. Are these choices consistent with the assumption that agents discount the future exponentially?

Exponential discount factors are most frequently used in economics. Thus, as in equation (1.1), in period t, the flow utility at time τ is discounted by the factor $\delta^{\tau-t}$. This has the following prediction: suppose an individual has the choice between consumption level c to be received n periods from now, and c' to be received $n+\Delta$ periods from now. Then, if he prefers c to $c' = c + x$ for a given n, then he should prefer c to c' for any n. Since c gives present value utility $\delta^n u(c)$ and c' gives present value utility $\delta^{n+\Delta} u(c')$, he chooses c if and only if $\delta^n u(c) \geq \delta^{n+\Delta} u(c')$ or, equivalently, $u(c) \geq \delta^{\Delta} u(c')$. Hence, what matters is the time lag between dates, but not how far in the future they are. If individuals make their choices in a way that satisfies this property, they are said to be dynamically consistent.

This is not how people (or some animals) choose (Ainslie [1991]). In the above experiment the time lag between receiving \$50 or \$100 is 2 years in both cases. Thus, if people prefer \$100 in 6 years to \$50 in 4 years, their behaviour can only be consistent with exponential discounting if they also prefer \$100 in 2 years to \$50 immediately. But the experimental data shows otherwise.

The choices of above are inconsistent with exponential discounting. The choices of the individuals can be described by a slightly more complicated discount function, called a hyperbolic discount function (see e.g. Laibson [1997]). This is of the form:

$$u_t(c_t) + \beta \sum\nolimits_{\tau \geq t+1} \delta^{\tau-t} u_\tau(c_\tau), \tag{1.6}$$

with $\beta \leq 1$. Hence, the future is discounted by an additional term β relative to the present. This distorts choices that involve the present and leads to interesting predictions. For $\beta = 1$ the hyperbolic function reduces to the exponential case. For $\beta < 1$ discounting is present-biased.

To illustrate how this functional form describes choices: preferring \$50 now to \$100 in 2 years implies:

$$u(50) > \beta \delta^2 u(100). \tag{1.7}$$

Preferring \$100 in 6 years to \$50 in 4 years implies:

$$u(50) < \delta^2 u(100). \tag{1.8}$$

In the second inequality, β cancels out because both dates are in the future. The term β does not cancel out in the first inequality because one of the possibilities is paid out immediately. Obviously, equations (1.7) and (1.8) can hold simultaneously for a small enough value of β.

Consider now an individual who needs to save so that in five years he can buy the overly expensive car he wants so desperately. He is predetermined to set part of his monthly paycheck away on another account. Sadly enough, he is a hyperbolic discounter. Every month when his paycheck is added on his regular account, he has to make the trade-off of between saving the required bit or consume it rightaway. Overweighing the present, he is too tempted to consume. After five years, the savings account is still empty.

This individual would be much benefitted by having to his disposal a commitment device to save. Fortunately, there are opportunities for him. Illiquid assets provide a form of commitment (Laibson [1997]). Investments in illiquid assets that are subject to a penalty for early withdrawal turn the cost-benefit ratio in favour of saving. Of course, plenty other forms of commitment can be explained within the same framework, such as self-imposed deadlines, fixing appointments well in advance to prevent endless postponements, putting the cookies box on the highest shelf, and moving the (very annoying) alarm clock far away from the bed. Heroic figures would even tie themselves to the mast in order to self-commit.

1.3.5 Maximization

The last element in program (1.1) is that agents maximize their utility (given constraints and information etc.). Do agents really maximize? Some studies show that they do not.

The existence of such non-maximizing agents is nicely illustrated in a study by Camerer et al. [1997] of the taxi cab drivers of New York City. Rather than maximizing revenues per hour by working longer hours on a rainy day with many clients, and shorter hours on less profitable days, they tend to stop working after reaching a certain target level of earnings. This behaviour is opposite to that predicted by maximization.

Target earnings and other aspiration levels are found in many other studies as well (see Diecidue and Van de Ven [2003] for a survey). Managers often do not choose a project that maximizes expected profits but they rather aim at reaching a target level (see Payne et al. [1980], [1981]). Similarly, portfolios are

often constructed on the basis of the likelihood that a certain target return will be met. Some farmers are known to grow safe crops until their subsistence level is guaranteed, and to grow more risky crops beyond that level (see Lopes [1984]).

In an experimental study by Loomes [1998], it is also shown that individuals do not maximize payoffs. Individuals had to divide 20 green and 20 white balls over two bags: A and B. Every bag should contain twenty balls but the individual could freely decide on the shares of green and white balls in each bag. Now, the individuals knows that a lottery will take place that selects bag A with probability 0.65 and B with remaining probability 0.35. From this bag, a ball will be randomly taken and if the ball is green a sum of £20 will be paid to the decision maker. If a white ball is picked, the decision maker earns nothing. Virtually all models of decision making predict that the decision maker puts all green balls in bag A, thereby maximizing the probability of payoff (i.e. 0.65)(see Loomes [1998]). As it turned out, most people choose to put 13 green balls in bag A and 7 in bag B . This matches the probability ratio that each bag will be chosen. Apparently, the individuals used simple heuristic rules (divide the green balls in proportion to the probability that the bag will be chosen) instead of maximization behaviour, even though this reduces the probability of gaining £20 with more than ten percent (from 65 percent to 54.5 percent).

1.4 Applications

Most of the foregoing results were descriptive violations of standard economic assumptions. However, one would also like to know why we find these violations. For example, why are individuals overconfident, fair to others, work less with a higher bonus, or do they ignore information? Exciting insights in these aspects result from combining some of the above elements.

One of the most powerful results in the recent literature is the assumption of imperfect self-knowledge. Bénabou and Tirole have used this assumption in several domains. In this section I highlight some of their ideas.

Carrillo and Mariotti [2000] and Bénabou and Tirole [2002b] explain ignorance of information by dynamic inconsistency and imperfect self-knowledge. Consider an agent who has to decide the next period (period 2) whether to undertake a task or not. Undertaking the task is costly. The benefits of undertaking are received in period 3 and depend on the ability of the agent. It is assumed that the agent has

imperfect self-knowledge and does not know exactly what his ability with respect to this task is. The individual can, however, learn about his ability in period 1 at no costs. All the information he collects in period 1 is known in the future, but he can only acquire the information in period 1 (for example by doing a related test in this period). Finally, the individual is dynamically inconsistent: his discounting function is hyperbolic (see (1.6)).

Suppose in period 1 the agent expects a net gain of undertaking the task and hence that he is willing to undertake it. Normally, he could gain by acquiring information. It may turn out that he comes to know that he is of low ability in which case he better refrains from undertaking the task. However, he can also lose from more information. Consider behaviour in period 2. In this period, the agent has to decide to undertake or not, at some costs. But with hyperbolic discounting, he puts extra weight on the current period. He inflates the importance of the costs and may no longer be willing to undertake the activity. There is therefore a potential dilemma: in period 1 the agents aims at undertaking, but in period 2 he may reconsider. This creates a 'time-inconsistency region'.

Will he acquire information in period 1? Not if there is a high probability that the information reveals that he is in the 'time-inconsistency region' where he prefers to undertake from the current viewpoint but will reconsider in the next period. On the other hand, he will not ignore information if the probability is high that he is of low ability. In that case ignoring information is too costly because he would undertake the task even though he should not.

Dynamic inconsistency combined with imperfect self-knowledge creates opportunities for ignorance of information for strategic reasons. Ignoring information keeps self-confidence high, and high-self-confidence makes persistence more likely. It also explains phenomena like self-handicapping such as drinking before an exam. Drinking alcoholic beverages has the same function as ignoring information. It makes the outcome of the test unreliable and it therefore reveals less information. Finally, it learns us something about why individuals tend to memorize achievements in a selective way: successes tend to be recalled whilst failures tend to forgotten. Similarly, successes tend to be attributed to one's self whereas failures are likely to be attributed to others (Baumeister [1998]). These self-serving biases in memory can help keeping self-confidence high, and minimize the temptation to give up along the way.

In another context, Bénabou and Tirole [2001] combine dynamic inconsistency and imperfect self-knowledge to explain compulsive behaviour. Suppose you have incomplete knowledge about your ability to cooperate, in the sense that you do not know how present-biased your discounting is (you are uncertain about β in (1.6), a measure of willpower). Sometimes you are involved in short term relationships, for example in a restaurant where the waiter serves you well in the expectation of a good tip. At other times you end up in long term relationships. In relationships that are likely to be short-term of nature, there is a big temptation to break up the relationship. You are better off leaving the restaurant, that you will probably never visit again, without leaving a tip. Long term relationships always pay off if you sustain them long enough. If you have no strong bias to the present (high β), you will succeed in sustaining the relationship. With a discount rate strongly biased to the present (low β), you are tempted to give up the long term relationship making you overall worse off.

In long term relationships you are not exactly sure about your ability to cooperate. If you knew you had a discount rate strongly biased to the present, you also knew that a potential long term relationship will not last. There would be no reason to get involved in a long term relationship. If, on the other hand, you knew that you do not have a strong tendency to overweigh the present, you would be able to sustain long term relationships with a high payoff.

A forward looking agent may reason as follows. If I manage to cooperate even in short term relationships, I will recall later that I must have no strong bias to the present. So I also must be able to sustain a long term relationship. Even though this person has no direct benefits from cooperating in short term relationships, he shows this compulsive behaviour to avoid that he will later be afraid to get involved in long term relationships. This may explain tipping behaviour.

As a last application I discuss crowding-out of intrinsic motivation. As explained earlier (section 1.3.2), many people are intrinsically motivated to undertake an activity. That is, monetary rewards are not always necessary to induce people to work. It may still be the case that people are not enough intrinsically motivated. Economists generally solve this problem by implementing an incentive system based on monetary rewards, to increase motivation. Surprisingly, many experiments from psychology show that these extrinsic rewards often undermine motivation. This result extends to the workplace. It is therefore suggested that monetary rewards crowd-out intrinsic motivation, for a variety of reasons. For

example, rewards imply competition between workers which discourages some of them (see in particular Kohn [1993] for a survey, and Frey [1997b] for an early treatment in economics).

Intrinsic motivation combined with, again, the imperfect self-knowledge framework can also shed light on crowding out of motivation (Bénabou and Tirole [2002a]). The elementary idea is as follows. Suppose that in a principal agent relationship, the agent has imperfect knowledge over his own ability to do a certain task. The only thing the agents gets is a signal, which is correlated with his ability to do this task, but only imperfectly. The principal wants him to do this task, and is aware of the ability of the agent. If he knows the agent has low ability, he reckons that this agent probably has low self-confidence. Thus, in order to motivate the agent, he has to give a high bonus. However, the agent realizes that he gets such a high bonus because the principal knows he is of low ability. The bonus is therefore also a (bad) signal about his ability. Consequently, the high bonus lowers self-confidence even more.

In the equilibrium of the above game, a high bonus lowers self-confidence. The bonus motivates the agent to work in the short run. But once removed, the agent ends up with a lower self-confidence and will be less motivated to work than before (see Bénabou and Tirole [2002a] and chapters 6 and 7 for details).

1.5 Discussion

> "*How strange and confusing are people's conceptions! Sometimes, when you think about it, you don't know whether to laugh or cry. Today it occurred to me that self-sacrificing love is nothing more than an extreme form of egoism.*" Alexander Herzen, *Who is to blame*?

The examples given in section 1.3 show violations of the standard assumptions in economics. A growing literature combines psychology and economics to formulate alternative assumptions that describe the data better. Here are some of my (I am afraid unorganized) views on the field of psychology and economics.

First and foremost, combining insights from psychology and economics clearly leads to interesting results, as section 1.4 and a bunch of other contributions to the literature make clear. Patterns of behaviour can be explained that could otherwise not have as easily and realistically been explained by the standard economic assumptions. Ignorance of information is a good example.

Second, it is worthwhile to note that in many cases most of economic methodology is being maintained. Bénabou and Tirole [2001] for instance, drop the assumptions of exponential discounting and perfect self-knowledge, but retain the idea of maximizing agents playing Bayesian equilibrium strategies. Likewise, altruism can still be modeled using the individualistic approach of economic methodology. Perhaps literally (as suggested by Herzen) but at least in the "as if" sense.

However, some disclaimers are in place. Some of the examples to illustrate violations may be constructed for this purpose or at least constructed to make the violations most visible. It is not always made clear how sensitive the experimental results are to variations in the payoff structure. This makes it difficult to generalize the results.

Moreover, the economic man can still be useful in a (conditionally) normative way. Even if it does not describe people how they actually behave, it is still a valuable framework for analyzing how people should behave in order to achieve certain ends. Thus, I tend to agree with Luce and Raiffa who relatedly discuss the use of game theory:

> "It is crucial that the social scientist recognise that game theory is not descriptive, but rather conditionally normative. It states neither how people do behave, nor how they should behave in an absolute sense, but how they should behave if they wish to achieve certain ends" (cited by Zwick et al. [1999, 7]).

The same can be said about the economic paradigm, which gives directions for how to behave conditional on agreement with the underlying assumptions.

Furthermore, it is not always useful to make more realistic assumptions about individual behaviour. In many cases, though certainly not all, making more realistic assumptions drastically increases the complexity of analysis and this is not always outweighed by an increased accuracy of predictions. Sometimes, it does not matter at all. Fehr and Schmidt [1999] have showed that even though agents behave as inequality averse in some experiments, in other experiments choices are exactly as predicted by standard economic assumptions (i.e. selfish). Smith [1962] showed experimentally that in a double auction market, prices converge to the competitive equilibrium.

Of course, it is still interesting to know when sentiments like inequality aversion affect outcomes and when they do not, so this argument does not imply that

psychological sentiments should immediately be disregarded. What it does mean is that, depending on the context, it is sometimes no sacrifice not to include such sentiments in the model. A bit paradoxically, more work on psychological sentiments is needed in order to know when including these sentiments is not needed.

Relatedly, one has to bear in mind that, taken as an "as if" approach, the economic paradigm is still a rough approximation of reality. In this context, Roth [1995] rightly remarks that "To the extent that utility maximization is viewed as a useful approximation of behaviour, it can't be easily displaced by counterexamples, since approximations always admit counterexamples" (Roth [1995, 78]). Roth continues by arguing that it is nevertheless still valuable to know the conditions under which the approximations break down. Is the Allais paradox (discussed in section 1.3.3) an anomaly and sensitive to the parametrization, or can it be generalized? In this thesis, I have tried to focus on cases where such breakdowns occur and where it seems to me that a rough approximation does not suffice.

It is also worthwhile to note the following. It seems that many psychological phenomena have two sides: an intrinsic value and a strategic role. Donations to charity are made out of love, but also to gain approval. The balance between those sides is a delicate matter. Assuming an intrinsic value for a sentiment often suggests that a shortcut approach is taken, and that the more fundamental motivations are ignored. Or, as Güth [1995, 342] puts it:

> "Very often this [explaining anomalies] is done by including additional arguments of utilities (...) Doubtlessly a lot can be learned from such attempts to explain experimental phenomena, especially when they are based on well accepted motivational forces. Very often, this type of research resembles, however, a neoclassical repairshop in the sense that one first observes behaviour for a certain environment and then defines a suitable optimisation or game model which can account for what has been observed."

However, it is easy to tip the balance too much in favour of strategic reasons. Something should not be too easily dismissed as an intrinsic value. Evolutionary forces can result in intrinsic values such as fairness and other emotions (Frank

[1985]). Here, research from other disciplines, notably psychology and biology, is in particular useful.

In this thesis, and more generally in the field of economics and psychology, there is a relatively intensive interaction with experiments both from economics and psychology. The advantage of these experiments are that situations are well controlled, and that they give much more insight in individual behaviour than aggregated data. Because of its importance, it is necessary to be aware of the shortcomings. Therefore, as a final consideration, I would like to point out some of the limitations of experimental economics.

First, experimental results are sometimes very sensitive to the framing and wording, and hence one should be cautious in generalizing the results. Secondly, it is by no means obvious that results can be directly translated into out-of-laboratory situations. Being fair in the ultimatum game is not the same (and certainly does not imply) that these people are also fair in "comparable" real life situations. Thirdly, subjects often have relatively little time to learn the game and understand the consequences. Experiments quickly become too complicated to be understood within the available time frame[5]. Time constraints pose a natural limit on the complexity of games. Of course, in real life there is also not always enough opportunity to learn, so this argument does not always hold. Fourth, experiments often use a relatively small sample and are not often replicated (Rubinstein [2001]). The latter is due to the fact that replications are unlikely to be published. Thus, although experiments have the advance of creating nicely controlled situations, and provide us with microdata, their shortcomings should be reminded. In this thesis, I have tried to borrow evidence from experiments which results seem robust, and otherwise to mention where more replications and investigations are welcome.

1.6 Overview of the thesis

1.6.1 Main themes

This thesis considers various psychological sentiments that are implemented in economic models. The purpose of this is to enrich economic models to account for

[5] It is therefore no surprise that the results in experiments are sensitive to things as whether or not a payoff table is provided in the instructions (Charness et al. [2001]).

behaviour that is observed in reality but is normally not predicted by standard economic models. The two central themes of the thesis will be to explain why people give and how people react to monetary incentives. The main departures from (or perhaps better: extensions to) standard economic assumptions are the inclusion of social preferences, imperfect self-knowledge, and rationalization of behavior (rather than rational behavior).

Gift-giving

Gift-giving is of interest because at first sight it seems inferior to efficient market trade but gift-giving is nevertheless still widely observed. In chapter 2, I present a survey on possible motivations why people give. I argue that two properties of gift-giving deserve special attention. First, a gift almost never stands on its own but is almost always followed by a countergift. This is called reciprocity. Second, gift-giving seems inadequate in the sense that it rarely maximizes the receiver's surplus, as a cash gift would, according to standard microeconomic theory. Chapter 3 then focuses on one possible motivation behind gift-giving that can explain the two phenomena of chapter 2, namely the demand for social approval.

Chapter 4 puts the analysis in a more broad perspective by contrasting the institution of gift-giving to that of the market. It is argued that the market need not necessarily crowd out gift-giving even though it may be a more efficient institution. Chapter 5 takes an even more positive view on gift-giving by arguing how, when properly designed, the market mechanism may become more efficient if it is complemented by gift-giving. The focus in this chapter is on the welfare aspects of labeling. It is argued that the same motivations as behind gift-giving may account for the willingness of people to pay price premiums for socially desired goods, e.g environmentally friendly goods. This partly solves the information problem.

Rewards

The second central theme is how consumers react to monetary incentives, where attention is paid to effects of rewards that are usually not taken into account. Economists normally assume that more monetary incentives have a positive and monotonic effect on people's behaviour: a reward (bonus, subsidy) for an action motivates people to take that action. The reason is that the focus is normally on

the *direct* impact of rewards on efforts. The direct effect is due to a preference for money, or more generally, consumption goods, by agents. Then if, for example, a reward is conditioned on performance, an agent is more willing to make efforts to obtain the reward. However, there are also *indirect* effects of rewards. For example, rewards interact with other motivations (a desire for approval, say), or it signals information (such as the perceived ability of the agent).

Including the indirect effects of rewards on behaviour has interesting consequences. For instance, there is evidence that the positive relationship between monetary rewards and behaviour does not always hold as such. Under some circumstances, rewarding behaviour leads to decreased motivation (see e.g. Deci and Ryan [1985], Kohn [1993]). By examining indirect effects of rewards, better understanding of the relation between rewards and motivation is gained.

Chapter 3 argues that if people care about social approval, it may well be that subsidizing gift-giving may reduce gift-giving rather than enhance it. A more positive result is obtained in chapter 6. Here, it is explained why principals may give a bonus that is not specified in a contract. Normally, in a relationship with a finite number of periods and no contract, there will be no bonus in equilibrium. However, things change if it is assumed that the agent is not perfectly informed about his ability. In equilibrium, a reward may signal high ability, which increases self-confidence and motivation. Finally, chapter 7 examines the effects of subsidies when people try to rationalize their behaviour, and it is found that higher subsidies have a less profound long-run effect than smaller subsidies.

1.6.2 Detailed overview

The chapters are roughly organized according to the two themes giving and rewards, although some of the chapters combine these two themes. Chapters 2 to 5 focus on gift-giving, and chapters 3, 6, and 7 focus on the effects of rewards. Of course, this distinction is a bit artificial, since in a loose sense rewards are also gifts.

I now give a somewhat more detailed description of the chapters' contents as a reading guide. All chapters can be read independently.

Chapter 2 is a first examination of why people give. Many motivations are possible: altruism, exchange, fairness, signalling, and social approval. A more detailed look at the properties of gifts reveal that not all motivations are equally likely as explanations.

The most elementary properties are reciprocity and inadequacy. Reciprocity means that nice behaviour is rewarded by nice behaviour (positive reciprocity), just as stingy behaviour is being punished (negative reciprocity). It is indeed observed that gifts are not one-way transfers as is often thought. Most of the time, gifts are reciprocated by return gifts. By inadequacy I mean that generally gifts do not maximize the receiver's utility. According to standard microeconomics arguments, cash-gifts are preferred to gifts in kind by receivers, but in reality cash-gifts are relatively rare.

All motivations are scrutinized keeping these elementary properties in mind. Altruism is likely to play a role for gift-giving, but not in those instances where gifts are inadequate. Gifts as exchange can only be sustained for a sufficiently long time horizon. Fairness explains charity to some extent. Fairness does not easily explain all the experimental data. Social approval can explain some of this, and also inadequacy and reciprocity (see chapter 3). Signalling explains gifts for situations with information asymmetries. If a person is not sure about the trustworthiness of other players, the other player can signal to him that he is indeed trustworthy by making a gift. It is also possible that a gift signals to the giver himself that he is trustworthy and that he will be able to sustain long-term relationships.

The chapter ends with a discussion where I argue for a hybrid explanation. For example, people give not so much because they are fair, but because they like to appear as fair and receive approval for being fair. Furthermore, I argue that in order to design efficient institutions, it is important to know what the motivations behind gift-giving are.

Chapter 3 examines social approval as a motivation to give in more detail. First I show that many individuals care about social approval and status. Then I argue that these two elements together explain the basic elements of gift-giving: reciprocity and inadequacy. The reasoning is as follows. For a gift, social approval is awarded. For the receiver, this means a loss in status. This gives incentives to give back (reciprocity). The first giver may want to prevent the receiver from giving back to keep his status high, and therefore gives in kind, making it more expensive for the receiver to return a gift. I also argue that part of charity can be explained by a demand for social approval. This explains why many individuals donate more when their donations become public knowledge, and also why

donations tend to be more densely distributed near the boundaries of published categories.

There is an interesting relationship with the other theme of this thesis: crowding-out of motivation. Rewarding gifts reduces the sacrifice. This is likely to reduce approval for a gift. Hence, if people are motivated to give because social approval is rewarded, monetary rewards may demotivate to give because less approval will be received.

Then, in *chapter 4* a macroeconomic perspective is taken on gift-giving. If one assumes that gift-giving is a result of the desire to exchange goods, then it seems plausible that the market mechanism will take over all gift-exchanges in the end. As the market grows in size, it becomes more efficient and gift-exchange becomes a poorer alternative.

In this chapter it is however argued that gift-giving is not only an exchange mechanism but also adds symbolic utility to an exchange. For example, approval is obtained in a gift-exchange relationship as argued in chapter 3. Symbolic utility is not generated by the market because the latter is an anonymous institution. This makes that gift-giving will not be crowded-out by the market mechanism. It is possible that gift-exchange is sustained even though the market mechanism is more efficient.

In the foregoing chapter it is argued that utility is derived from giving for various reasons, among which social preferences. This perspective is also taken in *chapter 5*. It is assumed that people are willing to pay a price premium for goods that are produced with methods that have less social externalities. Examples are goods produced with environmentally friendly production methods, that avoid child labour, or where fair wages are paid to employees.

Problematic is that consumers cannot distinguish production methods by examination of end products. Hence, they are not willing to pay a price premium for goods that claim to have less social externalities of production. Producers are therefore not willing to invest in more costly production methods.

Two ways are examined to improve upon the imperfect information equilibrium: standards and labels. If the government imposes a standard on production this means that all consumers are forced to buy goods according to this standard. Labels are certificates with information that are voluntary. This allows discrimination between consumers. Consumers with a high willingness to pay consume the labeled goods. Consumers with a low willingness to pay consume the unla-

beled goods. It is shown that labels can lead to a higher welfare compared to standards for an interval of consumer heterogeneity. For sufficiently low or high heterogeneity, standards lead to a higher welfare.

Chapter 6 studies the effects of rewards on self-confidence. This chapter builds on the work by Bénabou and Tirole [2002a] who found that a bonus can signal low ability or a high task difficulty, thereby decreasing self-confidence. Their focus is on rewards that are specified in a contract. By contrast, in this chapter it is assumed that the outcome is only perfectly observable to the principal. This makes a contract impossible. It is then shown that a principal may want to give an unexpected bonus anyway.

This chapter offers an explanation why principals reward unexpectedly. A crucial assumption is that the principal has more information than the agent about the outcome of the task. Thus, the theory is more likely to apply in relationships where the agent is in his learning phase: a child who is learning the piano, or an employee undertaking a task for the first time. Another possibility is that the agent performs a small task which is part of a bigger project. If the principal can judge what the individual contributions from all tasks are, then he can determine whether a specific agent has been successfull or not. The agent himself may not be able to make a good judgement about the value of his specific project because there are too many interactions going on.

An unexpected reward signals a good performance, and raises self-confidence. On its turn, a higher self-confidence increases motivation in the next period. In this way, the chapter offers an explanation why discretionary (that is, not contracted for) rewards are sometimes given. The reward brings good news to the agent, and motivates him to continue.

In the last chapter, *chapter 7,* I consider changes in preferences. While changing behaviour by appropriate incentives has been subject to extensive investigation within the field of economics, changing preferences has been left largely unexplored. Social psychologists, on the other hand, have paid considerable attention to the formation of preferences, trying to demonstrate that they are not stable. Nevertheless, the preference changes are in many cases quite predictable once we take up the idea that people first make a rational choice, and then seek to rationalize their choice afterwards. Hence, in this chapter both incentives and attitude changes are taken into account.

The idea is that people experience an unpleasant feeling (called cognitive dissonance) created by an inconsistency between behaviour and attitudes (you smoke even though you believe it causes lung cancer). To reduce this unpleasant feeling, individuals often try to rationalize their behaviour, for example by focusing on certain arguments congruent with their behaviour or by disregarding information that is incongruent.

As an application, I consider the effects a subsidy will have on the consumption of goods that have social externalities, e.g. environmental friendly goods ("green products"). The main result is that a low subsidy stimulates a positive attitude change towards the subsidized good but a high subsidy does nothing to the attitudes of people. This fits the experimental evidence. I therefore conclude that high subsidies are 'too much of a good thing': they affect current behaviour but fail to affect attitudes and therefore future behaviour.

2

The Economics of the Gift

"One dollar and eighty-seven cents. That was all." So the Christmas story by O'Henry begins. The main character, Della, had, with great pains, been able to save one dollar and eighty-seven cents for a Christmas gift. As this was in her opinion not enough, she went to the shops to sell her possession she was most proud of: her hair. With the money earned, she could just afford a splendid chain for her husband's watch, the only object of value he possessed. When the door opened, her husband stared at her with a peculiar expression, bedazzled from what he saw. He had just bought her a Christmas gift as well: a set of combs worshipped by Della. And he had bought it by selling his watch. Two foolish young people had sacrificed all their treasures for gifts that had no purpose. Were these two young people foolish or, as O'Henry himself thought, did they give the wisest gifts of all gifts given?

2.1 Introduction

Historically, exchange has been – and still is – one of the most fundamental objects of study by economists. It is, for instance, one of the basic ingredients in general equilibrium theory and modern theories of economic growth. Without exchange,

[0]I am indebted to Jeffrey James, Theo van de Klundert, and Sjak Smulders for very useful comments and suggestions.

no specialization is possible. Without specialization, it is hard to imagine how economies would ever grow rich. It is therefore of great importance to understand the functioning of a society. The logical starting point for that is to understand how exchange is organized.

If we were to give a very crude historical account of exchange, one could probably distinguish three phases, seemingly characterized by an ever higher degree of security and efficiency. In the first stage, exchange relied on gift-giving to organize societies. If we look at today's primitive societies, we indeed see a heavy reliance on gift-giving. Since a gift is thought of as a one-way voluntary transference of property, it is not particularly efficient nor is a full exchange secured. A little further in history, one would observe barter trade. Still inefficient, it is secure in the sense that it is a true exchange, not only a one-way transfer. In the last phase we find the most advanced institution to organize exchange, one that is ubiquitous in developed countries: the price system. The price system is particularly efficient in allocating goods by avoiding the need for a double coincidence of wants, something that is not accomplished with barter trade. It also allows for a much greater degree of specialization.

Seen from this perspective, gift exchange should not be of much importance in today's market oriented economies. The extravagance and importance of gift-giving in primitive societies[1] is therefore primarily studied by anthropologists, and not so much by economists. However, viewing gift-giving as a primitive mode of exchange does not do enough justice to this complex institution. For instance, it does not explain why the tribal economies which are oriented towards gift exchange have not been destroyed but sometimes even flourished in the presence of the –supposedly superior– market economy (Gregory [1989]). The 'efflorescence of gift exchange' thesis, by which it is meant that gift exchange has not suffered under the impact of market economies[2], is therefore considered as a valid description of modern exchange economies.

Fortunately, there have been a number of recent contributions by economists which acknowledge the more complex role of gift-giving in modern market-oriented economies, be it somewhat hidden in specific settings. Akerlof [1982] for instance, argues that the amount of time that an employee works in excess of the minimum

[1] Camerer [1988, p. 180].

[2] The term is borrowed from Gregory [1989]. He relates it to the impact of colonization which is broadly interpreted here as the introduction of a market economy.

requirement can be seen as a gift. Rabin [1993] considers gift-giving equilibria as situations where fairness considerations lead to cooperative behavior. More clear-cut examples include birthday, business, and Christmas gifts, voluntary labour, and donations to funds. The amount of these gifts in terms of income is sizeable: money spent on gifts alone by households already accounts for 3-4% of income (Prendergast and Stole [2001]). Charity donations make up another 2% of income in the US (Andreoni [2001]).

In this chapter, I intend to survey the economics literature on gift-giving. Other motivations besides the wish to accomplish a trade are discussed. This is done with respect to two recurring themes in the literature. One of them is the claim by many anthropologists that although gifts appear to be voluntary, they create in fact an obligation to the receiver to reciprocate the gift. The other is the finding by sociologists that it is very often the case that gifts in kind are preferred to cash gifts, something that may be regarded as a bit disturbing from an economics standpoint. I examine to what extent each particular motivation to give can or should account for these themes, and whether it is an efficient institution as compared to the market mechanism.

The setup is as follows. First, in section 2.2 some characteristics of gift-giving are discussed. Different approaches based on motivations are discussed in section 2.3. Each approach is examined on its potential of explaining the characteristics as mentioned in section 2.2. A general discussion and some conclusions are provided in the final section.

2.2 The Gift

> *"To say, here I am. To do something. To give. This is what it means to be a human spirit."* Levinas, *Ethics and Infinity.*

There are probably as many occasions for gift giving as there are relationships. In addition, each relationship is characterized by its own particular demands on how the gift is given. Should the gift be unwrapped at the spot? Is even a quick look into the envelope inappropriate? It is therefore not evident what these different kinds of gifts have in common: a Christmas gift is evidently unlike a business present, and neither do an end-of-the-year bonus and a charity donation resemble each other a lot. There are, in my view, nevertheless some essential characteristics

of gifts. Two of these in particular form the backbone of the discussion throughout the chapter: reciprocity and adequacy. These elements play an important role in most of the gift exchanges.

Reciprocity

At first sight it seems quite natural that a gift is voluntary in nature. Still, anthropologists stress that although voluntary on guise, factually gifts have strong reciprocal properties (Mauss [1925], Codere [1950]). One has not even only a duty to give, but also to receive and to return. The extravagance of gift-giving in primitive societies is underlined by the fact that a failure of accomplishing one's obligations to reciprocate often eventuates in warfare and the loss of dignity.[3] It is therefore often thought that reciprocal behavior is necessarily connected with gift exchange. Mauss called reciprocity one of the "human bedrocks on which society is built" [quoted by Arnsperger [2000, 72]. Or according to Binmore [1998, 24]: "Love and Duty are *not* the cement of modern societies ... the mechanism is *reciprocity*" (his emphases).

According to Camerer [1988] however, it is "especially misleading to assume that modern gift-giving must be reciprocal". It is indeed reasonable not to assume that it is a necessary aspect. Consider for example the case of blood giving. The giving of blood is not directed to specific individuals but to an anonymous agent, as carefully remarked by Arrow [1972]. Gifts or donations of this kind can by assumption not elicit reciprocal gifts, albeit this not immediately signifies that non-reciprocity is also unlikely to occur in personal relationships. But consider the higher effort of workers above minimum firm standards. This is not always reciprocated by the firm in the form of higher wages or bonuses (see Akerlof [1982]). If we take this behavior as a gift of the worker to the firm, then reciprocity is not connected with fairly personal relationships either. The correct conclusion would be that gifts are not necessarily reciprocal in nature. If we are to explain the existence of gift-giving, we also have to explain why certain kinds of gifts are given with a reciprocal intention and why others are not.

Adequacy

[3] This occured for example among the Kwakiutl. It should be noted however that their use of warfare mostly refers to warfare directed to an individual and not so much between nations. For a detailed description of the Kwakiutl, see Codere [1950] and chapter 3.

Consider the following two quotes. According to Camerer [1988, 198]: "A deliberate cash gift is a polite way of saying, we care about you less". And Douglas and Isherwood put it even more to our imagination by writing: "...in our society the line between cash and gift is ... carefully drawn. It is all right to send flowers to your aunt in the hospital, but never right to send the cash they are worth" (Douglas and Isherwood [1978, 58]).

One wonders why it is so bad to make a gift in cash. Standard microeconomics arguments tell us that it can never be worse to get money rather than a specific good. The reason is simple: with a cash-transfer it is in principle always possible to buy the same good as the intended in-kind transfer. Moreover, if existent, a more preferred good may be purchased instead. Whenever an in-kind transfer forces the recipient to consume more of that particular good than he would have done with a cash transfer, the recipient prefers a cash gift (see e.g. Mankiw [1998]). Because gifts in kind weakly lowers the recipient's utility relative to a cash-gift, I call them *inadequate.*

Besides the literary example from the introduction to this chapter, there is ample empirical evidence of inadequate gift-giving. Calculations by Walfdfogel [1993] show that for Christmas gifts, recipients valued the gift by 10-30% less than what the givers had spent on them. An extreme example of inadequacy is found among a tribe in Canada (the Kwakiutl) where some of the gifts are worthless to the receiver (see also the introduction to chapter 3).

For sake of completeness, I should add that there are some notable exceptions to the rule that gifts in kind are inadequate. First, it may be the case that the gift is more expensive for the recipient than for the giver. For example, a souvenir brought from abroad cannot be purchased from an amount of cash that equals its retail price, as the recipient would have to incur transportation costs. Another example is when the recipient has incomplete or imperfect knowledge about his own preferences, such as when he is not aware that this particular good existed so that he could never have bought this good with a cash-gift even though he derives great pleasure from it (see also Camerer [1988]). Finally, a gift can help recipients to stick to self-control. Thaler [1999] considers the example of a couple who cannot afford to spend more on wine than $10 on average. To retain self control, they may decide never to spend more than $20 on a bottle, even though they prefer to have a $30 bottle occasionally. A bottle of $30 as a gift may be greatly appreciated since they are able to enjoy a fine wine without giving up

self-control. Although interesting in its own right, in the rest of this chapter I assume that gifts in kind are inadequate.

The challenge, then, is to find theories of gift-giving that are capable of unifying these dimensions of gift-giving. This is the purpose of the next section.

2.3 Motivations to give

Being familiar with the characteristics of gifts, I next review some approaches in the literature and determine the potential explanatory power of each of them. The aim of this section is to assess the different, sometimes competing, models of gift-giving with regard to the characteristics mentioned. In order to structure the discussion I classify the different models based on their underlying assumptions with regard to motivation. To that end, I distinguish between exchange, altruism, fairness, social approval, and signalling.

2.3.1 Gifts as exchange mechanism

Probably the most obvious approach lending support for gift-giving is to think of gift-giving as accomplishing an exchange between agents. Above all, gifts are found most profoundly in primitive societies. And indeed, it is not unreasonable to assume that at least initially gifts served as a way to separate production from consumption. In this way consumption could be diversified and production could be increased through specialization. The market economy can in this way be interpreted as a more efficient way of exchange, one where gifts are replaced by the use of money. Indeed, Kranton [1996a] argues that this is the case. In her model, agents choose between reciprocal (gift) exchange and market exchange. Since the market is characterized by a thickness externality –more agents on the market reduces search costs– eventually all gift exchange relationships vanish whenever the market size exceeds a threshold level.

While intuitively appealing, the model of Kranton [1996a] cannot account for the coexistence of gift exchange and market exchange.[4] If contemporary markets are so efficient as we think they are, why do people still partly stick to gift ex-

[4] This is not entirely right. The model is able to predict market size for which gift exchange is sustainable. But the model cannot explain how evolution got us in this equilibrium except for some shocks that can be responsible for this. If we start in a gift exchange relationship and some agents find it attractive to enter the market, then the model predicts that *all* agents enter the market.

change? Is it not just cheaper to buy all goods and services at the market? Of course, one reason could be that some products cannot be efficiently produced on the market. Another line of reasoning is provided in chapter 4. There, I argue that gift exchange contains a social interaction element that is valued in itself by the trading agents. Quite often there is a need for mutual sympathy and recognition. These are suppressed entirely in the formal anonymous markets usually studied (Bowles [1998]). But mutual sympathy is rooted in human nature, as is so breathlessly described in Kropotkin [1904]. Thus workers develop sentiment for their co-workers and institution (Akerlof [1982]) and gifts "symbolize and convey meaning" (Camerer [1988, 181]).

In the terminology of Khalil [1997], gift exchange provides symbolic utility on top of substantive utility. A good consumed therefore gives its ordinary substantive utility –in a market exchange as well as in a gift exchange relationship – and on top of that the agent experiences symbolic utility but only if the trade has been accomplished in a gift exchange relationship.

This symbolic utility has to be explained in somewhat more detail. Let the valuation ratio refer to the ratio of utilities that one experiences in a gift exchange relationship and on the market. It is suggested in chapter 4 that the valuation ratio is dependent on the market size in two directions. First the valuation ratio tends to increase as the market gets larger. This is so because mutual sympathy and recognition are lacking in anonymous market exchange relationships, making sympathy more valuable.

However, there is also a tendency for the valuation ratio to decline. This idea builds on the literature on cognitive dissonance in psychology. People have a resistance to change that is lower if more people are supporting a certain view. If agents have to decide whether to stay in their personal gift exchange relationship or to enter the market, then the decision to enter the market gets easier with a larger market size; in essence if more people are supporting the same view.

It is argued that these two opposing tendencies are likely to result in a valuation ratio that is first declining and then increasing in the market size. Under appropriate conditions, this model predicts that the market can become efficient enough to attract part of the population. But as the market gets larger, the valuation ratio becomes larger (lack of sympathy becomes more and more oppressing) and the people who stayed in the gift exchange will decide not to enter the market after all. They stay even though the market has become more efficient due to the

larger size. This can be a stable equilibrium, no agent having the incentive to switch regimes. As a result part of the population is involved in market exchange and part of the population in gift exchange.

The model described is interesting in itself since it argues that the focus of economics should not be a one-sided inquiry into the market as a possible exchange mechanism. In addition, the model can explain a number of things mentioned in the previous section.

First it is able to explain the *seemingly* inadequacy of gifts by taking sympathy into account. For example, it can be that the market provides the same good at lower costs. If people still consume the good within their gift exchange relationship then this points to an inadequate gift. The reason is that part of the utility is neglected; symbolic utility. Substantive utility is higher in the market (more goods at the same costs) but the market provides no symbolic utility. Hence, on net gift exchange is preferable. If in reality we only look at substantive utility, then the gift seems inadequate. If we take into account symbolic utility, there is no matter of inadequacy. Once we take this properly into account we are able to explain the sustainability of gift exchange.[5]

Secondly, gifts have an obligatory element to reciprocate. It is even part of the motivation to reciprocate gifts. If some agent does not return, the relationship ends and both enter the market.[6] As a consequence, an important class of gift-giving, namely charity, cannot be explained by exchange as a motivation to give as this usually takes place anonymously and without a countergift. Moreover, the model is not able to explain why some gifts have no reciprocal character or how we can trust people in short run relationships, something that is resolved in section 2.3.5.

2.3.2 Altruism

Another motivation for gift-giving, perhaps a more natural one in the eyes of people from countries with well developed markets, is to consider the idea that persons have altruistic feelings towards each other. Within the economic method-

[5] There can still be inefficiency in that everybody could be made better off if all people would enter the market or all would stay in their gift exchange. This is due to the existence of externalities that are present in the model.

[6] This is partly due to the assumed tit-for-tat strategy of the players. But it seems that this or any such strategy where the cheater is ultimately punished is reasonable.

ology, this can be modeled as individuals having either preferences for the consumption level or the utility level of other individuals. The structure of a utility function that represents such preferences is given by $U_i = U_i(x_i, U_j(x_j))$, where x_i is the consumption level of person i. If person j has altruistic feelings for person i as well, there is an infinite regress: $U_i = U_i(x_i, U_j(x_j, U_i(x_i, U_j(...))$. The regress easily becomes an unbounded process but Becker [1974] shows an example where it is not. For example, consider the utility function:

$$U_i = x_i^{\alpha} U_j^{\beta}. \tag{2.1}$$

Then the reduced form of the utility function follows straightforwardly by substitution and is given by:

$$U_i = x_i^{\frac{\alpha}{1-\beta^2}} x_j^{\frac{\alpha\beta}{1-\beta^2}}. \tag{2.2}$$

Clearly, this is finite for $\alpha \geq 0$ and $0 \leq \beta < 1$.

Altruistic feelings will take care of a redistribution such that an optimal balance results between personal consumption and consumption of the other. If the endowment of a particular individual is in his view relatively too high, he can gain by giving some of it to the others. Existence of gift giving can therefore be rationalized.

There are several aspects of this model with respect to efficiency that are noteworthy. First, the equilibrium allocation is generally not efficient because neither one of the players acts like a social planner despite their altruistic feelings. To see this, consider player i being altruistic towards j but not vice versa. Note that player i maximizes U_i by setting x_i such that:

$$\frac{dU_i}{dx_i} = \frac{\partial U_i}{\partial x_i} + \frac{\partial U_i}{\partial U_j}\frac{\partial U_j}{\partial x_j}\frac{dx_j}{dx_i} = 0, \tag{2.3}$$

whereas a social planner would set x_i such that:

$$\frac{dU_i}{dx_i} + \frac{dU_j}{dx_i} = \frac{\partial U_i}{\partial x_i} + \frac{\partial U_j}{\partial x_j}\frac{dx_j}{dx_i}\left[1 + \frac{\partial U_i}{\partial U_j}\right] = 0. \tag{2.4}$$

In other words, the social planner counts player j twice: once because player j has his own utility, and one more time because player i derives utility from him. In general, this means that the optimal choice of x_i by player i differs from that by the social planner.

Within a family context, Becker [1974] has shown that an altruistic head of the family internalizes externalities within the other selfish family members by the

appropriate transfers (the 'rotten kid theorem'). However, this efficiency result is somewhat special (Bernheim and Stark [1988]). In particular, altruism can create inefficiencies, such as is the case in the 'Samaritan's dilemma'. This dilemma concerns the problem that if a recipient knows that he will be helped out by an altruist, he has less incentives to, say, self-discipline himself by saving money for the future (Bernheim and Stark [1988] provide a more detailed discussion).

Another, related, efficiency result is obtained by Kranich [1994]. Suppose some players have preferences over the entire allocation of the economy, rather than just one's own consumption level. This can be due to altruism, but also to a preference for fairness. In this case, Kranich proves that the set of Pareto-efficient equilibria is a subset of the set of all gift equilibria. In other words, the equilibrium that results when agents can freely redistribute endowments need not be Pareto-efficient. This can possibly be caused by the public good character that gifts can take. One can think of a case with three players. Would players 1 and 2 both give to a third player everybody may be better off, but if either one of the players gives then it is not profitable anymore for the other to give[7].

Let us now consider the question whether altruism can account for reciprocity. With only one good, the answer is negative.[8] If we consider more goods, however, one can easily see that a recipient may indeed have an incentive to reciprocate. If the endowments are sufficiently different between persons, they may all gain by redistributing, very much like the logic of international trade models. With altruistic feelings, this redistribution may be accomplished without further motivations, since each player gains indirectly by giving part of his or her endowment to another person who would be made better off.

Concerning adequacy, however, we see that altruism as a motivation to give is incongruent with giving in kind. Any cash gift would make the recipient better off[9], without changing one's own consumption, and it therefore necessarily increases one's own utility. Hence we conclude that altruism alone is not a good

[7] Goldman [1978] puts less restrictions on preferences and finds that in that case the reverse also holds: Pareto-efficient equilibria need not be gift-equilibria. It is for example possible that a gift from person 1 to person 2 may decrease the welfare of person 3 (because he cares, say, about the consumption level of person 1), hence moving away from a Pareto-efficient situation.

[8] This is true because in equilibrium it must be the case that if player 1 gives to player 2, player 2 has no incentive to give back. For suppose he has, then his utility from giving some of the good to player 1 must increase. But then player 1's utility should also increase, since the utility of player 2 increases and his own consumption as well. This would contradict player 1 playing an equilibrium strategy.

[9] With the exceptions mentioned earlier.

candidate to explain most of the situations where gift-giving takes place, as it cannot explain the widely observed inadequacy of gifts. Still, there is at least one important situation where altruism cannot be excluded, namely that of charity. Empirical studies are indeed supportive of the view that altruistic motives lie behind charity, although these studies at the same time demonstrate that altruism alone cannot be the unique motivation (see section 2.3.4 for more on this).

2.3.3 Fairness

As a third motivation for gift-giving, the focus in this section is on fairness considerations. This approach recently got most attention in the literature. It successfully accounts for a broad range of experimental games by assuming that people not only care about their own monetary payoffs, but also about the distribution of payoffs and the intentions that other players have.

To see the power of this approach, consider the ultimatum game. In this game, the proposer gets to offer a share of a certain amount to the responder. If the responder accepts, the responder gets the share and the remainder is for the proposer. May he reject, both players end up with nothing.

The game-theoretic prediction of this game is easy to see if only purely selfish players are assumed to participate. Since the responder is always better off accepting any positive amount than rejecting it, the proposer offers the smallest possible positive amount. This prediction is, however, clearly refuted by the data. Most offers by the proposer are nowhere near zero. Typically, they are between 40% and 50% of the amount of money (Fehr and Schmidt [1999]).

Several authors have suggested that the data can be fruitfully explained by assuming that people are not purely selfish but have more 'social preferences' (Charness and Rabin [2002], Bolton and Ockenfels [2000], Fehr and Schmidt [1999]). There are at least two important components to these preferences. Let us first focus on the first component. The first component consists of a preference for equality. This explains why proposers give more than the smallest possible offer in the ultimatum game. They may consider the spread between payoffs too large if they offer the smallest possible amount. By giving some of it to the responder, they reduce the inequality a bit.

The same model can also explain reciprocity. In the so called gift-exchange game, the proposer can give an amount of money from his endowment to the responder in the first round. The amount offered is then, say, doubled. After

that, in the second period, the responder has a chance to make a countergift. In this game, both players can be made better off by a gift and a countergift. The structure with only selfish preferences leads to an equilibrium with no gifts. With inequality aversion, Fehr and Schmidt [1999] show that gift-giving can be explained, something also found in the data.

The strongest evidence that people care about the distribution of outcomes is in my opinion the result of another variant of the ultimatum game. In the dictator game, the responder gets again to propose a share to the responder, but this time the responder has no say in the outcome: every proposal is directly implemented. It is found that some of the amount of money is still directed towards the responder (see Bolton and Katok [1998]). This result is particularly strong as in this game no strategic effects on the part of the proposer should be present: the responder has no power anyway.

The second component of social preferences is the part of intentions. People not only care about outcomes, but also about how these outcomes are realized. This is clear from for instance the game in figure 2.1 taken from Falk et al. [1999].

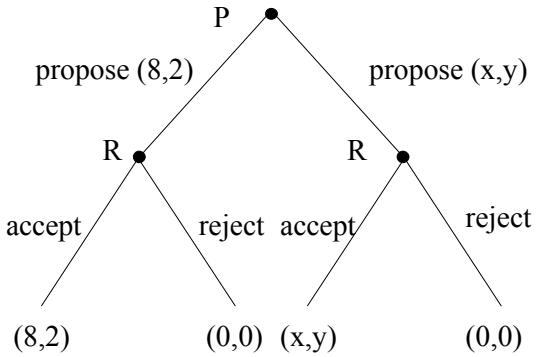

FIGURE 2.1. A variant on the ultimatum game

The game is a variation on the ultimatum game and goes as follows. In the first stage the proposer can propose the offers $(8, 2)$ or (x, y), e.g. $(8, 2)$ means 8 for the proposer, 2 for the responder. The values of (x, y) differ among experiments. In the second stage (after observing the offers of the proposer) the responder can accept, in which case each player gets the proposed offer, or rejects, in which case none of the players gets anything.

In the first variant of the experiment (x, y) was set to $(8, 2)$. Note that the proposer can in this case only make the offer $(8, 2)$. Clearly, if the responder only cares about monetary payoffs he will accept the offer giving him a payoff of 2 rather than rejecting and getting nothing. But in fact, in 20 percent of the cases the offer *is* rejected. A possible explanation can in fact indeed be inequity aversion. Accepting would yield a higher payoff but also an increased inequality between the two players.

This is, however, not the complete story. Another variant was played in which (x, y) was set to $(5, 5)$. The proposer can in this case offer $(8, 2)$ or $(5, 5)$. The situation for responders remained unchanged for those whose proposer sticked to the offer of $(8, 2)$. He still ends up with either 2 when accepting or 0 when rejecting, and the inequality is the same for him as in the first variant. The prediction is therefore that if he accepted (rejected) in the first variant then he should accordingly also accept (reject) the offer $(8, 2)$ in the second variant. But apparently something did change for the responder because the rejection rate increased from 20 to 45 percent when $(8, 2)$ was proposed. The explanation of Falk et al. [1999] is that intentions matter as well. In the first variant proposers had no choice but to offer $(8, 2)$. What matters is that in the second variant they *could* have chosen to propose $(5, 5)$ but they did not. As a consequence, the bad intentions of the proposer were punished by the responders. Therefore, they conclude, both unfair outcomes and unfair intentions matter.

Reciprocity is replicated in many experiments, either in the form of rewarding good behavior (positive reciprocity) or by punishing bad behavior (negative reciprocity). Whether this is due to responders who try to change outcomes or to responders who try to reward or punish intentions, in both cases the possibilities for cooperation are increased. Recall however also the negative result from section 2.3.2 that if people have preferences over the distribution of outcomes, a Pareto-efficient situation need not be reached.

Less obvious is how inadequacy can be explained by fairness models. If gifts are a means to reduce income inequalities, then the most efficient way to do so is to give adequate gifts (unless, of course, players only have the possibility to give inadequate gifts or nothing). Also, if the intention is to reward good behavior, it does not make sense to give inadequately since this makes the reward more costly.

The in my view most problematic part of social preferences as an explanation for giving behavior is due to a variant of the ultimatum game introduced by Güth and Van Damme [1998]. In their variant, a third player is added. In the first stage the proposer proposes an allocation for all three players. In the second stage, the responder accepts or rejects. The third player remains inactive. One of the main results is that only marginal offers were proposed to the dummy player. Although Bolton and Ockenfels [1998] have shown that this is not necessarily incompatible with their inequality aversion model, it is my personal belief that this experiment shows that the proposer was not intrinsically motivated by fairness considerations. I come back to this point in the general discussion at the end.

2.3.4 Warm glow & social approval

Searching for motivations for charity-giving, Andreoni [1989] assumes that people have a taste for giving. He reasons that, if people have purely altruistic motives for donating to a public good, they should only care about the total supply of it. His "egoists" and "impure altruists", on the other hand, do not only care about the supply of the public good that they donate for, but also experience a "warm glow" from having "done their bit." (Andreoni [1989, 1448]). Thus in his model contributions to a public good are made not only for the benefits of public good supply but also for experiencing the (egoistic) warm glow feeling.

The distinction between pure and impure altruists is a useful one and an attempt is made to test the hypothesis whether or not people donate to public goods out of purely altruistic motives. Interestingly, it is found that generally people are impure altruists: part of their motivation is attributed to the warm glow feeling (Andreoni [1990], [1993], Bolton and Katok [1998]).

Here I go one step further in trying to explain *why* people get this warm feeling from giving. The basic hypothesis, for which there is ample evidence, is that people are searching for social approval. The following passage that is taken from the *Fable of the Bees* illustrates this neatly:

> "That a man with small skill in physick and hardly any learning, should by vile arts get into practice, and lay up great wealth is no mighty wonder, but that he should so deeply work himself into the good opinion of the world as to gain the general esteem of a nation, and establish a reputation beyond all his contemporaries, with no other qualities but a perfect knowledge of mankind, and a capacity of making the most out of it, is something extraordinary." (B. Mandeville [1714, 262].

This passage concerns the, at that time illustery, Dr. Radcliffe, who gained the general esteem of a nation by donating his wealth to Oxford University. As he was aware of, making gifts is a way to be approved by others. The warm glow is therefore due to the social approval received and not so much for the act of giving itself. That is, the gift is a means to get a warm glow and has no intrinsic value in its own in this respect.

Since social approval as a motivation to give is the subject of chapter 3, only the main ideas are briefly discussed here. The important premises are that people get approval for a gift and they care about this, and that there is a status element in the approval domain in the sense that people want to get more approval than others. It is thus argued that people not only want to be admired, but also want to be more admired than others (see Holländer [1990]).

With these building blocks, chapter 3 tries to account for reciprocity and adequacy. This is done in a sequential gift-giving game with two players. In the first period, player 1 has the opportunity to transfer some of his endowment to player 2. Player 2 gets a chance to make a transfer to player 1 in the second period. Both players have a utility function of the form (for current purposes slightly simplified):

$$u_i = u_x(x_i) + u_s(s_i - s_j), \tag{2.5}$$

where x_i is the consumption level of player i, s_i his social approval obtained, and s_j the social approval obtained by player j. Total utility consists of a consumption and an approval part. Social approval is increasing in the size of a gift, i.e. a transfer of x.

Solving this model, one can see that player 1 has an incentive to make a gift to player 2 in order to gain approval. This has two effects on the utility of player 2. First, it lowers his marginal utility of consumption. Second, because player 2 is

now behind in the race for approval, his marginal utility of approval has increased. Both effects give incentives to player 2 for making a countergift, explaining reciprocity.

In some cases, the countergift of player 2 is appreciated by player 1 because his consumption increases. However, he also gives in in terms of net approval. If this latter effect dominates, player 1 may have an incentive to give inadequate gifts. To see this, note that an inadequate gift has less impact on the utility of consumption for player 2, resulting in a higher marginal utility of consumption. This makes a countergift more expensive, resulting evidently in a smaller countergift.

The analysis of chapter 3 also sheds light on the efficiency, or rather inefficiency, of gift-giving. Because players are involved in a race to be more approved than the other, gift-giving is generally inefficient. This results in a standard example of a prisoners' dilemma-like game. Both would be better off keeping their endowments, but each player has an incentive to deviate from this situation. When both deviate, they all lose since net approval sums up to zero and resources are wasted.

Another inefficiency results from the strategic power of player 1. Although he does take into account the behavior of player 2, he does not take it into account in a socially efficient way but only insofar as it concerns himself. That is, he ignores the effect of his behavior on player 2's utility.

Note also the positive result. In a situation where no explicit contracts can be written down, and implicit contracts are infeasible, a mutually beneficial exchange can still take place. Due to the need for approval, players give part of their endowment away. Their gift is related to, but not necessarily dependent on whether other players will give something back. The desire for approval does not, of course, guarantee an exchange whenever there are mutual benefits possible, but at least creates some opportunities.

It is not immediately obvious how the social approval approach is capable of explaining charity since the benevolences of these gifts are most often unobservable to the recipient. Recall however that social approval was obtained from the fact that a gift was made, not necessarily to the person who shows approval. Surely it is often socially approved to donate to developing countries by the people from the developed countries. Similarly, you can give blood which is anonymous for the receiver, but by telling your friends of your act it is not a truly anonymous gift to everybody. Charity funds seem to acknowledge this aspect by helping to

make donees visible. This can, for instance, be done through the provision of coffee mugs or label pins with the funds' name, or by publishing the names of the donees. There is indeed evidence that some people are sensitive to such actions (see Harbaugh [1998], Andreoni and Petrie [2000]) and in addition that this is related to the need for social approval (Satow [1975]).

2.3.5 A gift as signalling device

As argued in section 2.3.4, there is more to giving than altruism alone. In chapter 3 I elaborate in much more detail on social approval as a motivation to give, trying to make a case for it. There certainly seems something to it, as it explains many of the characteristics of gift-giving for a wide range of circumstances, including "anonymous" donations. However, as it stands, the model misses one important element: why should one give to get approval from others?

Social approval may be an important motivation but is perhaps only a derivative of even deeper more fundamental desires. It seems easy to come up with candidates for which one could get approval: status, being kind, honest, trustworthy, etc. However, it remains to be seen why a gift is necessary to get approval. Why does the recipient not approve of me without a need of giving? Of course, one can argue that the gift itself is approved. A gift itself is a kind act indeed, approval worthwhile. Another interesting line, taken up in this section, would be to assume that there is an information asymmetry which can be partly or completely resolved by making a gift. The recipient does not approve of me because he cannot tell what kind of person I am without making a gift. For example, being fair may be approved by others. But it is not obvious to judge who is fair or not from the outside. Still, a person who gives may signal that she is fair just by doing this. Because information asymmetries are so common in everyday life, Barkow [1989, 100] has even coined humans 'impression managers' rather than decision makers.

There are many more possibilities for gifts having the role of signalling something, including income, fairness, and even your own personality. This section reviews the role and scope of gifts as a signal. I start with a relatively detailed description of an early contribution by Camerer [1988]. This makes clear how a gift can be used as a signal. The other variants are then more easily understood. It should be stressed beforehand that gifts as signals need not be contrary to the

other motivations mentioned, but may rather be complementary. I come back to this, together with a general discussion, in the next section.

A gift signals trustworthiness

To enhance cooperation, it is necessary that agents trust each other. In section 2.3.1 it is already being discussed how a gift exchange can take place. There, the agents could trust each other to complete an exchange by the threat of losing the trading partner in future exchanges. This is only possible insofar as agents are not too impatient and if the time horizon is sufficiently long. Camerer [1988] has shown how a gift itself can signal trust, making even short-run cooperation feasible.

The essential mechanism is the following. It is assumed that there are two (groups of) players. Exchange can be realized between those groups. Each group consists of two types of players: trustworthy players and cheaters.[10] The fractions of these types in the groups are known but the individual type cannot be observed directly. The main difference between the types is their payoff. Trustworthy players are resistant to cheating. They would feel ashamed if they did, lowering their payoff. Cheaters on the other hand find it profitable to cheat, they have no feelings of shame or guilt whatsoever. The problem now is that an agent cannot know beforehand if his trade partner can be trusted. If he is of the trustworthy type the deal will work out fine, but if he cheats, the payoff will be considerably lower for the befooled agent than if he would not have traded at all.

What does the model predict? One result is that if the fractions of honest players in both groups is large enough, then the honest players will trade at the risk of being cheated. A more interesting result however is that even if the fractions of honest players are low, trade can still occur. The chances of meeting an honest trade partner are low, but if there is a possibility of giving a signal of trustworthiness then this does not need to be so much of a problem. The signal is to make a gift. This strategy can be explained as follows. If the fraction of cheaters players is high, then without gifts nobody would be trading. The payoffs are in this case not very large, but trading would on average be even worse for the honest players. Now, if a honest player makes a gift of a size that a cheater

[10] Nothing depends on the assumption of two different groups. One may also interpret them as two single players who are with some probability a cheater or a trustworthy person.

is not willing to make because it is not profitable enough for the latter, then this is a signal to the other player that this person is honest. And if both players are honest, then they can trade. In principle the cheater can make a gift as well, but if the gifts made by the honest person are expensive enough, the cost of the gift does not compensate for the payoff by cheating.[11] If such a separating equilibrium exists, honest players can signal the trustworthiness of trade partners by inspection of the size of the gift received.

Example. As an illustration of the above, consider the following example. There is a honest player ($H1$) who wants to trade with another person on the market. There are two players that he can trade with but he doesn't know which one of them is the honest ($H2$) and which one is a cheater (C) (he meets each player with probability 0.5). He has to decide whether to invest (I) or not (N). After investing or not, the trade partner makes a decision. Players that do not meet a trading agent have a payoff 0. The rest of the payoffs are as in the matrices below.

		$H2$	
		I	N
$H1$	I	$6/6$	$-10/5$
	N	$5/-10$	$0/0$

		C	
		I	N
$H1$	I	$6/1$	$-10/2$
	N	$5/-10$	$0/0$

As one can see, both the honest and the dishonest players are worst off when they invested but their trade partner did not. Moreover, both types prefer not to invest if the other doesn't. The difference is that the honest player prefers to invest if the other invests whereas the cheater prefers not to invest if the other invests. (An economic example may be two persons trading where cheating is beneficial for both in pecuniary terms, but where honest players have a sense of guilt outweighing the pecuniary payoffs and the cheaters have no sense of guilt.) It is readily seen that the dominant strategy for the cheater is not to invest.[12] What should $H1$ do? If he invests, he meets with probability $1/2$ $H2$ who then also invests and with probability $1/2$ C who does not invest. His expected payoff

[11] Thus a one-shot Prisoner's Dilemma (PD) game does not satisfy the assumptions needed for a separating equilibrium. A necessary assumption is that honest players gain by cooperating with another honest player and lose by cooperating with a cheater. By the structure of a PD-game, cooperating is never a best-response no matter what the type of the other player.

[12] Mixed strategies are not considered here as they are not equilibrium strategies (see the appendix in Camerer [1988]).

is therefore -2. His expected payoff by not investing equals 0. As a result, the honest player will not invest and consequently $H2$ experiences expected utility of 0 by also not investing. What $H2$ can do is to make a gift to player $H1$ before $H1$ decides.
Suppose that he decides to give an amount of 3 to $H1$. The net payoffs are then given in the left part of the matrix below:

		$H2$	
		I	N
$H1$	I	$9/3$	$-7/2$
	N	$8/-13$	$3/-3$

		C	
		I	N
$H1$	I	$9/-2$	$-7/-1$
	N	$8/-13$	$3/-3$

If C does not do the same then it is obvious for $H1$ what to do. Now he knows that if he gets a gift after meeting his trading partner, then the other is honest and so he should invest. Note that the expected utility of $H2$ is now equal to 1.5 which is still an improvement for him even considering the costs of the gift. What remains to be shown is that the cheater does indeed not make a gift. Consider the right part of the matrix above. If the cheater makes the gift of 3, then whatever the strategy of $H1$ is, he is worse off than the payoff of 0 when he didn't make the gift. The best thing he can do is therefore indeed not to make a gift. □

Obviously this model is able to explain the existence of gift giving. It can also account for some other aspects of gifts mentioned in section 2. It can explain one-sided gift giving. If the fraction of honest players in group 1 is large and if group 2 consists mainly of cheaters, then an equilibrium can be that honest players of group 2 must make a gift to signal their honesty, but for group 1 it is not necessary to give. The extreme case is where all players in group 1 and only one player in group 2 are honest. Obviously the players in group 1 do not have to give to signal whereas the player in group 2 does have to give. This can explain why gift-giving is not always reciprocal in nature and also when it is: if in both groups the number of cheaters is relatively large.

Additionally, with a slight modification the model is able to explain inadequate gift-giving. Recall that the gift must be large enough to make it unprofitable for the cheaters to give. But if gifts are adequate, under some circumstances it can still be profitable for cheaters to give since they benefit a lot from the gift they receive. This would make it impossible to distinguish the honest players from the

cheaters. Whenever the cheaters find it relatively more profitable also to signal, gifts should be more inadequate or else they fail in their aim.

Example (ctd.) Suppose for simplicity that the honest player $H1$ in the above game also makes a gift. This does not change the strategy of the cheater since both his payoffs from playing N or I are now increased. The adequacy also does not matter in this case since it does not change the strategy. But suppose that another pre-stage is constructed. In this pre-stage, some entering costs must be paid. If the entering costs, T, are paid then the rest of the game is as in the above example. If the entering costs are not paid by a player, then he is not allowed to play the second stage and both will not invest. The purpose of the entering costs in the pre-stage is that the adequacy of the gift can now influence the strategy that will be played by the cheater. To see this, consider a gift size of x. Only a fraction δ of the gift adds to the payoff of the other player. The parameter δ is a measure of the adequacy of the gift. The cheater now has to choose to pay the entry costs or not, and if he pays the entry costs then he has to decide whether to give or not. We know that in equilibrium, once entry costs are paid, he will not give (otherwise the signal is useless). But he may still pay the entry costs and then collect the possible gift of the other. This would give him an expected payoff of $.5(-T + \delta x)$. To prevent him from doing this, the entry costs must be such that it is not profitable for the cheater to enter the second stage in the first place: $T > \delta x$ or $\delta < T/x$. As a result, the adequacy may not be too high. □

There are some other models in the same spirit where gifts are taken to be signals of the willingness to cooperate. Carmichael and MacLeod [1997] derive Nash-equilibria where inadequate gift-giving signals the right intentions for long-term cooperation.[13] Bolle [2000] presents a similar model as Camerer and explicitly derives how adequate gifts should be. Kranton [1996b] derives a strategy for the formation of relationships by incurring a cost at the beginning of a new relationship and gradually increasing the level of exchange. In Iannaccone [1992] it is tried to explain sacrificial behavior within social clubs. Again, sacrifices are inadequate gifts that signal the good intentions of the players. By demanding a gift from members of a social club that offers a good which is anti-congestible[14] free riding behavior is prevented.

[13] In their setup gifts are necessarily inadequate.

[14] Anti-congestible indicates that each member's participation increases benefits for the other members, contrary to congestible club goods where the benefits decrease with larger utilization of others.

A gift signals income

People care about consumption goods and a gift may help them to cooperate and exchange goods. A gift need not necessarily signal trustworthiness though. To see this, note that people not only care about consumption but equally about status (see section 3.2.1). One main source of status is, no doubt, the wealth of a person. The exact amount of wealth that a person possesses is however often not directly assessable. Glazer and Konrad [1996] show that the level of wealth can be demonstrated by making gifts. The mechanism is like above. Both the poor and the rich care about status. For the poor, however, the marginal utility of status is lower relative to that of consumption. A gift is then a credible signal of a certain amount of wealth since the poor are not willing to spend as much resources on status enhancing activities.

A gift signals what you know about the recipient

Also of interest is the idea put forward by Prendergast and Stole [2001] that a person derives utility from knowing how well they are believed to know the preferences of the other. This can be an important element in friendships. Besides this, they assume that people are altruistic. The altruistic motive is a reason to make a gift to the other. The interesting aspect is what kind of gift they give. The choice is between a cash gift or a gift in kind. By making a gift in kind he reveals to what extent he knows the recipient's preferences. If the good is desired by the recipient, he will believe that the giver was aware of his preferences, something that the giver appreciates. A utility loss arises when the gift was improperly chosen, making the recipient believe that the giver is unaware of his preferences[15].

Suppose that the giver is not completely informed about the recipient's preferences. If he now gives in kind, he risks a loss of utility because he may end up buying the wrong gift, deriving no utility from knowing the recipient's preferences and because the recipient's utility is lower than a cash gift. If he believes he knows the recipient's preference quite well, he may nevertheless be willing to take this risk. Hence, a gift in kind signals the belief of the giver that he knows

[15] A somewhat related model is that by Ruffle [1999]. Instead of deriving utility from knowing the recipient's preferences, he assumes that utility is derived from surprising the recipient, e.g. by making a larger gift that was expected. This makes it a so called psychological game because beliefs enter the payoff function. Interestingly, it explains why you should not observe wish lists all the time since it would exclude the possibility of surprise.

the recipient's preferences. Notice that the story at the outset of this chapter fits this interpretation particularly well.

There are also situations in which the information asymmetry concerns the payoffs for the giver. In a principal-agent relationship, it often happens that the principal is more informed than the agent about the difficulty of a task or the ability of the agent (Bénabou and Tirole [2002a]). A variant of this is studied in chapter 6. There, the principal observes the outcome of a task, in essence whether it has been successful or has resulted in a failure, whereas the agent only gets to know this information through an imperfect signal. Neither the principal nor the agent have any direct assessment possibilities of the agent's ability. It is shown how a gift (a bonus in this case) by the principal may signal a success, which indicates that the agent is more likely to be of high ability. A gift then stimulates the agent to continue making efforts, because he is more self-confident after a reward.

A gift reveals your self

So far, the informational asymmetry existed in the receiving party having imperfect information about the giver. Relatedly, the giver may have imperfect knowledge over his own preferences. Although the human organism has the capacity to be conscious about its *self,* it is by no means the case that each person is totally aware of his personality. It is also the case that the self is not directly perceived but built up through experience about one's own behavior or reflection by others. For example, Baumeister [1998, 684] writes: "Consider what is involved in knowing that you are shy. You notice that you always get nervous in the presence of others, and you prefer to avoid large social gatherings and meeting new people; these observations permit the conclusion that your are shy". You can only know that you are shy through encounters with others. In this way, you learn something about your self by reflecting on your own behavior.

With imperfect information about the self, there is room for 'self-reputation'. Bénabou and Tirole [2001] provide an interesting application. Their assumption involves imperfect knowledge about one's own willpower. A person may have a strong or weak willpower, but his type is *ex ante* unknown to him. In a repeated setting, past behavior may (partly) reveal the type. By not giving in to temptations today, the individual may later draw inferences that he must have strong willpower.

Within this framework, Bénabou and Tirole try to explain tipping behavior. The individual may get involved in profitable long term relationships. Cooperation in many of these situations would be beneficial. However, the individual also has short term interests conflicting with long term cooperation. Only if the individual can think of himself as having the willpower to sustain long term relationships will he get involved in these. The individual may therefore wish to show strong willpower in short term relationships as well, to signal to himself that he can sustain long term relationships. Tipping can, according to Bénabou and Tirole, be regarded as a compulsion: the individual is so motivated to signal a strong willpower to himself that he is even willing to show cooperativeness in situations where any direct long term gains are absent.

A gift reveals a strategy

Whereas in the above references signals serve to reveal one's *type*, there is also a variant of signalling that reveals one's *strategy*. Here, it suffices to consider only one type. Consider the battle-of-the-sexes. The essence of the game is that both players benefit only by choosing the same strategy. Each player has a preference for one strategy over the other and these preferences are different for both players. However, choosing different strategies makes both players worse off. Now suppose that both players can actually make a worthless gift to the other. This is in fact a reinterpretation of the example taken from Van Damme [1989] where there is an opportunity to burn a certain amount of money. By forward induction, Van Damme shows that a gift is made with positive probability. The gift serves as a credible threat of playing a particular strategy.

Generally, in accordance with empirical data, gifts as signals should be inadequate or need at least not necessarily be adequate. For example in the games by Camerer [1988] and Kranton [1996b] inadequacy is a prerequisite. In the model of Van Damme [1989] inadequate as well as adequate gifts are possible[16]. Reciprocity can also be explained by most signalling models. This is a strong point of the signalling approach. An obvious restriction of this approach is that it is related to informational asymmetries. This is likely to be most relevant for relationships

[16] This is simple to prove. The argument in Van Damme [1989] (see in particular his fig. 5) is independent of what the other player gets. The strategy of each player takes the other player's strategy as given and as a result the adequacy of the gift does not matter.

that are short in nature or are in their beginning phase. Business gifts may belong to this category, Christmas gifts less so.

2.4 Discussion

The main focus of this chapter has not been to provide a unique unified theory of gift-giving. Rather, it aims at exposing competing theories and to evaluate them on their explanation power of accounting for the two characteristics reciprocity and adequacy. The possibility of coexisting motivations for gift-giving should not be disregarded. Perhaps the motivations differ between different kinds of gifts, different people, or different time periods.

The chapter did, however, shed some light on which motivations to look for in a wide range of situations. Gifts with the purpose of exchange can only be expected when players have a sufficiently long horizon and where the players are familiar to each other. Altruism cannot explain inadequate gifts, although it is likely to partly explain charity. Social approval can to a certain extent explain charity, although the scope is somewhat limited: at least some people should get to know about the act. Fairness may explain charity as well, and describes many experimental games relatively well. Signalling explanations are powerful to explain short-run or beginning relationships. They seem not applicable to long run relationships, where information asymmetries have dissolved over time.

I would also like to argue for a hybrid model of gift-giving: one where different motivations act and interact simultaneously. Emotions and strategical actions may both play a role. What I have in mind here is for example that people are not intrinsically fair but still like to appear to other people as fair, and accomplish that by giving a signal that one is fair. This combines the different motivations fairness, a need for approval, and signalling. This particular example comes also out of the experiment by Güth and Van Damme [1998], where a third dummy player is added to the ultimatum game (see section 2.3.3). Recall that one of the main results is that only marginal offers are proposed to the dummy player. They conjecture that the proposer was not intrinsically motivated by fairness considerations but that they do not want to *be* fair, but rather want to *appear* fair, e.g. to prevent a rejection by the responder.

Knowing the motivations helps to design efficient incentive systems. There is one classical example where misinterpreting the motivations to give led to a wors-

ening of the situation. When Titmuss [1970] examined the blood market, he found that when it was tried to stimulate blood donations to reduce the shortage of blood, by giving monetary compensations, this resulted in less blood donations rather than more. If altruism were the only explanation of blood donation (which is not very likely, see section 2.3.4) the decision to compensate would be understandable. If other factors play a role, such as a need for approval, the effects of monetary compensations on these motivations should not be disregarded. Indeed, in chapter 3 I show that monetary compensations may decrease the willingness to give by obstructing possible approval (less approval is rewarded if the sacrifice is less).

The blood market is not the only example. In another example taken from Gneezy and Rustichini [2000] the other way around is also found, namely that the parents that arrived too late to collect their children at day-care centers increased in number after a fine was imposed (the gift is here the additional time that was spent by the employees). In the interpretation of opening markets, the creation of a market is in this case the cause of crowding out of the gift, something that is indeed often observed (see for instance Frey [1997a,b], and Frey and Jegen [2002] but also the discussion in Arrow [1972]). This conclusion is similar in spirit to that of Holländer who concludes that it may well be the case that "the opening of a market (...) reduces voluntary contributions" (Holländer [1990, 1165]). I refer to section 3.2.4 for a more detailed discussion about the interaction between markets and gifts. As one can see already, the appropriate incentive mechanism takes into account other motivations than selfishness. A name in the records may have more impact than a reward.

A final note on where all these motivations come from. Carmichael and MacLeod [1997] have shown that gifts as a signal at the beginning of a relationship can lead to cooperation. They also showed that this institution is evolutionarily stable. The intuition is as follows. A gift is a signal of trustworthiness and has to be given at the beginning of any new relationship. If both agents in a partnership conform to this custom they only have to give once and stay in the partnership forever. A free rider, on the other hand, is detected after one period and has to search for a new relationship again. This cannot be a strategy that often pays off, since for every new relationship a gift has to be incurred.

Other authors have shown that many other emotions such as anger and altruism can also be evolutionarily stable (see e.g. Güth and Kliemt [1994], [1998]).

Emotions like anger may result in lower current payoffs, since punishing is usually also costly to the punisher, but may still induce behavior that is evolutionarily successful. Defectors are deterred away by the threat of punishments. This explains such sentiments as having emerged from selection pressures, which has resulted in modern institutions such as gift-giving.

3

The Demand for Social Approval as a Motivation to Give

The Kwakiutl were once one of the major tribes of the Norhwest Coast. They were quite wealthy, often even richer than many of the other settlers in that area (see Codere [1950]). They have intrigued anthropologists not in the least because of their ceremonies. The best known ceremony is the so called potlatch, a gift-giving ceremony. During the potlatches, many blankets and large amounts of copper were being given away to other tribesmen. However, this was done in a way that seems particularly inefficient: many of the gifts are destroyed on the spot. The anthropologist Mauss observed that during these potlatches "they go as far as the purely sumptuary destruction of wealth" (Mauss [1926, 6], witness also the following quote:

> *(...) she ordered one of her kinsmen to tow it [copper] to sea behind a canoe and to cut it adrift in deep water and let it sink. "This is my gift to you, O chief." (Drucker and Heizer [1967, 105]).*

Although less visible and certainly less extreme, the same behaviour can also be found in gift exchanges that take place in modern societies. Rather than giving

---

[0] A shorter version of this chapter appeared in the *Journal of Institutional and Theoretical Economics*, 2002 vol. 158 (3), 464-482 under the title: The Demand for Social Approval and Status as a Motivation to Give. I am indebted to Michèle Belot, Jeffrey James, Luuk van Kempen, Fieke van der Lecq, Sjak Smulders, participants at the European Economic Association Conference (Lausanne, 2001), and especially to Theo van de Klundert and two anonymous referees for very valuable comments and suggestions on earlier versions.

in cash, gifts are mostly in kind, thereby destroying some of the monetary value to the receiver. This chapter tries to shed some light on the motivations behind such, and other related behaviour.

3.1 Introduction

Gift-giving has mainly interested anthropologists because it has been taken as a primitive mode of exchange. Relying both on a double coincidence of wants and on the existence of trust between the agents, gift-giving clearly seems to be inferior to the market mechanism. Yet, gift-giving is still widely observed even in countries with well-developed markets. Indeed, gift-giving sometimes even flourished in the presence of a market economy (Gregory [1989]).

The difficulty is not to explain gift-giving *per se*. One can simply attribute utility to the act of giving, or, in the terminology of Andreoni [1990], a warm glow feeling. Rather, there are some stylized facts related to gift-giving that have to be explained but are puzzling from a standard economics point of view. These stylized facts include reciprocity, inadequate giving, and a negative correlation between monetary compensations for gifts and the level of gift-giving. Reciprocity refers to the observation that, although voluntary on guise, gifts appear in fact to have strong reciprocal properties. Inadequacy points to the fact that gifts should be in cash in order to maximally satisfy the preferences of the receiver, but often they are not. Finally, it is observed that gift-giving is sometimes reduced after compensation is offered. Neither one of these stylized facts can easily be explained by standard economic arguments.

The main point to this chapter is to explain gift-giving by means of a demand for social approval and status; two factors deeply rooted in human nature. The idea that approval motivates gift-giving is in my opinion intuitively an appealing hypothesis. Public goods experiments show that familiarity with the identity and actions of other players leads to significantly higher contributions (Andreoni and Petrie [2000], Gächter and Fehr [1999]). Earlier Satow [1975] has shown that the increase in donations in public conditions as compared to private conditions is strongly correlated to the individual's need for approval.

Although not entirely neglected in the economics literature,[1] the concepts of social approval and status have, above all, found an eminent place in sociology and social psychology. It plays the same part there as money does in economics: recognition by others is regarded as a primary source of satisfaction (see for example Homans [1961]; Kenrick, Neuberg, and Cialdini [1999]; and Coleman [1990]). Social approval in this sense appears to be a functional substitute for money. Hence, a transaction that is unequal in monetary terms – as a gift is – can in principle still be in balance as long as approval is awarded.

Gift-giving indeed appears to be a virtue and a source of prestige (Polanyi [1957], Schwartz [1967]). The way in which approval is obtained is however complex. Evidently, approval will be higher the more the gift is valued. But it also turns out to depend on the sacrifice incurred by the giver. Moreover, approval seems to be closely linked with status. Recognition and status are often mentioned in one and the same breath (for example in Schwartz [1967, 7], and Harsanyi [1969, 523]). Taking these factors into account, the stylized facts become natural implications of the model.

Other implications of taking into account the taste for social approval follow from relating gift-giving to the market institution. Since the market is in its purest form an anonymous institution, no social approval is obtained in a market exchange. On the other hand creates the market incentives to maximize adequacy. Gift-giving as an exchange mechanism does allow for acquiring approval but in general fails to maximize adequacy. Hence, in choosing between a gift-exchange or a market exchange the trade-off to be made is that between approval and adequacy. However, spontaneous order does not necessarily lead to the most efficient institution because the links between the institutions are shot through with externalities (Dasgupta [2000]). Unfortunately, trying to correct for any inefficiencies with a standard economic tool like money compensations does not always resolve this and may even have the opposite effect of worsening the situation. This is in line with what the model is able to predict and with ample empirical evidence.

Roughly, the basic line of argument runs as follows. In the model, two individuals are playing a sequential move gift-giving game. Each of them has preferences for both a consumption good and social approval. Player 1, then, makes a gift for which approval is awarded. But lagging behind in the status race for wanting

[1] See Akerlof [1997], Holländer [1990], and the references therein.

more approval than the other, player 2 finds it profitable to react by making a countergift. This explains the so often found reciprocal behavior. Inadequacy is just as easily explained. In trying to prevent player 2 from catching up in terms of status, player 1 can deliberately devalue his gift by giving in kind rather than in cash. This makes it more expensive for player 2 to reciprocate. Finally, note that a compensation reduces the sacrifice needed to make a gift. This is positive in terms of consumption, but it is also likely to reduce the awarded social approval for the gift. Whereas with standard assumptions on preferences a compensation should increase gift-giving, with a preference for approval it is ambiguous whether compensation increases or reduces gift-giving in equilibrium.

The setup is as follows. The subsequent section is concerned with deriving the basic properties of gift-giving from a simple model. Building on empirical as well as experimental evidence, an extensive account is given on what the social approval function should look like and what kind of consequences it has for the way economists think about the workings of the market institution. Section 3.3 relates the model to the existing literature. Finally, section 3.4 concludes.

3.2 Social approval, status, and gift-Giving

"Gratitude is bestowed on a giver." Aristotle*, Nichomachean Ethics.*

3.2.1 The basic model

In this section, the tentative explanations of the stylized facts are made more precise. The model is highly stylized and is only meant to be suggestive in explaining how social approval can affect decisions. I believe however that the key insights are not sensitive to the specification of the model and that they would survive in a more general framework.

In the model there are two players. Player 1 is the fist mover in a sequential gift-giving game, and player 2 follows. Their decision variable is the amount of time spent doing volunteer work (l_i^v). The rest of their available unit of time ($1 - l_i^v$) is devoted to working in the market sector at wage w_i^m. With the income that is earned on the market, $w_i^m(1 - l_i^v)$, the consumption good x can be purchased, which is available at unity price.

The time spent doing volunteer work is a gift and contributes to the other person's consumption level. For simplicity, it is assumed that the gift is the same

consumption good x. The gift increases the other person's consumption by the amount of $\delta_i l_i^v$. This way, the parameter δ can be interpreted as a measure of *adequacy*. By adequacy is meant how the receiver's utility of consumption is increased relative to the costs incurred by the giver. A more adequate gift increases the receiver's utility of consumption more given the costs incurred by the giver. Note that standard microeconomics arguments tell us that it can never be worse in terms of utility to get a gift in cash rather than in kind. As a result, a cash-gift is in general more adequate than a gift in kind. This can be translated back into the model as follows. If $\delta_i = w_i^m$ then the gift is exactly identical to giving money. Hence, this case is interpreted as if it were a cash-gift. If $\delta_i < w_i^m$, the gift is worth less than the cash-equivalent. The latter case is interpreted as a gift in kind.[2]

In sum, total consumption of good x by player i is given by:

$$x_i = w_i^m(1 - l_i^v) + \delta_j l_j^v. \tag{3.1}$$

If utility is only derived from consumption, as is usually assumed in economics, neither one of the players will give. Whatever player 1 gives to player 2, it is optimal for player 2 not to make a countergift. Foreseeing this behavior, player 1 does not make a gift as well.

The main departure from standard models is the inclusion of a preference for social approval as well as for status. There are good reasons to do so. The worry about status and the 'thirst for approval' are among the most recurring themes in anthropology (Wright [1994]). They thus seem to be deeply rooted in human nature. The existence of suchlike sentiments possibly has emerged from selection pressures because emotive motivations indirectly induce behavior that is evolutionary successful (see Güth and Kliemt [1994], [1998]). Each of these emotions is explained in more detail below.

Social Approval

The first building block is to take into account the taste for approval. It is beyond any doubt that a preferences for approval exists. Man is a 'social being' whose economy is submerged in his social relationships. Polanyi [1957, 46] considers this

[2] If $\delta_i > w_i^m$, a gift is worth more to the receiver than its cash-equivalent. As said, this is not the standard case, but one can think of examples where it is a possibility, for instance when the receiver has incomplete knowledge over his own preferences.

to be the "one conclusion [that] stands out more clearly than another from the recent study of early societies."[3] Indeed, the preference for social approval may very well be as important as the preference for consumption goods (Harsanyi [1969], Sugden [1989], Dasgupta [2000]). But it seems that the actual obtaining of approval is a rather complex process. In order not to complicate matters any more than is necessary for current purposes, the focus is on two elementary properties. First, the higher the value of the gift as judged by the receiver, the more the gift is approved. The social approval function for player i should consequently be increasing in δ_i.[4] Second, approval is increasing in the sacrifice made by the giver. Sacrifice is likely to be something relative to what you earn. A gift of one dollar by the poor is approved more than the same donation by a millionaire. This is exactly the behavior that Pruitt [1968] finds. In his experiment, more reward was provided by the receiver if the giver had sent out 80% of his endowment of \$1 than if he had sent out 20% of his endowment of \$4, presumably because the sacrifice is larger in the former case. The simplest measure of sacrifice is the wage level. The higher the wage, the more consumption is forgone in order to give. So, a higher wage corresponds to a higher sacrifice. This amounts to a social approval function that is increasing in w_i^m. Further empirical evidence for this is found by Robben and Verhallen [1994].

Status – Why humans get ulcers

> *"Men have an immoderate love of pleasure, influence, prestige, power – in a word, wealth."* F. Bastiat*, Economic Sophisms.*

It is conventional in economics to assume that a higher income generates a higher well-being. When psychologists try to measure the happiness of people they indeed find such a relationship on the individual level. Paradoxically, however, the average satisfaction level is remarkably stable over time, despite significant increases in per capita income (see for example Frank [1997]).

It seems that the disappointing increase in happiness is closely related to a question that occupied Hume more than two centuries ago: How can it be that

[3] For a similar account in modern sociology see Coleman [1990].

[4] This seems evident but may in fact not be entirely trivial. In some cases a very large gift may actually cause embarrassment thereby decreasing approval, as in the case one would get diamonds on a first date. I do not pursue this point any further here.

our impressions of the same object can at one time be an admiration of its bulk, and at another to despise it for its littleness? Hume found the solution in a careful examination of human passions:

> "So little are men governed by reason in their sentiments and opinions, that they always judge more of objects by comparison than from their intrinsic worth and value" (Hume [1886, 158]).

Comparison is the keyword here. Much like a certain mountain looks big next to a smaller mountain yet only little next to a bigger one, the same amount of wealth is pleasurable if you have more of it than your neighbours, but becomes frustrating when you have less. In other words, the apparent solution to the paradox in question can be found in the supposition that people deeply care about status. That is, they certainly do care about wealth, but only insofar as it increases their prestige, a form of status. This role of wealth, also present in the quote by Bastiat at the outset of this section, implies that the increased well-being of an individual from a higher income dissolves once the other people in her reference group reach the same level of prestige that she was privileged to enjoy before. A general increase of the wealth of a nation has thus no profound effect on the general happiness of people.

The idea that people care about status has often been recognized before in the economic literature[5]. The reason for that is clear, as already as far back as Veblen [1899/1953, 80] it is acknowledged that in many articles "... the traces of conspicuous waste, or at least the habit of ostentation, usually become evident on a close scrutinity". Darwinian anthropologists have put forward the hypothesis that worries about status is one of the most recurring patterns in all cultures (see Wright [1994]). Furthermore, research in biology confirms that up to the present day humans are equipped with a hard-wired preference for status. The achievement of status involves physiological consequences. For example, it turns out that more of the neurotransmitter serotonin (which is, as it happens, also used in many antidepressants) is being secreted when humans (or animals) are in the position of a leader (see Wright [1994] and Frank [1985] for a more elaborate

[5] For an early contribution that employs status as an assumption: see Leibenstein [1950]. More recent contributions include Cooper et al. [2001] and Corneo and Jeanne [1997]. Van Kempen [2003] points out that even poor people have status concerns. See also the references therein and in the main text of this chapter for more contributions.

treatment). In his *Why zebras don't get ulcers*, Sapolsky [1998] even goes as far as associating sudden cardiac deaths (and human ulcers, for that matter) with major stressors such as the loss of status, and cites several studies that are indeed congruent with this view. Evolutionarily speaking, all this makes sense: a high level of status gives better access to resources, such as food, and increases the probability of successful reproduction as it attracts partners (Frank [1985], Wright [1994]).

Much of the research done has focused on domination-oriented status, and mostly among animals. But, according to Barkow [1989], as a consequence of selection contemporary human prestige has a more symbolic nature than the agonistic dominance found among primates. Prestigious objects or acts signal an increased ability and willingness to make paternal investments in offspring.

Giving can be a source of status as well, in the form of prestige. This is for example the case in the consumption domain. Wealth increases prestige. When wealth is not so visible to outsiders, gifts can enhance status by giving a signal about the wealthiness of a particular person.[6] I, however, follow Holländer [1990] in employing status effects in the approval domain. Holländer [1990] cites several studies in which it is argued that people not only want to be admired, but also want to be more admired than others. This is also experimentally verified by Gächter and Fehr [1996]. They find that social approval decreases in the average level of contributions by other subjects and point to the similarity with status effects. Implicitly, Forge [1972] finds this also to be the case in the more archaic societies in New Guinea when he remarks that a gift exchange is never perfectly balanced so that if the recipient fails to outdo the giver, it is the giver who gains and the receiver who loses. Or, in the more explicit words of Gregory [1989, 110] "... gift exchange necessarily introduces status inequalities".[7]

Preferences

Based on these building blocks, the following measure for approval is proposed:

$$s_i = \beta_i l_i^v - \alpha\beta_j l_j^v, \qquad \alpha \geq 0. \tag{3.2}$$

[6] See Glazer and Konrad [1996] for a formal model.

[7] And how many others will not be as tempted as Ng is to give as expensive gifts as the schoolmates of one's child receive? See Ng [1997].

Every unit of voluntary labor is weighted by the function $\beta_i = \beta(\delta_i, \tilde{w}_i)$, which is increasing in both of its arguments: the weight is larger if the value of gift is higher (δ_i) and if the sacrifice ($\tilde{w}_i$) is greater, as measured by the forgone wage of the giver (w_i^m). The parameter α in equation (3.2) reflects the degree of status orientation.

Finally, preferences are represented by a utility function that is for simplicity additive in consumption and net social approval:

$$u_i = u_x(x_i) + u_s(s_i). \tag{3.3}$$

It satisfies the usual assumptions with respect to x and s: u'_x, $u'_s > 0$ and u''_x, $u''_s \leq 0$, where (double) primes denote first (second) derivatives. At this point it should be noted that in practice the degree of status orientation is not uniform among people (see Wright [1994]) and neither is the intensity of the need for approval (Satow [1975]). Including such heterogeneity would not alter the basic arguments of this chapter and is for that reason left for future research.

Equation (3.3) is a possible representation of the theory of social behavior that Homans has in mind when he speaks of 'social behavior as exchange' (Homans [1958, 606]). In his view, social behavior is an exchange of goods, including non-material ones such as approval. Now note that although an exchange of material goods can be accomplished in any social setting, the accomplishment of approval is tied to social interaction. That the utility derived from a good depends on the social context is a simple extension of Lancaster's theory of consumption (Lancaster [1966], Hirsch [1976, 85]). In the view of Lancaster [1966], goods as such are not the direct object of utility. Instead, utility is derived from characteristics and goods are bundles of characteristics. The environment of exchange can be seen as one of those characteristics. Compare also Bowles [1998, 87] who notes that "the terms on which [people] are willing to transact depends on the perceived relationship."

This amounts to a useful interpretation of equation (3.3) in that it captures various exchange regimes characterized by different intensities of social interaction, with the strictly anonymous market exchange as the extreme where no social interaction exists at all ($u_s = 0$, although practically speaking even the market is characterized by some social interaction rather than by complete anonymity). Hence, in studying market trade there are good reasons to neglect the role of any such sentiments since they are ruled out by anonymity. But things are dif-

ferent when studying gift-giving where, contrary to abstract markets, trade is not anonymous. Gift exchange is above all a social relationship (Gregory [1989]). In such a case the taste for approval strongly influences behavior. Suggestive in this respect is the recent finding by Gächter and Fehr [1999] that in experimental settings some minimal social familiarity generates a significant rise in cooperation.

3.2.2 Sequential move equilibrium

The game is solved by backward induction. The problem of player 2 is to solve for:

$$l_2^{v*} \in \arg\max_{l_2^v} u_2(l_1^v, l_2^v), \tag{3.4}$$

for any given l_1^v. Ignoring parameters, this gives a reaction function of player 2 of the form $l_2^{v*} = f_2(l_1^v)$. Player 1 takes this behavior of player 2 into account and therefore solves for:

$$l_1^{v*} \in \arg\max_{l_1^v} u_1(l_1^v, f_2(l_1^v)). \tag{3.5}$$

The solutions to (3.4) and (3.5) constitute an equilibrium. Throughout the focus is on an interior solution since this is the only interesting case.

Reciprocity

The first result is obtained from the properties of player 2's response function $f_2(l_1^v)$. Since this function relates the optimal gift of player 2 to the gift of player 1, it predicts whether reciprocal behavior should be observed or not. Note first that, as argued earlier, without a preference for approval ($u_s = 0$) the optimal response of player 2 is not to make a gift, independent of the gift by player 1. Without a preference for approval, the only effect of a gift would be a loss of consumption. Player 1 foresees that $f_2 = 0$, so that it is also for him optimal not to give. No gift-giving occurs even if there are mutual gains of giving ($\delta_1 > w_2^m$ and $\delta_2 > w_1^m$). The cause of this is that player 1 has to rely on a countergift by player 2, but he has no reason to trust player 2 on this. An explicit or implicit contract can of course solve this problem of trust. But reciprocal gift-giving is also observed in situations where no explicit contract exists, or at most an incomplete one, and where an implicit contract is not credible. Mauss [1925/1980], for example, portrays gift-giving among tribes, and finds that reciprocity is one of the basic elements. And Akerlof [1982] gives a description of labor contracts, where wages

are set above the minimum acceptable standard and these are reciprocated in the form of higher efforts. These results of reciprocal gift-giving without complete explicit contracts are replicated in many laboratory experiments (see for example Fehr and Schmidt [1999] and Bolton and Ockenfels [2000]). In many cases these were one-shot games so that reputation cannot enforce credible implicit contracts either (see for example Cooper et al. [1996], and Gächter and Falk [1999]). The point of all this, is that apparently contracts are not always necessary to induce reciprocal gift-giving. As it turns out, a preference for social approval is enough.

That gifts can be positive in equilibrium is quite obvious: all one needs to assume is that the gain in net social approval outweighs the loss of consumption in utility terms. More interesting is to account for reciprocity. This is established in the following proposition:

Proposition 1 (Reciprocity): *With a preference for social approval, the optimal gift of player 2 is positively related to (i) the gift of player 1* ($\partial f_2/\partial l_1^v > 0$)*, and (ii) the degree of status orientation* ($\partial f_2/\partial \alpha > 0$).

Proof. By definition of f_2: $\frac{\partial u_2(l_1^v, f_2(l_1^v,\cdot),\cdot)}{\partial l_2^v} \equiv 0$. Differentiating both sides with respect to l_1^v gives: $\frac{\partial^2 u_2(\cdot)}{\partial l_2^v \partial l_1^v} + \frac{\partial^2 u_2(\cdot)}{\partial l_2^{v2}} \cdot \frac{\partial f_2(l_1^v,\cdot)}{\partial l_1^v} = 0 \Leftrightarrow \frac{\partial f_2(l_1^v)}{\partial l_1^v} = -\left.\frac{\partial^2 u_2(\cdot)/\partial l_2^v \partial l_1^v}{\partial^2 u_2(\cdot)/\partial l_2^{v2}}\right|_{l_2^{v*}}$. By the second order condition of maximization, the denominator of the right hand side is negative, so that $sign\frac{\partial f_2(l_1^v)}{\partial l_1^v} = sign\frac{\partial^2 u_2(\cdot)}{\partial l_2^v \partial l_1^v} = sign$
$\left[u_x'' \frac{\partial x_2}{\partial l_2^v}\frac{\partial x_2}{\partial l_1^v} + u_s'' \frac{\partial s_2}{\partial l_2^v}\frac{\partial s_2}{\partial l_1^v}\right]\Big|_{l_2^{v*}}$. The first term in brackets is strictly positive. The second term is zero if $\alpha = 0$, in which case $\frac{\partial s_2}{\partial l_1^v} = 0$, and positive if $\alpha > 0$. Hence, at the optimum $sign\frac{\partial f_2(l_1^v)}{\partial l_1^v} > 0$. Similarly, $sign\frac{\partial f_2}{\partial \alpha} = sign\left.\frac{\partial^2 u_2}{\partial l_2 \partial \alpha}\right|_{l_2^{v*}} > 0$, proving the second part. ■

If the players care about approval, they behave in a reciprocal manner. This result in itself (i.e. (i) in Proposition 1) holds in principle for every degree of status orientation (including $\alpha = 0$), but is strengthened by a higher degree of status orientation (see (ii)). This partly solves the trust problem. Player 1 has now reason to believe that player 2 makes a countergift, simply because it is in player 2's own benefit to do so. This does not mean that player 2 *always* reciprocates the gift of player 1, just like an altruist not always gives something. What it means is that the more player 1 gives, the more he is likely to get something back. The intuition behind this result is the following. The gift of player 1 increases the consumption level of player 2, thereby decreasing marginal

utility of consumption. Insofar as status plays a role, the gift also decreases the net social approval received by player 2, thereby increasing the marginal utility of net social approval. Because a countergift has opposite effects on the consumption and the net approval level, a gift by player 1 creates incentives for player 2 to give as well.

The foregoing has concentrated on reciprocity in the sense of rewarding generous gifts with countergifts, also called *positive* reciprocity. The other side to the coin is *negative* reciprocity, by which it is commonly understood that players who are stingy are punished for being so. For instance, low offers in ultimatum games are often rejected, making both players worse off.

The terms 'rewarding' and 'punishing' seem to imply a (fairness) norm to which players are expected to comply to. Positive deviations are rewarded, negative ones punished. Rejections in anonymously played ultimatum games and other public good games with the option to punish is indeed best explained by feelings of resentment, negative reciprocity in other words (Fehr and Gächter [2000]). This has not much to do with approval. Nonetheless does a demand for social approval and status give some hints as to what will happen in those games when there would be some familiarity among the players. First, players would have the opportunity to get approval for punishing stingy players, strengthening the incentives for punishments. Second, apart from feelings of resentment, players might be less willing to accept low offers in the ultimatum game for fear of a loss of status. Since player 2 can either accept (A) or reject (R) in the ultimatum game, his gain in utility from accepting the gift of player 1 is[8]:

$$\Delta u_2 = u_2(A) - u_2(R) \approx \delta_1 l_1^v u'_x - \alpha\beta_2 l_1^v u'_s, \tag{3.6}$$

which is the sum of the gain in consumption and the loss in net approval. To accept but not being able to reciprocate means a sure loss in net approval, hence the observation that "Charity is still wounding for him who has accepted it" (Mauss [1925/1980, 65]). For high levels of status orientation this status loss outweighs the consumption gain and makes rejection sensible. Note that the receiver would have preferred to give back ($f_2 > 0$) but is by construction forced to $l_2^v = 0$ (with $l_1^v > 0$), or to refrain from any gifts at all (both $l_1^v = l_2^v = 0$).

[8] To derive equation (3.6), note that $\Delta u_2/\Delta l_1^v \approx \partial u_2/\partial l_1^v = \delta_1 u'_x - \alpha\beta_1 u'_s$ and $\Delta l_1^v = l_1^v - 0$ since the reference point is no gifts.

Interestingly, and perhaps related, Mauss [1925] also points out that the root of the word gift is the same as that of poison. Harm befalls on the recipient of a gift, as he is challenged to catch up in the competitive race for status. This imbalance that is created by a one-sided gift is also nicely captured in the Indian verb quoted in Brigham [1991, 300]: "Why do you hate me? I've never even helped you".

A cautious remark should be made here. A rejection itself can be perfectly rationalized within the current model. But it turns out that while low offers are sometimes rejected, high offers are normally accepted. It is easily checked that with $u_s'' < 0$ this behavior is inconsistent. This indicates that the approval function is more complex than the one operated, for instance, one that is convex above a reference point and concave below. Although most of the results would go through under this extension, it leaves open the question how such a reference point is determined, and is therefore left for further research.

One would thus expect that rejections are more often observed in ultimatum games where there is familiarity among the players. Translated to out-of-laboratory situations, one would expect a needy person to more easily accept a gift from an anonymous source than from a person he can identify, since the latter introduces a loss in net social approval. Finally, one would expect that if possible, players would rather want to give something back. At least for the latter statement there is already some evidence (see Kenrick, Neuberg, and Cialdini [1999]) but it would certainly be worth to investigate these and related issues more thoroughly.

Adequacy

The presence of a taste for approval secures that a gift is usually reciprocated. This allows an exchange to take place without the support of contracts. But in the current gift-giving game, the exchange is only an intermediate variable with approval as the ultimate objective. It is interesting to see how this role of exchange affects the properties of the gift. This relates to a puzzle that concerns the adequacy of gifts. A cash-gift is more adequate than a gift in kind. But on many occasions, it is quite unusual to give cash. Research by Webley and Wilson [1989] is illustrative of this. Their questionnaires reveal that subjects prefer to give presents rather than money. Likewise, Caplow [1982] finds that in the data he collected less than 9% of the Christmas gifts were in cash.

This behaviour is puzzling at first sight because of the assumption that gifts serve to gain approval and presumably less adequate gifts are less approved. But on closer inspection, there is a rationale behind this behavior. Player 1 likes getting a countergift in that it increases his consumption. But he dislikes it in that it decreases his net approval. Clearly, if the gift of player 2 would be a little bit lower, he can be either better or worse off, depending on the relative decrease in consumption as compared to the relative increase in net social approval. This means that there are incentives for player 1 to manipulate the gift of player 2. Suppose he indeed wants to lower the gift of player 2. According to the following lemma, he can do this by decreasing the adequacy of his own gift.

Lemma 1 : *Player 2 reciprocates more if he receives more adequate gifts, that is,* $\partial f_2/\partial \delta_1 > 0$.

Proof. With the introduction of the second decision variable for player 1, δ_1, the optimal gift of player 2 is dependent both on l_1^v and δ_1: $l_2^{v*} = f_2(l_1^v, \delta_1)$ (ignoring parameters). By definition of f_2: $\frac{\partial u_2(l_1^v, f_2(l_1^v,\delta_1),\delta_1)}{\partial f_2} \equiv 0$. Differentiating both sides with respect to δ_1 gives, after rearranging: $\frac{\partial f_2}{\partial \delta_1} = -\left.\frac{\partial^2 u_2/\partial l_2 \partial \delta_1}{\partial^2 u_2/\partial l_2^{v2}}\right|_{l_2^{v*}}$. By the second order condition the denominator is negative, so that $sign\frac{\partial f_2}{\partial \delta_1} = sign\left.\frac{\partial^2 u_2}{\partial l_2^v \partial \delta_1}\right|_{l_2^{v*}} > 0$. ■

The intuition behind this intermediate result is that by lowering the adequacy of his gift, player 1 contributes less to the consumption level of player 2, keeping the marginal utility of consumption for him at a high level. This makes it more expensive for player 2 to give. Besides this effect, it also increases the net status of player 2, decreasing the marginal utility of social approval. Both effects reduce the incentives for player 2 to make a gift. The behaviour by the kwakiutl of destroying gifts (see the introduction to this chapter) can also be understood along to these lines.

To derive the desired result formally, I say that the gesture is relatively important if the approval rate is not very sensitive to the value of the gift: i.e. $\partial \beta_i/\partial \delta_i$ is low. In this case, it is the *act of giving* that counts, and not so much the gift itself. This is not so unreasonable. As Holländer [1990, 1161] puts it: "we generally approve of cooperative behavior even if it does not make us significantly better off."

With the foregoing definitions, I obtain the following result:

Proposition 2 (Adequacy): *With a preference for social approval, player 1 has an incentive to reduce the adequacy of his gift* ($du_1/d\delta_1 < 0$) *if the degree of status-orientation* (α)*is sufficiently high and the gesture is relatively important* ($\partial\beta_1/\partial\delta_1$ *low).*

Proof. Player 1's utility is given by: $u_1 = u_1(l_1^v, f_2(l_1^v, \delta_1), \delta_1)$. The total derivative is given by: $du_1 = \left[\frac{\partial u_1}{\partial l_1^v} + \frac{\partial u_1}{\partial f_2}\frac{\partial f_2}{\partial l_1^v}\right] dl_1 + \left[\frac{\partial u_1}{\partial \delta_1} + \frac{\partial u_1}{\partial f_2}\frac{\partial f_2}{\partial \delta_1}\right] d\delta_1$. The term in the first brackets is zero by the first order condition. Hence: $\frac{du_1}{d\delta_1} = \left[\frac{\partial u_1}{\partial \delta_1} + \frac{\partial u_1}{\partial f_2}\frac{\partial f_2}{\partial \delta_1}\right]\Big|_{l_1^{v*}}$. The first term in brackets is the direct effect and is positive $\partial u_1/\partial\delta_1 = (\partial\beta_1/\partial\delta_1) l_1 u_s'$: a higher δ_1 increases net approval at the rate $(\partial\beta_1/\partial\delta_1) l_1 u_s'$. This term vanishes as $\partial\beta_1/\partial\delta_1$ becomes smaller. The second term in brackets in the indirect effect due to the response of player 2. Since by Lemma 1 $\partial f_2/\partial\delta_1 > 0$, the sign of the indirect effect is given by the sign of $\frac{\partial u_1}{\partial f_2} = \delta_2 u_x' - \alpha\beta_2 u_s'$. Using the FOC and the additional assumption that $w_1^m > \delta_2 f_2'$ it is easy to show that the sign is negative for $\alpha > (\beta_1/\beta_2)\cdot(\delta_2/w_1^m)$. Hence, the total effect is negative if $\partial\beta_1/\partial\delta_1$ is low and α is sufficiently large. The assumption that $w_1^m > \delta_2 f_2'$ merely states that gifts require a sacrifice and are made to gain approval, not to increase consumption. This seems reasonable when one speaks of social approval as a motivation to give. ■

Let me reflect for one moment on the results so far. In the introduction to this chapter, I have described the gift-giving ceremony of the Kwakiutl: the Potlatch. Here, I would like to argue that the need for social approval and status may be an accurate description of their gift-giving practices. Many of Mauss' observations which often clearly point to status concerns, including his observations that "Face is lost forever if [a return gift] is not made" (Mauss [1926, 41] and that large gifts are made by men "in order to outdo their rival" (Mauss [1926, 6][9]. Also congruent with a demand for social approval and status is his mentioning of the fact that gifts are often ostentatious. Interestingly, we also saw that the Potlatch is characterized by an extreme degree of inadequate gift-giving, in line with proposition 2. According to Godelier [1996, 56], the goal of this is explicitly to "make it difficult or impossible to give back the equivalent". The broken copper is seen as a victory much like as a killing of a rival would have been. It is a

[9] It is in this respect also interesting to note that Mauss [1926] traces the origin of the word "gift" back and finds that is has the same roots as poison.

fight for prestige, a war that is fought with property as weapons, to paraphrase Codere [1950]. The last thing you want is to provide your enemy with weapons. The destruction of copper does exactly prevent that.

Efficiency

The incentive to give inadequate gifts has warranted researchers to conclude that gift-giving creates a deadweight loss. Subsequent calculations accredited to Christmas gifts a deadweight loss in the range of 10-30% of their value (Waldfogel [1993]). Yet, one cannot conclude from inadequacy itself that gift-giving creates an inefficiency, in the sense that everybody could be made better off. In the research done by Waldfogel [1993], and also in the follow-up studies by Solnick and Hemenway [1996] and Waldfogel [1996], it is tacitly assumed that the giver is indifferent between giving in cash and in kind. However, as found in questionnaires (see e.g. Webley and Wilson [1989]) and in line with Proposition 2, givers are not at all indifferent but clearly prefer to give in kind. Hence, inadequacy does not imply inefficiency. However, the following proposition nevertheless shows that in equilibrium gift-giving is inefficient.

Proposition 3 (Efficiency): *There exists a feasible distribution of endowments that is a Pareto-improvement upon the gift-giving equilibrium.*

Proof. The proof follows in a fairly straightforward way from proposition 1. Note that the reaction function f_2 of player 2 is an upward sloping curve in the (l_2^v, l_1^v) plane. Recall that this curve is the set of points where player 2's utility is maximized for any given l_1^v. Thus, the line f_2 cuts each of player 2s indifference curve at its maximum. Hence, the indifference curve has slope zero everywhere along f_2. Player 1 chooses a point on this curve that maximizes his utility. This is necessarily a point where the slope of his indifference curve is tangent to the reaction function of player 2. Hence, in equilibrium the slope of the indifference curve is positive for player 1. Since the slopes of the indifference curves differ between the players, there exists a point where both could be made better off. ■

The cause of this inefficiency is twofold. First, the preference for status introduces an inefficiency. This is clear: both players put resources in the race for status, but only the net result counts. This is most clearly seen for the case where $\alpha = 1$. Would both players spend the same amount on gifts, none of them would derive

utility from it with regard to social approval, since net social approval would equal zero. Second, player 1 acts strategically by taking the optimal response of player 2 into account. He does not, however, take into account the effects on player 2s welfare. To derive the social optimum, he would have to take into account the welfare of both players.

3.2.3 Simultaneous move equilibrium

So far, the focus has been on the sequential move gift-giving game. The reason is that the inclusion of time is necessary to have a meaningful interpretation of reciprocity. After all, reciprocation is a notion of giving *back*. But in an important class of gift-giving, namely charity, gift-giving occurs anonymously and therefore necessarily without any countergift.

Quite obviously does the absence of countergifts not preclude gift-giving. The difficulty therefore lies in the anonymous aspect of charity. Can approval be obtained in such a setting? The answer, perhaps surprisingly, is sometimes yes. Status is often awarded on the basis of events that do not directly affect the person who awards status (Coleman [1990, 130]). The receiver may very well be anonymous, but the giver can make his act known to other persons than the receiver. This is most clearly described by Schwartz [1967, 2] who states that "it is common knowledge that men present themselves publicly by the conspicuous presentation of gifts. Generous contributions to charity have always been a source of prestige in the United States." This is acknowledged by fund raisers by providing people with "I gave" stickers to be affixed to the front door (Schwartz [1967]) or coffee mugs, lapel pins, and bumper stickers and so on (Andreoni and Petrie [2000]). In this way, donors can signal their contributions.

Clearly, if approval is indeed the motivation of charity then behavior should depend heavily on whether or not the possibility exists to make the act of giving publicly known. In a clever designed experiment Andreoni and Petrie [2000] have tested for this. They found that, if faced with the opportunity, virtually all subjects that donated choose to be publicly known as having done so. Moreover, average contributions in last rounds are more than twice as high as in anonymous settings. Satow [1975] has found a similar effect in an experiment where subjects could voluntarily donate part of their earnings to a research fund. Not only did they donate significantly more to the research fund when they were observed by the experimentator (public condition) than when they were unobserved (private

condition), the amount donated was also significantly related to their need for approval, as measured by a 'social desirability scale'. Subjects with a high need for approval donated significantly more than those with a low need for social approval. Furthermore, there was also a significant interaction effect: under the public condition the effect of a high need for approval was greater than under the private conditions. The main results are reproduced in the table below. Unfortunately, I am not aware of any other experimental study that relates gift-giving to the need for approval, so the evidence remains circumstantial.

TABLE 3.1

Mean Percentage of Earnings Donated to the Research Fund

Need for approval	Private condition	Public condition
Low need for approval	4.45	17.27
High need for approval	3.11	38.71

Source: Satow [1975].

As a last piece of evidence, I report an empirical study by Harbaugh [1998]. He found that when charities report names of contributors in categories rather than in exact amounts, a large proportion of donations is at the boundaries of those categories. Social approval as a motivation predicts this: it is impossible for the public to discriminate between different donations within each category. The donor can therefore as well contribute the lowest amount within a certain category.[10]

3.2.4 Gifts and markets

Thus far the focus has been on the properties of gift-giving. In this subsection these findings are related to another exchange mechanism: the market.

Recall once again that gifts in kind are usually taken to be a social waste because the recipient could have achieved a higher utility level from an equally costly cash-gift. Problematic in this respect is that in the environment of gift-giving, individual rationality sometimes prescribes to give in kind rather than in cash (see

[10] Though a reasonable alternative explanation is that framing effects play parts. Note also that the significance of this evidence is undecisive as there may be many more institutions that do not publish names.

proposition 2). The market mechanism does not suffer from this tendency. This has warranted researchers to conclude that, compared to the market, gift-giving creates a deadweight loss. Subsequent calculations accredited to Christmas gifts a deadweight loss in the range of 10-30% of their value (Waldfogel [1993]).

In the research done by Waldfogel [1993] and in the follow-up studies (Solnick and Hemenway [1996] and Waldfogel [1996]) it is tacitly assumed that the giver is indifferent between giving in cash and in kind. However, as found in questionnaires (see e.g. Webley and Wilson [1989]) and in line with Proposition 2, givers are not at all indifferent but clearly prefer to give in kind. This has consequences for the measurement of the deadweight loss. True, turning to the market prevents inadequacy. But insofar as the market is an anonymous institution, it does not realize this valuable psychological sentiment called approval. It is for that reason not obvious that gift-giving creates a deadweight loss. The trade-off to be made is that between adequacy and approval. In neglecting the latter part of this trade-off one tends to overestimate the size of the deadweight loss, if there is any.

Of equal interest is how the market institution interacts with gift-giving. There are instructive examples that show how the opening of a market leads to a crowding-out of gift-giving (Hirsch [1976], Yellen [1990]). There are hints that this crowding-out is detrimental to social welfare. The keypoint is that the links between the market institution and gift-giving are shot through with externalities (see in particular Kranton [1996a] and Dasgupta [2000], and chapter 4). As Hirsch [1976, 78] puts it, "social relationships do not, by their nature, have the character of private economic goods." Indeed, when the !Kung tribe gained access to markets which were superior in the supply of goods, they abolished gift-giving and at the same time retreated from their social life, eroding the cohesion of society (Yellen [1990]).

Yet, there is also another side to the coin. Where the market is an efficient institution, gift-giving can be a hindrance. Kranich [1994] points out that when gifts are permitted, an efficient equilibrium need not be reached. For instance, gift-giving acquires the character of a public good in case agents have preferences over the economy's entire allocation of income rather than just their own income. As another example, recall that due to moral hazard the market does not provide full insurance. Nonmarket insurers (such as a family) may therefore want to supplement the market. Arnott and Stiglitz [1991] then show that when these nonmarket insurers have no better information than market insurers, reciprocal

assistance not only crowds out market insurance but is also harmful and therefore dysfunctional.

On the more positive side can gift-giving be a useful complement to the market institution in case the latter fails. Consider the supply of a public good. The market creates no incentives to contribute at all because individually it does not pay off in terms of consumption. Social approval at least gives *some* incentives, although there is nothing that guarantees that gift-giving for the sake of obtaining approval continues up to the point where the supply is optimal. Gift-giving can complement the market to a certain extent, though additional incentives are likely to be needed in order to reach the optimal amount of gifts.

In providing these additional incentives, it is probably not unfair to say that the effects on social approval are usually neglected when it comes to institutions being recipients. The right way to stimulate giving then seems obvious: compensate donors directly for the efforts. This is exactly the measure that was expected to alleviate the shortage of blood donations. However, quite unexpectedly to many, the actual change in the supply of blood turned out to be negative (see Titmuss [1970]). Other studies that show this pattern emerges in different settings as well include Gneezy and Rustichini [2000] and Frey [1997a], [1997b] (see also chapter 7). Yet, this kind of behavior is only surprising insofar as the object of human behavior is material gain. A higher compensation makes you richer and it makes giving relatively cheaper. For that reason one should indeed, if anything, increase one's gift. On the other hand is there the effect on social approval. Rewarding gift-giving reduces sacrifice, and as argued before, there is evidence that sacrifice is positively correlated with social approval. Hence, rewarding gift-giving deprives the agent from the opportunity to realize social approval. This latter effect reduces the incentives to give and makes the end outcome ambiguous.

Consider therefore the case where the receiver does not make a countergift but gives a monetary compensation instead. Thus, set $l_2^v = 0$ and denote by w^v the compensation the institution offers for each unit of voluntary labor. The forgone wage of giving is now $\tilde{w}_i \equiv w_i^m - w^v > 0$. I say that sacrifice is relatively important if the approval rate is sensitive to the sacrifice made: i.e. $\partial\beta_i/\partial\tilde{w}_i$ is high.

Proposition 4 (Compensation): *With a preference for social approval, the optimal gift size of player 1 is decreasing in compensation* ($dl_1^v/dw^v < 0$) *if the*

absolute elasticity of u'_s is less than unity and if sacrifice is relatively important ($\partial\beta_1/\partial\tilde{w}_1$ high).

Corollary 1 *If compensation crowds out gift-giving then welfare is reduced: $\partial u_1/\partial w^v < 0$ if $dl_1^v/dw^v < 0$.*

Proof. With $l_2^v = 0$, player 1's utility function becomes (ignoring fixed parameters): $u_1 = u_1(l_1^v(w^v), w^v)$. Differentiating at the optimum with respect to w^v gives: sign of $\frac{dl_1^v}{dw^v}$ is equal to the sign of $\frac{\partial^2 u_1}{\partial l_1^v \partial w^v} = u'_x \frac{\partial^2 x_1}{\partial l_1^v \partial w^v} + u''_x \frac{\partial x_1}{\partial w^v}\frac{\partial x_1}{\partial l_1^v} + u'_s \frac{\partial^2 s_1}{\partial l_1^v \partial w^v} + u''_s \frac{\partial s_1}{\partial w^v}\frac{\partial s_1}{\partial l_1^v}$. The first two effects are positive. Since $\frac{\partial^2 s_1}{\partial l_1^v \partial w^v} = -\frac{\partial \beta_1}{\partial \tilde{w}_1}$ the last two effects can be combined into $-\frac{\partial \beta_1}{\partial \tilde{w}_1}(u'_s + \beta_1 l_1 u''_s)$. Clearly, $\frac{\partial^2 u_1}{\partial l_1^v \partial w^v} < 0$ requires that $u'_s + \beta_1 l_1 u''_s > 0$ (which is true for an absolute elasticity of u'_s smaller than one), and that $\frac{\partial \beta_1}{\partial \tilde{w}_1} > \xi \equiv \left[u'_x - \tilde{w}_1 l_1 u''_x\right] / \left[u'_s + \beta_1 l_1 u''_s\right]$. This proves Proposition 4. To prove Corollary 1, note that $du_1(l_1^v(w^v), w^v)/dw^v$
$= (\partial u_1/\partial l_1^v)\cdot(dl_1^v/dw^v) + (\partial u_1/\partial w^v) = \partial u_1/\partial w^v|_{l_1^{v*}}$. Hence, at the optimum, utility decreases in compensation if $\partial u_1/\partial w^v = l_1^v u'_x - (\partial\beta_1/\partial\tilde{w}_1) l_1^v u'_s < 0 \Leftrightarrow \partial\beta_1/\partial\tilde{w}_1 > u'_x/u'_s$. With the above definition of ξ we have $\frac{\partial \beta_1}{\partial \tilde{w}_1} > \xi > u'_x/u'_s$ so that if beta is high enough such that compensation reduces gift-giving, then compensation also reduces welfare. ■

The intuition is again straightforward. If sacrifice is relatively important then every unit of voluntary labor is valued a lot less when sacrifice is a little bit lower. This reduces the incentives to give if compensation is offered. There is also a positive effect on social approval since a lower β reduces s and this increases marginal utility of giving. But this latter effect is always dominated by the former effect as long as the absolute elasticity of u'_s is smaller than one. Furthermore, by Corollary 1 the deprivation of social approval is in this case enough to reduce utility.

A result similar to proposition 4 is obtained by Holländer [1990] who argues that the *opening* of a market reduces incentives to give to a public good. If the market (or government) provides some of the public good, then the marginal benefits of the public good declines and under his specification this means that less approval is obtained for each contribution.

It should be emphasized that crowding-out is not a general phenomenon, but Frey and Jegen [2002] conclude that there exists compelling empirical evidence that it does occur on some occasions. To appreciate the significance of this to

its fullest extent, note that gifts in the form of bequests can affect the long-run growth of an economy (see Galor and Zeira [1993]). The insight gained is therefore of importance, so much the more Frey [1997b] argues that in addition there are spill-over effects.[11] For example, spill-over effects to other activities or spill-overs over time may occur. Thus, the social approval that can be obtained for giving this particular good can influence the approval for giving some other good. Or, not unreasonably, social approval is time dependent.[12] Giving an inadequate gift reduces social approval now, but persistently giving inadequate gifts reduces social approval even more in the future. Albeit these are interesting extensions of the model, it would be too speculative to elaborate upon them at the present state. More should be known about the exact determinants and shape of the social approval function, as well as its dynamics. Nonetheless, the conclusion can be drawn that one should be very careful with the implementation of market incentives in non-market institutions. This can result in unintended consequences that are harmful to social welfare.

3.3 Related literature

The idea of social approval seems intuitively plausible. It is therefore no surprise that social approval takes a prominent place in other social sciences. This makes it all the more remarkable that the economics literature is mostly silent on this. The works that are most closely related are that of Holländer [1990], Frey [1997a], [1997b] and Andreoni [1990]. But there are some notable differences. Holländer relies on the seemingly unintuitive assumption that sacrifice is negatively correlated with approval. Compensation would then increase gift-giving, contrary to many empirical findings. Frey [1997a], [1997b] and Andreoni [1990] use other emotive concepts that are very close to approval, namely *intrinsic motivation* and *warm glow* respectively. Both concepts attribute the act of gift-giving to the fact that people derive utility from the act of giving *per se*. To a certain extent these concepts are compatible with the current framework.[13] The advantage of the cur-

[11] The approach of Frey [1997a], [1997b] to explain crowding-out is taken up in Section 3.3.

[12] Some of the evidence reported in Frey [1997b] as well as in Gneezy and Rustichini [2000] indicates persistence over time.

[13] One difference is that the concepts of intrinsic motivation and warm glow feeling also work in a strictly anonymous setting whereas social approval does not. But see the discussion in Section 2.3 where it is argued that even charity is often not truly anonymous.

rent framework is that an attempt is made to further uncover the exact channel through which the intrinsic motivation or a warm glow feeling is obtained. This leads to a better insight in what the properties of gift-giving and the gift itself will be. For example, if social approval is indeed the real source of the warm glow, then the assumption of Andreoni [1990] that the warm glow is independent of the subsidy is doubtful since a subsidy reduces sacrifice. This doubt is strengthened by the fact that this assumption rules out a negative relation between subsidy and gift-giving, contrary to the empirical findings. Neither do the theories of Andreoni [1990] and Frey [1997a], [1997b] in their current state have anything to say about the optimality of the adequacy of the gifts even though this property is of critical importance in making welfare judgements. Central in the theory of Frey [1997a], [1997b] is that intrinsic motivation is crowded out by extrinsic motivation. Although this may be costly, the extrinsic motivation perhaps results in a higher level of adequacy.

There is by now a fast-growing literature that tries to explain the properties of gift-giving through other causes than social approval, see in particular chapter 2. These include altruism, signaling, and fairness. Alike the current approach, each of these alternative theories has its own deficiencies. For example, if altruism is the motivation to give as suggested by Becker [1974], then the higher the utility of the recipient the better, so that inadequate gifts cannot be explained. Neither can crowding-out be explained, since a compensation would make gifts less effective so more would be given to obtain the desired distribution. On the other hand, with social approval compensation creates costs to the giver since he is deprived from some approval and this cannot be costlessly compensated by giving more.

In case information is asymmetric, gifts can also act as a signal. Exemplary in this respect is the work by Camerer [1988]. In his model, honest players can make gifts to signal themselves as being trustworthy. Cheaters do not find it profitable to make gifts. As a corollary, this explains the inadequacy of gifts, because with adequate gifts cheaters would participate in the gift-giving game anticipating that they will collect valuable gifts. However, a compensation would increase gift-giving since it would make current gifts less effective as signals. More should be given to reveal one's type.

The fairness approach has probably gotten most attention in the literature (see especially Fehr and Schmidt [1999], and Bolton and Ockenfels [2000]). It successfully accounts for a broad range of experimental games by assuming that people

not only care about their own monetary payoffs, but also about the distribution of payoffs and the intentions that other players have. The theory is especially productive in explaining reciprocity: just think of the saying "an eye for an eye, a tooth for a tooth." Inadequacy seems not to be easily explicable: if one gives to make the distribution more equal, then the most effective way to do this is to give as adequately as possible. People would therefore prefer to give adequate gifts and only give inadequately if no better alternative is at someone's disposal, e.g. if the only choice they face is to give inadequately or not to give at all. For similar reasons should a compensation increase gift-giving: more should be given to achieve the desired equality in the income distribution.

As it stands, no single theory can explain everything. The current approach should therefore be considered as being complementary to the existing literature, not as an alternative theory. Future research will aim at synthesizing the different theories, for example by arguing that people are intrinsically motivated by principles of fairness and at the same time also want to be seen as such, that is, to be approved by others for being fair.

3.4 Concluding remarks

The purpose of this chapter is to argue that at least part of gift-giving can be sensibly understood as the result of a desire for social approval. At the cost of a loss in predictive power, the analysis is kept rather general. The motivation for that is simply because little is known about the exact shape of the approval function. However, under weak but nevertheless (at least in my view) reasonable assumptions I have been able to derive some qualitative properties that I think are not sensitive to the specifications of the model.

The foregoing shows that social approval may indeed play a role in people's behavior. This is not to say that social approval is the unique motivation next to one's own payoff or that social approval plays a role in every kind of exchange. Indeed, it is very probable *not* the unique motivation. The gift-giving behavior found in truly anonymous games cannot be attributed to a taste for social approval. But insofar as it does explain part of the behavior, its implications are far reaching. Wherever social approval is important the superiority of the market institution to gift-giving must be reconsidered, despite the fact that gift-giving elicits inadequate gift exchanges. And where the market is considered to be su-

perior but does not arise from spontaneous order, the incentive schemes based on price incentives as designed by typical economic arguments should be rethought so as to take into account the effects they have on social approval. A better solution might be to make the gifts more visible to outsiders, rather than to give monetary compensations.

Prior to that, more research is needed to assess when social approval is a main driving spirit, what the exact determinants are, and furthermore to substantiate the role of inadequacy of gifts as a strategic variable. Perhaps a promising direction to go in is to design experiments that not only reveal the identity or actions of players, but in addition communicate more precise information on the approval of players. Because the propositions are stated in terms of how social approval is affected by different actions, information on the approval allows the propositions to be more specifically tested. More concretely, one possibility is to include an index of how each player approves the gift. This allows players to reveal their appreciation of the gift. It is then possible to see how the other player responds to changes in that index. If this index is designed carefully, the propositions in this chapter lend themselves to being tested. Proposition 2 can be tested by making the adequacy of the donation in a public goods games a choice variable. This is in the spirit of Goeree, Holt, and Laury [2002]. They test the effects of changing the *external rate* of the contribution. The external rate is the rate at which other players profit from your gift, the adequacy level in other words. In their experiments the external rate is a given variable for the subjects. Finally, Proposition 4 can be tested by examining whether the existence of crowding out is related to the approval obtained in a way corresponding to the conditions stated.

4

On the Viability of Gift-exchange in a Market Environment

In the harsh Kalahari Desert in Africa, a group of people known as the !Kung lived for a long time by the principles of reciprocity and sharing. These principles provided enough security not to live on the brink of starvation, despite the apparent difficult circumstances. Then, the access to the market caused an influx of money and goods. This was partly responsible for the eroding traditional values and social cohesion. To which extent can we expect that, over time, access to the market crowds out reciprocity?

4.1 Introduction

Reciprocal exchange in its pure form can be observed in special places where the market is not strong enough to break personal connections. On the other hand, there are fascinating stories by anthropologists showing how reciprocal exchange arrangements vanish when tribes encounter markets. This chapter studies the relationship between market exchange and reciprocal exchange.

In an interesting paper Kranton [1996a] shows that in order to become beneficial, markets need enough participants to reduce search costs. Therefore, reciprocal exchange may survive if initially the proportion of the people that engage

[0]This chapter is coauthored by Theo van de Klundert. We are indebted to Jan Boone, Lans Bovenberg, Patrick François, Richard Nahuis, and Sjak Smulders for useful comments on an earlier version of the paper.

in market exchange is not too large. There are other determinants which may tip over the balance in favor of market exchange, such as a lack of trust. In reciprocal exchange people have to trust each other, because production and consumption are separated over time. Moreover, reciprocal exchange may involve a less elaborate division of labor.

Kranton [1996a] borrows evidence documented by Yellen [1990] that describes how the !Kung tribe abandoned reciprocal exchange once they encountered the market economy of Botswana. But reciprocal exchange is not limited to tribal communities nor to some corners of the economy but is an element of importance in developed market economies as well. Reciprocity or gift exchange is an essential aspect of culture. There are social and moral dimensions to economics so to say (see for instance Etzioni [1988]). Even Adam Smith already knew that "moral sentiments" are important, something also recognized by Kenneth Arrow : "ethical behavior can be regarded as a socially desirable institution which facilitates the achievement of economic efficiency in a broad sense" (Arrow [1972, 354]). People have a sense of belonging to society at large, and care for social interactions. This induces cooperative behavior in different guises. Self-interest seeking behavior is a *sine qua non* for coordination in the economy, but there are limits to opportunistic behavior. People want to be respected by others. To a certain extent respect follows from success in the accumulation of wealth, but there are limits to respectfulness in this sense. Mutual aid and sympathy are important values of their own. For this reason producers may take pride in the quality of the product they deliver, workers may be motivated to do a good job, and people in general may take account of each other's interests in different situations. Experiments in economic settings indeed confirm that people may behave different from what standard neoclassical economics predicts. Take for instance, Gächter and Fehr [1999] who find that social approval and social familiarity generate a significant rise in cooperative behavior (see also chapter 3).

Casual observation learns that people are concerned about a loss of commitment in recent times. This is often associated with commodification, meaning that markets are expanding into almost every territory of human life. In our interpretation this means that the market system crowds out reciprocal exchange, bringing about a lack of close personal relationships. The questions to be answered are then the following. What are the main factors causing such a crowding-out? Is reciprocal exchange in the sense of moral behavior completely wiped out or are

there conditions such that both regimes can coexist in a long-run equilibrium? To answer these questions we apply the analytical framework based on Kranton [1996a] and Diamond [1982] in a modified way, and with a different interpretation. In the model used, people can make a deliberate choice between two regimes: a reciprocal relationship or market exchange. In reciprocal exchange agents value social behavior as they have a sense of belonging to society at large. Applying the terminology introduced by Khalil [1997], it can be argued that agents produce goods from which they derive substantive utility, but sticking to the moral codes of the group provides also satisfaction in the form of symbolic utility. Initially we assume that the terms of trade between substantive and symbolic utility are deteriorating over time. Extending the model later on, it will be argued that the terms of trade typically depend in a non-linear way on the fraction of agents engaged in reciprocal exchange.

In the regime of market exchange agents behave as prescribed in neoclassical theory. If the fraction of agents engaged in pure market activities rises, the market operates more efficiently. People are then less hampered by tradition and can seize every opportunity for making a profit. The division of labor can be exploited more fully. The set of goods produced generally differs from that in case of reciprocal exchange. This has an impact on substantive utility. The idea that market efficiency is related to the number of people in the market is characteristic for search models. We generalize this idea by assuming that more encounters between people induce more production by leveling organizational and institutional restrictions. However, both ideas can be modelled in the same manner.

The model applied here differs in a number of aspects from that in Kranton [1996a]. First, in modelling the search externality we follow the original set-up of Diamond [1982], [1984] more closely. Second, to simplify further we assume that goods are produced at constant cost instead of introducing a distribution from which agents draw randomly. Finally, it should be observed that we not only obtain corner solutions as in Kranton. Assuming that the terms of trade between regimes depend on the fraction of agents engaged in the market system we find interior solutions with agents operating under different regimes.

The rest of this Chapter is organized as follows. In section 4.2.1 we first specify what is meant by sympathy and substantive utility. This sets the stage for modelling reciprocal exchange in section 4.2.2. Market exchange is discussed in section 4.2.3. In section 4.3 we consider equilibrium solutions under different assumptions

with respect to the shape of symbolic utility. Complete commodification obtains in section 4.3.1. Interior solutions with partial commodification are considered in section 4.3.2. The role of discounting is scrutinized in section 4.3.3. Welfare considerations are taken up in section 4.4. The chapter closes with concluding remarks in section 4.5. Proofs of propositions are deferred to the appendix.

4.2 Exchange mechanisms

In this section we describe the formal model. The general setup is one in which each agents starts in a situation in which he is involved in a personal relationship with one other agent. A key feature of such a reciprocal exchange relationship is the element of trust. Production and consumption are typically separated over time. Contrary to the market where money serves as a medium of exchange, no such security exists in a reciprocal relationship. It is therefore possible that agents who did produce last period see their relationship end without having the possibility of consuming in return. Although this favors the existence of markets, market exchange has disadvantages of its own. The market is typically characterized by anonymous agents, without relation-specific commitments, and search costs have to be made in order to find a trading partner. On the other hand is it generally acknowledged that the market is capable of supplying a larger array of goods.

Before turning to a detailed exposition we present an outline in figure 4.1. This facilitates the reading of the subsequent sections. The arrows denote possible flows. Agents start in a reciprocal exchange relationship. They can end their relationship and enter the market as unemployed. The distribution of agents in each period is given by the fractions r, u, and m that correspond to the different states as in figure 4.1. The market is characterized by a search process that describes how agents become employed with a good, sell that good and become unemployed again. This search process is determined by the technical parameters a and b to be explained later on.

We start with an exposition of reciprocal exchange and postpone a treatment of the market part to the next section. First however, we discuss the agents' preferences in more detail.

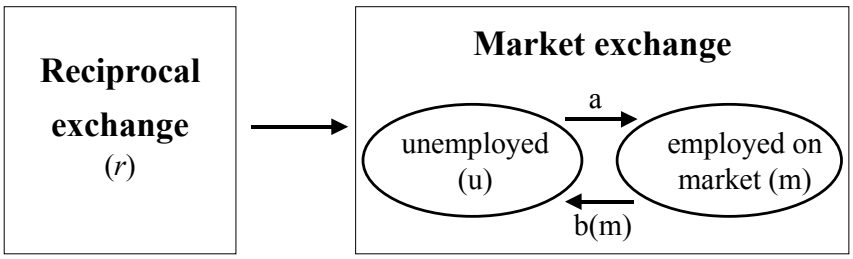

FIGURE 4.1.

4.2.1 Substantive and symbolic utility

Following Khalil [1997] we distinguish between *substantive* utility (ordinary tastes for material goods) and *symbolic* utility (tastes for selfhood and alike). Consumers derive substantive utility from a basket of goods, which may differ across regimes. Substantive utility for each representative consumer is denoted by x_i, where the index $i = r, m$ denotes reciprocal or market exchange. Symbolic utility can be taken into account by introducing extended preferences (see for instance Becker [1996]). Here we take a short-cut by introducing a mark-up (θ_i) on substantive utility. Extended utility is then defined as:

$$y_i = \theta_i x_i, \ \theta_i \geq 1, \ i = r, m. \tag{4.1}$$

It is assumed that under market exchange extended utility coincides with substantive utility ($\theta_m = 1$). Reciprocal exchange conveys symbolic utility as people value social interaction as such ($\theta_r > 1$). The above set of equations can be combined to:

$$y_r = \theta y_m, \tag{4.2}$$

where $\theta \equiv (\theta_r x_r)/x_m$. As substantive utility is fixed x_r and x_m can be treated as constants. It seems plausible to assume that θ_r, and therefore θ, depends on the

fraction of people engaged in market exchange, m. If the number of people in gift exchange declines it will be harder to uphold a sense of community spirit. As the market regimes expands, individualism spreads and solidarity may become less attractive. These considerations lead to a negative relation between the valuation ratio θ and the fraction of people in the market regime. However, according to Adam Smith it can be maintained that the need for mutual sympathy and respect is deep-rooted and cannot be suppressed entirely. A similar view based on biological principles is expressed in Kropotkin [1904]. It is not unreasonable to assume that the ramification of this becomes more distressing as the market gets to dominate exchange relations. This could imply a positive relation between θ and the market size after some threshold level of market participation has been passed. Thus we have that $\theta = \theta(m)$ is first decreasing and after some critical point, say m', increasing, because the market becomes "too large".

More specifically, the U-shaped $\theta(m)$-curve can be seen as the result of several opposite forces. For instance, as more people leave the regime of reciprocity the cost (in disutility terms) of changing beliefs for the remaining people decline. This idea builds on the literature on cognitive dissonance in psychology. Following Festinger [1957] it can be stated that changing beliefs induces a negative arousal. Moreover, the resistance to change crucially depends on the difficulty of finding people who support the new belief. Therefore, the more people are already in the market regime the easier it becomes for people to switch regimes. It is in particular hard for the first few individuals that are to switch. This is shown by the dashed downward sloping RC-curve (resistance to change) in figure 4.2 .

Furthermore, as the market system expands the social deficit becomes more important. As argued in Bowles [1998] markets are characterized by impersonality and ephemerality of contact. But people also want to socialize. There is a need for mutual sympathy and recognition. If reciprocal exchange is relatively large market participants may have the feeling that the social deficit can be easily repaired. The more people are available as potential candidates to socialize with, the easier it is to change back to reciprocal exchange. Therefore, the opportunity costs (in disutility terms) of switching to the market regime increase as the market system becomes dominant. The social deficit is felt more heavily for a larger market share. This gives rise to the dashed upward sloping SD-curve (social deficit) in figure 4.2 . Under appropriate conditions the summation of both curves may lead to an U-shaped function $\theta(m)$ as illustrated by the bold curve in figure 4.2. To analyze

different possibilities, the equilibrium solution discussed in section 4.3.1 will be based on a monotonic decreasing θ-curve. This case eventuates in the corners as the only solutions. The U-shaped θ-curve will be introduced in section 4.3.2. As it turns out, this opens the possibility of interior solutions.

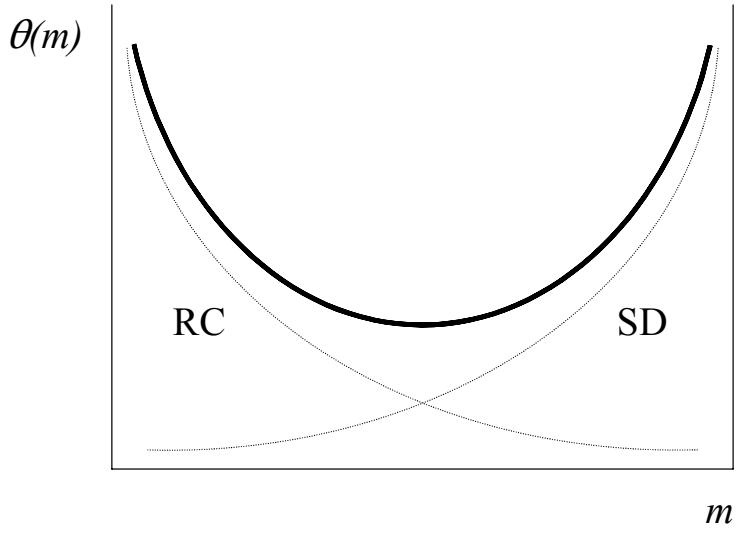

FIGURE 4.2.

It should be noted that there is a close parallel with the descriptive approach of Fukuyama [1957]. In his latest book Fukuyama describes "the great disruption" of the social system starting somewhere in the sixties under influence of the upcoming information technology. The change of a traditional industrial society towards an economy dominated by the service sector leads to a certain disorientation and as a consequence of this to a decay of moral values. But according to the author things have changed lately. The reason is that human nature is geared towards cooperation and reciprocity. Fukuyama bases his view on the biological approach going back to Kropotkin [1904]. The reconstruction of the social order is reflected in declining criminality statistics and a more positive evaluation of social relations in systematic surveys. Such a change may lead to a new equilibrium with a certain amount of gift exchange in the current market economy. In our model it is the upward sloping branch of the $\theta(m)$ curve which reflects the

views of Fukuyama and others on the viability of the moral system in modern times.

4.2.2 Reciprocal exchange

The timing in a reciprocal exchange relationship is as follows. Each two periods a complete reciprocal gift exchange can be accomplished. In the first period, one of the agents produces first and the other consumes the good. In the second period they switch roles. Whenever they consume, they derive utility y_r, and whenever they produce they bear a cost in disutility terms of c[1]. The discount factor for the next period equals $\delta = e^{-\rho}$, where ρ is the subjective discount rate[2]. Agents are infinitely lived.

Let V_{rp} and V_{rc} denote the lifetime discounted utility of the agent that is involved in reciprocal exchange and starts as producer and consumer respectively. The agent starting as a producer incurs a cost (c) and expects to be in the position of a consumer next period:

$$V_{rp} = -c + \delta V_{rc}. \tag{4.3}$$

Similarly, the lifetime discounted utility of the agent that starts as consumer, V_{rc}, is given by:

$$V_{rc} = y_r + \delta V_{rp}. \tag{4.4}$$

It is further assumed that each agent ends up being first consumer or producer with equal probability. Expected discounted lifetime utility of staying in a reciprocal exchange relationship is then given by $V_r = \frac{1}{2}(V_{rp} + V_{rc})$ or, after substitution and rearranging terms:

$$V_r = \frac{y_r - c}{2(1-\delta)}. \tag{4.5}$$

The element of trust is introduced by the possibility of ending the relationship. An agent can decide to consume first, but not to produce in return. He then consumes and enters the market. Whenever he does, the other agent will of course take notice of being cheated and as a punishment he will end the relationship[3]. As

[1] For simplicity we assume that the complete production of an individual is exchanged. A more realistic extension would be that individuals exchange only part of their production.

[2] A natural restriction on the discount factor is that $\delta \in (0,1)$: future revenues are valued positively but less than current revenues.

[3] This tit-for-tat is only one of many possible strategies. The strategy is common in the microeconomic literature. The classic defence is given in Axelrod [1984]. More interestingly, the strategy is defended by Aristotle on principles of justice: "Now if proportionate equality between the products be first established, (...) then

a consequence, both agents will have to enter the market. Obviously, the agent will deliberate upon producing in order to sustain the relationship or to cheat upon his partner. Whether the agents decide to cheat and enter the market or stay in the reciprocal exchange relationship is dependent on the derived utility of being in the market regime, V_u (to be specified later on), and of the derived utility in the reciprocal exchange relationship, V_r. The following definition shows a Nash-equilibrium constraint under which both agents will decide not to cheat (a derivation is given in the appendix).

Definition 1 *(enforceability) A reciprocal exchange relationship is enforceable if* $-c + \delta V_r(\cdot) \geq V_u(\cdot)$.

The constraint is more likely to be satisfied when the discount factor is high (low subjective discount rate) or when the market size is small. The higher the future is valued, the less beneficial it is to cheat and so reciprocal exchange is more easily enforceable. Alternatively, we can interpret the subjective discount factor as a measure of trust. If the discount factor is low, then the faith in getting consumption in return is poor. The lower the discount factor, the harder it is to uphold the relationship. The market, on the other hand, is characterized by increasing returns to scale. The larger the market, the more easy it is to find a partner to trade with. This implies that the value of trading on the market is positively dependent on the market size[4]. The enforceability constraint is therefore harder to satisfy at larger market sizes. Detailed comparative statics are provided in section 3.3. We now turn our attention to the determination of the value of entering the market.

4.2.3 Market exchange

Agents that have decided to enter the market do not have a fixed trading partner. They enter the market "unemployed", that is with no goods, and have to search for production possibilities. This will be represented by a Poisson process with arrival rate a. They can either accept or not accept the production opportunity. After they have found and accepted one they bear the same disutility costs c as in reciprocal exchange. Being employed they still have to search someone to trade

reciprocation take place (...) but if this is not done, the bargain is not equal, and intercourse does not continue." (Aristotle [1994, 283]).

[4] From (4.9) introduced later in the text, it is easily proved that V_u is indeed (weakly) increasing in m.

with. They find someone with probability $b(m)$. Having found a trade partner they exchange and derive utility y_m.[5] The market is characterized by increasing returns to scale in the form of an externality. As the fraction of the population on the market, m, gets larger, average search time decreases or equivalently $b^{'}(m) > 0$. Recall that r, u, m are the fractions of the population in reciprocal exchange, unemployed on the market, and employed on the market respectively (see also figure 4.1). If we normalize total population to unity then $u = 1 - r - m$. The flow dynamics in the market can then be described by the differential equation:

$$\dot{m} = a(1 - r - m) - b(m) \cdot m. \tag{4.6}$$

Here, $\dot{m} \equiv dm/dt$. The fraction of people on the market increases with the number of agents finding a production possibility and decreases with the number of agents accomplishing their exchange.

In the remainder of the chapter the focus is on steady state solutions. Steady states are marked by a constant distribution of agents over states, hence we have a constant rate of employment: $\dot{m} = 0$. Based on this assumption we can derive the value equations of the agents on the market. Let V_u and V_m denote the discounted lifetime utility of being unemployed and employed respectively. Under the assumption of a steady state, $dV_u/dt = dV_m/dt = 0$ and the value equations are given by:

$$\rho V_u = a(V_m - V_u - c), \tag{4.7}$$

$$\rho V_m = b(m)(y_m + V_u - V_m). \tag{4.8}$$

This set of equations can be rewritten in terms of V_u. Since unemployed agents on the market always have the possibility of not accepting a production possibility, due to high costs, V_u is always nonnegative[6]:

$$V_u = \max\left\{\frac{a}{\rho}\left[\frac{b(m)(y_m - c) - \rho c}{\rho + b(m) + a}\right], 0\right\}. \tag{4.9}$$

definition 1 is now completely determined by equations (4.2)-(4.9). Together with the steady state condition that $\dot{m} = 0$ this describes the long-run equilibrium.

[5] The market is typically characterized by the use of money. Diamond [1984] explicitly takes money into account by distinguishing between unemployed, buyers, and sellers. Here we assume that buying and selling simultaneously take place. None of the results are sensitive to this assumption.

[6] We exlude negative values of V_u by assuming that agents can choose for complete idleness with $V_u = 0$.

4.3 Equilibrium

Based on the steady state assumption, this section explores the consequences of different assumptions on the shape of the valuation ratio $\theta(m)$. We stay close to the disquisition in the introduction and section 4.2.1. Thus in section 4.3.1 we examine the steady state solutions when the valuation ratio θ is monotonically decreasing in the market size and in section 4.3.2 we allow for an increasing part of the θ-curve for large market sizes. Finally, comparative statics are provided in section 4.4. Throughout we use the notion of a short-run equilibrium whenever $\dot{m} = 0$, and of a long-run equilibrium when $\dot{m} = 0$ and, in addition, definition 1 is satisfied.

4.3.1 Complete commodification

Where markets replace social relations we speak of commodification. Complete or full commodification indicates a situation in which markets expand to an extent where all social relations are abolished. Under incomplete or partial commodification social relations are still embedded in the community, but are partly driven out by the existence of markets. In this section we consider the case where θ is monotonically decreasing in the market size (for reasons given in the introduction and section 4.2.1). As it turns out, this is a situation where, if any, *full* commodification results.

Figure 4.3 depicts this case[7]. The E_r and E_u curves show respectively the LHS and the RHS of the enforceability constraint in definition 1. Therefore, whenever the E_r lies above the E_u curve reciprocal exchange is enforceable, and for market sizes where E_u lies above E_r agents maximize expected utility by entering the market. The value of being unemployed increases in the market size and is therefore upward sloping. The value of reciprocal exchange depends on the valuation ratio $\theta(m)$ and is therefore downward sloping. The curves are drawn for the range $[0, \tilde{m}]$. These are the possible short-run equilibrium market sizes. The upper limit is given by $\tilde{m}$; the maximum possible market size for which no agents are involved in reciprocal exchange, i.e. $r = 0$.[8] The critical value of the market size at which

[7] The following set of parameters is used to obtain the figure: $\{a, b, c, g, h, y_m, \delta\} = \{.8, .6, 1, -.1, .3, 8, .8\}$ where we assumed that the probability of finding a trading partner on the market is linear in m: $b(m) = b \cdot m$ and the subjective valution is given by $\theta(m) = h + gm$. This set of parameters assures that both exchange mechanisms are enforceable for some market sizes for a value of θ in the range: $\underline{\theta} < \theta < \overline{\theta}$. (see section 3.3).

[8] Thus from equation 4.6 and the definition of short-run equilibrium $\tilde{m}$ is such that $a(1 - \tilde{m}) = b(\tilde{m}) \cdot \tilde{m}$.

the enforceability constraint holds with equality is given by m_1. At all market sizes $m \leq m_1$ reciprocal exchange is enforceable. Whenever the market size for some reason exceeds the critical size m_1 the market size is too large for reciprocal exchange to be enforceable and eventually all agents will enter the market. This cannot be a long-run equilibrium. The only possible long-run equilibria are when economy ends up in a corner solution[9]. As appears from equation (4.6) the long-run equilibrium in case people opt for the market regime equals $\tilde{m}$. Thus we see that for some market sizes reciprocal exchange is enforceable and that for larger market sizes the economy will converge to a market exchange economy. It is clear from the picture that if the E_u curve is everywhere above the E_r curve then reciprocal is never enforceable no matter what the market size is. Similarly, if the E_u curve is everywhere below the E_r curve then reciprocal is always enforceable.[10]

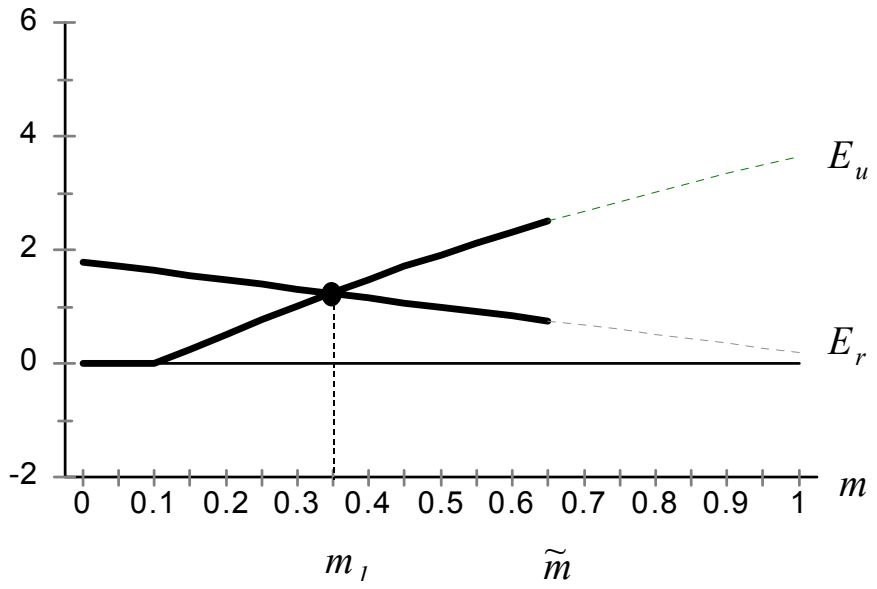

FIGURE 4.3.

[9] Strictly speaking, it is possible that the economy sticks to its initial market size (possibly at a positive level) or converges to a market size of $\tilde{m}$.With slight abuse of notation, we speak of corner solutions even if the initial market size is positive.

[10] However, $V_u = 0$ at $m = 0$ and if E_r is everywhere below E_u then E_r must be negative. But V_r can still be positive and the case is therefore meaningful.

4.3.2 Partial commodification

In the previous section, the valuation of reciprocal exchange was assumed to be monotonically decreasing in the market size. However, as mentioned earlier (notably in section 4.2.1) it is more likely that the valuation depends on the market size in a slightly more sophisticated way. Despite the decreasing disutility costs of changing beliefs, the feeling of a social deficit that cannot be repaired becomes distressingly oppressive and the wish for sustaining existing reciprocal relationships becomes increasingly weighty. Thus we have that $\theta(m)$ is first decreasing in m and after some point increasing in m. The E_r curve will for that reason behave similarly because of its dependency on $\theta(m)$. Figure 4.4 below depicts this case[11].

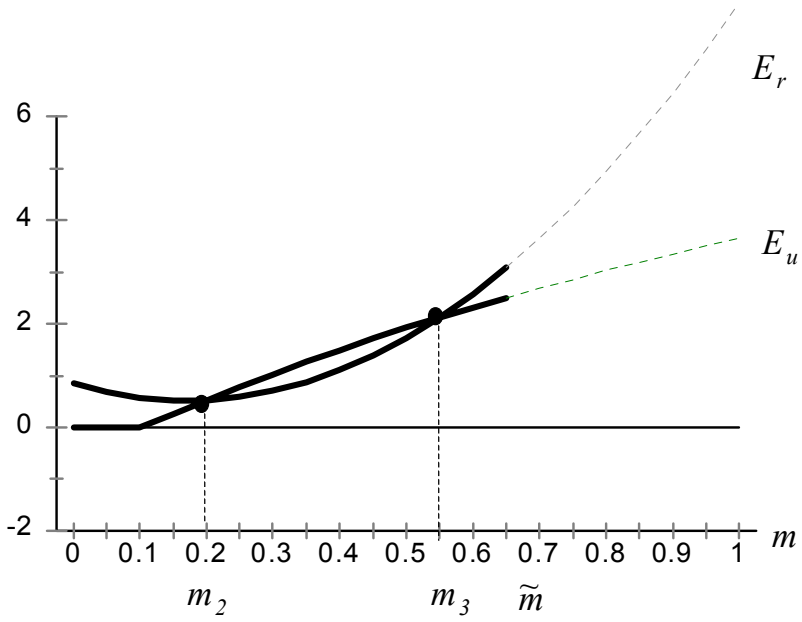

FIGURE 4.4.

An interesting feature of the variable subjective valuation is the possibility of multiple interior equilibria. For example, in figure 4.4 there are two interior equilibria, of which one is stable (m_3). For small market sizes below m_2 the search costs are too high to enter the market. However, at larger market sizes the search costs are lower and the subjective valuation of the market is higher. This is true for

[11] Here $\{a, b, c, g, h, j, y_m, \delta\} = \{.8, .6, 1, -.24, .24, .7, 8, .8\}$ where $b(m) = bm$ and the subjective valution is given by $\theta(m) = h + gm + jm^2$.

all market sizes between m_2 and m_3. Reciprocal exchange is no longer enforceable and agents enter the market. But contrary to the case in the previous section where θ is monotonically decreasing and where eventually everybody would be engaged in market exchange, there is a point m_3 at which the market is so large that agents have relatively high preferences for reciprocal exchange again. At this point, the market will stop growing and part of the population will stay in their reciprocal exchange relationship. An interior long-equilibrium solution is therefore obtained, potentially explaining why gift exchange continues to exist in the contemporary environment that became more market oriented until now.

4.3.3 Valuation, patience, and viability.

Next, we turn to comparative statics. In general, reciprocal exchange can be enforceable for some small market sizes but not for large markets. The extent to which reciprocal exchange is enforceable is first of all obviously depending on the relative valuation (θ) of reciprocal exchange. Likewise, it depends on the discount rate (δ) which can alternatively be interpreted as a measure of trust (see also section 4.2.2). The next two propositions characterize the general relation between the two variables.

Proposition 5 *For every $\delta \in (0,1)$ there exist positive $\underline{\theta}(\delta)$ such that $\forall\theta(m) \geq \underline{\theta}(\delta)$ reciprocal exchange is enforceable and $\forall\theta(m) < \underline{\theta}(\delta)$ reciprocal exchange is not enforceable. $\underline{\theta}(\delta)$ is first decreasing in δ and then, if $\theta(m) > \theta(0)$ for some m, possibly increasing in δ.*

Proposition 6 *For every $\delta \in (0,1)$ there exist positive $\overline{\theta}(\delta)$ such that $\forall\theta(m) \geq \overline{\theta}(\delta)$ market exchange is not enforceable and $\forall\theta(m) < \overline{\theta}(\delta)$ market exchange is enforceable. $\overline{\theta}(\delta)$ is first decreasing in δ and then possibly increasing in δ.*

Proof. All proofs of the propositions appear in the appendix. ■

Based on these propositions, figure 4.5 represents the typical shape of the $\underline{\theta}$-curve and the $\overline{\theta}$-curve for all possible combinations of θ and δ.

In this figure, we plotted the areas for which reciprocal exchange is not enforceable for any market size, for which it is enforceable at positive market sizes, and for which it is enforceable at any market size. The latter implies that market exchange is not enforceable at any market size. Thus for example, at e_1 reciprocal exchange is not enforceable at any market size ($\theta < \underline{\theta}$), whereas at e_3 market ex-

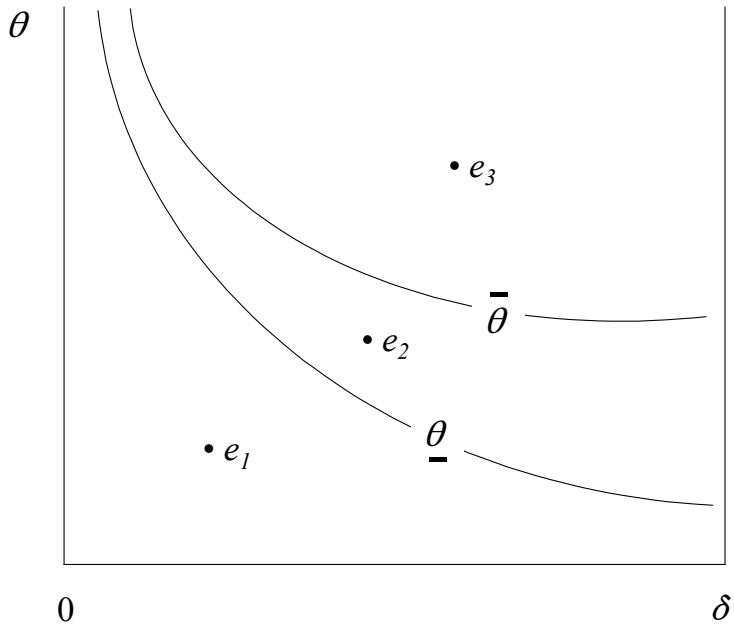

FIGURE 4.5.

change is not enforceable at any market size ($\theta > \overline{\theta}$). The point e_2 is an in-between case where at some market sizes, but not all, reciprocal exchange is enforceable, and at some market sizes market exchange is enforceable. The latter case is exactly the one which figures 4.2 and 4.3 are based on. In contrast, e_3 for instance would describe the case where the E_u curve lies entirely above the E_r curve.

As can be seen from Figure 4.5, at higher discount factors reciprocal exchange is enforceable at lower values of θ. In other words, when the discount factor or the measure of trust is low, cheating is relatively profitable and the subjective valuation of the reciprocal good must consequently be high for reciprocal exchange to be enforceable. Exactly the reverse is true for the case of market exchange, i.e. at higher discount factors, it becomes less likely that market exchange is enforceable[12].

Here we have a resemblance with optimal contract theory. For example in Baker et al. [1994] a firm has to choose an optimal bonus system. They can either rely on an objective but imperfect performance measure or on an unbiased but not objectively measurable variable. In the former case the firm can rely

[12] The interpretation that for high discount rates the $\underline{\theta}$ and $\overline{\theta}$ curves can be increasing is somewhat complicated and is deferred to the appendix.

on an explicit contract (we loosely interpret this as the money reward in case of the market exchange), in the latter only on an implicit contract (i.e. based on trust). They show that under appropriate conditions, the optimal contract is a combination of the implicit and the explicit contract. But whenever the discount rate is sufficiently high only implicit contracts should be used whereas for sufficiently low discount rates one should rely on explicit contracts only.

4.4 Welfare

In this section we take a welfare perspective by comparing the efficiency of stable equilibria. We follow Kranton [1996a] by taking the weighted discounted lifetime utility of agents on the market as the measure of comparison, but none of the results hinge on this. The weights are the shares of the employed and unemployed agents. Thus, we have:

$$V_w \equiv \left[1 - \frac{m}{1-r}\right] V_u + \frac{m}{1-r} V_m. \tag{4.10}$$

Evidently, $V_u \leq V_w \leq V_m$. Since both V_u and V_m are increasing in the market size, so is V_w. We now state:

Proposition 7 *It is possible that there exist stable equilibria which are Pareto-dominated by other equilibria. As a consequence, inefficient market sizes can be sustained.*

Proposition 7 relies on a rather strong criterium of Pareto-optimality, namely where state i is socially preferred to state j if in state j no agent is worse-off than in state i and at least one is better off, without taking into account the possibility of income redistributions. This is stronger than the Kaldor-Hicks criterium and the Pareto-criterium where income redistribution is allowed. Clearly, under the weaker versions of Pareto-optimality proposition 7 is as well satisfied.

The rational behind proposition 7 is intuitively clear. The reason that dominated stable equilibria can be maintained is caused by the existence of an external effect and a coordination failure. In essence, if reciprocal exchange is initially large, search costs on the market are high, even though search costs would be low if the market was large. Similarly, if the market starts out large, search costs are low and reciprocal exchange relationship may not be enforceable even though if

the market was small they would be preferable in value terms. Besides this thick-market externality there is also a coordination failure. People do not take into account the full value of a reciprocal trade in deciding between market and reciprocal exchange. A social planner would value reciprocal exchange by its present value: V_r. But individuals only take into consideration the value of reciprocal trade for which it can be trusted upon that the relationship can be maintained: $-c + \delta V_r$ (see definition 1). The gap between those values can be considered as a coordination failure caused by mutual distrust. As a consequence, if the market starts out large, there can be a degree of trust that is insufficient to maintain the reciprocal exchange relationship even though if it could be maintained it would be superior to market exchange in value terms. Proposition 7 is illustrated below by means of a graphical treatment (see the appendix for the formal conditions).

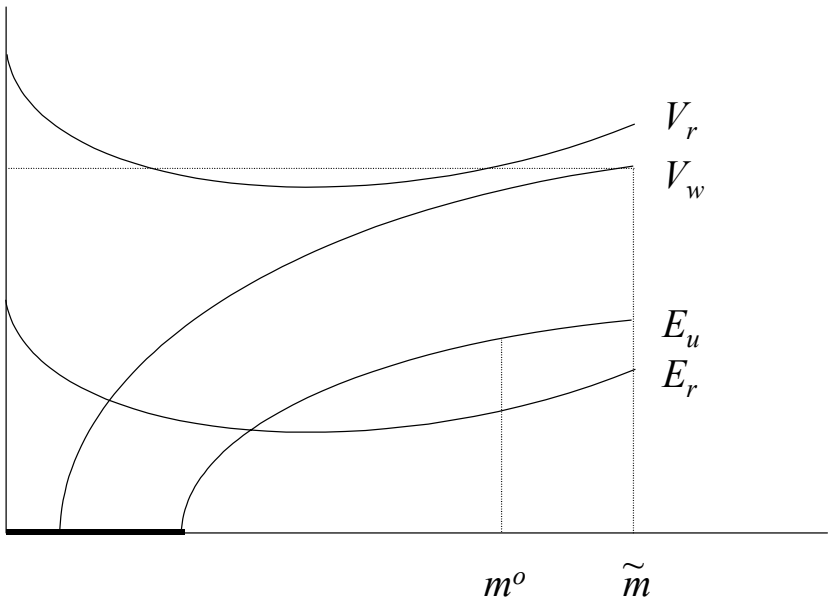

FIGURE 4.6.

In Figure 4.6 the initial market size is denoted by m°. At m°, definition 1 is not satisfied, and the economy tends to move to $\tilde{m}$, the maximum sustainable market size ($r = 0$)[13]. It is immediately seen that at $\tilde{m}$ the value of being in

[13] It is evident that the economy could as well converge to a stable interior equilibrium such as m_3 in figure 4.

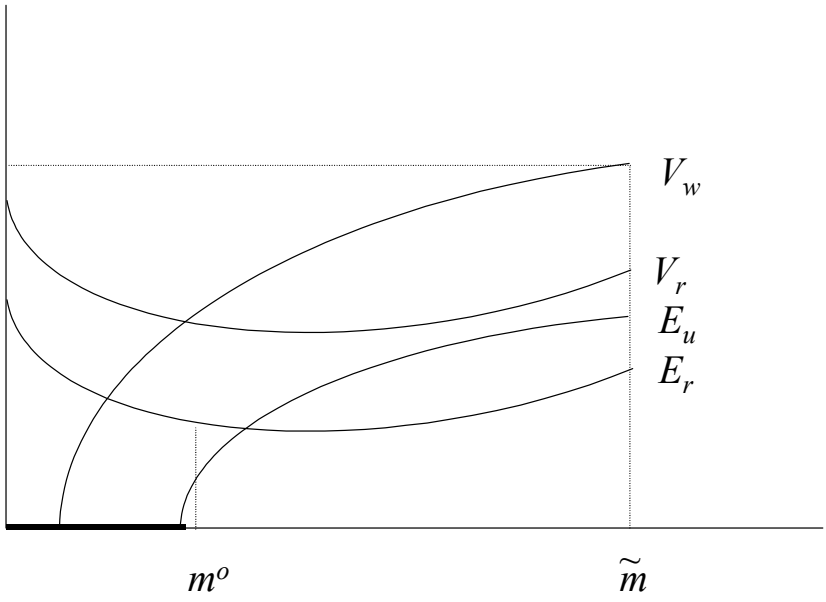

FIGURE 4.7.

a reciprocal exchange relationship is higher than that of being on the market: $V_r > V_w$. Therefore, if all agents would be involved in reciprocal exchange they would *all* be better off. Figure 4.7 illustrates the case where reciprocal exchange is sustainable although everyone could in principle be better off by entering the market.

The preceding analysis shows that individual agents need not necessarily select the socially most efficient exchange mechanism nor that there is any tendency towards a Pareto-optimal equilibrium. There are of course ways to change incentives towards the first best solution, such as a subsidy on entering the market or lowering search costs in one way or another. Due to the hysteresis present such a subsidy need only be a temporary one. Stimulation to form reciprocal relationships seems to be of a more difficult order. A one-sided inquiry into the market mechanism clearly would disguise aspects of critical importance.

4.5 Concluding remarks

Many economic activities imply a certain gift-dimension. In the literature gifts are discussed from different perspectives (see chapter 2). In this chapter we hold

the view that the logic of the gift contains some notion of reciprocity. There are no free gifts in case the institution of gift-giving is considered as a coordination mechanism to cope with the state of nature. Market exchange and gift exchange have therefore something in common: the reference to the notion of reciprocity. But reciprocal exchange is based on personal relationships and trust whereas market exchange is based on anonymity and money.

Here we assume that market exchange and reciprocal exchange can be made commensurable by extending preferences. Gift-giving, and thus reciprocal exchange, renders symbolic utility as people value the idea of belonging to a group or society at large. Symbolic utility and substantive utility, which relates to traditional economic activities, are different components of the extended utility function.

Starting from these premises the main question to be answered is under which circumstances reciprocal exchange is viable and not crowded-out by the market regime. Such a form of crowding-out can be conceived as a complete commodification of society. Following Kranton [1996a] reciprocal exchange is modelled in a strict manner. Agents expect a counter-performance. Market exchange is modelled as a search process where in case of matching goods are exchanged immediately. The model is closed by introducing an enforceability condition showing under which conditions agents defect in case of reciprocal exchange. The condition critically depends on the market size and on the discount factor. The latter has its resemblance in the optimal contract literature once we loosely interpret market exchange as an explicit contract and a reciprocal relation as an implicit contract.

It is shown that complete commodification obtains if symbolic utility declines as the relative size of the market increases. However, reciprocal exchange may survive for low subjective discount rates which are in some sense indicative for a high level of trust. There is another reason why full commodification may not be the equilibrium outcome. Symbolic utility may rise if reciprocity becomes scarce as the market takes over. In the end people may be aware of a social loss and revalue reciprocal exchange accordingly. As a result commodification will stop at some point and there will be an interior equilibrium solution with market exchange and reciprocal exchange coexisting.

From a welfare point of view, full or partial commodification is not necessarily superior to reciprocal exchange. It is shown that agents need not necessarily select

the exchange mechanism that is Pareto-optimal. This result is related to the search externality in market exchange and the coordination failure in reciprocal exchange. If the market is large initially, search costs are low and the economy may converge to an inefficient outcome. Similarly, the economy may converge to an inefficient outcome if the coordination failure with respect to the choice of reciprocal exchange is important.

What seems most urgent in additional research is a more flexible way of modelling reciprocal exchange to cope with the different aspects of gift-giving. In particular, it may be rewarding to shed some light on intergenerational gifts, where reciprocity in the usual sense is out of reach or where the time interval is extremely long. The latter is for example the case in family relationships where parents take care of their children in the hope of mutual care once they themselves become dependent on their children's readiness to support them. In addition it may well be more realistic to assume that some goods are more suited to be produced on the market (e.g. bread), others to be offered in a reciprocal relationship (e.g. baby-sitting), or in a combination of these (e.g. insurance, see Arnott and Stiglitz [1991]. Individuals can be assumed to be heterogeneous and spend their time in different proportions over the two exchange mechanisms. Another interesting question emanates from the paradox formulated by Etzioni [1988, 250]: "The more people accept the neoclassical paradigm as a guide for their behavior, the more the ability to sustain a market economy is undermined". The paradox suggest that market exchange becomes less efficient if morality is on the retreat. This need not be true, as our analysis suggests, but it certainly deserves serious consideration.

4.6 Appendix

In this appendix we first show that definition 1 characterizes a Nash-equilibrium and subsequently proof propositions 5 to 7. Where no confusion can arise we simply write θ instead of, for example, $\theta(m)$.

Proof that definition 1 is a Nash-equilibrium.
Consider the payoff matrix below. Players are denoted by P and C, the agent that produces and the one that consumes first respectively. They can either be honest, H, or cheat, D (defect). The pay-offs are in utility terms. As an example, consider P playing honest and C cheating, then P produces the good, bearing cost in utility terms of c and enters the market the next period after finding out of being cheated, δV_u. C consumes y_r, does not produce in return and enters the market getting δV_u. Thus in the upper-right cell we have the pay-offs $-c + \delta V_u$ and $y_r + \delta V_u$. Consider then the strategies $\{H; H\}$. This can, by construction, only be a (weak) Nash-equilibrium if it is in both players advantage not to deviate: $-c + \delta y_r + \delta^2 V_r \geq \delta V_u$ and $-c + \delta V_r \geq V_u$. Fortunately, we can show that first inequality is implied by the second. Note that $\frac{1}{(1-\delta)} y_r \geq \frac{1}{2(1-\delta)}(y_r - c_r) = V_r$ (see equation 4.5). Then $\delta y_r + \delta^2 V_r \geq \delta V_r$. If now $-c + \delta V_r \geq V_u$ (second constraint) then it is surely the case that $-c_r + \delta y_r + \delta^2 V_r \geq V_u \geq \delta V_u$ which proves the first inequality. Finally note that $\{D; D\}$ is always a (weak) Nash-equilibrium. ■

$P \downarrow; C \rightarrow$	H	D
H	$-c + \delta y_r + \delta^2 V_r; y_r - \delta c + \delta^2 V_r$	$-c + \delta V_u; y_r + \delta V_u$
D	$\delta V_u; \delta V_u$	$\delta V_u; \delta V_u$

Proof of Propositions 5 and 6.
We try to find values of θ for which definition 1 is satisfied with equality: $-c + \delta V_r = V_u$. Label this equality D1. For this value agents are indifferent between reciprocal exchange and entering the market given the market size. For higher values of θ they prefer reciprocal exchange, for lower values they prefer market exchange. In particular we try to find values of θ, say $\underline{\theta}$, for which the equality is satisfied but will not be satisfied for any lower θ for *any* market size. It is instructive first to consider the case where θ is (weakly) monotonically decreasing in m and then generalize the results, as in the main text. The advantage of this is that if reciprocal exchange is enforceable at all, it is certainly enforceable at

$m = 0$ since V_r is decreasing in m and V_u increasing in m. Thus, we can focus on $m = 0$. Then D1 reads, by restricting V_u to $\mathbb{R}^+$:

$$\frac{\frac{1}{2}\delta(\theta y_m + c) - c}{1 - \delta} = 0. \tag{4.11}$$

Since we put no restrictions on θ, given any $\delta \in (0, 1)$ there exists a $\underline{\theta}$ such that the equality holds (and is in this particular case easy to find). (As a marginal comment, note that as $\delta \to 1$, the numerator of the LHS of (4.11) approaches $\frac{1}{2}(\theta y_m - c)$ which equals zero for some $0 < \theta \leq 1$.). Since the LHS of D1 is increasing in δ, $\underline{\theta}$ is lower for higher values of δ.

Generalizing the argument, we see that the RHS can be positive for markets $m > 0$, but also that $\theta(m)$ may be larger than $\theta(0)$. So even if the equality holds at $m = 0$, it may well be that the LHS is larger than the RHS at positive market sizes. The critical value of θ (i.e. $\underline{\theta}$) can therefore be lower than the value of θ for which the equality holds at $m = 0$. The caveat in the generalization is, however, the fact that V_u is now increasing in δ as well. At $m = 0$, $V_u = 0$ no matter what the rate of time preference is. This remains true up to the market size where V_u becomes strictly positive. For the range of values for which V_u is strictly positive, V_u is also increasing in δ. Thus both sides can be increasing in δ at some market sizes, and the relation between $\underline{\theta}$ and δ becomes ambiguous. We stress however that for low values of δ, $V_u = 0$, and hence there exists an interval where $\underline{\theta}$ and δ are negatively correlated. The intuition behind this increasing part is that because on the market costs are made before revenues, for sufficiently low discount rate the present value is negative, whilst in a reciprocal exchange with some probability you consume before you produce and so even for low (but positive) discount rates expected gains are positive for some valuation ratio.

As a special case, if $\theta(m) \leq \theta(0)\ \forall m$ (in other words θ is nonincreasing), then $\underline{\theta}$ can be determined by inspection of $m = 0$ alone (since if then reciprocal exchange is not enforceable at m, and since V_u is nondecreasing and V_r nonincreasing in the market size, then it is not enforceable at any m). Since at $m = 0$, $V_u = 0$, if the discount rate increases a little, V_u remains zero but V_r increases so $\underline{\theta}$ unambiguously declines. But note that such an unambiguously declining $\underline{\theta}$-curve is only a special case and that it is not directly related to the shape of $\theta(m)$.

The same line of argument can be used to derive proposition 6. Here the aim is finding the θ such that market exchange is just enforceable at one particular market size, and not for any higher θ. We first try to find $\overline{\theta}$ in the case of non-

increasing θ so that the focus can be restricted to $m = 1$. We do not state the proof here. ■

Proof of proposition 7.
The proof consists of showing that it need not be contradictory to have a stable equilibrium that is inefficient in the sense that it is Pareto-dominated by another stable equilibrium. Denote the two equilibria under investigation by m_i and m_e (the subscripts stand for inefficient, efficient). Let the initial point be the inefficient equilibrium m_i. Three cases are to be considered: 1. m_i is at one of the corners of the economy. If $m_i = 0$ then it is stable if $-c + \delta V_r \geq V_u$. 2. If $m = \tilde{m}$ then it is stable if $-c + \delta V_r > V_u$. 3. m_i is an interior solution. It is stable if $-c + \delta V_r = V_u$ and if $\delta \frac{dV_r}{dm_i} \geq \frac{dV_u}{dm_i}$.

We have to show that the following set of equations need not be inconsistent:

$$V_r(m_e) \geq V_r(m_i), \tag{4.12}$$
$$V_w(m_e) \geq V_w(m_i). \tag{4.13}$$

Additionally, the enforceability constraints have to be satisfied as indicated at m_i and m_e. Consider for example the case where m_i is the corner solution $m = 0$ and m_e is an interior stable solution that Pareto-dominates the corner solution. Thus we have:

$$V_r(m_e) \geq V_r(0), \tag{4.14}$$
$$V_w(m_e) \geq V_w(0), \tag{4.15}$$
$$-c + \delta V_r(0) \geq V_u(0). \tag{4.16}$$
$$-c + \delta V_r(m_e) = V_u(m_e) \tag{4.17}$$

The second inequality is naturally satisfied. (Indeed, since $V_w'(m) > 0$, the only case where an equilibrium can be dominated by a smaller market size is where $m = 0$ since otherwise all remaining market participants would lose some welfare. *Except*, of course, when nobody stays). Since we put no restrictions on θ the first inequality can be satisfied as well (not, however, when $\theta(m)$ is nonincreasing in the market size). Combining the (in)equalities we see that as long as V_r is increasing over the interval $[0, m_e]$ (but remember that it may be decreasing in the first stage and increasing thereafter), but not as fast as V_u there is no inconsistency and it cannot be ruled out that $m = 0$ is indeed inefficient. Other cases can be analyzed

in a similar manner and are omitted. The result of possible inefficiency is easily extended to the nonweighted case for V_m, V_u, and V_w all behave in a similar way (namely increasing in the market size). ■

5

Welfare Gains of Labeling with Heterogeneous Consumers

Diamonds are a girl's best friend. But not only her's: they are no less a rebel's best friend. Rebel armies use guns to force labourers to dig holes and earn hundreds of millions of dollars by selling the "blood stones". To stop these activities, many countries have now signed to support a system of certificates for diamonds without a conflicting history. When does this lead to welfare gains?[1]

5.1 Introduction

Implicit in the standard formulation of the fundamental welfare theorems is that the characteristics of commodities are observable to all market participants (Mas-Colell et al., [1995]). However, in many cases the consumer cannot observe all characteristics of a specific good, such as the safety or the quality level. Due to these informational asymmetries, the consumer is not willing to pay price premiums for different goods. No distinct markets can therefore exist for goods that are differentiated with respect to unobservable characteristics. As is well known, this can have dramatic consequences for the efficiency of the market mechanism. Famous in this respect is the market for lemons as described by Akerlof [1970].

---

[0] This chapter is coauthored by Theo van de Klundert. We are thankful to Richard Nahuis and Sjak Smulders for helpful comments.

[1] See *The Economist* [2003] and *BBC News* [2000].

An in particular interesting class of goods where some characteristics are unobservable is that of goods with social externalities of production. For example, some production methods have damaging effects on the environment, involve child labour, or rely on what are perceived as unfair wages. These externalities are not taken into account by the producers and are therefore not reflected in the consumer price. And even though it seems that many consumers are in principle willing to pay premiums for the use of production methods that do not, or to a lesser extent, involve such social externalities, they cannot observe from the end product which production method has been used. Hence, goods with social externalities of production fall into the class of asymmetric information. As in the case of the lemons market, without improving the consumers' information, there is little hope that the equilibrium will be efficient.

In this chapter, it is argued that goods with social externalities of production pose an even more severe problem to the efficiency of markets than most other goods with unobservable characteristics. In other cases where some characteristics are not directly observable, producers can often still signal them to the uninformed party. For example, the price, advertisements or warranties can sometimes provide a credible signal to the consumer that the product in question is of high quality (Tirole [1988]). However, such signaling strategies often do not exist for goods with social externalities of production (see the next section).

Lacking the possibilities of the usual market responses to informational asymmetries, the government can decide to intervene. One obvious way is to impose a standard on production methods. Under some circumstances this can be welfare improving. However, during recent years government regulation is increasingly relying on information provision to alter behavior (Magat and Viscusi [1992]). One example is labeling. Labels are certificates issued by a third party that provide credible information about the contents of a product. By now there exists a variety of such labels that concern, among other things, the environment, working conditions, fair trade, and child labour. Should they wish to do so, consumers can contribute to a reduction of social externalities by buying these labeled products.

The crucial difference between standards and labels is the fact that labels are voluntary. Firms can decide whether or not to apply for the label, and adjust their production technology conform the requirements. In contrast, standards imposed by the government are mandatory for all firms in the market. Labeling schemes therefore allow for more flexibility in the choice of production technology. This

has the advantage that labels serve the consumers' needs better than standards when consumers are heterogeneous in their willingness to pay for reducing the social externality. The scope of differentiation is however limited by the costs of passing on informational contents to the consumers. These costs consist of designing the label, screening costs of the firm made by the third party, and the costs of information acquisition by the consumers, each of which can be significant. Given this trade-off between flexibility and costs, the aim of this chapter is to examine under which conditions labeling is preferred over standards.

The setup is as follows. The next section first reviews background information concerning markets with asymmetric information and some of the aspects of labeling. In section 5.3, the model is described and a derivation of the welfare under imperfect information is given, together with the optimal standard and label technology. Section 5.4 presents the main proposition where a comparison is made between the welfare level under a standard and under a labeling policy. The chapter ends with a discussion and conclusion.

5.2 Background

5.2.1 General background

In this chapter the class of goods is considered whose production methods directly affects the well-being of the consumers in the economy. This may for example occur because of the emissions it generates, or because the consumers are altruistic towards other people who are in some way affected by the production (emissions in their neighborhood, the wage or working conditions of involved employees etc.). This class of goods will be referred to as goods with social externalities of production, in short social externalities.

It is now being recognized that some consumers are willing to pay a price premium for goods whose production methods involve less social externalities (see for example Kirchhoff [2000]). Their motivations to pay more can stem from as diverse reasons as a 'warm glow' feeling from donating (Andreoni [1989]), a sense of equity (Fehr and Schmidt [1999]), a feeling of guilt or responsibility (Frank [1988]), or the bestowal of social approval (Holländer [1990], chapter 3). Despite the willingness to pay a premium, however, producers are discouraged from supplying goods with less social externalities. As pointed out in the introduction, consumers cannot distinguish by inspection from the end-products which pro-

duction method has been used. A producer with a more expensive production method that generates less externalities has to charge a higher price for his product in order not to make losses. The higher price is however not always a credible signal to consumers of a more expensive production method: producers with a cheaper production method can raise their price and pretend to have an expensive production method as well. Clearly, if the low cost producers have an incentive to imitate the high costs producers, this cannot be an equilibrium.

The above arguments are reminiscent to markets where different producers offer a variety of quality levels but the quality difference is unobservable to the consumers. The literature on industrial organization has studied in detail when and how it would still be possible for producers of high quality to give a signal to the consumers that their quality is high. The conditions under which this is possible depend on the type of goods under consideration.

In Tirole [1988] a distinction is made between search goods, experience goods, and credence goods. In case of search goods, the consumer can determine the quality in advance by simple inspection. Here, no signalling is needed at all. With experience goods, the quality is only revealed after consumption. Possible signalling strategies include advertisements and introductory prices (see Milgrom and Roberts [1986]). The basic mechanism is that consumers will return to the producer if they find out that the quality is good. A producer of high quality goods can then, for instance, give low introductory prices, knowing that the consumers will return, and reap the lost profits back later. A producer of low quality goods cannot imitate this strategy, since low introductory prices mean lower short-run profits and this loss will not be made up in later periods. In effect, a low introductory price signals that the quality must be high. Finally, credence goods offer almost no opportunity to learn the quality, even not after using the product for a long time. In this case, signalling strategies break down because consumers stay uninformed about the quality even after consuming the good. Hence, there is no particular reason why consumers should return to this particular producer, and the strategy to give up short run gains in exchange for long run gains loses credibility.

The class of goods that is considered in this chapter belongs to the latter group of goods, in essence credence goods, since the consumer cannot learn anything about the production method either on beforehand nor by experience, a view shared for example by Kirchhoff (2000). Labeling may in this case be a good

alternative to communicate the information of the production method. We now turn to a brief overview of the practice of labeling.

5.2.2 Labeling

Labeling provides information. A label can be a statement about the production method of a good. For example, it can state that fair wages are paid or that no CFKs are used. Another possibility is that labels are "trademarks" that are associated with certain contents, like some so called 'fair trade' labels are associated with fair wages. Of course, labeling is a very crude way of passing on information. For instance, the statement that 'fair wages' are paid leaves a lot of discretion to the actual wages that are paid. However, the contents of such a message can be subject to legal restrictions.

In principle, labels do not have to be different from brands, like in the case of the fair trade label. However, when a lot of small firms are active on the market, they may want to signal their environmental quality collectively to target a broader public. One way to achieve this is to participate in collective labeling, separated from the individual brands. Collective labeling can also be demand-driven. Consumers have to invest time in studying the labels, as they have cognitive limitations in absorbing information. Not only do they have to recognize the labels, they also have to understand what they say, for example what it means that no CFKs are used. Several studies show that the limitation on human cognition to process information from labels is significant. For example, in one study concerning labels containing hazard warnings, it is found that as the amount of information is increased, the consumer's recall of other information on the product's label declines, and furthermore that label clutter easily leads to problems of information overload (Magat and Viscusi [1992]). Naturally, consumers will only incur these learning costs when these are outweighed by the benefits, which is more likely when there are only few labels. This favours collective labeling.

Labeling is voluntary. An important feature of labels is that firms voluntarily participate. A standard, on the other hand, is enforced by the government. The crucial difference therefore is that under a labeling scheme, some firms may decide not to participate at all, to serve that part of the market which values quality less. With minimum standards all firms are required to meet the specified standard. This latter observation gains importance when consumers are heterogeneous in their willingness to pay.

Certification is performed by a third party. The intervention of a third party, possibly the government, ensures credibility. Consumers that buy labeled goods must be able to trust the certificate otherwise they buy non-labeled goods.
Examples. There are numerous of instances where labels are used. Most labels concern the environment, working conditions or fair trade. Some of the more well known examples include the German *Blue Angel* and Scandinavian *Nordic Swan* eco-labels, and the international *Fair-trade* label, to name just a few. Another more recent attempt is the *Kimberley process*, which has as purpose to certify diamonds. The latter label is an attempt to prevent rebel armies from using revenues to buy weapons (*The Economist* [2003]).

5.3 Description of the model

This section describes the general features of the model.
Producers. The focus is on goods that have social externalities of production. There are many producers competing on a market of perfect competition. Each producer can choose a production technology. The available production methods are characterized by a 'social responsibility index' s. One may think of s as representing investments in green technologies or the wage level paid to employees. Henceforth we refer for simplicity to s as the quality level. It is supposed that s belongs to $[\underline{s}, \infty)$. A higher s is interpreted as more investments in *reducing negative* social externalities tied to production and is therefore more costly in terms of production.

The unit costs of producing a good of quality s is denoted by $c(s)$ and this function is assumed to be convex. In order to derive closed form solutions, it is assumed that the unit cost is quadratic in quality:

$$c(s) = \alpha s^2, \qquad \alpha > 0. \tag{5.1}$$

The timing of the game will be that firms first simultaneously choose their quality levels and then they simultaneously set their prices.
Consumers. As pointed out in the previous section, consumers may care about the social externality because it directly concerns them (as with effects on the environment) or because they are altruistic towards the victim (as with child labour and underpaid workers). To model this, we consider a vertically differentiated product space where each consumer consumes either one or zero units of

the good. Let the consumers' preferences be given by:

$$\begin{cases} U = \theta s - p & \text{if he buys at a quality level of } s \text{ at price } p, \\ U = 0 & \text{if he does not buy.} \end{cases} \tag{5.2}$$

In this case, a higher θ implies a higher willingness to pay for a higher quality, which represents more concern for the production method. This can be reasonably attributed to a higher income level but surely also to other exogenous factors, such as education.

Consumers are heterogeneous in the sense that they have a different willingness to pay to reduce social externalities, e.g. because their incomes differ. The parameter θ appears to distinguish between different types of consumers. It is assumed that θ is uniformly distributed on $[\underline{\theta}, \bar{\theta}]$ with density $f(\theta)$, c.d.f. $F(\theta)$, and $\underline{\theta} \geq 0$. Without loss of generality, assume that their total mass is unity. Also define θ^a to be the arithmetic average type: $\theta^a \equiv (\underline{\theta} + \bar{\theta})/2$. Furthermore, to cut down on the many possible cases and in order to focus on the interesting ones, the following assumption on the taste parameter is made:

Assumption 1 $\underline{\theta} > \alpha \underline{s}$.

This assumptions states in effect that when the lowest possible quality is supplied at its competitive price, the market is covered, in the sense that even the lowest type is willing to purchase a good on the market.

Welfare. For the comparative statics we rely on interpersonal comparable utilities and define welfare as the sum of total consumer and producer surpluses. By the assumption of perfect competition, no producer surplus exists in equilibrium. Hence:

$$W = \int_{\underline{\theta}}^{\bar{\theta}} U(s, \theta) dF(\theta). \tag{5.3}$$

We now examine the welfare level under different policies.

5.3.1 Imperfect information

As a first benchmark, consider the case where consumers have no way to determine the quality of the product. As the firms cannot discriminate themselves from each other, it will be clear that there can only be one price in equilibrium. Furthermore,

competition implies that this price equals the marginal cost of supplying at quality s: $p(s) = c(s)$. Moreover, it will be clear that given the demand functions in the second stage, producers have an incentive to cut on quality in the first stage. Consumers foresee this and (rightly) expect the minimum possible quality level $\underline{s}$.

Note that assumption 1 implies that $\underline{\theta}\,\underline{s} > \alpha\underline{s}^2 = p(s)$, so that under imperfect information the market is covered. This leads to the following proposition:

Proposition 8 *In the imperfect information equilibrium all firms supply a good of quality $\underline{s}$ at price $p = \alpha\underline{s}^2$. Total welfare is given by $W^I = \theta^a\underline{s} - \alpha\underline{s}^2$.*

Proof. All proofs are in the Appendix. ■

5.3.2 Standards

Suppose next that the government wants to improve on the imperfect information equilibrium by imposing standards, i.e. the government requires the quality to be at least $\sigma \in [\underline{s}, \infty)$.

It is assumed that the consumers are aware of the standard. As in the case of imperfect information, consumers also rightly foresee that producers will supply the good at the minimum quality allowed, that is σ. Competition leads again to marginal cost pricing.

The consumer who is indifferent between buying or not has a taste parameter $\tilde{\theta}$ such that $\tilde{\theta}\sigma = p = c(\sigma)$. All consumers with a taste parameter equal or greater than $\tilde{\theta} = \tilde{\theta}(\sigma)$ buy the good, all the others do not. If the government wants to maximize welfare, it should therefore choose the standard according to:

$$\sigma^* \in \arg\max_{\sigma} W^s(\sigma) = \int_{\tilde{\theta}(\sigma)}^{\bar{\theta}} (\theta\sigma - c(\sigma))dF(\theta). \tag{5.4}$$

The first-order condition to this problem is, by Leibniz's rule, given by:

$$\int_{\tilde{\theta}(\sigma)}^{\bar{\theta}} \frac{\partial}{\partial\sigma}(\theta\sigma - c(\sigma))d\theta - (\tilde{\theta}\sigma - c(\sigma))\frac{\partial}{\partial\sigma}\tilde{\theta}(\sigma) \leq 0, \tag{5.5}$$

with equality if $\sigma > \underline{s}$. The second term drops out since this is the utility of the consumer who is indifferent between buying or not buying and he has zero utility

by definition. When integrated out, one can write the first order condition as:

$$(\bar{\theta} - \tilde{\theta}) \left[\frac{1}{2}(\bar{\theta} + \tilde{\theta}) - 2\alpha\sigma \right] \leq 0, \tag{5.6}$$

with equality if $\sigma > \underline{s}$. It can be shown that at the optimum the indifferent consumer is not the highest type: $\tilde{\theta} \neq \bar{\theta}$. The intuition behind this is that no standard should be chosen such that no type consumes. This would give a welfare equal to zero, whereas welfare would be strictly positive for a standard equal to $\underline{s}$. Hence, the term in brackets must be equal to zero in equilibrium. If $\tilde{\theta} > \underline{\theta}$ then equation (5.6) has solution $\sigma^* = \bar{\theta}/(3\alpha)$. This, implies that $\tilde{\theta} = \bar{\theta}/3$ which is only compatible with $\tilde{\theta} > \underline{\theta}$ if $\bar{\theta} \geq (3/2)\theta^a$. For all other values of $\bar{\theta}$, it must be that $\tilde{\theta} = \underline{\theta}$ (covered market) and $\sigma^* = (\bar{\theta} + \underline{\theta})/(4\alpha)$.

Proposition 9 *Suppose the government imposes a minimum standard σ on s. Then, if $\bar{\theta} < (3/2)\theta^a$ the optimal standard is given by $\sigma^* = \max\{\underline{s}, \theta^a/(2\alpha)\}$ and if $\bar{\theta} \geq (3/2)\theta^a$ then the optimal standard is given by $\sigma^* = \max\{\underline{s}, \bar{\theta}/(3\alpha)\}$. Welfare is equal to to $W^s = (\theta^a)^2/4\alpha$ for $\sigma^* = \theta^a/(2\alpha)$ and $W^s = \bar{\theta}^3/(27\alpha(\bar{\theta} - \theta^a))$ for $\sigma^* = \bar{\theta}/(3\alpha)$.*

For low values of $\bar{\theta}$ all consumers consume the standard. The optimal standard and welfare are in this case only dependent on the average taste parameter θ^a. An equal increase in $\bar{\theta}$ and decrease in $\underline{\theta}$, hence keeping the average constant, has thus no effect on the optimal standard or welfare. For higher values of $\bar{\theta}$, it does not pay off to make the low types consume. The optimal standard is increasing in $\bar{\theta}$ because this means that the high types are willing to pay more.

Notice that social welfare with optimal standards is always weakly higher than under imperfect information. This must be true since the government can always reach the same welfare level of imperfect information by setting the minimum standard equal to $\underline{s}$. However, in many case can the government do strictly better than under imperfect information by increasing the minimum standard above what is minimally feasible for producers. Note however, that a minimum standard can be, but is not always a Pareto-improvement upon the imperfect information equilibrium. The consumer of type θ' is indifferent between the minimum quality and a standard if $\theta'\underline{s} - \alpha\underline{s}^2 = \theta'\sigma - \alpha\sigma^2$ or, provided that $\sigma > \underline{s}$:

$$\theta' = \alpha(\underline{s} + \sigma), \tag{5.7}$$

and trivially at $\sigma = \underline{s}$. The minimum standard σ lowers welfare for all types $\theta < \theta'$. In particular, if $\underline{\theta} < \theta'$ then some consumers are worse off under the minimum standard. Their consumer surplus decreases because they now pay a higher price that does not outweigh the quality increase for them. Some of them, namely for whom $\theta < \alpha\sigma$, even stop consuming.

Lemma 2 *The introduction of a minimum standard is a Pareto-improvement if and only if* $\sigma \leq (\underline{\theta}/\alpha) - \underline{s}$.

It is easy to show that a higher $\underline{s}$ or α create less opportunities for Pareto-improvements. To relate Pareto-improvements to heterogeneity consider the following measure of heterogeneity:

Definition 2 *Consumers are more heterogeneous if the taste parameter of the highest type,* $\bar{\theta}$*, is higher and the average taste parameter,* θ^a*, is kept fixed.*

Thus, consumers are more heterogeneous if the spread of types $(\bar{\theta} - \underline{\theta})$ increases keeping the average constant. With this definition, the the following proposition is obtained:

Proposition 10 *The optimal minimum standard is a Pareto-improvement if and only if consumers are not too heterogeneous in their taste parameter, that is, if:* $\bar{\theta} \leq \frac{3}{2}\theta^a - 2\alpha\underline{s}$.

It follows that if consumers are 'sufficiently' heterogeneous, the optimal minimum standard benefits the high type consumers and hurts the low type consumers.

5.3.3 Labeling

The latter proposition is interesting, because it lays bare the limits of standards: with a minimum standard, all consumers are forced to buy at or above the minimum standard, or refrain from buying, including those who have no specific interest in buying high quality. In this section we discuss another mechanism mentioned in the previous section that also provides information about the quality but leaves the option of supplying at low quality: labeling.

Characteristic about labeling is that it is voluntary: each firm can freely decide whether or not it wants to carry a certain label and conform to the label's specified standards. The government (or another third party) will set up controls to guarantee that certified firms keep up to the standards in agreement with the

label's content. Consumers, on their side, can freely choose whether or not to purchase a labeled product. If they decide to buy a labeled product, they can read the quality from the label's description. If they purchase a non-labeled product, they infer that it must be of the lowest possible quality.

We already pointed out that there are several cost factors specific to a labeling policy. First, consumers are limited in their cognitive abilities and therefore have to make some costs in order to process the information carried on the label. Second, producers may have to restructure their production process, design a label, and draw public attention to their labeling policy through advertising campaigns and social responsibility reports. Of course, in order to ensure credibility of the label's contents, the certifying third party has to involve in costly monitoring of the firms' production process. Such monitoring costs are equally likely for the case of standards though.

It may be assumed that the costs of informing consumers in particular are relevant and for the current purpose we therefore restrict attention to these. We limit the number of possible labels to one and assume that the additional costs, b, over normal production costs for carrying a label take the following form:

$$b(\tau) = \beta(\tau - \underline{s}), \tag{5.8}$$

where τ is the quality of the label as specified in the certificate. Hence, it is assumed that no costs are made for a label that specifies the same technology as the minimum possible technology ($b(\underline{s}) = 0$) and furthermore that costs are increasing in the quality level. This positive relationships between the cost and quality of a label may for example be because a higher quality level means a more complicated technology making it more difficult to well inform the consumer. How accurate this specification of costs of information exactly is, is unclear, but this specification helps keeping the analysis tractable and does not affect the main results in an important way.

Suppose therefore that the government introduces a label for products that are at least of quality τ. Because of the marginal cost pricing by firms, the price of a labeled good will be equal to:

$$p(\tau) = c(\tau) + b\,(\tau)\,. \tag{5.9}$$

The type who is indifferent between buying the labeled and the non-labeled product is implicitly defined by $\widehat{\theta}\underline{s} - \alpha\underline{s}^2 \equiv \widehat{\theta}\tau - \alpha\tau^2 - \beta(\tau - \underline{s})$, or, provided $\tau > \underline{s}$:

$$\hat{\theta} = \alpha(\tau + \underline{s}) + \beta. \tag{5.10}$$

No types refrain from consumption in this case: all types $\theta < \widehat{\theta}$ consume the non-labeled product of quality $\underline{s}$, whilst all types $\theta \geq \widehat{\theta}$ consume the labeled product of quality τ. Hence, welfare W^l is given by:

$$W^l(\tau) = \int_{\underline{\theta}}^{\widehat{\theta}(\tau)} (\theta\underline{s} - \alpha\underline{s}^2)dF(\theta) + \int_{\widehat{\theta}(\tau)}^{\overline{\theta}} (\theta\tau - \alpha\tau^2 - \beta(\tau - \underline{s}))dF(\theta). \tag{5.11}$$

The government then sets the quality of the label as to maximize welfare: $\tau^* \in \arg\max W(\tau)$. The first order condition is given by:

$$\frac{\partial W^l}{\partial \tau} = \int_{\hat{\theta}(\tau)}^{\overline{\theta}} (\theta - 2\alpha\tau - \beta))dF(\theta) \leq 0, \tag{5.12}$$

with equality if $\sigma > \underline{s}$. (Note that by the rule of Leibniz one should also take the changes in the limits of the integral into account. However, this is a change in the marginal consumer who is by definition indifferent, and hence these terms cancel out.) To rule out uninteresting cases, we make an additional assumption:

Assumption 2 $\theta^a \geq 2\alpha\underline{s} + \beta$.

This assumption is sufficient to rule out that the optimal quality of the label will be so high that no consumer finds it interesting to consume the label. Solving the first order condition leads us then to the following proposition:

Proposition 11 *Suppose the government certifies producers who produce at least a quality of τ. Then the optimal quality required for a label is given by $\tau^* = \max\{\underline{s}, (\theta^a - \beta)/(2\alpha)\}$ if $\bar{\theta} < (3/2)\theta^a - \alpha\underline{s} - (1/2)\beta$ and $\tau^* = \max\{\underline{s}, (\overline{\theta} + \alpha\underline{s} - \beta)/(3\alpha)\}$ if $\bar{\theta} \geq (3/2)\theta^a - \alpha\underline{s} - (1/2)\beta$.*

Like with a standard, if $\bar{\theta}$ is relatively low all types consume the labeled product and the optimal label is independent of $\bar{\theta}$. The optimal quality of the label is decreasing in the cost parameter β. This is so because when the cost parameter β increases, by (5.8) the costs of a label increases in quality and so a lower quality is preferred. When $\bar{\theta}$ increases further it is beneficial to increase the quality of the

label. High types are willing to pay a lot while low types can still enjoy a quality level of $\underline{s}$. The optimal quality is again decreasing in costs β, and, similar to the case for standards, increasing in $\bar{\theta}$. Furthermore, if the cost parameter α is higher or the minimum available quality is higher, so is the optimal quality. A higher α or $\underline{s}$ makes the nonlabeled product more attractive for low types. Now that more types will prefer the low quality, the high types can be made a bit better off by increasing the quality of the label.

The expression of the welfare is a bit cumbersome in this case, and is therefore suppressed. But note again that the introduction of a label always weakly increases welfare. However, contrary to imposing a standard, a labeling scheme is always a Pareto-improvement upon the imperfect information equilibrium. This is true, because consumers still have the option of consuming a good of quality $\underline{s}$ but their choice is enriched with products that supply quality τ^*.

Before coming to the main proposition, it is interesting to discuss how the optimal label relates to the optimal standard. First note that when everybody consumes the label or standard, the optimal label quality, $(\theta^a - \beta)/(2\alpha)$, is lower than the optimal standard, $(\theta^a)/(2\alpha)$. This makes sense: since the costs of a label are increasing in the quality, the optimal quality is lower if the costs are higher. For zero costs ($\beta = 0$) the two cases are identical.

Secondly, when not everybody would consume the label or standard, the optimal label is again lower than the optimal standard if the costs are high, but can be higher if the cost parameter α is high enough. The intuition is that if α is high, a high standard is bad because everybody pays high costs. But with a label, only the higher types will consume the label and these are the ones that are willing to pay most, making it attractive to set the quality relatively high.

Finally, note that the threshold value of $\bar{\theta}$ for which not everybody consumes the label or standard is lower under the labeling policy, where this threshold is given by $(3/2)\theta^a - \alpha\underline{s} - (1/2)\beta$, than under standards, where the threshold is $(3/2)\theta^a$. With a fixed average θ^a, a higher $\bar{\theta}$ means a lower corresponding $\underline{\theta}$. Increasing the quality of a label means that low types will consume $\underline{s}$. Increasing the standard means pushing the low types out of the market. Pushing consumers out of the market is worse than making them consume a nonlabeled product, which should therefore be prevented. Hence, the standard is chosen such that low types stay longer in the market.

5.4 Standards or labels?

We are now ready to formalize the idea mentioned in the introduction that the trade-off between a standard and a label is that of costs versus flexibility in the producers' choices of technology. One the hand brings the label a cost to the consumer in terms of a higher price. It is clear that if the costs parameter β is very high, a standard is to be preferred over a label, everything else equal. Also, when the cost parameter is very low, a label is unambiguously to be preferred. Indeed, if $\beta = 0$, then the welfare under a label is weakly higher than under a standard, since the government can set $\tau = \sigma^*$ in which all consumers who want to can achieve the same utility level under a labeling scheme than with standards, and some can achieve higher utility by consuming quality $\underline{s}$ instead of the standard (or refrain from consuming). For a particular range of heterogeneity, the welfare of labeling is in this case strictly higher.

The other part of the trade-off is that labeling is more flexible in providing quality in the customers' needs. Intuitively, labeling leads to higher welfare than a standard if consumers are heterogenous in their taste parameter. This intuition turns out to be only partly correct. In a few steps, we show the following result: labeling is better than standards only for an interval of heterogeneity. For either small heterogeneity or high heterogeneity, a standard is to be preferred.

First, for notational convenience we introduce some more notation:

Notation 1 $\eta \equiv (3/2)\theta^a - \alpha\underline{s} - (1/2)\beta$.

The parameter η is the threshold for which all types are willing to buy the label. Consider then the following result:

Lemma 3 *If $\bar{\theta} < \eta$ then $W^l \leq W^s$ with strict inequality for $\beta > 0$.*

If $\bar{\theta} < \eta$, the optimal standard and label would be such that all types would consume the standard or the labeled product respectively. In essence, no consumer would refrain from consuming or buy the unlabeled good. The intuition is straightforward. If $\beta = 0$, the optimal standard and label coincide: $\sigma^* = \tau^*$. Clearly welfare would be identical. If the cost parameter β increases, but still all consumers buy the label, the welfare under a labeling policy must necessarily be lower. They still prefer the label to $\underline{s}$, but the costs decrease their utility.

Next, consider:

Lemma 4 *If* $\eta \leq \bar{\theta} < (3/2)\theta^a$ *and* $\beta = 0$ *then* $W^l > W^s$.

For values of $\bar{\theta}$ within this range, the welfare of standards is independent of heterogeneity and that of labeling is increasing in heterogeneity. As $W^l = W^s$ at $\bar{\theta} = \eta$ for $\beta = 0$ this lemma follows straightforwardly. As a final intermediate step assume that $\bar{\theta} \geq (3/2)\theta^a$ and consider:

Lemma 5 *Let* $\Delta W(\bar{\theta}, \beta) \equiv W^l - W^s$. *Then under assumptions 1 and 2, and with* θ^a *fixed,* $\Delta W_{\bar{\theta}}(\bar{\theta}, \beta) \geq 0$ *for* $\bar{\theta} \leq \xi$ *and* $\Delta W_{\bar{\theta}}(\bar{\theta}, \beta) < 0$ *for* $\bar{\theta} > \xi$, *where* $\xi \equiv \theta^a + \frac{1}{3}\sqrt{3}\sqrt{(3\theta^a(\theta^a - 2\alpha\underline{s} - \beta) + (2\alpha\underline{s} + \beta)^2}$.

(A subscript denotes a partial derivative.) Thus, up to a certain threshold level of $\bar{\theta}$, ξ, the welfare of labeling is increasing faster in heterogeneity than that of standards. Beyond this threshold level, the reverse holds. Finally note that W^l is always decreasing in β. Putting things together, the following result is obtained:

Proposition 12 *If heterogeneity is small, welfare under standards is always weakly higher than under a label and strictly higher for any* $\beta > 0$. *For intermediate heterogeneity there exists an interval where labels outperform standards provided that the cost parameter* β *is not too high. For high heterogeneity standards outperform labels again if* β *is high enough.*

Although not all the thresholds levels for which labels are welfare improving compared to standards have been derived explicitly, the picture is clear in qualitative terms. For intermediate levels of heterogeneity labels are better than standards. The size of this interval increases as the cost parameter β decreases.

The result of proposition 12 is pictured in the figure below. The figure is drawn for a positive cost parameter β. Up to the point η, welfare of standards and labeling are both constant in heterogeneity, though welfare of labeling is lower. After this point, the welfare of labeling starts to increase and surpasses that of standards. The difference in welfare $\Delta W \equiv W^l - W^s$ starts to increase. After the point ξ the difference decreases. The curve shifts upward for lower values of β and would be positive or zero over the whole domain for $\beta = 0$. The curve shifts downward for higher values of β, and the domain where ΔW is positive ceases to exist for high enough values of β.

The result that standards are better for sufficiently high heterogeneity is perhaps somewhat counterintuitive and deserves some more attention. Under a minimum standard scheme, consumers have the choice between buying the product

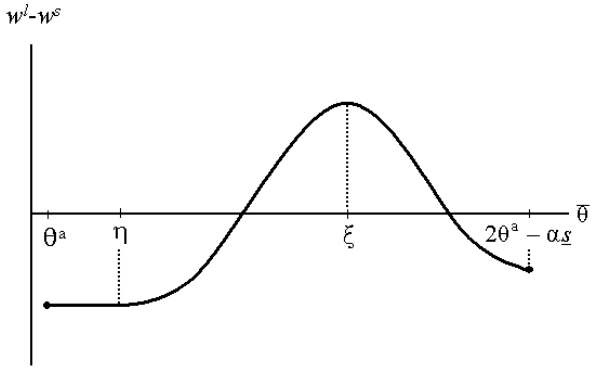

FIGURE 5.1.

at the quality of the standard, or refrain from buying. The consumer who is indifferent is characterized by $\tilde{\theta}$. Under a labeling scheme, the choice is between a labeled and a nonlabeled product. The consumer who is indifferent is characterized by $\hat{\theta}$. Consider any $\bar{\theta} \geq \xi$. Under that condition, $\tilde{\theta} < \hat{\theta}$. It follows that $(\bar{\theta} - \tilde{\theta}) > (\bar{\theta} - \widehat{\theta})$. In words: the fraction of consumers who consumes the highest available technology is greater under a standard $(\bar{\theta} - \tilde{\theta})$ than under a labeling policy $(\bar{\theta} - \widehat{\theta})$. This is intuitive: under a labeling scheme the alternative is a nonlabeled good which is still better than the alternative under a minimum standard (refraining from buying). Thus, a higher fraction will consume the highest quality under a minimum standard because the alternative is worse. Then, when the heterogeneity increases, both the optimal standard and the optimal label increase at the same rate $1/(3\alpha)$. Under the standard, however, a larger fraction of the consumer benefits from this increase. This effect is becoming more important for higher values of $\bar{\theta}$, when heterogeneity increases.

5.5 Discussion and conclusions

In this chapter we have formalized some of the basic trade-offs between labeling and standards. Conditions under which labeling is welfare improving are derived.

We now discuss the importance of the main underlying assumptions, and the consequences of relaxing them.

The key assumption underlying the model is that some consumers are willing to pay a price premium for goods that involve less social externalities of production. This is a necessary requirement for labeling to be an effective instrument. Although there is evidence that consumers are indeed willing to pay a price premium, the evidence is still somewhat mixed (Bjørner et al.[2002]).

Assumption 1 is a relatively innocent assumption. Allowing for uncovered markets only changes the baseline welfare level W^I but qualitatively speaking it does not affect any of the conclusions.

Assumption 2 ensures that there is a scope for government intervention in the first place. If assumption 2 does not hold, then the optimal label is such that all consumers buy the unlabeled good. In this case, no intervention is needed other than possibly a standard. Hence, this assumptions is a first check for the relevance of policy analysis.

We end the chapter by suggesting extensions to the model that seem necessary for a well founded policy and that we intend to incorporate in future studies.

First of all, goods with social externalities of production often have a public good character. Next to say a feeling of guilt of buying these products, they may also care about the externality per se. The case of polluting production methods belongs to this category. All people may care about a cleaner environment, but there are clear incentives for free-riding. Even if a group of consumers is willing to contribute to reduce the externality by buying labeled goods, there will likely also be a group of consumers who prefer the nonlabeled goods. The latter group imposes a negative externality on society. When the group of nonlabeled product buyers is large, a labeling policy is hardly effective in reducing the externality. Standards circumvent this problem as it does not allow for consumers who wish not to contribute.

Second, it is interesting to take note of the results of chapter 7 that deals with self-serving biases in information procession. In the current chapter we have assumed that some consumers may for instance experience a feeling of social irresponsibility when they buy a good with high social externalities, inducing a feeling of guilt. It is thus argued that they are willing to pay a price premium to avoid any feelings of guilt.

If this is related to the results of chapter 7, it becomes clear that consumers may nevertheless be motivated to stay ignorant about labeled products. It is clearly best for the consumers to buy a cheap good without reflecting on the negative social externalities that the product brings along. There is ample evidence from psychology that many people indeed process information in a biased way. Moreover, this seems to be mostly going on when their self-concept is at stake, such as when people behave in a way that does not support their view of being decent (Aronson [1988]). This is also likely to be the case for some of the products considered in this chapter. The taste parameter itself is then a function of efforts in gathering information, which on its turn is depending on the costs of acquiring information. The more information is gathered about the wrongdoing of a production method, the higher the willingness to pay for a product with a higher social responsibility index. A labeling policy may well be very effective as it raises the costs of ignoring the information much more than, say, a tax would.

A second relation with chapter 7 is the following. In this chapter, it is analyzed which policy instrument is better given the tastes of people. Thus, the feelings of guilt make a labeling policy possible. Chapter 7 on the other hand, learns us something how to foster such feelings because it endogenizes the willingness to pay in relation to subsidies.

Third, the assumption of perfect competition may be relaxed to the case where firms have some market power. This can have interesting consequences. Take for example the case where there are only two firms. Under imperfect information and a minimum standard, Bertrand price competition drives profits to zero. However, if one firm supplies a labeled product and the other a non-labeled product, then their products are differentiated and they exercise some market power (for a more elaborate treatment see Bansal [2002]). In this way, labeling has a negative effect on efficiency by increasing market power which may or may not be outweighed by the larger array in the variety of goods.

Fourth, the impact on trade has not been discussed. One common theme in the discussion on the implementation of labeling is that it imposes hidden trade barriers (see in particular Keyzer [2002]). If consumers are willing to pay premiums for clean technology, then this may well be in the advantage of producers in developed countries who have better knowledge about clean technologies.

Finally, it is assumed that the government is perfectly informed about the production technology. However, it may be hard for the government (or any other

certifying party) to obtain precise information on all the stages of the production process, especially if many small producers are involved. Uncertainty on the governments side can have a great impact on the optimal policy instrument (see Baumol and Oates [1988]). Furthermore, consumers may become skeptical about paying price premiums if the contents of a labeled product is open to debate (see Mason [2002]).

5.6 Appendix

This appendix contains all the proofs and lemmas in the main text with the exception of the proofs of propositions 8 and 10 and of lemmas 2 and 4 which follow from straightforward substitution and/or are elaborated upon in the main text and are for that reason suppressed.

Proof of proposition 9.
Most of the proof is contained in the main text or is straightforward. To complete the proof, note that W^s is continuously decreasing in σ between the unconstrained optimum (unconstrained in the sense of ignoring the constraint $\sigma \geq \underline{s}$) and the point where $\tilde{\theta}$ would equal $\bar{\theta}$, and constant thereafter. Hence, if the unconstrained optimum is not within the interval $[\underline{s}, \infty)$ then the constrained optimum is indeed at $\underline{s}$. ■

Proof of proposition 11.
Integrating out equation (5.12) gives:

$$(\bar{\theta} - \widehat{\theta})((1/2)(\bar{\theta} + \widehat{\theta}) - \beta - 2\alpha\tau) = 0, \tag{5.13}$$

as the first order condition. If $\bar{\theta} = \widehat{\theta}$ all consumers purchase the nonlabeled good and welfare is as under imperfect information. With $\bar{\theta} \neq \widehat{\theta}$ and using equation (5.10) we can see that the first order condition (5.13) is satisfied for $\tau = (1/(3\alpha))(\bar{\theta} + \alpha\underline{s} - \beta)$ and consequently $\widehat{\theta} = (1/3)(\bar{\theta} + 4\alpha\underline{s} + 2\beta)$. Hence, $\bar{\theta} = \widehat{\theta}$ occurs at $\bar{\theta} = 2\alpha\underline{s} + \beta$. For any smaller $\bar{\theta}$ the optimal quality would be such that even the highest type would consume the nonlabeled product but this case is ruled out by assumption 1. With the foregoing expressions we can also see that $\underline{\theta} = \widehat{\theta}$ occurs at $\overline{\theta} = \frac{3}{2}\theta^a - \alpha\underline{s} - \frac{1}{2}\beta$. In this case, substitution of $\underline{\theta} = \widehat{\theta}$ into equation gives the optimal quality $\tau = (\theta^a - \beta)/(2\alpha)$. These are the unconstrained optimal qualities. Since W^l is continuously decreasing in τ between the unconstrained optimum (with respect to $\underline{s}$) and constant after $\widehat{\theta} = \bar{\theta}$, the constrained optimum is indeed $\underline{s}$ if the unconstrained optimal τ is not element of $[\underline{s}, \infty)$. ■

Proof of lemma 3.
For $\bar{\theta} \leq \eta$, $W^s = (\theta^a)^2/(4\alpha)$. It is straightforward to show that in this case $W^l = (1/2)\tau^*(\bar{\theta}+\underline{\theta}) - \alpha\tau^{*2} - \beta(\tau^* - \underline{s})$. This can be rewritten as $W^l = (\theta^a - \beta)^2/(4\alpha) + \beta\underline{s}$.

For $\beta = 0$, $W^l = W^s$. For any $\beta > 0$, $W^l < W^s$ since by assumption 2 it must be the case that $2\theta^a > 4\alpha\underline{s} + \beta$. ■

Proof of lemma 5.
1. $\bar{\theta} \leq 3\underline{\theta} \Leftrightarrow \bar{\theta} \leq (3/2)\theta^a$. First note that it follows immediately from proposition 9 that W^s is independent of both β and $\bar{\theta}$. It is also straightforward to show that W^l is decreasing in β. We continue by proving that W^l is improving in $\bar{\theta}$.

All types $\theta \in [\underline{\theta}, \widehat{\theta})$ consume the unlabeled product of quality s, all types $\theta \in [\widehat{\theta}, \overline{\theta}]$ consume the labeled product of quality τ. Hence, the welfare under the optimal label τ^* is given by:

$$W^l = \int_{\underline{\theta}}^{\widehat{\theta}} \frac{1}{\overline{\theta} - \underline{\theta}}(\theta\underline{s} - \alpha\underline{s}^2)d\theta + \int_{\widehat{\theta}}^{\overline{\theta}} \frac{1}{\overline{\theta} - \underline{\theta}}(\theta\tau^* - \alpha\left(\tau^*\right)^2 - \beta(\tau^* - \underline{s}))d\theta. \quad (5.14)$$

We examine the behavior of W^l around the optimum with respect to changes in $\bar{\theta}$. By the implicit function theorem we have (ignoring other parameters):

$$\left.\frac{dW^l(\overline{\theta}, \tau(\overline{\theta}))}{d\overline{\theta}}\right|_{\tau=\tau^*} = \frac{\partial W^l(\cdot)}{\partial\overline{\theta}}. \quad (5.15)$$

In the rest of the proof we suppress the condition that $\tau = \tau^*$ but it is assumed to hold throughout. Taking the partial derivative of W^l with respect to $\overline{\theta}$ using the rule of Leibniz (and recalling that $\underline{\theta} = 2\theta^a - \overline{\theta}$)gives:

$$\begin{aligned}\frac{\partial W^l}{\partial\overline{\theta}} &= \int_{\underline{\theta}}^{\widehat{\theta}} -\frac{1}{2(\overline{\theta} - \theta^a)^2}(\theta\underline{s} - \alpha\underline{s}^2)d\theta + \frac{\partial\widehat{\theta}}{\partial\overline{\theta}} \cdot \frac{1}{2(\overline{\theta} - \theta^a)}\left[\widehat{\theta}\underline{s} - \alpha\underline{s}^2\right] \quad (5.16)\\ &+\frac{1}{2(\overline{\theta} - \theta^a)}\left[\underline{\theta}\underline{s} - \alpha\underline{s}^2\right] + \int_{\widehat{\theta}}^{\overline{\theta}} -\frac{1}{2(\overline{\theta} - \theta^a)^2}(\theta\tau - \alpha\tau^2 - \beta(\tau - \underline{s}))d\theta\\ &+\frac{1}{2(\overline{\theta} - \theta^a)}\left[\overline{\theta}\tau - \alpha\tau^2 - \beta(\tau - \underline{s})\right]\\ &-\frac{\partial\widehat{\theta}}{\partial\overline{\theta}} \cdot \frac{1}{2(\overline{\theta} - \theta^a)}\left[\widehat{\theta}\tau - \alpha\tau^2 - \beta(\tau - \underline{s})\right].\end{aligned}$$

The second and the last term of the RHS cancel out against each other. Integrating out results in:

$$\begin{aligned}\frac{\partial W^l}{\partial\overline{\theta}} &= -\frac{1}{2(\overline{\theta} - \theta^a)^2}[\widehat{\theta} - \underline{\theta})(\frac{1}{2}\underline{s}(\widehat{\theta} + \underline{\theta}) - \alpha\underline{s}^2)] + \frac{1}{2(\overline{\theta} - \theta^a)}\left[\underline{\theta}\underline{s} - \alpha\underline{s}^2\right] \quad (5.17)\\ &-\frac{1}{2(\overline{\theta} - \theta^a)^2}[\overline{\theta} - \widehat{\theta})(\frac{1}{2}\tau(\overline{\theta} + \widehat{\theta}) - \alpha\tau^2 - \beta(\tau - \underline{s})]\\ &+\frac{1}{2(\overline{\theta} - \theta^a)}\left[\overline{\theta}\tau - \alpha\tau^2 - \beta(\tau - \underline{s})\right].\end{aligned}$$

After some manipulations, this can be rewritten as:

$$\frac{\partial W^l}{\partial \overline{\theta}} = \frac{\tau - \underline{s}}{2(\overline{\theta} - \theta^a)^2} \left[\frac{1}{2}(\widehat{\theta} - \overline{\theta})^2 + (\widehat{\theta} - \theta^a)(\overline{\theta} - \alpha(\underline{s} + \tau) - \beta) \right]. \tag{5.18}$$

It is useful to proceed as follows. We distinguish between $\overline{\theta} < 3\theta^a - 4\alpha\underline{s} - 2\beta$ and $\overline{\theta} \geq 3\theta^a - 4\alpha\underline{s} - 2\beta$. The former case implies that $\widehat{\theta} < \theta^a$, the latter case that $\widehat{\theta} > \theta^a$. For each case we determine when $\widehat{\theta}$ hits the boundary. If $\widehat{\theta} < \theta^a$ and $\widehat{\theta}$ hits the boundary, then it must be the case that $\widehat{\theta} = \underline{\theta}$. (It cannot be the case that $\widehat{\theta} = \overline{\theta}$ since $\overline{\theta} \geq \theta^a$ and $\widehat{\theta} < \theta^a$). Similarly, if $\widehat{\theta} > \theta^a$ and $\widehat{\theta}$ hits the boundary, then it must be the case that $\widehat{\theta} = \overline{\theta}$. The thresholds are given by $\overline{\theta} = \frac{3}{2}\theta^a - \alpha\underline{s} - \frac{1}{2}\beta$ (which follows from setting $\widehat{\theta} = \underline{\theta}$) and $\overline{\theta} = 2\alpha\underline{s} + \beta$ (following from setting $\widehat{\theta} = \overline{\theta}$). Note that as a consequence of Assumption 2, $3\theta^a - 4\alpha\underline{s} - 2\beta > \frac{3}{2}\theta^a - \alpha\underline{s} - \frac{1}{2}\beta$. Note furthermore that Assumption 2 also implies that $\overline{\theta} > 2\alpha\underline{s} + \beta$ which rules out the possibility that $\widehat{\theta} = \overline{\theta}$. We are left with the following three different (exhaustive) cases:

Case (i): $\overline{\theta} < \frac{3}{2}\theta^a - \alpha\underline{s} - \frac{1}{2}\beta$.

In this case, $\widehat{\theta} = \underline{\theta}$. Substituting into equation (5.18) gives

$$\frac{\partial W^l}{\partial \overline{\theta}} = \frac{\tau - \underline{s}}{2(\overline{\theta} - \theta^a)^2} \left[\frac{1}{2}(\underline{\theta} - \overline{\theta})^2 + (\underline{\theta} - \theta^a)(\overline{\theta} - \alpha(\underline{s} + \tau) - \beta) \right], \tag{5.19}$$

which is equal to:

$$\begin{aligned}
\frac{\partial W^l}{\partial \overline{\theta}} &= \frac{\tau - \underline{s}}{2(\overline{\theta} - \theta^a)^2} \left[2(\underline{\theta} - \theta^a)^2 + (\underline{\theta} - \theta^a)(\overline{\theta} - \alpha(\underline{s} + \tau) - \beta) \right] \qquad (5.20) \\
&= \frac{(\tau - \underline{s})\,(\underline{\theta} - \theta^a)}{2(\overline{\theta} - \theta^a)^2} \left[2(\underline{\theta} - \theta^a) + (2\theta^a - \underline{\theta} - \alpha(\underline{s} + \tau) - \beta) \right] \\
&= \frac{(\underline{\theta} - \theta^a)}{2(\overline{\theta} - \theta^a)^2} \left[(\tau - \underline{s})\,(\underline{\theta} - \alpha(\underline{s} + \tau) - \beta) \right] \\
&= \frac{(\underline{\theta} - \theta^a)}{2(\overline{\theta} - \theta^a)^2} \left[\left(\underline{\theta}\tau - \alpha\tau^2 - \beta(\tau - \underline{s}) \right) - \left(\underline{\theta}\underline{s} - \alpha\underline{s}^2 \right) \right] \\
&= \frac{(\underline{\theta} - \theta^a)}{2(\overline{\theta} - \theta^a)^2} \left[U(\tau, \underline{\theta}) - U(\underline{s}, \underline{\theta}) \right] \\
&\geq 0.
\end{aligned}$$

The last inequality follows by a revealed preference argument, i.e. the fact that the lowest type (weakly) prefers the label.

Case (ii): $\frac{3}{2}\theta^a - \alpha\underline{s} - \frac{1}{2}\beta \leq \overline{\theta} < 3\theta^a - 4\alpha\underline{s} - 2\beta$.

For values of $\overline{\theta}$ within this interval, $\hat{\theta} = \left(\overline{\theta} + 4\alpha\underline{s} + 2\beta\right)/3$. Recall also that the optimal label is in this case given by $\tau^* = \left(\overline{\theta} + \alpha\underline{s} - \beta\right)/3\alpha$. Before proceeding, note that the two conditions on $\bar{\theta}$ imply that $\bar{\theta} > 2\alpha\underline{s} + \beta$ otherwise the case does not exist. Substituting $\hat{\theta}$ and τ^* into equation (5.18), and expanding all terms gives:

$$\begin{aligned}\frac{\partial W^l}{\partial\overline{\theta}} &= \frac{\tau - \underline{s}}{2(\overline{\theta} - \theta^a)^2}\left[\frac{4}{9}\overline{\theta}^2 - \frac{4}{9}\alpha\underline{s}\overline{\theta} - \frac{2}{9}\beta\overline{\theta} - \frac{8}{9}\alpha^2\underline{s}^2 - \frac{8}{9}\alpha\beta\underline{s}\right. \\ &\quad \left. -\frac{2}{9}\beta^2 - \frac{6}{9}\overline{\theta}\theta^a + \frac{12}{9}\alpha\underline{s}\theta^a + \frac{6}{9}\beta\theta^a\right].\end{aligned} \tag{5.21}$$

This can also be written as:

$$\frac{\partial W^l}{\partial\overline{\theta}} = \frac{\tau - \underline{s}}{2(\overline{\theta} - \theta^a)^2}\left[\frac{4}{9}(\overline{\theta} - 2\alpha\underline{s} - \beta)\left(\overline{\theta} - \left(\frac{3}{2}\theta^a - \alpha\underline{s} - \frac{1}{2}\beta\right)\right)\right]. \tag{5.22}$$

By inspection of the assumptions and the note we made before, it is then easily seen that this expression must be positive.

Case (iii): $\overline{\theta} \geq 3\theta^a - 4\alpha\underline{s} - 2\beta$.

We now have $\hat{\theta} = \left(\overline{\theta} + 4\alpha\underline{s} + 2\beta\right)/3$ again. This time, rewrite equation (5.18) as:

$$\begin{aligned}\frac{\partial W^l}{\partial\overline{\theta}} &= \frac{1}{2(\overline{\theta} - \theta^a)^2}\left[\frac{1}{2}(\tau - \underline{s})(\widehat{\theta} - \overline{\theta})^2\right. \\ &\quad \left. +(\widehat{\theta} - \theta^a)\left[(\overline{\theta}\tau - \alpha\tau^2 - \beta(\tau - \underline{s}) - (\overline{\theta}\underline{s} - \alpha\underline{s}^2)\right]\right]\end{aligned} \tag{5.23}$$

which is equal to:

$$\frac{\partial W^l}{\partial\overline{\theta}} = \frac{1}{2(\overline{\theta} - \theta^a)^2}\left[\frac{1}{2}(\tau - \underline{s})(\widehat{\theta} - \overline{\theta})^2 + (\widehat{\theta} - \theta^a)\left[U(\tau, \overline{\theta}) - U(\underline{s}, \overline{\theta})\right]\right]. \tag{5.24}$$

By a revealed preference argument, it is easy to see that it must be the case that $U(\tau, \overline{\theta}) - U(\underline{s}, \overline{\theta}) > 0$ since $\overline{\theta} > \widehat{\theta}$. Under the assumptions made we also have $\widehat{\theta} - \theta^a > 0$. Hence, we conclude that in this case $\partial W^l/\partial\overline{\theta} > 0$.

2. $\bar{\theta} > 3\underline{\theta} \iff \bar{\theta} > (3/2)\theta^a$. Note first that it cannot be the cast that $\tilde{\theta} > \underline{\theta}$ and $\hat{\theta} = \underline{\theta}$. Hence, if the market for standards is not covered, then also not with labels. This on its turn means that we can substitute for $\hat{\theta}$ in equation (5.18) giving us equation (5.22) for $\partial W^l/\partial\overline{\theta}$. It is easy to show that:

$$\left.\frac{\partial W^s}{\partial\overline{\theta}}\right|_{\sigma=\sigma^*} = \frac{2\overline{\theta}^2(\bar{\theta} - \frac{3}{2}\theta^a)}{27\alpha(\bar{\theta} - \theta^a)^2}. \tag{5.25}$$

Subtracting this from (5.22) gives us (again suppressing notation that we examine behavior around the optimum):

$$\frac{\partial(W^l - W^s)}{\partial\bar{\theta}} = \chi\left[(\bar{\theta} - 2\alpha\underline{s} - \beta)^2\left(\bar{\theta} - \left(\frac{3}{2}\theta^a - \alpha\underline{s} - \frac{1}{2}\beta\right)\right) - \bar{\theta}\left(\bar{\theta}^2 - \frac{3}{2}\theta^a\right)\right], \tag{5.26}$$

where $\chi \equiv 2/(27\alpha(\bar{\theta} - \theta^a)^2)$. This can, after some manipulations, be rewritten as:

$$\frac{\partial(W^l - W^s)}{\partial\bar{\theta}} = -\chi(2\alpha\underline{s} + \beta)\left[\frac{3}{2}\bar{\theta}^2 - 3\theta^a\bar{\theta} + 3\alpha\underline{s}\theta^a + \frac{3}{2}\beta\theta^a - (2\alpha\underline{s} + \beta)^2\right]. \tag{5.27}$$

It is then straightforward to show that:

$$\frac{\partial(W^l - W^s)}{\partial\bar{\theta}} = 0 \Leftrightarrow \tag{5.28}$$

$$\bar{\theta} = \xi \equiv \theta^a \pm \frac{1}{3}\sqrt{3}\sqrt{3\theta^a(\theta^a - 2\alpha\underline{s} - \beta + (2\alpha\underline{s} + \beta)^2}. \tag{5.29}$$

Since $\bar{\theta} \geq \theta^a$ we only need to consider the positive root. For any $\bar{\theta} \lesseqgtr \xi$, $\partial(W^l - W^s)/\partial\bar{\theta} \gtreqless 0$. This proves the lemma. ■

Proof of proposition 12.
This follows in a straightforward manner from lemmas 3, 4, and 5. ■

6

Rewards, Self-confidence, and Motivation: The Hidden Rewards of Rewards

6.1 Introduction

Many real-life situations concern relationships where no complete explicit contracts can be written down. Moreover, many situations do not allow the possibility of implicit contracts either. Incomplete contracts, however, can often give strong incentives to shirk (Williamson [1985], Fehr and Gächter [2000]). This raises some questions such as: why are efforts rewarded, and why are they made in the first place? A particularly intriguing question is why spontaneous rewards are sometimes given. This latter question is taken up in this chapter by examining the role of self-confidence and its effect on motivation.

There are several possible channels through which motivation can be stimulated. Most of economic theory is built on the assumption that monetary rewards motivate agents to make efforts. The channel through which rewards motivate is straightforward: agents care about money, and hence they are more willing to make efforts if this increases the probability of payments[1]. Social psychologists have identified another channel through which motivation can be stimulated, namely through changes in self-confidence. Whenever an agent is more confident

[0] This chapter is coauthored by Anton Souvorov. We thank Jean Tirole for helpful ideas. Part of the research was done during my stay at GREMAQ at the University of Toulouse 1, that was supported by the ENTER exchange and a Marie-Curie fellowship.

[1] More generally, rewards can give non-monetary benefits – payments in kind, promotions, recognition. What matters is that rewards directly affect the agent's utility.

about his ability to succeed, he is more likely to try to undertake a task. Thus, if one succeeds in making the agent more self-confident, his effort increases and a successful outcome is more likely.

The primary aim of this chapter is to study both channels in one simple model. Although both monetary rewards and self-confidence are elements in economic models, the self-confidence effect has rarely been studied: in most conventional models with asymmetric information, the agent has full information about his ability. Consequently, there is no role for the principal to give signals to the agent. By contrast, in this chapter it is assumed that the agent has incomplete information about his chances to attain a successful outcome, for example because he undertakes a task that is new to him. In this case, it is shown that monetary rewards can give credible signals about these chances and therefore influence the level of self-confidence.

Interaction effects between rewards and self-confidence are also examined in a recent paper by Bénabou and Tirole [2002a]. Their focus is on promised bonuses that are specified in a contract. The focus in this chapter is on discretionary rewards. Such unexpected bonuses are interesting in their own respect, because why would someone give a bonus that is not expected or specified in a contract anyway? A possible answer is that such rewards can increase self confidence. Or, in the words of Bénabou and Tirole [2002a, 22]:

> "Rewards that are discretionary (not contracted for) may well boost the agent's self-esteem or intrinsic motivation, because (...) the worker or child learns from the reward that the task was considered difficult (and therefore that he is talented), or that the supervisor is appreciative of, proud of, or cares about his performance – and that it is worth repeating it. (...) And receiving the reward is good news, because the agent initially did not know how to interpret his performance."

The main idea of the model in this chapter is that there is an agent who is only willing to make efforts if he has enough self-confidence that he will succeed in the task. However, it is difficult for him to assess the outcome. By giving a bonus, the principal can give a signal to the agent that his efforts resulted in a successful outcome last period. Thus, bonuses increase motivation in our model in two ways: first of all, they raise the agent's self-confidence, and he realizes that it

is worthwhile for him to continue working hard; secondly, foreseeing that bonuses will be given more frequently after successful outcomes, the agent works harder in the first place. Altogether, this can explain why sometimes "unexpected" rewards are given, even in a game with a finite number of rounds[2]. These results are in line with actual behavior in existing markets (Akerlof [1982]) as well as in laboratory experiments (Fehr and Gächter [2000]).

Some social psychologists have stressed that rewards need not necessarily increase self-confidence. In their view, rewards sometimes have "hidden costs" (e.g. Kohn [1993]). The hidden costs of rewards are the possible negative effects on self-confidence. There are two main causes for this. First, rewards can be perceived as controlling, thereby undermining self-determination. Second, they can carry an informational content which can be negative. For instance, in the model by Bénabou and Tirole [2002a], promising a high bonus in case of success may give a signal to the agent that the task in question will be difficult and that he is unlikely to succeed. For as long as rewards are given, the agent is motivated to make efforts. In the meantime, rewards lower his self-confidence. At the moment that rewards are withdrawn, the self-confidence effect persists and induces less motivation than before. Interestingly, there is a considerable body of evidence showing exactly this pattern (see Deci and Ryan [1985]).

The hidden cost component of rewards have not been found in studies where unexpected rewards are used (e.g., Deci, Koestner and Ryan [1999]). This is predicted by the basic model of this chapter: in any equilibrium, the principal has no incentive to give a discretionary bonus that decreases motivation in the next period. Thus, the rewards bring good news. Besides, they cannot be controlling since they are discretionary. However, with a slight modification, the model can replicate a negative correlation between rewards and self-confidence. One key condition is examined when a reward can lower self-confidence in equilibrium. This is the case when effort and ability are substitutes. Now, the bonus is a signal that the outcome was a failure. This reduces his self-confidence in his ability, which he compensates by making more efforts. However, even though the bonus reduces self-confidence, it stimulates the agent to work harder because he realizes he must make efforts in order to succeed.

[2] Unexpected in quotation signs, because in equilibrium the agent foresees the payment of a bonus with some probability.

The rest of the chapter is organized as follows. Section 6.2 presents the model and discusses the main assumptions. The derivations of equilibria are given in section 6.3. Section 6.4 describes some qualitative properties of the equilibrium. The results are discussed in section 6.5. After that, the results are related to the results of related literature in section 6.6. Finally, section 6.7 concludes.

6.2 The model

6.2.1 Preliminaries

This section describes the general setup of the game. Some of the basic ideas are closely related to the work by Bénabou and Tirole [2002a] and Souvorov [2003]. Where assumptions differ, this is made clear in the text.

In the game, there are two players: a principal and an agent. There is a finite number of periods, for simplicity set to two. The case with an infinite horizon has been studied briefly by Bénabou and Tirole [2002a] in a slightly different model.

The agent has to decide on his effort level. He chooses to make efforts or not: $e \in \{0, 1\}$. If the agent undertakes the task, i.e. $e = 1$, he incurs a cost of c in terms of disutility. Depending both on effort, e, and ability, θ, the outcome of the effort can be either a success, S, or a failure, F. The probability of success is given by:

$$prob(S \mid e) = e\theta. \tag{6.1}$$

In other words, ability and effort are complements. No effort induces a failure with certainty. In case of success, it yields a payoff equal to V to the agent, and W to the principal. A failure yields a payoff equal to zero for both. Both parties are risk-neutral and the agent is protected by limited liability.

The principal has to select a reward policy. In each period, he can offer a bonus $b \in \mathbb{R}^+$ to the agent.

6.2.2 The main assumptions

Most of the above description is relatively standard for a principal-agent game. The model, however, departs from most conventional models in several respects. Each of these are discussed in more detail.

Imperfect self-knowledge. Although it is not a usual assumption, the idea that people have only imperfect knowledge about their personal characteristics is plau-

sible (see for instance Bénabou and Tirole [2002b]). First because retrospective evaluations of past utilities are known not always to be reliable (Kahneman [1994]). Thus, based on retrospection, people make incorrect estimates about how they will feel about certain matters. Moreover, some situations are new to people. In this case, they do not have enough information about themselves to infer their ability. Someone who tries to quit smoking for the first time is unlikely to be able to guess how persistent he will be. This requires some learning, but learning opportunities are usually limited.

Imperfect self-knowledge in the current context means that people are not perfectly informed about their ability. They cannot foresee their ability to make a success out of it. The task they have to undertake is for example relatively new to them, or they have forgotten how well they did on this or a comparable task in the past. They do form an estimate about their ability. Based on this estimate, they form an estimate of their chances to succeed, which represents their self-confidence.

To make things concrete: suppose that the agent can be either one of two possible types, a high type with ability θ_H or a low type with ability $\theta_L < \theta_H$. His prior on being a high type is given by ρ. His self-confidence is then given by:

$$\rho\theta_H + (1-\rho)\theta_L. \tag{6.2}$$

Clearly, self-confidence is increasing in ρ. In the remainder of the chapter the parameters θ_H and θ_L are kept fixed, and with slight abuse of terminology, self-confidence is identified with the parameter ρ.

Non-contractibility of the bonus. The principal has the possibility to give a reward b to the agent. However, a crucial assumption in the model will be that the outcome is not observable to the agent or an outside party. The outcome is therefore private information to the principal. It follows that a reward contingent on the outcome cannot be specified in a contract, because the agent or third party would not be able to verify the truthfulness of the principal's claim. That is, the principal can always report a failure and no party can contest this claim.

The non-contractibility is one of the main departures from the model of Bénabou and Tirole [2002a]. They have analyzed the case where a contract can be written that specifies the bonus in advance. Of course, they also have to assume that the output is verifiable to the agent. The case of noncontractibility is interesting because first, in reality there are many situations where the bonus is indeed

noncontractible, and second, evidence from experiments show that the relation between bonuses and motivation differs depending on whether or not a bonus is specified in advance.

Intrinsic motivation. Even though no contract that specifies a bonus can be written, it is still assumed that agents have a motivation to make efforts. Economists usually takes rewards as the motivation to work. According to Frey [1997], many psychologists emphasize that the motivation to undertake a task can come from within the person. If they are motivated without apparent reward or environmental control, they are said to be intrinsically motivated. In the words of Deci and Ryan [1985, 43]:

> "Intrinsic motivation is the innate, natural propensity to engage one's interests and exercise one's capacities, and in so doing, to seek and conquer optimal challenges. Such motivation emerges spontaneously from internal tendencies and can motivate behavior even without the aid of extrinsic rewards or environmental controls".

It is undisputed that people are intrinsically motivated to do certain things: playing football, solving a puzzle, the list is endless. An assumption in this chapter is that people are indeed motivated for the task they have to undertake, even if they get no current rewards. This is not a completely innocent assumption. Even if people are intrinsically motivated to perform certain tasks, it does not follow that they are intrinsically motivated to do all possible thinkable tasks. However, the assumption is not crucial in the sense that the agent may also be motivated for expected rewards in the future, despite the absence of current rewards. The model allows for both interpretations.

The motivation of the agent is modeled as the value V in the model. To make things interesting, one additional assumption has to be made on V, namely that it cannot be directly observed. In other words, it is assumed that the agent is motivated to do a task for which the benefits come later in life. Thus, one can interpret V as the discounted value of payoffs later in life, be it extrinsic or intrinsic. The agent may be a pupil learning to play the piano. First, he needs to practice all kind of chords, a rather dull activity. The reward only comes when he is able to play a decent piece. The agent may also be a student studying for an exam, or writing an essay, not pleasant tasks for many people. His benefits

may be to get a job afterwards that he really likes. Or he may be a worker, who undertakes the task with the prospect of getting a promotion afterwards.

Asymmetric information. As explained, the agent is not sure about his ability to bring the task to a successful end. Moreover, the focus in on situations where even afterwards he does not get to know directly for sure whether it was a success or a failure. He only gets an imperfect signal about the outcome. On the other hand, the principal is able to observe the outcome. For instance, the pupil learning to play piano cannot really tell whether he is talented after a few sessions, but the principal can tell, having seen many pupils trying before this pupil. The same is true for the student, whose grade will only be imperfectly informative about his ability. This is certainly the case where the grade is dependent on the subjectivity of the teacher, as with an essay. For a worker, it may be the case that this is the first time he undertakes the task, or that his task is only a small part of a bigger whole he is part of, so that he is not able to judge the outcome based on his own information only.

Note that this assumption is contrary to most conventional principal-agent models, where the agent has more information rather than less. For example, in the classic job-market signalling model of Spence [1973] it is the agent who knows his ability, whereas the principal only knows the distribution of abilities among the population.

The private signal that the agent gets is given by $\sigma \in [0,1]$. This signal has a conditional distribution $G(\sigma \mid y) = G_y(\sigma)$ and density $g(\sigma \mid y) = g_y(\sigma)$, where y is the outcome of the task: $y \in \{S, F\}$. A higher σ is interpreted as good news in the sense that it is more likely that a success has occurred. To capture this idea, it is assumed that the likelihood function $l(\sigma)$ with

$$l(\sigma) \equiv \frac{g_S(\sigma)}{g_F(\sigma)}, \tag{6.3}$$

is an increasing function in σ. This is the monotone likelihood ratio property (MLRP).

The next section examines equilibrium behavior of the principal and agent. To focus on interesting cases, the following additional assumptions are made.

Assumption 3 *Were the agent to know his type, then he would only undertake the task without a bonus if he is a high ability type:* $\theta_L V < c < \theta_H V$.

As will be demonstrated shortly, no bonus is offered by the principal in the second period. If assumption 3 did not hold, then either the agent would never work in period 2, or he would always work, independent of his self-confidence. In both cases, there is no role for the principal to increase self-confidence. Thus, no bonus would be given in the first period either.

Furthermore, the following restriction is put on the likelihood ratio:

Assumption 4 *The likelihood ratio $l(\sigma)$ is continuous in σ and has full support on $[0, +\infty)$. Furthermore, the monotone likelihood ratio property (MLRP) is satisfied: $l(\sigma)$ is everywhere increasing in σ.*

The full support assumption simplifies matters. It is also used by Bénabou and Tirole [2002a] and Souvorov [2003] in related settings. The MLRP is an essential assumption in many models with asymmetric information.

Assumption 5 *In period 1, the agent undertakes the task: $e = 1$ in period 1.*

This last assumption is made to focus on the interesting aspect of the model, which is the behavior of the agent in period 2. Although conditions can be derived under which $e = 1$ is an equilibrium strategy in period 1, not much insight is gained from doing that.

6.2.3 Timing and summary of the game

Each period is divided in two subperiods. In each first subperiod, the agent decides to make effort or not: $e \in \{0, 1\}$. Effort costs c in terms of disutility. At the end of the first subperiod, the principal observes the outcome ($y \in \{S, F\}$) and the agent receives a private signal σ about the outcome. A success occurs with probability $e\theta$ and gives a payoff V to the agent and W to the principal. In the second subperiod, the principal determines his reward policy $b \in \mathbb{R}^+$.

Note also the following: at the beginning of the game, both the principal and the agent have the same prior ρ that the agent is of the high type, θ_H. To simplify, both of them observe effort. The signal σ is private information to the agent, but the conditional distribution functions are common knowledge.

6.3 Equilibrium behavior

Consider then the behavior by the agent if he does not know his type. First note that no bonus is ever offered to him in period 2. The intuition is simple: the bonus is costly to the principal, and since it can have no impact on strategies played in the past, he should never give a bonus at the final stage of the game. Therefore, given his posterior on being a high type ρ', the agent works in period 2 if and only if:

$$[\rho'\theta_H + (1-\rho')\theta_L]\, V \geq c. \tag{6.4}$$

His posterior on being a high type is depending on his prior of being a high type, the private signal, and the bonus he received in the first period. If the agent did not work in period 1, he receives no additional signals and his posterior remains at his prior ρ. Suppose then that the agent did work. Suppose also, quite generally, that the principal offers a reward b_S with probability x_S after success, and with probability x_F after a failure. Then, after a bonus b_S and a signal σ, the agent updates his prior on being a high ability type to ρ_S with:

$$\frac{\rho_S}{1-\rho_S} = \frac{\rho}{1-\rho}\frac{\theta_H g_S(\sigma)x_S + (1-\theta_H)g_F(\sigma)x_F}{\theta_L g_S(\sigma)x_S + (1-\theta_L)g_F(\sigma)x_F}. \tag{6.5}$$

Based on expression (6.4), the following proposition can be derived:

Proposition 13 *Then there exist threshold levels of his initial self-confidence $\tilde{\rho}_S$ and $\tilde{\rho}_F \geq \tilde{\rho}_S$ such that if the agent worked in the first period, was given a bonus b_S and observed a private signal σ, in the second period the agent: (i) never works if $\rho < \tilde{\rho}_S$, (ii) always works if $\rho \geq \tilde{\rho}_F$, and (iii) if $\rho \in [\tilde{\rho}_S, \tilde{\rho}_F)$, works if and only if he has sufficiently good news (σ high enough).*

Proof. All proofs are collected in the appendix. ■

In other words, if the agent has a sufficiently low self-confidence, he will not work in period 2 whatever the signal or bonus he gets. If he has sufficiently high self-confidence, he will work in period 2 even for the worst possible signal or bonus he could get. For intermediate levels of initial self-confidence, he is sensitive to the news he gets. If he gets good news, he will work in period 2. If he gets bad news, he will not work in period 2.

All the specific thresholds are derived in the appendix. However, it is interesting to note the following. Suppose that his initial self-confidence is in the intermediate

range: $\rho \in [\tilde{\rho}_S, \tilde{\rho}_F)$. He will work if and only if after observing b_s, the signal σ exceeds $\tilde{\sigma}$, with:

$$l(\tilde{\sigma}) = \frac{x_F}{x_S} A(\rho). \tag{6.6}$$

Here, A is a parameter depending on initial self-confidence[3]. It is easy to show that this parameter is decreasing in initial self-confidence. Thus, the threshold signal $\tilde{\sigma}$ is decreasing in initial self-confidence (for a given principal's policy). This increases the set of private signals for which the agent will work in period 2. In sum, for a higher initial self-confidence, it becomes more likely that the agent will work in period 2. Note furthermore that the threshold signal is decreasing in the probability that the bonus is paid after a success (x_S), which makes it more likely that the outcome was a success, and increasing in the probability that the bonus is paid after a failure (x_F), which makes it more likely that the outcome was a failure.

Consider now the behavior by the principal in period 1. For $\rho \notin [\tilde{\rho}_S, \tilde{\rho}_F)$, it should be clear that the principal gives no bonus in period 1, since he is not able to influence behavior in period 2. From here on, the focus will therefore be on $\rho \in [\tilde{\rho}_S, \tilde{\rho}_F)$. For these values of ρ, the principal may try to signal through a bonus that the agent was successful in period 1. The reason is that for this interval of initial self-confidence, it can happen that the agents works and is successful, but receives a private signal that is below the threshold $\tilde{\sigma}$. This possibility is less likely if the principal increases the probability of giving a bonus b_S after success, or decreasing the probability of giving a bonus b_S after failure.

However, there are typically many perfect Bayesian equilibria in this game. This is a common feature of signalling games. Some of these equilibria are less reasonable than others, primarily because out-of-equilibrium beliefs are unrestricted. To reduce the set of possible equilibria, it is useful to restrict the set of possible out-of-equilibrium beliefs. Here, a relatively standard refinement is applied, which is the *Never a Weak Best Response* (NWBR) refinement for signalling games, which is, in the current context, equivalent to the *universal divinity* criterion (see Cho and Kreps [1987] and Fudenberg and Tirole [1991]). A general definition is given

[3] For the values of ρ in the interval from $\tilde{\rho}_S$ to $\tilde{\rho}_F$, $A(\rho)$ is equal to the ratio of the expected loss from working after a failure to the expected gain from working after success. In the absence of any intermediate information the agent would work if and only if $A(\rho) \leq 1$.

in the appendix. Reformulated for the current model, the refinement requires the following:

Assumption 6 *Fix an equilibrium outcome. For any out-of-equilibrium bonus* $\hat{b}$*, let* $\bar{\sigma}_F(\hat{b})$ *be the agent's reaction*[4] *to this bonus that makes the principal indifferent between the expected payoff following* $\hat{b}$ *(the agent playing* $\bar{\sigma}_F(\hat{b})$*) and his expected equilibrium payoff if a failure has occurred. Then, if for this reaction* $\bar{\sigma}_F(\hat{b})$ *the principal strictly prefers his expected equilibrium payoff to the expected payoff from deviating to* $\hat{b}$ *if a success has occurred, then the agent should believe that a failure occurred after getting* $\hat{b}$*. If for this reaction* $\bar{\sigma}_F(\hat{b})$ *the principal would strictly prefer to deviate to* $\hat{b}$ *if a success has occurred, the agent should believe that a success has occurred after getting* $\hat{b}$*.*

The intuition behind the assumption is that if the principal wants to deviate to an out-of-equilibrium bonus $\hat{b}$ after success for any agent's reaction making him want to deviate after failure, the agent should believe that success is (infinitely) more likely. Similarly for the failure. That is, given an equilibrium outcome, if the principal expects $\bar{\sigma}_F(\hat{b})$ after deviating to an out-of-equilibrium bonus, and he would strictly prefer to deviate after a success but be indifferent after a failure, the agent should believe a success has occurred.

Under the assumptions made, the equilibrium is unique (see the appendix). Depending on initial self-confidence, the equilibrium is either a pooling equilibrium or a semi-separating equilibrium. In the pooling equilibrium, the principal always offers the same reward. In the semi-separating equilibrium, the principal always offers the same reward after success, but randomizes between two rewards after a failure. These two equilibria are examined in detail below.

Before continuing, the following lemma will prove to be helpful. Define $\tilde{\sigma}$ as the threshold signal such that the agent works for all $\sigma \geq \tilde{\sigma}$. Note that in equilibrium this threshold signal depends on the bonus because the bonus provides additional information. One can therefore write $\tilde{\sigma} = \tilde{\sigma}(b)$. Then:

Lemma 6 *In any equilibrium, for* $b_1 > b_2$*, it must be that* $\tilde{\sigma}(b_1) < \tilde{\sigma}(b_2)$*.*

In other words, a higher bonus increases the likelihood of effort. This is easy to see. Suppose b_1 and b_2 are equilibrium bonuses but $\tilde{\sigma}(b_1) > \tilde{\sigma}(b_2)$. Then a

[4] Somewhat loosely, we shall call $\bar{\sigma}(b)$ (or $\sigma^*(b)$) the agent's strategy to work after getting bonus b if and only if his signal exceeds $\bar{\sigma}(b)$.

lower bonus increases the likelihood of effort. Clearly, this makes the principal unambiguously better off so that b_1 could not have been an equilibrium bonus.

6.3.1 A pooling equilibrium

Suppose that there is a *pooling* equilibrium, with $x_S = x_F = 1$. Thus, the same bonus bonus $\tilde{b} = b_s$ is always given independent of the outcome. Suppose also that given this bonus, the agent only works for signals exceeding $\tilde{\sigma} > 0$, where $l(\tilde{\rho}) = A$ is determined by (6.6). Denote by $\hat{\theta}_F$ and $\hat{\theta}_S$ the estimates by the principal of the agent's chances to succeed in period 2, conditional on failure and success in the first period. Thus:

$$\hat{\theta}_y = prob(S \text{ in period } 2 \mid y \text{ in period } 1). \tag{6.7}$$

It is assumed that the agent tried in period 1. The exact probabilities are given in the appendix. The expected payoffs for the principal is then given by:

$$\hat{\theta}_y(1 - G_y(\tilde{\sigma}))W - \tilde{b}. \tag{6.8}$$

Assume that the principal deviates from the equilibrium strategy, and offers a bonus $\hat{b} = \tilde{b} + \varepsilon$. Let $\hat{\sigma}$ be the agent's reaction to this offer which makes the principal indifferent between deviating or not after failure:

$$\hat{\theta}_F(1 - G_F(\tilde{\sigma}))W - \tilde{b} = \hat{\theta}_F(1 - G_F(\hat{\sigma}))W - \hat{b}, \tag{6.9}$$

or

$$\varepsilon = \hat{\theta}_F(G_F(\tilde{\sigma}) - G_F(\hat{\sigma}))W. \tag{6.10}$$

For instance, the left-hand side of equation (6.9) gives the expected profit of the principal by paying $\tilde{b}$ knowing that a failure has occurred. The probability of success in the second period is given by $\hat{\theta}_F$ and the probability that the agent works is given by $1 - G_F(\tilde{\sigma})$.

The question then is: can given this bonus $\hat{b}$ and the corresponding $\hat{\sigma}$, the bonus $\tilde{b}$ be an equilibrium? This depends on the beliefs of the agents after observing $\hat{b}$. If the agent believes that a bonus $\hat{b}$ is given after a success, the principal would be able to achieve an increase in the probability of effort at arbitrarily small cost ε. Obviously, the principal then has incentives to deviate to $\hat{b}$, and $\tilde{b}$ cannot be an equilibrium in this case. According to assumption 6 (NWBR), it must be that:

$$\hat{\theta}_S(1 - G_S(\tilde{\sigma}))W - \tilde{b} \geq \hat{\theta}_S(1 - G_S(\hat{\sigma}))W - \hat{b}. \tag{6.11}$$

If this inequality did not hold, the agent should believe that a success has occurred according to the NWBR assumption. This cannot be an equilibrium. In other words, if the principal is indifferent between a bonus $\hat{b}$ and $\tilde{b}$ after a failure, he should not be better off with a bonus $\hat{b}$ after a success. This is a necessary condition.

Equations (6.9) and (6.11) combined yield:

$$\hat{\theta}_S(G_S(\tilde{\sigma}) - G_S(\hat{\sigma})) \leq \hat{\theta}_F(G_F(\tilde{\sigma}) - G_F(\hat{\sigma})). \tag{6.12}$$

Dividing both sides by $\tilde{\sigma} - \hat{\sigma}$ (note that $\hat{\sigma}$ is necessarily smaller than $\tilde{\sigma}$ for a positive ε) and taking the limit $\varepsilon \to +0$ one gets $\hat{\theta}_S g_S(\tilde{\sigma}) \leq \hat{\theta}_F g_F(\tilde{\sigma})$ or:

$$l(\tilde{\sigma}) \leq \frac{\hat{\theta}_F}{\hat{\theta}_S}. \tag{6.13}$$

Conversely, assume that (6.13) is satisfied and consider a possible deviation $\hat{b} = \tilde{b} + \varepsilon$. Since the MLRP implies that[5]:

$$\frac{G_S(\tilde{\sigma}) - G_S(\hat{\sigma})}{G_F(\tilde{\sigma}) - G_F(\hat{\sigma})} < l(\tilde{\sigma}), \tag{6.14}$$

for any $\hat{\sigma} < \tilde{\sigma}$, condition (6.13) together with assumption 6 are sufficient to insure that the agent will believe in failure after receiving any out-of-equilibrium bonus $\hat{b}$ and so the principal has no incentive deviate to a higher bonus. In the appendix we prove that the equilibrium bonus is zero so the principal cannot deviate to a slightly lower bonus. In sum, for the proposed pooling equilibrium to exists, a necessary and sufficient condition is (6.13).

For later reference, note that condition (6.13) and the condition that $l(\tilde{\sigma}) = A$ imply that $A \leq \hat{\theta}_F/\hat{\theta}_S$.

6.3.2 A semi-separating equilibrium

Consider next a *semi-separating* equilibrium where the principal offers $\tilde{b}_S$ after success and randomizes between $\tilde{b}_S$ and $\tilde{b}_F$ after failure. In this case, a bonus $\tilde{b}_F$ is only given in case of failure and therefore perfectly reveals a failure. It follows that

[5] To prove this: suppose $l(x) < l(z)\ \forall\ z \in [x, y]$. Then, since $l(z) \equiv g_S(z)/g_F(z)$, $g_F(z)l(x) < g_S(z)$. This implies $\int_x^y g_F(z)l(x)dz < \int_x^y g_S(z)dz$. Integrating out yields: $l(x) < [G_S(y) - G_S(x)] / [G_F(y) - G_F(x)]$. Similarly, for $l(z) < l(y)\ \forall\, z \in [x, y]$, it follows that $[G_S(y) - G_S(x)] / [G_F(y) - G_F(x)] < l(y)$. Note in particular that for $y > 0$, $G_S(y)/G_F(y) < l(y)$.

$\tilde{b}_F = 0$, since there is no reason for the principal to incur a cost when conveying a negative signal.

Next note that in equilibrium the principal must be indifferent between $\tilde{b}_S$ and $\tilde{b}_F$ after a failure, otherwise he would not be willing to mix. In other words:

$$\hat{\theta}_F(1 - G_F(\tilde{\sigma}_S))W - \tilde{b}_S = 0, \tag{6.15}$$

where $\tilde{\sigma}_S$ is the threshold signal for which an agent works after a bonus $\tilde{b}_S$. Note that after $\tilde{b}_F = 0$ the agent does not work (recall that $\theta_L V < C$) so the payoff for the principal is zero. This condition determines the bonus $\tilde{b}_S$.

Moreover, the principal should not want to deviate to a bonus slightly above or below $\tilde{b}_S$. Following similar logic as above, the principal does not want this in case the agent would believe a failure occurred after observing the deviation from the equilibrium bonus. Suppose first that the principal deviates to $\hat{b}_S = \tilde{b}_S - \varepsilon$. Assumption 6 implies that if

$$\hat{\theta}_F(1 - G_F(\hat{\sigma}_S))W - \hat{b}_s = 0, \tag{6.16}$$

it must be that

$$\hat{\theta}_S(1 - G_S(\hat{\sigma}_S))W - \hat{b}_S \leq \hat{\theta}_S(1 - G_s(\tilde{\sigma}_S))W - \tilde{b}_S \tag{6.17}$$

Again, if this inequality did not hold, the agent should believe that a success has occurred by the NWBR assumption. This cannot be an equilibrium.

Together, these conditions imply:

$$l(\tilde{\sigma}_S) \geq \frac{\hat{\theta}_F}{\hat{\theta}_S}. \tag{6.18}$$

(Recall that a deviation to a smaller bonus is considered, hence $\hat{\sigma} > \tilde{\sigma}$).

Now suppose $\tilde{\sigma}_S = 0$. The above condition implies $l(0) = 0 \geq \hat{\theta}_F/\hat{\theta}_S$. However, $\hat{\theta}_F/\hat{\theta}_S > 0$ so this case can be ruled out. Then consider $\tilde{\sigma}_S > 0$, so that the agent does not always work after a bonus $\tilde{b}_S$. Then, the principal should also not be willing to deviate to a higher bonus $\hat{b}_S = \tilde{b}_S + \varepsilon$ to separate the success outcome. Like in the case of pooling, this implies:

$$l(\tilde{\sigma}_S) \leq \frac{\hat{\theta}_F}{\hat{\theta}_S}. \tag{6.19}$$

Hence, combining (6.18) and (6.19) gives:

$$l(\tilde{\sigma}_S) = \frac{\hat{\theta}_F}{\hat{\theta}_S} \tag{6.20}$$

as the only possibility. The agent's reaction is given by

$$l(\tilde{\sigma}_S) = \tilde{x}_F A. \tag{6.21}$$

Condition (6.20) determines $\tilde{\sigma}_S$, (6.21) defines $\tilde{x}_F$ and (6.15) determines $\tilde{b}_S$. Finally note that conditions (6.20) and (6.21) imply that $\hat{\theta}_F/\hat{\theta}_S \leq A$ since $\tilde{x}_F \leq 1$.

6.4 Rewards, self-confidence, and motivation

The following is the main proposition of the chapter, which relates the reward to initial self-confidence and motivation in period 2[6]:

Proposition 14 *Under assumptions (3)-(6), there always exists is a unique continuation equilibrium depending on the initial self-confidence. In particular, there exists a threshold level ρ^* such that for values $\rho \in [\tilde{\rho}_S, \tilde{\rho}_F)$:*

(i) for $\rho < \rho^*$, the unique equilibrium is a semi-separating equilibrium in which the principal always offers a bonus $b_S = \hat{\theta}_F(1 - G_F(\tilde{\sigma}_S))W$ after success ($x_S = 1$), and randomizes between b_S and $b_F = 0$ after a failure with probabilities $x_F = \frac{l(\tilde{\sigma})}{A(\rho)}$ and $1 - x_F$ respectively. The threshold $\tilde{\sigma}_S$ is positive and determined by $l(\tilde{\sigma}_S) = \frac{\hat{\theta}_F}{\hat{\theta}_S}$, and $\tilde{\sigma}_F = 1$.

(ii) for $\rho \geq \rho^*$, the unique equilibrium is a pooling equilibrium where no bonus is ever offered. The threshold $\tilde{\sigma}$ is positive and determined by $l(\tilde{\sigma}) = A(\rho)$.

The threshold level ρ^* is determined by the ρ for which[7]

$$\hat{\theta}_F(\rho^*)/\hat{\theta}_S(\rho^*) = A(\rho^*). \tag{6.22}$$

[6] The proposition states the equilibrium conditions for values of ρ such that $\rho \in [\tilde{\rho}_S, \tilde{\rho}_F)$. Recall that it was already established that for values of $\rho \notin [\tilde{\rho}_S, \tilde{\rho}_F)$ no bonus is ever offered. For $\rho < \tilde{\rho}_S$, $\tilde{\sigma}_S = \tilde{\sigma}_F = 1$ (the agent never works) and for $\rho \geq \tilde{\rho}_F$, $\tilde{\sigma}_S = \tilde{\sigma}_F = 0$ (the agent always works).

[7] The existence of a point ρ^* is obvious: $A(\rho)$ decreases from infinity to 0 on the interval $[\tilde{\rho}_S, \tilde{\rho}_F]$ and $r(\rho) = \frac{\hat{\theta}_F(\rho)}{\hat{\theta}_S(\rho)}$ is positive and bounded away from 0 on this interval. Implicitly we assume uniqueness as well. For this it is sufficient (but not necessary) to require that $r(\rho)$ be non-decreasing on the relevant interval. Multiple ρ^*s would slightly modify the proposition in a straightforward manner.

That the unique equilibrium for sufficiently high initial self-confidence is a pooling equilibrium with no bonus is intuitive: if self-confidence is high, the agent is likely to work in the second period, and it becomes too costly for the principal to signal a success. The main point of the proposition is however that there is a region where the principal does have an incentive to give a bonus, and that this bonus increases self-confidence. In this region, the agent is relatively unlikely to make efforts in the second period. In this case, the principal has an incentive to make a costly signal to the agent to make clear to him that a success has occurred.

It is also possible to show that for $\rho < \rho^*$ the probability of a reward increases in initial self-confidence. Since $\hat{\theta}_F/\hat{\theta}_S$ is increasing in ρ, so must $l(\tilde{\sigma}_S)$. It is then easily seen that x_F must be increasing, since $A(\rho)$ is decreasing in ρ. The probability that an agent works, on the other hand, is decreasing, as $\tilde{\sigma}_S$ increases in ρ. The change in the size of the bonus for a higher initial self-confidence is ambiguous. There are two opposing effects: first $\hat{\theta}_F$ increases, since a higher probability of a high type increases the success of the agent in the second period. Secondly, $1 - G_F(\tilde{\sigma}_S)$ decreases since $\tilde{\sigma}_S$ increases. The total effect depends on the relative sizes of these two opposing effects.

The size of the bonus in the region where self-confidence is relatively low, that is $\rho < \rho^*$, is proportional to the payoff for the principal in case of success, W. This means that the scope of applications is not limited to situations where the stakes are high for the principal. For example, it would be enough if the principal derives a small benefit from observing a successful performance of the agent, say out of altruistic feelings. For smaller stakes, the corresponding equilibrium bonus will be lower.

6.5 Discussion

When contracts are absent in a relationship, one easily ends up with the argument that no bonus will ever be given, and neither that efforts will be made. The cause is the strong backward induction argument: the agent knows that the principal has no incentive to give a reward in the last period and so he makes no effort, after which the principal realizes that rewarding in the before-last period makes no sense, and so the agent will not work in that period either, and so on until the very first period. This chapter sheds some light on why rewards and efforts may be observed after all.

There is no shortage of empirical and experimental evidence that shows the existence of rewards which are not conditioned on performance, and also that there is a positive relationship between rewards and efforts. Akerlof [1982], for instance, has noted that labour markets can often be characterized as gift exchange relationships. The employers give wages above the minimum wage, and workers make more efforts than is required. Laboratory experiments show the same positive relationship between wages and efforts, even in the absence of explicit performance incentives (Fehr and Gächter [2000]). Deci and Ryan [1999] survey the psychological literature on rewards and intrinsic motivation. They present some studies which find a positive effect, although not all studies which used unexpected rewards find a positive effect, and on average they find no significant relation[8].

The positive relationship between rewards and efforts is also called positive reciprocity. Chapters 2 and 3 provide a more detailed exposition and extend the phenomenon to other environments. These chapters give other explanations of this relationship. For example, it is advanced that people are reciprocal by virtue of their fair nature: they are driven by the moral obligation to reward generous behavior by generous behavior (see e.g. Falk et al. [1999] and section 2.3.3). Another possibility is that people care about social approval and that generous behavior elicits generous behavior (see chapter 3). The current chapter adds another explanation to the existing literature, by focusing on the role of self-confidence.

Obviously, the proposed mechanism can only be valid as long as the main assumptions are satisfied. An important assumption is that the principal has more information about the expected payoffs than the agent. This makes the theory more applicable to situations where agents are in their learning phase: at school or at new jobs. A second important assumption is the sorting condition that is implicit in the model. The principal must obtain a higher marginal benefit from rewarding an agent after a success than after a failure. Otherwise, the principal would be tempted to reward the agent after a failure as well, disturbing the proposed equilibrium.

[8] However, in these studies it is not clear to the participants what the benefits of the experimentator are. The setup of the experiments do therefore not completely fit the current model.

The model is also extendable to other situations with asymmetric information. For example, it extends to situations where the agent is unsure about his own payoff rather than his ability. Another possibility is that the agent cares about the principal's payoff (e.g. through altruism), but is unaware of how much utility the principal derives from his effort.

6.6 The hidden costs of rewards

So far the focus has been on how noncontracted rewards stimulate motivation. By contrast, the papers by Bénabou and Tirole [2002a] and Souvorov [2003] study how rewards can *decrease* motivation. As argued earlier, the main difference with their approach is that they consider bonuses that are specified in a contract. To sketch the argument: if a principal observes that the agent has low ability, he also expects him to have low self-confidence. He therefore proposes a high powered contract which specifies a high reward contingent on success to motivate the agent anyway. This makes the agent realize that he must be of low ability, which lowers his self-confidence. Whenever rewards are withdrawn, the agent will be less motivated.

The mentioned papers are initiated because there is much evidence that this effect occurs. For example, in one experiment it is found that children who were paid for engaging in an activity, showed less interest in the activity once rewards were withdrawn than children who were never rewarded for the activity (Lepper et al. [1973])[9]. Deci and Ryan [1999] survey the literature and find that such crowding-out of motivation also comes out of a meta-analysis of more than one hundred earlier studies. In sum, there is a rich body of experiments showing that there are hidden costs of rewards[10].

How much motivation is crowded out depends to a great extent on the nature of the reward (Deci and Ryan [1985]). For example, rewards contingent on performance have an effect on motivation, but the effect of rewards contingent on fulfilling the task are less profound. No such an effect has been found for experiments where rewards were unexpected, although the evidence is mixed and the results are small and on average insignificant.

[9] Interestingly, when the reward was unexpected, there was a slight (yet insignificant) increase in interest.
[10] See also Kohn [1993], Deci and Ryan [1985], Frey [1997], Frey and Jegen [2002], and chapters 2, 3 and 7.

It is possible to replicate the negative correlation between rewards and self-confidence in the current model. In the model of section 6.2, the principal has more to gain from rewarding an agent after a success, and only then it is worth making a costly signal. Thus, the reward is good news for the agent and raises his self-confidence. However, so far it is assumed that efforts and ability are complements, so that the principal indeed wants to increase the agent's self-confidence. There are, however, also cases where a higher self-confidence would be a bad thing from the principal's viewpoint, and the results of the model would be reversed: the principal would have an incentive to offer a reward to signal a failure, hence reducing self-confidence.

Consider for example the case where the agent has to perform a task that has again only two possible outcomes: a failure or a success. Now assume that the ability and effort are substitutes rather than complements. For instance, a student may be intelligent enough to pass an exam without any efforts. Less gifted students can compensate their lack of ability by making more efforts and study hard. In any case, the parent would rather make the student work hard to avoid any risks of failure. By giving a reward he could say: "Look, here's a reward to show that you have failed, you'd better work hard next period to pass." Although this lowers his self-confidence, it would increase his motivation. Hence, no hidden costs of unexpected rewards should be expected[11].

6.7 Conclusions

Studies by psychologists show that rewards can undermine motivation. This has stimulated economists to examine the effects of rewards in more detail. One in particular interesting contribution is the paper by Bénabou and Tirole [2002a] where the focus is on the role of self-confidence. This chapter also examines the effect of rewards on self-confidence. The focus is on rewards that are not specified in contracts. This answers several questions, such as why discretionary rewards are used and how they can stimulate motivation.

[11] Consider for instance the following version of the model. The agent's payoff is given by $\lambda V - ec$, with $\lambda = 1$ if and only if $e + \theta_i \geq \varphi$. Thus, $\lambda = 1$ is a success and needs either effort or a high ability. Assume that $\varphi < 1$, and $V > c$ so that effort is enough to pass. However, an able agent can pass without effort whereas a low ability agent cannot: $\theta_H > \lambda > \theta_L$. If no effort is made in the first period, and the outcome was not a success, the principal may want to signal through a bonus that the agent has low ability and should make efforts in the second period.

Rewards can motivate agents by signalling a success, increasing self-confidence. Rewards can also decrease self-confidence in the special case where efforts and ability are substitutes. In accordance with the empirical and experimental evidence, no negative relation between rewards and motivation is found for such unexpected rewards.

This chapter is only one of the first few attempts to formally study the interaction between rewards and self-confidence. Future work should generalize some of the assumptions. For example, the role of shirking is central in many principal agent models. Here, it is simply assumed that effort is observable. Another straightforward extension is to include more than two periods. This would shed light on the dynamics of rewards and self-confidence in the spirit of Souvorov [2003]. For instance, in a richer framework, the agent may benefit from pretending his confidence is low in order to get a bonus. In addition, people may come to expect another reward once they received one in a previous period. This may complicate the analysis somewhat as this changes their reference point. Furthermore, it would be interesting to examine the relation between promised bonuses (specified in contracts) and discretionary bonuses. Promised bonuses can undermine motivation but at least specify when a bonus will be given. Discretionary bonuses increase motivation to work after a bonus, but insofar as they are unexpected they cannot motivate agents in the first period. Thus, perhaps there is something such as an optimal mix of different kinds of rewards, consisting both of announced as well as unexpected bonuses.

6.8 Appendix

Proof of proposition 13.
With the strategy of the principal is to give b_S after success with probability x_S and after failure with probability x_F, the agent updates his prior ρ to ρ_S after a bonus b_S. The expressions for ρ_S is given in the text (equation (6.5)). Consider an agent who observes b_S. He works if and only if:

$$[\rho_S \theta_H + (1-\rho_S)\theta_L]\, V \geq c. \tag{6.23}$$

This can be written as:

$$\frac{\rho_S}{1-\rho_S} \geq \frac{c/V - \theta_L}{\theta_H - c/V} \equiv \phi. \tag{6.24}$$

Define ρ' to be ρ_S such that (6.24) holds with equality. Then, using (6.5), one gets:

$$\frac{g_S(\sigma)}{g_F(\sigma)} = \frac{x_F}{x_S} \frac{\phi(1-\rho')(1-\theta_L) - \rho'(1-\theta_H)}{\rho'\theta_H - \phi(1-\rho')\theta_L} \equiv \frac{x_F}{x_S} A(\rho'). \tag{6.25}$$

The denominator of the RHS is zero for:

$$\rho' = \tilde{\rho}_S \equiv \frac{\phi\theta_L}{\phi\theta_L + \theta_H}. \tag{6.26}$$

The numerator of the RHS is zero for:

$$\rho' = \tilde{\rho}_F \equiv \frac{\phi(1-\theta_L)}{\phi(1-\theta_L) + (1-\theta_H)}. \tag{6.27}$$

The sign of the derivative of the RHS of (6.25) is equal to the sign of $(\theta_L - \theta_H)\phi < 0$. Furthermore, it is easy to show that $\tilde{\rho}_S < \tilde{\rho}_F$. Finally, note that:

$$\lim_{\rho \uparrow \tilde{\rho}_S} A(\rho) \to -\infty, \tag{6.28}$$

and

$$\lim_{\rho \downarrow \tilde{\rho}_S} A(\rho) \to +\infty \tag{6.29}$$

In sum, for $\rho < \tilde{\rho}_S$ and $\rho \geq \tilde{\rho}_F$ the RHS of (6.25) is negative and decreasing, and for $\tilde{\rho}_S < \rho < \tilde{\rho}_F$, the RHS of (6.25) is positive and decreasing. Since $g_S(\sigma)/g_F(\sigma) \geq 0$, there is no signal for $\rho \notin [\tilde{\rho}_S, \tilde{\rho}_F)$ such that (6.25) holds with equality. For $\rho < \tilde{\rho}_S$ the agent never finds it profitable to work. For $\rho > \tilde{\rho}_F$, the

agent always finds it profitable to work. For $\rho \in [\tilde{\rho}_S, \tilde{\rho}_F)$ the agent works for any signal $\sigma > \tilde{\sigma}$ with:

$$\frac{g_S(\tilde{\sigma})}{g_F(\tilde{\sigma})} = \frac{x_F}{x_S} A(\rho). \tag{6.30}$$

Since it is assumed that $g_S(\sigma)/g_F(\sigma)$ has full support on $[0, +\infty)$, the threshold signal $\tilde{\sigma}$ always exists. Note that thresholds $\tilde{\rho}_S$ and $\tilde{\rho}_F$ do not depend on the bonus b_S.■

Proof of proposition 14.
In proving the proposition, we first introduce some intermediate results in the form of lemma's and a corollary. At the end we give the precise definition of the NWBR assumption, because this requires some additional notation of it's own.

Lemma 7 *If bonus $\tilde{b}$ is given in equilibrium with probability $\tilde{x}_S > 0$ after success and with $\tilde{x}_F > 0$ after failure, and $\tilde{\sigma}$ is the agent's reaction to the bonus (i.e. the agent works if and only if he received a signal above $\tilde{\sigma}$), then $l(\tilde{\sigma}) = \frac{\tilde{x}_F}{\tilde{x}_S} A(\rho)$ and either*

- *$\tilde{b} > 0$ and $l(\tilde{\sigma}) = r(\rho)$ or*
- *$\tilde{b} = 0$ and $l(\tilde{\sigma}) \leq r(\rho)$,*

with $r(\rho) \equiv \hat{\theta}_F / \hat{\theta}_S$.

Proof. $l(\tilde{\sigma}) = \frac{\tilde{x}_F}{\tilde{x}_S} A(\rho)$ determines the agent's optimal reaction to the principal's policy unless the agent would find it worthwhile to always work when offered $\tilde{b}$, i.e. $\tilde{\sigma} = 0$. In the latter case we should have $l(0) \geq \frac{\tilde{x}_F}{\tilde{x}_S} A(\rho)$ in contradiction with Assumption 4 which states $l(0) = 0$.

When $\tilde{b} > 0$, for the principal not to be able (and *a fortiori* willing) to signal that the agent has succeeded by deviating to $\tilde{b} \pm \varepsilon$ for a small $\varepsilon > 0$, it must be the case that $l(\tilde{\sigma}) = r(\rho)$ (see the analysis of the pooling equilibrium in section 6.3.1). When $\tilde{b} = 0$, only deviations to $\tilde{b} + \varepsilon$ are relevant so the requirement reduces to $l(\tilde{\sigma}) \leq r(\rho)$. ■

Lemma 8 *In equilibrium only one bonus is offered after success.*

Proof. Assume that b_1 and $b_2 > b_1$ are offered after success with positive probability, and $\tilde{\sigma}_1$ and $\tilde{\sigma}_2$ are the corresponding agent's reactions ($\tilde{\sigma}_1 > \tilde{\sigma}_2$ by Lemma

6). The smaller bonus, b_1, must be offered after failure with a positive probability (otherwise the agent would always work after b_1 and the principal would never give the larger one, b_2). For the principal not to be willing to separate the successful outcome by offering $b_1 + \varepsilon$, it must be that $l(\tilde{\sigma}_1) \leq r(\rho)$. Then,

$$b_2 - b_1 = \hat{\theta}_S(G_S(\tilde{\sigma}_1) - G_S(\tilde{\sigma}_2)) < \hat{\theta}_F(G_F(\tilde{\sigma}_1) - G_F(\tilde{\sigma}_2)). \tag{6.31}$$

The equality in (6.31) comes from the principal's indifference between b_1 and b_2, and the inequality follows from $l(\tilde{\sigma}_1) \leq r(\rho)$ and MLRP and implies that that the principal strictly prefers to give b_2 rather than b_1 after a failure – a contradiction. ■

Corollary 2 *At most two different bonuses are offered with positive probability in equilibrium. There are three potential types of equilibrium:*

A. pooling – the same bonus offered to both types;

B. semi-separating – the principal always gives $\tilde{b}_S$ after success and randomizes between $\tilde{b}_S$ and $\tilde{b}_F \neq \tilde{b}_S$ after failure;

C. separating – the principal always gives $\tilde{b}_S$ after success and $\tilde{b}_F \neq \tilde{b}_S$ after failure.

Each possible type of equilibrium is considered below.

A. Pooling equilibria

Some bonus $\tilde{b}$ is always given to the agent ($x_F(\tilde{b}) = x_S(\tilde{b}) = 1$), who works when observes a signal σ above $\tilde{\sigma} = \tilde{\sigma}(\tilde{b})$.

Case A.1: $\tilde{b} = 0$ and $\tilde{\sigma} = 0$ (the agent always works).
For this equilibrium to occur it must be the case that

$$l(0) \geq A(\rho). \tag{6.32}$$

By the full support assumption for the likelihood ratio, this is equivalent to $\rho \geq \tilde{\rho}_F$.

Case A.2: $\tilde{b} = 0$ and $\tilde{\sigma} > 0$ (the agent does not always work).

See section 6.3.1 in the main text. Note that:

$$\hat{\theta}_F = \frac{\rho(1-\theta_H)\theta_H + (1-\rho)(1-\theta_L)\theta_L}{\rho(1-\theta_H) + (1-\rho)(1-\theta_L)}, \tag{6.33}$$

$$\hat{\theta}_S = \frac{\rho\theta_H^2 + (1-\rho)\theta_L^2}{\rho\theta_H + (1-\rho)\theta_L} \tag{6.34}$$

Case A.3: $\tilde{b} > 0$ and $\tilde{\sigma} > 0$ (the agent does not always work).
As in case A.2, the principal should not want to deviate in case of success. Now the condition for the principal's inability to separate a success is more stringent:

$$l(\tilde{\sigma}) = \frac{\hat{\theta}_F}{\hat{\theta}_S}. \tag{6.35}$$

Indeed, by the same reasoning as in the previous case it easy to show that $l(\tilde{\sigma}) \leq \hat{\theta}_F/\hat{\theta}_S$ is necessary and sufficient for the principal's inability to separate the success outcome by increasing $\tilde{b}$ by ε. According to similar logic, $l(\tilde{\sigma}) \geq \hat{\theta}_F/\hat{\theta}_S$ is a necessary and sufficient condition for the principal not being able to signal success through reducing the equilibrium bonus by ε. This leaves (6.35) as the only possible case.

Besides (6.35), there is another condition for this equilibrium: the principal should not wish to separate the failure outcome. In the previous case it was satisfied trivially because increasing a bonus in order to decrease the probability of effort is clearly not a good idea. Now, when the principal can reduce her equilibrium bonus, this requirement may be restrictive. When (6.35) is satisfied, any deviation will be interpreted by the agent a signal of failure. Hence, the optimal deviation is to pay no bonus at all. This deviation will not to be profitable if and only if

$$\hat{\theta}_F(1 - G_F(\tilde{\sigma}))W - \tilde{b} \geq 0$$

because the agent will not work if convinced in failure (remember that $\rho < \tilde{\rho}_F$).

Thus, in this non-generic case (i.e. when $l(\tilde{\sigma}) = \hat{\theta}_F/\hat{\theta}_S$ for $\tilde{\sigma}$ determined from $l(\tilde{\sigma}) = A$) there is a continuum of pooling equilibria with $\tilde{b} \in [0, \hat{\theta}_F(1-G_F(\tilde{\sigma}))W]$ satisfying the NWBR criterion.

Case A.4: $\tilde{b} > 0$ and $\tilde{\sigma} = 0$ (all agents work).

Clearly, this case requires $l(0) \geq A(\rho)$. By the full support assumption for the likelihood ratio, this is equivalent to $\rho \geq \tilde{\rho}_F$. It also requires $l(0) \geq \hat{\theta}_F/\hat{\theta}_S$, which can be proved by a a line of reasoning similar to the other cases above.

Semi-separating equilibria.
Case B.1: Failure outcome semi-separated. See section 6.3.2 in the main text.

Case B.2: Success outcome semi-separated.
This equilibrium is impossible: Assume that the principal pays $\tilde{b}_F$ after failure and randomizes between $\tilde{b}_S$ and $\tilde{b}_F$ after success; the agent then always works after $\tilde{b}_S$ and works when $\sigma \geq \tilde{\sigma}_F$ after $\tilde{b}_F$. Clearly $\tilde{b}_S > \tilde{b}_F$ – otherwise the principal would always pay $\tilde{b}_S$, and $\tilde{\sigma}_F > 0$ – otherwise the principal would always pay $\tilde{b}_F$. For the principal not to be able to separate a successful outcome by a bonus $\tilde{b}_F + \varepsilon$ for arbitrarily small ε it must be that

$$l(\tilde{\sigma}_F) \leq \frac{\hat{\theta}_F}{\hat{\theta}_S} \tag{6.36}$$

In case $\tilde{b}_F > 0$, this should hold with equality since the principal should also not be able to separate the success outcome by decreasing the bonus by ε. Furthermore, the principal should be indifferent between $\tilde{b}_S$ and $\tilde{b}_F$ after a success otherwise he should not be willing to randomize:

$$\hat{\theta}_S(1 - G_S(\tilde{\sigma}_F))W - \tilde{b}_F = \hat{\theta}_S W - \tilde{b}_S, \tag{6.37}$$

But (6.36) implies that if the principal is indifferent between giving $\tilde{b}_S$ or $\tilde{b}_F$ after success, she should prefer to give $\tilde{b}_S$ after failure – a contradiction. Indeed, the MLRP property of $l(\sigma)$ implies that

$$\frac{G_S(\tilde{\sigma})}{G_F(\tilde{\sigma})} < l(\tilde{\sigma}). \tag{6.38}$$

But then if the principal is indifferent between $\tilde{b}_S$ and $\tilde{b}_F$ after success, he prefers $\tilde{b}_S$ after a failure:

$$\hat{\theta}_F(1 - G_F(\tilde{\sigma}_F))W - \tilde{b}_F < \hat{\theta}_S W - \tilde{b}_S. \tag{6.39}$$

Hence $\tilde{b}_F$ cannot be an equilibrium bonus.

Case B.3: Both outcomes semi-separated

This equilibrium cannot occur. It implies that the principal must be indifferent between $\tilde{b}_S$ and $\tilde{b}_F$ both after a failure and after a success. It is easy to show that this is not compatible.

Separating equilibria.
In a separating equilibrium the principal gives $\tilde{b}_F = 0$ after a failure (again, there is no sense to incur any cost to send a negative signal) and $\tilde{b}_S > 0$ after success. The agent always works after $\tilde{b}_S$ and never works after $\tilde{b}_F$. For this pair of bonuses to be an equilibrium, the principal should not strictly prefer to give $\tilde{b}_S$ after failure:

$$\hat{\theta}_F W - \tilde{b}_S \leq 0. \tag{6.40}$$

If (6.40) were a strict inequality, then the principal could reduce $\tilde{b}_S$ by a small ε so that (6.40) would still be satisfied and according to he NWBR criterion the agent should believe success has occurred: the set of the agent's reactions that make the principal indifferent between giving 0 and $\tilde{b}_S - \varepsilon$ after failure – the empty set – is strictly included in the set of reactions that make her indifferent between $\tilde{b}_S$ and $\tilde{b}_S - \varepsilon$ after success. Hence, $\tilde{b}_S$ is uniquely determined:

$$\tilde{b}_S = \hat{\theta}_F W. \tag{6.41}$$

It must also be the case that the principal cannot separate the success outcome by a bonus slightly lower than $\tilde{b}_S$. A now familiar condition for this is

$$l(0) \geq \frac{\hat{\theta}_F}{\hat{\theta}_S}. \tag{6.42}$$

To see this, assume that the agent's reaction $\hat{\sigma}$ to an out-of-equilibrium bonus $\hat{b}$ is such that the principal is indifferent between deviating to $\hat{b}$ after a failure or not:

$$\hat{\theta}_F(1 - G_F(\hat{\sigma}))W - \hat{b} = 0. \tag{6.43}$$

Then we need the principal not to be willing to deviate after success:

$$\hat{\theta}_S(1 - G_S(\hat{\sigma}))W - \hat{b} < \hat{\theta}_S W - \tilde{b}_S, \tag{6.44}$$

or, using $\tilde{b}_S = \hat{\theta}_F W$

$$\hat{\theta}_S G_S(\hat{\sigma}) > \hat{\theta}_F G_F(\hat{\sigma}). \tag{6.45}$$

For (6.45) to be satisfied for all $\hat{\sigma}$, a necessary and sufficient condition is (6.42).

In sum, all equilibria other than those of proposition 14 require either that $l(0) \geq \hat{\theta}_F/\hat{\theta}_S$ or $l(0) \geq A(\rho)$ or both. Since $\hat{\theta}_F/\hat{\theta}_S$ and $A(\rho)$ are both positive for $\rho \in [\tilde{\rho}_S, \tilde{\rho}_F)$ and since $l(0) = 0$, these can be ruled out. The only exception is case A.3 which exists only for the point where $\hat{\theta}_F/\hat{\theta}_S = A(\rho)$ which has measure zero and is therefore for simplicity ignored. ■

Never a Weak Best Response in signalling games.
The following is a condensed presentation of the NWBR criterion. See Fudenberg and Tirole [1991] for more details.

Suppose the principal chooses action a_1 and the agent a_2. This notation facilitates comparison with the literature. We say that a principal is of type t if he has observed outcome t. Then, fix an equilibrium, an out-of-equilibrium action a_1, and let $u_1^*(t)$ be the principal's expected payoff in the proposed equilibrium if he is of type t. Then define the set $D(t, T, a_1)$ to be the set of mixed-strategy best responses α_2 to action a_1 and beliefs concentrated on T (a subset of the set of all possible types Θ) that make type t strictly prefer a_1 to his equilibrium strategy:

$$D(t, T, a_1) = \cup_{\mu:\mu(T|a_1)=1}\{\alpha_2 \in MBR(\mu, a_1) \text{ s.t. } u_1^*(t) < u_1(a_1, \alpha_2, t)\}, \quad (6.46)$$

for beliefs μ over different types.The set $D^o(t, T, a_1)$ is defined as the set of mixed-strategy best responses that make type t indifferent:

$$D^o(t, T, a_1) = \cup_{\mu:\mu(T|a_1)=1}\{\alpha_2 \in MBR(\mu, a_1) \text{ s.t. } u_1^*(t) = u_1(a_1, \alpha_2, t)\}. \quad (6.47)$$

Then, a type-action pair (t, a_1) can be deleted under the NWBR criterion if:

$$D^o(t, \Theta, a_1) \subset \cup_{t' \neq t} D(t', \Theta, a_1). \quad (6.48)$$

Type-action pair (t, a_1) can be pruned if the (sequential) equilibrium response to the out-of-equilibrium action a_1 that makes t indifferent between his expected equilibrium outcome and a deviation to a_1, makes some other type strictly prefer to deviate to the out-of-equilibrium action. In other words, the agent is assumed to believe that it is infinitely more likely that a_1 has come from some other type t'.

7

Optimal Subsidies with Rationalizing Agents: Subsidize Enough but Don't Subsidize Too Much

"He suddenly recalled how he had once in the past been asked, "Why do you hate so and so, so much?" And he had answered them, with his shameless impudence, "I'll tell you. He has done me no harm. But I played him a dirty trick, and ever since I have hated him." F. M. Dostoyevsky, *The Brothers Karamazov*, 1879.

7.1 Introduction

The typical economic approach is to take preferences as given and then study how certain incentives can alter behavior in a desired way. Psychologists on the other hand, have paid considerable attention to the formation of preferences, trying to demonstrate that they are not stable. Preferences are not stable indeed, sometimes even capricious (see the next section). According to psychologists, the formation of preferences is nevertheless in many cases quite predictable.

While changing behavior by appropriate incentives has been subject to extensive investigation within the field of economics, changing preferences (or attitudes, or tastes) has been left largely unexplored. In this chapter both effects are taken into account. I wish to argue that a subsidy on environmental friendly goods

[0] I am indebted to Riccardo Calcagno, Theo van de Klundert, Eloic Peyrache, Karim Sadrieh, Sjak Smulders, participants at the European Economic Association (Lausanne, 2001), the ENTER Jamboree (Toulouse, 2002) and seminar participants at Tilburg for very helpful comments.

(green products) not only influences behavior but also attitudes towards such goods. Although subsidies have a straightforward effect on buying behavior – a higher subsidy giving more incentives to buy that good – the relationship between subsidies and attitudes is surprising. The main result is that a low subsidy stimulates a positive attitude change towards the subsidized good. Yet, a high subsidy does nothing to the attitudes of people. This result fits the experimental evidence in the psychological literature well, reporting a negative relationship between rewards and attitude change (Aronson [1988]). It follows that after the removal of a high subsidy behavior and attitudes are as in the status quo. But a low subsidy induces attitude changes that can persist even after withdrawing the subsidy. I therefore conclude that high subsidies are 'too much of a good thing': they affect current behavior but fail to affect attitudes and therefore future behavior.

The leading example of this chapter is the consumption of environmental damaging goods, but the model applies to other instances of goods with social externalities as well, such as production methods that involve child labour or cruelty to animals (see also Chapter 5 for more on goods with social externalities). Typical of environmentally damaging products (as well as the other examples) is that the externalities tied to these goods are considered to be socially or morally undesirable. Consuming these goods therefore poses a threat to our self-concept of being decent people. Psychologists argue that in such a case consumers experience an unpleasant feeling. To reduce this unpleasant feeling, they will either try to change their behavior, or – by rationalizing their choice – their attitudes[1].

That people can and do change their behavior is undisputed. But, as said, they can change their attitudes as well. There are several ways to accomplish attitude changes (see Baumeister [1998]). For example, Frey [1997] argues that when the costs of following principles of environmental ethics are high, people find a lot of reasons why they should desist from doing so. They can for instance highlight the argument that their effect of buying environmentally damaging goods is negligible on the environment anyway. The mere fact that you contribute marginally to a bad environment does not make you an indecent person, does it? Or you may recall that article in the newspaper saying that the damaging effects on the environment are staggeringly exaggerated. And if that is not enough, you can

[1] In psychological terminology: the bad feeling is caused by cognitive dissonance – a discrepancy between simultaneously held cognitions (beliefs, attitudes, opinions). Originated by Festinger [1957] the theory has been found to be mostly relevant in situations where one's self-concept is violated by one's behavior (Aronson [1988]).

always buy yourself the book *"The Skeptical Environmentalist"* by Bjørn Lomborg, which plainly denies any signs of an impoverishing environment. Just try to ignore those other articles claiming the opposite. Possibilities abound. It should therefore come at no surprise that self-serving biased information processing is indeed systematically found (see the next section).

The benefits of changing attitudes towards environmental damaging products is to take away the associated unpleasant feeling. This goes at some psychological and/or psychic costs: information has to be gathered and mental efforts have to be put into rationalizing choice. It is precisely the impact of subsidies on the cost-benefit structure that drives the main result of this chapter. In words (formalized in the next section): both a high as well as a low subsidy on green products might induce a change in behavior towards buying these goods. But it does not in both cases pay off to change your attitude. A high subsidy provides enough justification of itself to buy green products. Even if you think that it does not help the environment, it is still attractive to buy these goods due to their low price. A low subsidy, on the other hand, does not give such a justification of itself. Why did you until recently refuse to buy green products (perhaps claiming that buying them has no significant merit) but consume them now that they are slightly subsidized? That does not make sense. But it would start making sense if you concentrate on the argument that, though each individual as such does not contribute anything noteworthy, for the society as a whole it would make a huge difference if everybody would behave the same way. Be supportive of Kantian principles of morality. And, think of the social approval you will get for the noble act of buying green products (see Chapter 3). Accordingly, the people who received a relatively small reward are likely to revise their attitude most.

The rest of the chapter is organized as follows. In the next section I discuss in more detail the basic premises of the model. Some evidence for the model is provided in section 7.3. The formal model is presented in section 7.4, together with the results. In the subsequent section I briefly discuss the robustness of the results. Finally, section 7.5 discusses the results, relates them to the existing literature, and concludes.

7.2 Preference management

When a person deliberately harms someone else, he experiences an unpleasant feeling: sentiments of guilt. This unpleasant feeling stems from a discrepancy between held beliefs ("I am a decent person") and actual behavior ("I harmed her"). In the psychological literature this feeling has been labeled *dissonance* (see also footnote 1). For ease of comparison with the existing literature, in the remainder I stick to this terminology.

It is not so controversial to think that people experience dissonance when their behavior imposes negative externalities on society. More controversial is that, to reduce this unpleasant feeling, people can manipulate their preferences in a way that serves their interest best. In other words, that people would 'rationalize' choice. It is one of the most basic assumptions of theories of rational choice that preferences are stable. Theories of habit formation depart from this by assuming instead that future tastes are dependent on current consumption, for example as a consequence of addiction or learning. These theories therefore allow for changing tastes over time, but typically in a slow manner and in the direction of getting to appreciate what you consumed in the past. Hence, habit formation does not account for self-serving changes in tastes. Rationalizing choice, on the other hand, demands a more rapid change in preferences and more flexibility in the direction it goes.

The existence of rapid changes in preferences is supported by a sizeable body of research showing a picture of a remarkably labile nature of preferences (see Slovic [1991]). Illustrative of this are the following two well-known examples.

- Endowment effects: once goods are part of one's endowment, the valuation immediately increases sharply. This effect is present even if the subjects are made familiar with the object on beforehand, thereby excluding learning arguments as an explanation (Loewenstein and Adler [1995], Thaler [1980]).

- Framing: another well established phenomena is the sensitivity of choice to the way that a choice problem is formulated. For example, the valuation of a gamble is sensitive to whether the outcomes are framed as gains or losses relative to the status quo (Tversky and Kahneman [1992]).

The unstable nature of preferences paves the way for manipulating preferences in a way that is beneficial to oneself (though not necessarily consciously). Thus, not

surprising on retrospective grounds, Slovic [1991] reports the existence of "preference management": preferences are constructed on the spot by "...discarding nonessential differences, adding new attributes into the problem frame in order to bolster one alternative, or otherwise restructuring the decision problem to create dominance and thus reduce conflict and indecision." (Slovic [1991, 500]). In a similar vein, Bénabou and Tirole [2000] speak of "awareness management": people reframe performance by remembering successes, forgetting failures, and by trying to convince themselves that the act was not so bad. Thus, although "... the individual updates his beliefs according to broad Bayesian principles ... it is also widely recognized that information acquisition and belief updating are subject to self-serving biases." (Bénabou and Tirole [2000, 2]).

From the marketing literature it is clear that firms have acknowledged that people rationalize choice. To give two examples, according to Dibb et al. [1997, 108], buyers seek positive information to justify their choice. They claim that "motoring journalists often note with amusement that car shows and exhibitions are frequented by consumers who have just recently purchased a new car". Berkowitz et al. [1994, 144] take the advertising campaign by Buick as an example, which had as message "Aren't you really glad you bought a Buick". Firms know that people seek to justify their choice afterwards, and they give response to this desire.

These findings are incorporated in the model of section 7.4. For a more elaborate exposition of evidence of preference management I refer to Bénabou and Tirole [2000] and Rabin [1995][2].

7.3 Evidence from psychology

The theory of cognitive dissonance roughly boils down to the idea that people rationalize their choice. The idea is not limited to goods with social externalities. In fact, despite its relative simplicity, there is a wide range of applications (Aronson [1988], Akerlof and Dickens [1982]). This section describes some experimental

[2] Other contributions to economics that try to explain the findings by social psychologists by formalizing the theory of cognitive dissonance include the following. The focus of Akerlof and Dickens [1982] is on the purchase of safety equipment in an environment of uncertainty, Akerlof [1991] considers time-inconsistency, Dickens [1986] criminal behaviour, Rabin [1994] examines social norms, and Rabin [1995] moral behavior. James and Gutkind [1985] apply the concept to the conditional help provided by the IMF.

results. All of them are easily understood by assuming that people are motivated to rationalize their choice[3].

In a very early experiment by Brehm [1956], some women were asked to rate various household products. The experimenter selected two of the appliances that were presented wrapped to the women, and which were rated equally by the women. The women were then told that they could choose one of these two products as a reward for participation (without knowing on which base these were selected). Still wrapped, they were asked to evaluate these two products again. A systematic feature of the second evaluation relative to the first one, was that the chosen product increased in valuation, whereas the rejected product apparently became less attractive to these women.

Aronson and Carlsmith [1963] designed an experiment where children were asked to rate several toys in attractiveness and were then left alone. In one condition, the experimenter took the second rated toy with him. In a second condition, the toy was left with the child but the child was asked not to play with it, with the added mild threat that the experimenter would otherwise be annoyed. In the third condition, the threat was more severe, announcing that he would be very angry. No child played with the toy. After that, the experimenter returned and asked for another evaluation. As it turns out, the perceived attraction increased in conditions 1 and 2, where the child had enough external reasons not to play with the toy. In the mild threat condition, however, the attraction was less in the second evaluation. Apparently, the child had no really good justification for not playing with the toy and so tried to rationalize behavior by reasoning that the toy was not so attractive after all.

Festinger and Carlsmith [1959] report an experiment where subjects had to listen to a rather boring seminar. After the seminar, some were asked to tell the next participants that the seminar was going to be very interesting. Some were paid \$1 for this, others \$20. All participants were asked to rate the seminar. Interestingly, the subjects who were paid \$1 showed a much more positive evaluation record than the other groups. This could be expected. Some individuals were asked to tell a lie, and nobody rejected this request. Likely, telling a lie is not congruent with the self-image of being a decent person. This can create some dissonance. However, the group of people which received \$20 has a clear rational for lying:

[3] Most of the material in this section draws upon the expositions by Aronson [1988] and Brigham [1991].

they were paid considerably for this. Hence, the only group where the lie created dissonance was the group which only received $1 and they indeed revised their opinion most.

If we assume that people try to rationalize behavior, this has an interesting consequence for how people should recall things. Consider a situation where there are good reasons pro and contra an individual's position. From all the reasons pro and contra, some make sense and others are less plausible. Which ones would be best to recall for an individual? In order to rationalize the position taken, the best thing to do is to remember the sensible reasons congruent with the position taken, and the implausible arguments that are incongruent. In an experiment by Jones and Kohler [1959] exactly this pattern was found.

Three more things are worthwhile to note. First, dissonance seems to bring a real physiological arousal. An experiment by Croyle and Cooper [1983] shows that conditions creating high-dissonance situations show more skin conduction responses. According to them, this is a reliable indicator of physiological arousal. This means that the phenomenon of dissonance goes beyond subjective self-reports (Aronson [1988]). Relatedly, attitude change is not a superficial tendency of people to pretend they were changing (Baumeister [1998]).

Second, attitudes are, if properly measured, a reliable indicator of behavior (Brigham [1991]). In a famous study by Lepper, Greene and Nisbett [1973] children were asked to draw a picture using attractive magic markers. Some children were paid for this. In a second session, the children were left alone to play freely with the materials. Those children who were not paid before (and had therefore no good rational to participate other than that the magic markers were enjoyable in themselves to play with) were more likely to spend time with the materials. Hence, instead of measuring attitude changes, this study shows that also behavioral changes are induced. This is an important aspect from an economics point of view, since if changed attitudes do not lead to changes in behavior they are not very interesting from the perspective of choice theory.

Finally, the attitude changes are persistent. Between the first and second session of the experiment by Lepper et al. [1973] was a time span of several days. In a study by Freedman [1965] children were less likely to play with previously forbidden toys, even after nine weeks had passed.

7.4 A simple model

The aim of this section is to present a simple model that captures the foregoing observations. Still abstracting from preference management, subsection 7.4.1 introduces the unpleasant feeling that people experience when doing something threatening to their self-concept into the utility function. In subsection 7.4.2 I allow for preference management. Here, choice can be rationalized along the ways described in the introduction and the previous section. Although this clearly is a dynamic process, for simplicity the model is essentially static. It is my feeling however, that the main results go through in a more general setting. Because the ultimate aim of the chapter is to consider the effects of subsidies on efforts, this is subsequently studied in section 7.4.3.

7.4.1 Preferences and choice

The standard utility function

Consider an individual with income I whose preferences can be described by a quasilinear utility function. Later I explain how this specification simplifies the analysis considerably but it is noteworthy to mention that it is not crucial to obtain the results. Thus, let the following *standard utility*[4] function represent her preferences:

$$u(x, y) = x^{\alpha} + y, \qquad 0 < \alpha < 1. \tag{7.1}$$

Here, x and y are both bundles of goods. My concern here is only the bundle of x goods. Within the bundle of indivisible x goods a distinction can be made between x_1 (the environmental damaging good) and x_2 (the green product). These goods are in principle good substitutes but each good has its own advantages which will have to be weighed against each other. Thus, for example, if x is the product 'wood', then x_1 and x_2 may be tropical hardwood and certificated wood guaranteeing forest preservation, respectively. The parameter δ reflects their *relative attractiveness* as perceived by the individual in the following way:

$$x = x_1 + \delta x_2. \tag{7.2}$$

[4] Standard refers here to the fact that normally dissonance is not an argument in the utility function, not to the use of a quasilinear specification.

Hence,

$$u(x_1, x_2, y) = (x_1 + \delta x_2)^{\alpha} + y. \tag{7.3}$$

For simplicity, it is assumed that the x-goods are indivisible. Note that at most one of the x-goods will consumed. As a shortcut, where no confusion can arise I employ the notation $u_1 = u(x_1, 0, y)$ and $u_2 = u(0, x_2, y)$. Hence, u_i denotes the utility of consuming the bundle (x, y) knowing that good x_i is consumed.

The extended utility function

I now wish to incorporate the unpleasant feeling, or dissonance, into the framework. I propose the following *extended utility* function of an individual who chooses product x_j:

$$\hat{u}_j = u_j - d_j. \tag{7.4}$$

This specification gives credit to the thought that reducing dissonance, d, increases extended utility. Hence, reducing dissonance is a motivational factor. Cognitive dissonance is measured by the relative attractiveness (in standard utility terms) between the chosen alternative and the rejected good[5]. This seems reasonable. Buying environmental damaging products creates dissonance, but less so if their green counterparts are clearly inferior in quality. Let $d_{j,}$ denote dissonance if good x_j is consumed and good x_{-j} is rejected, where $j = 1, 2$. Then:

$$d_j = d(u_j, u_{-j}), \tag{7.5}$$

[5]Note carefully the use of 'rejected'. It is possible that another good is valued positively but is out of the budget restriction. In this sense, it is not rejected and causes no dissonance. Similarly, coerced regulations (such as some taxes) that cannot be avoided do not cause dissonance. This is known in the psychological literature as 'forced compliance'.

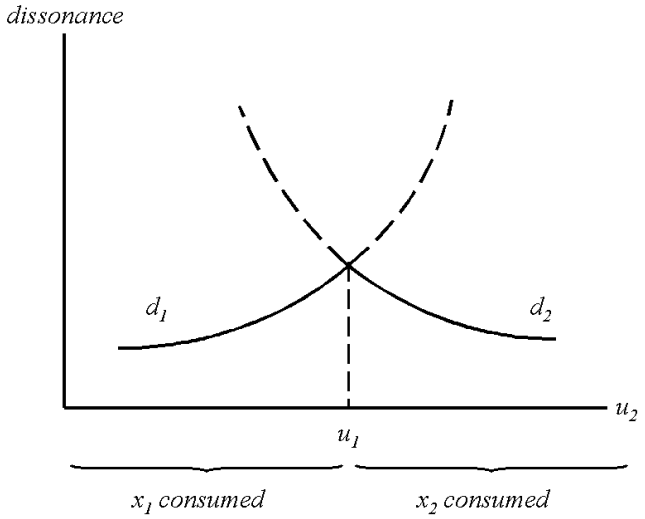

FIGURE 7.1.

with the properties:

$$\begin{aligned}
\frac{\partial d(u_j, u_{-j})}{\partial u_j} &< 0, \\
\frac{\partial d(u_j, u_{-j})}{\partial u_{-j}} &> 0, \\
\frac{\partial^2 d(u_j, u_{-j})}{\partial u_j^2} &> 0, \\
\frac{\partial^2 d(u_j, u_{-j})}{\partial u_{-j}^2} &> 0,
\end{aligned} \tag{7.6}$$

and furthermore that $u_j = u_{-j}$ implies $d(u_j, u_{-j}) = d(u_{-j}, u_j)$.These conditions are shown graphically in figure 7.1. Dissonance decreases as the chosen alternative gets better or as the rejected alternative gets worse. Dissonance is maximal at the point where both alternatives are equally attractive. (I therefore might as well have labeled individuals as indifference averse.)

At this point, it is worthwhile to emphasize that the way dissonance is incorporated is not in contradiction with standard rationality assumptions. If the standard utility function u represents the preference relation of the individual, then, with the assumptions on (7.5), so does $\hat{u}$.

Lemma 9 *$\hat{u}$ is a monotonic transformation of u.*

Proof. I want to show that $u_1 \lesseqgtr u_2$ is equivalent to $\hat{u}_1 \lesseqgtr \hat{u}_2$. Suppose first that $u_1 > u_2$ for any given y. Since $\partial d(u_j, u_{-j})/\partial u_j < 0$ and $\partial d(u_j, u_{-j})/\partial u_{-j} > 0$ we have the following chain: $d_1 = d(u_1, u_2) < d(u_1, u_1) < d(u_2, u_1) = d_2$. Hence, $\hat{u}_1 = u_1 - d_1 > u_2 - d_2 = \hat{u}_2$. The other cases are similar. ■

The intuition behind this result is the following. Suppose that good x_1 is preferred to x_2. Choosing x_1 causes some dissonance. But would x_2 be consumed, and consequently x_1 be rejected, then dissonance would be even higher. This can also be seen in figure 7.1. In the area left to the point u_1, good x_1 is preferred to x_2 ($u_1 > u_2$). As is clear from the figure, in this area the d_2-curve is everywhere above the d_1-curve. This means that dissonance of rejecting x_1 is higher than dissonance of rejecting x_2. Clearly, if both standard utility of x_1 is higher and dissonance is lower, then the extended utility of consuming x_1 is also higher.

This result can be used to derive the individual's choice without specifying d_j. I assume that the individual receives a lump-sum subsidy s from the government conditional on consumption of the green good, x_2. To simplify even more, I assume that both bundles of goods are consumed and that all prices are unity. Finally, $v(p, I, s)$ defines the standard indirect utility function that gives the maximum standard utility achievable at the prices, income and subsidies the individual faces. It is then straightforward to show the following choice behavior (see also figure 7.2):

Proposition 15 *There exists a subsidy level* $\bar{s} \equiv k_1 - k_2\delta^{\alpha/(1-\alpha)}$ *such that (i) for all* $s < \bar{s}$ *good* x_1 *is consumed and* $v = I + k_1$ *and for all* $s \geq \bar{s}$ *good* x_2 *is consumed and* $v = I + s + \delta^{\frac{\alpha}{1-\alpha}}k_2$*, where* k_1 *and* k_2 *are positive constants and (ii)* $\bar{s}_\delta < 0$*.*

Proof. Let $u(x, y) = x^\alpha + y$ and for simplicity $p_x = p_y = 1$. We can proceed in two stages. First, an optimal (x, y) combination can be determined and then the optimal (x_1, x_2). Since it is assumed that $y > 0$ and marginal utility of y is equal to 1, it must be that an additional unit of x gives a marginal utility of less than unity. Define $\varphi_i = \min \partial u(x, y)/\partial x$ s.th. $\partial u(x, y)/\partial x \geq 1$. Suppose x_1 is consumed. Then x_1 is consumed up to the point where $\alpha x_1^{\alpha-1} = \varphi_1$, or:

$$x_1^* = (\alpha/\varphi_1)^{1/(1-\alpha)}. \tag{7.7}$$

Consumption of good y is determined by the budget restriction:

$$p_y y^* = I - p_x x_1^* \Leftrightarrow y^* = I - (\alpha/\varphi_1)^{1/(1-\alpha)}. \tag{7.8}$$

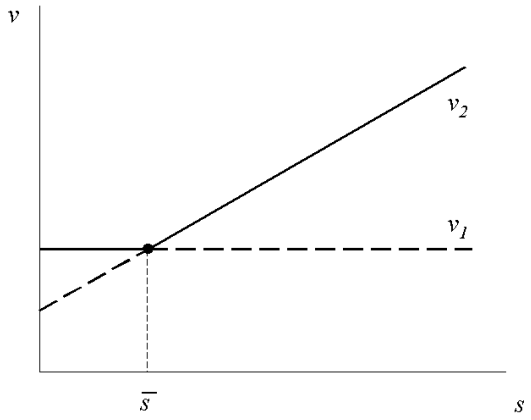

FIGURE 7.2.

Substitution of the demand function in the utility function gives the indirect utility function

$$v(p, I, s) = v(1, I, 0) = I + k_1, \tag{7.9}$$

where $k_1 \equiv (\alpha/\varphi_1)^{\alpha/(1-\alpha)} - (\alpha/\varphi_1)^{1/(1-\alpha)}$. ($k_1 > 0$ for $\alpha < 1$.) Next, suppose that x_2 is consumed. Then it follows that:

$$x_2^* = (\alpha/\varphi_2)^{1/(1-\alpha)}\delta^{\alpha/(1-\alpha)}. \tag{7.10}$$

The budget now includes the subsidy, hence in this case $p_y y^* = I + s - p_x x_2^*$. Indirect utility is in this case given by:

$$v(p, I, s) = I + s + k_2\delta^{\alpha/(1-\alpha)}, \tag{7.11}$$

where $k_2 \equiv (\alpha/\varphi_2)^{\alpha/(1-\alpha)} - (\alpha/\varphi_2)^{1/(1-\alpha)}$. The individual consumes good x_2 if and only if indirect utility is higher, i.e. if $I + s + k_2\delta^{\alpha/(1-\alpha)} \geq I + k_1$ and equality at the threshold subsidy level $\bar{s}$. Hence:

$$\bar{s} \equiv k_1 - k_2\delta^{\alpha/(1-\alpha)}. \tag{7.12}$$

Since $k_2 > 0$, $\bar{s}_\delta < 0$. ■

7.4.2 Rationalizing choice

> *The absence of any general principles or rules, either of personal or administrative morality, which made it possible for him either to agree or disagree with anybody according to what was wanted at the time.* L. Tolstoy, *Resurrection.*

Plagued by an unpleasant feeling, the individual then tries to rationalize her choice. By focusing on certain arguments while ignoring others, she restructures the problem as to convince herself that the choice she made was indeed the right one. This means that if she chose good x_1 she puts efforts in trying to find arguments that bolster the choice of this good, and deprive the attractiveness of the rejected good x_2. In case she chose good x_2, she does just the reverse[6].

In the model, the relative valuation of the two x goods is given by δ. The psychological or physical efforts, e, are therefore reflected in a change in δ. Let this change, $\Delta\delta$, be governed by:

$$\Delta\delta = \delta(\gamma e - (1-\gamma)e), \tag{7.13}$$

$$\begin{aligned} \text{with } \gamma &= 1 \text{ if } v_1 \leq v_2 \text{ and} \\ \gamma &= 0 \text{ if } v_1 > v_2. \end{aligned}$$

where $\delta(0) = 0$ and $\delta'(\cdot) > 0$. For simplicity, it is also assumed that $\delta''(\cdot) = 0$. The properties of γ assure that someone who chose x_1 decreases δ (making x_1 look relatively more attractive) and someone who chose x_2 increases δ (making x_2 look relatively more attractive), so that in both cases efforts indeed rationalize choice.

It goes without saying that, unlike the character of Tolstoy, for most people rationalizing choice comes at a cost, c. Information has to be gathered, and rehearsed to recall later on, principles of decency have to be given up, mental efforts have to be put into convincing oneself of one's rightness, and so forth. These costs put a limit to the ability to change your opinion. Hence,

$$c = c(e), \tag{7.14}$$

[6] Note that the timing is that first decisions are made and that they are rationalized only afterwards. This is in accordance with the psychological literature (see for example Festinger [1957]) but is not crucial to obtain the results in this paper.

where $c(0) = 0$, and with the usual assumptions $c'(e) > 0$ and $c''(e) > 0$.

In sum, the individual solves the following program:

$$e^* \in \arg\max_{e \geq 0} \hat{v}(e) - c(e), \tag{7.15}$$

where $\hat{v} = v - d$ is indirect extended utility. Given the assumptions on $\delta(e)$ and $c(e)$ there is a unique effort level e^* that optimizes attitudes. The next section examines some properties of the optimal effort level in relation to the subsidy level.

7.4.3 The effect of a subsidy

It is obvious that the optimal effort level will somehow be related to the subsidy level. The subsidy not only influences choice behavior and standard utility by changing the budget restriction, but ultimately also extended utility through its effect on dissonance. One may conjecture that the efforts put in rationalizing behavior pay off most when dissonance is most severe. In this subsection I show this conjecture to be true.

First, consider the effect of a subsidy on the level of dissonance. A low subsidy makes good x_1 relatively attractive to purchase, and a high subsidy does the same for good x_2. Both a low and a high subsidy in itself therefore provide a good rationalization of choice. In effect, low and high subsidies create little dissonance. It are the intermediate subsidy levels that create most dissonance, peaking at the point of indifference. This leads to the following lemma:

Lemma 10 *$\frac{\partial d(\cdot)}{\partial s} > 0$ for $s < \bar{s}$ and $\frac{\partial d(\cdot)}{\partial s} < 0$ for $s \geq \bar{s}$.*

Proof. Since v_1 is constant in s, and v_2 increasing in s (see prop. 1), it follows straightforward from the assumptions on d that $d(v_1, v_2)$ is increasing in s and $d(v_2, v_1)$ decreasing. ■

This immediately leads us to the conclusion that, without preference management, up to a certain point (that is, $\bar{s}$), increasing subsidies make individuals worse off by creating dissonance. When the subsidy exceeds this level, however, increasing subsidies makes individuals better off by enriching their consumption level and taking away dissonance. Thus:

Proposition 16 *$\frac{\partial \hat{v}(\cdot)}{\partial s} < 0$ for $s < \bar{s}$ and $\frac{\partial \hat{v}(\cdot)}{\partial s} > 0$ for $s \geq \bar{s}$.*

Proof. For $s < \bar{s}$, by proposition 15 $v(p, I, s)$ is constant in s and by lemma 10 dissonance is increasing in s. Hence, $\hat{v}(p, I, s)$ is decreasing in s. For $s \geq \bar{s}$, by proposition 15 $v(p, I, s)$ is increasing in s (see prop. 15) and (by lemma 10) d is decreasing in s. Hence, $\hat{v}(p, I, s)$ is increasing in s. ■

The existence of dissonance depresses extended utility. The threshold level $\bar{s}$ is the point where dissonance is most severe. It is also the point where the individual gains most from rationalizing her choice. Dissonance decreases at subsidy levels that are further away from this threshold, and so do the gains from rationalization. In terms of optimal efforts, this means that efforts are increasing in the subsidy level up to $\bar{s}$, whereas for any higher subsidy level efforts are decreasing:

Proposition 17 $\frac{\partial e^*}{\partial s} > 0$ *if* $s < \bar{s}$ *and* $\frac{\partial e^*}{\partial s} < 0$ *if* $s \geq \bar{s}$.

Proof. Let $\hat{v}'(e(s), s) \equiv \hat{v}(e(s), s) - c(e(s))$. By definition:

$$\left.\frac{\partial \hat{v}'(\cdot)}{\partial e}\right|_{e=e^*} \equiv 0. \tag{7.16}$$

Therefore at e^*:

$$\frac{\partial^2 \hat{v}'(\cdot)}{\partial e^2}\frac{\partial e}{\partial s} + \frac{\partial^2 \hat{v}'(\cdot)}{\partial e \partial s} \equiv 0, \tag{7.17}$$

which is identical to:

$$\frac{\partial e}{\partial s} = -\frac{\partial^2 \hat{v}'(\cdot)/\partial e \partial s}{\partial^2 \hat{v}'(\cdot)/\partial e^2}. \tag{7.18}$$

Since $\hat{v}'(\cdot)$ is maximal at e^*, the second-order condition already requires that $\partial^2 \hat{v}'(\cdot)/\partial e^2 < 0$. Consider first the case where $s < \bar{s}$. Define $\tilde{\delta} \equiv \delta^{\alpha/(1-\alpha)}$. We have that:

$$sign\frac{\partial e}{\partial s} = sign\frac{\partial^2 \hat{v}'(\cdot)}{\partial e \partial s} = sign - \frac{\partial^2 d(u_j, u_{-j})}{\partial u_{-j}^2} k\tilde{\delta}'(e). \tag{7.19}$$

The last equality follows from the fact that $\frac{\partial^2 v'(\cdot)}{\partial e \partial s} = 0$. Since $\gamma = 0$ we have that $\tilde{\delta}'(e) < 0$ and because $\frac{\partial^2 d(u_j, u_{-j})}{\partial u_{-j}^2} > 0$ we have that $\frac{\partial e^*}{\partial s} > 0$ as stated. The case where $s \geq \bar{s}$ is similar but now $\gamma = 1$ so that $\tilde{\delta}'(e) > 0$, and now $\frac{\partial^2 d(u_j, u_{-j})}{\partial u_j^2} > 0$ is a sufficient condition for $\frac{\partial e^*}{\partial s} < 0$. ■

Proposition 17 has the following consequences for the attitudes held. For low subsidy levels ($s < \bar{s}$) dissonance can be reduced by decreasing δ. The higher the subsidy, the more efforts are made and, consequently, the larger the attitude change. For high subsidy levels ($s \geq \bar{s}$) it is optimal to increase δ as compared

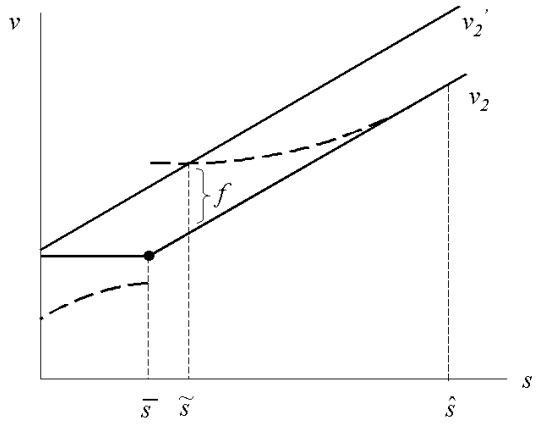

FIGURE 7.3.

to the initial beliefs. The higher the subsidy, the smaller the increase however. In this range there is an inverse relationship between the subsidy and the attitude change δ. This result is in accordance with much of the psychological literature (see for example Aronson [1988] and section 7.3).

Figure 7.3 shows the implications of proposition 17 for the standard utility function. The solid line v_2 represents standard utility at any given subsidy level. The dotted lines show the magnitude of the shift of the v_2-curve. For example, at point $\tilde{s}$, the standard utility function is shifted by the distance f so that the new standard utility function of consuming x_2 is given by v_2'.

Although strictly speaking the model is a static one, proposition 17 still gives some hints to interesting dynamic implications. Suppose that the objective of the government is to stimulate the consumption of green products. It may in principle do so by providing any subsidy greater or equal than $\bar{s}$. But proposition 17 suggests that the government does best by giving a subsidy exactly equal to $\bar{s}$. In this way, the individual is stimulated most to change his attitude towards the green product. The dynamic consequence is that this subsidy has the highest chance of inducing persistent consumption of green products after removal of the subsidy. A subsidy level of $\tilde{s}$, for instance, induces a shift of the v_2-curve by the distance f, enough to make the individual prefer the green product even at a

subsidy level of zero (see Figure 7.3)[7]. On the other hand, a higher subsidy level like $\hat{s}$ does nothing to the attitude of individuals and removing the subsidy means that the individual switches back to consumption of the environmental damaging good. In sum, a high subsidy is 'too much of a good thing': it affects current choice, but it fails to affect attitudes and therefore choice at any future time period. Note also that according to the same logic, subsidies that are below the threshold level may have undesired consequences in that people try to rationalize their choice by degrading the green good. This is especially the case for subsidies near the threshold level. Seen from a dynamic perspective, this means that it is better not to subsidize at all than to subsidize a bit. The point is that the stimulation of green goods can become unnecessarily very costly if the government fails to take into account changes in attitudes. Subsidies that are not very well targeted are costly: in a static as well as in a dynamical sense.

7.5 Discussion

Robustness

Proposition 17 is the main result of the chapter and it is therefore interesting to see whether or not it can be generalized to other than quasilinear specifications of the utility function. The interesting part of the proposition is that for subsidy levels exceeding $\bar{s}$, where there is a negative relationship between subsidies and efforts. The focus is therefore on this range of subsidies.

As usual, the sign of the change in effort as a response to a change in the subsidy is determined by the sign of the cross partial derivative of the objective function; $\partial^2(\hat{v}(e,s)-c(e))/\partial e\partial s$. In general, with $\hat{v}=v-d$ this is given by:

$$sign\frac{\partial^2(\hat{v}(e,s)-c(e))}{\partial e\partial s}=sign\left[\frac{\partial^2 v}{\partial e\partial s}-\frac{\partial^2 d}{\partial e\partial s}\right]. \tag{7.20}$$

Note that the first term measures the effect of a subsidy on the marginal benefits of efforts. In general, this effect is positive: at higher subsidy levels, the individual

[7] It is fairly easy to show that attitudes can persist after withdrawal of the subsidy. Consider an individual that faces subsidy level $\overline{s}$ and suppose that she makes efforts to increase δ to δ'. By proposition 15, the new threshold subsidy level is $\overline{s}'<\overline{s}$. Hence, there exists a range of subsidies smaller than $\overline{s}$ for which the individual continues to consume x_2.

spends more on good x_2 and this increases the marginal benefits of efforts. Under a quasilinear specification, however, all additional income is spent on the y-good and the effect drops out, strengthening proposition 17. Note furthermore that if an inverse relationship between efforts and subsidies is indeed observed, then it can only attributed to the dissonance term. Without dissonance, subsidies always work as a reinforcer of making efforts through the term $\partial^2 v/\partial e \partial s$.

Crowding-out

Crowding-out is an interesting phenomenon that is supported by a substantial amount of empirical and experimental evidence (see e.g. Deci and Ryan [1985], Frey [1997], Kohn [1993], and Chapters 1, 3, and 6). Crowding-out is said to occur at instances where money has a perverse effect on motivation. That is, rather than being encouraging, higher rewards reduce the motivation to undertake a activity. The common interpretation of this effect is that, although higher rewards create more external motivations to undertake an activity, at the same time it tends to destroy the intrinsic motivation people have.[8]

The reduction in intrinsic motivation has sometimes been given a dissonance reduction interpretation.[9] Seen from that perspective, sufficiently high rewards provide enough external justification to perform a task, whereas relatively low rewards can only justify efforts if people can convince themselves that the task is intrinsically motivating. Intrinsic motivation is in the latter case likely to be built up, whereas in the former case people feel no need to do so. Higher rewards thus result in lower intrinsic motivation than lower rewards, mimicking the result of the previous section (proposition 17). This result notwithstanding, cognitive dissonance theory cannot explain true crowding-out. Higher rewards may give less incentives to build intrinsic motivation, it does not give incentives to destroy it. If anything, one would expect that higher rewards give more justification to undertake the rewarded activity, even if insignificantly so (see however the next section for an explanation why higher rewards may undermine motivation after all).

[8] In fact, crowding-out is not confined to situations where rewards are provided. Other external motivations, such as punishments, can produce the same result.

[9] See for example Dickens [1986] who presents a formal model where an increase in punishments can lead to more criminal activity. His explanation is based on dissonance reduction.

Alternative explanations

It should at this point be noted that many of the experimental results put forward by psychologists in fact do have an appealing more traditional economic interpretation, namely, that the rewards are signals. To continue with the leading example of this chapter, suppose that the quality of the green good is known to be relatively low. This means that a relatively high subsidy is needed to make the green good equally attractive as its environmental damaging counterpart. If, on the other hand, the consumers cannot directly perceive the quality of the green good, they may infer from a high subsidy that the quality must be low. In a related paper, Bénabou and Tirole [2002] present a formal model where a higher bonus signals a more difficult task or lower ability to the agent who has to perform the task (see also Chapter 6). Under some conditions, this can lead to the same predictions as proposition 17. High subsidies signal a low quality, and once removed, the consumer is no longer willing to pay more for the green good. Low subsidies may signal a high quality and result in permanent consumption of the green good, even after withdrawal of the subsidy.

The reward as a signal is a reasonable alternative view, and it plays without doubt a role in many situations. There are, however, experiments where a clear signal is lacking and yet attitude changes still occur. For instance, in one study children were offered an opportunity to cheat but no payments were made whatsoever. They became more lenient towards cheating when they did not resist the temptation (see Aronson [1988]). In the experiment from section 7.3 where women were asked to rate several appliances, no new information was revealed during the experiment because the rewarded appliance was given wrapped to them until after their second rating. In the experiment by Lepper et al. [1973] where children were given the task to draw a picture, it was told to the children that a reward would follow just for drawing a picture, independent of the endresult. These rewards therefore carry no clear-cut information about their individual ability or task difficulty. The result of decreased interest was replicated in a follow-up study by Greene and Lepper [1974] where it was told to the children that *all* of them would receive a reward for doing the task, again revealing no information. Finally, the experiment by Jones and Kohler [1959] where people were asked which arguments they recalled, it is not obvious what kind of signal could have been given.

A competing alternative interpretation from psychology to cognitive dissonance theory is self-perception theory, developed by Bem [1972]. This theory, in essence, reverses the causality between attitudes and behavior. Self-perception theory states that people infer their attitude from their behavior (see also chapter 1 and section 2.3.5 of chapter 2 on self-signalling). Thus, the people in the experiment of Festinger and Carlsmith [1959] who received $1 for telling the next participants that the seminar is going to be a lot of fun, must have reasoned that, because they did this for just one dollar, they really must have liked the seminar.

Even though this theory has many merits in many situations, it gives in many instances ultimately the same behavioral predictions, and in some cases it is somewhat less convincing as an explanation. First, the theory assumes imperfect information about one's self, which is more reasonable for people who try to quit smoking for the first time than for people who have already tried many times before. It therefore depends on whether one believes that the people were capable of valuing the seminar directly or had to infer this from their behavior. Since the participants in this experiment were college men, one can safely assume that they were capable of directly assessing the seminar[10]. Furthermore, self-perception theory, as well as the signalling approach discussed above, assumes that no physiological arousal takes place in any condition. However, as pointed out earlier, such a physiological arousal has been measured to be present.

7.6 Conclusions

While psychologists have put great efforts in understanding the formation of preferences in order to explain changes in behavior, the economic approach insists on explaining any changes in behavior by changes in income, relative prices, and information. The purpose of this chapter is to show some of the consequences it can have when attitude changes have been taken care of as well. Under, in my view, fairly intuitive assumptions, I have been able to replicate the experimental finding that low rewards induce more attitude change than high rewards.

[10] It is not unreasonable that self-perception theory becomes more relevant if a significant time period has elapsed between the action and evaluation. Since retrospective evaluations of past utilities are known not always to be reliable (Kahneman [1994]), inferring preferences from past behaviour may be more accurate than from recalled utilities. I do not know of any study which has worked this out.

The focus of this chapter has been on a feeling of dissonance to explain the data. Others have stressed the informational aspects. Likely, both elements are present: informational signals and an unpleasant physiological arousal. Unfortunately, in many cases the literature seems insufficiently conclusive in establishing the relative weight of informational aspects. In my view, the results of this chapter should be seen as complementary to a signalling interpretation. The latter approach is taken up in Chapter 6.

8
Summary in Dutch

Beknopte inleiding in de onderwerpen

Vaak wordt er in economische modellen aangenomen dat mensen egoïstisch zijn en alleen om geld en goederen geven. In veel gevallen kunnen op basis van deze veronderstellingen goede verklaringen geboden worden voor hoe mensen zich gedragen. Maar er blijven ook veel onopgeloste vraagstukken over: waarom laten mensen fooien achter? Waarom zijn ze soms minder gemotiveerd om iets te doen wanneer er een beloning tegenover staat? Waarom verzamelen mensen zoveel informatie nadat ze een aankoop hebben gedaan in plaats van dat vooraf te doen? En waarom kopen mensen Max Havelaar koffie? Deze en andere vragen stel ik centraal in dit proefschrift.

Om een antwoord te geven op bovenstaande vragen verwerk ik onderzoek uit de psychologie in economische modellen. Zo stel ik bijvoorbeeld in navolging van de psychologische literatuur dat mensen sociale waardering willen krijgen voor hun gedrag. Andere concepten die ik overneem zijn bijvoorbeeld dat mensen hun gedrag graag rationaliseren en dat ze onvolledige informatie over hun eigen persoonlijkheid hebben. Door economische modellen uit te breiden met dit soort ideeën kan een verklaring gegeven worden voor veel gedrag dat afwijkt van standaard economische modellen.

Het proefschrift is ingedeeld in twee centrale thema's. Het eerste thema gaat in op de vraag waarom mensen aan elkaar of aan een liefdadigheidsinstelling

geven. Het tweede thema betreft hoe mensen reageren op beloningen. Eerst ga ik dieper in op deze twee thema's. Vervolgens geef ik een inhoudelijk overzicht van de individuele hoofdstukken.

Geven

De meeste mensen geven een behoorlijk bedrag weg aan familie, vrienden, of liefdadigheidsinstellingen. Het gaat hier om uiteenlopende dingen zoals verjaardagscadeaus en bloed, maar ook vrijwilligerswerk kan opgevat worden als een gift. Men is geneigd om te denken dat het hier om vrijwillige en vrijblijvende schenkingen gaat, maar niets is minder waar. In de werkelijkheid blijkt er een sterke sociale druk te bestaan om iets terug te geven nadat men iets ontvangen heeft. De ontvanger heeft als het ware een schuld uit staan die terug betaald moet worden. Het is dan ook niet helemaal verwonderlijk dat vergif de tweede betekenis van gift is.

Vanuit een economisch perspectief is het in eerste instantie verwonderlijk dat er van alles weggegeven wordt, en slechts een klein deel daarvan in de vorm van geld. Volgens de micro-economie kan geld nooit slechter zijn dan een cadeau. Immers, de ontvanger kan met dat geld hetzelfde cadeau kopen, of iets wat hij nog liever heeft. Maar als de ontvanger beter af is met geld, waarom geven we dan zo vaak cadeaus?

Beloningen

Het tweede thema betreft hoe mensen reageren op beloningen. Doorgaans wordt in de economie verondersteld dat een beloning in het vooruitzicht mensen motiveert. De reden hiervan is dat de nadruk meestal ligt op het directe effect van beloningen. Mensen geven om geld, dus het ligt voor de hand dat een beloning hen stimuleert om harder te werken. Er zijn echter ook indirecte effecten, soms met een tegengestelde werking. Dit blijkt uit experimenten die door psychologen zijn gedaan.

Een deel van die experimenten meet de zogeheten intrinsieke motivatie van mensen. Met intrinsieke motivatie wordt bedoeld dat mensen gemotiveerd zijn om bepaalde dingen te doen zelfs als er geen beloning tegenover staat of als ze niet onder controle staan. Een goed voorbeeld zijn kinderen die ijverig werken aan een puzzel. Het blijkt echter dat deze intrinsieke motivatie soms afneemt zodra een

beloning wordt verstrekt. De kinderen gaan in eerste instantie nog harder aan de slag, maar zodra de beloning weggenomen wordt zijn ze minder gemotiveerd dan voorheen. Een indirect effect van beloningen is dus het verdringen van intrinsieke motivatie. Andere experimenten laten zelfs zien dat de motivatie direct afneemt na het invoeren van een beloning. Zo nam het aanbod van bloeddonoren af nadat een compensatie werd verstrekt, en kinderen wisten minder geld op te halen bij een collecte nadat ze een kleine vergoeding kregen. Er zijn dus 'verborgen kosten van beloningen'.

Er is nog een reeks experimenten die een onverwacht effect van beloningen laat zien. Het gaat hierbij om de invloed van beloningen op voorkeuren van mensen. In de economie wordt verondersteld dat voorkeuren vastliggen. Psychologen hebben echter gevonden dat mensen iets meer gaan waarderen wanneer ze er een kleine beloning voor krijgen, terwijl dit effect niet gevonden wordt bij een hoge beloning. Zo woonden sommige mensen een uitgesproken saai seminar bij. Er werd hen gevraagd de volgende groep luisteraars te zeggen dat het allemaal zeer interessant zou worden. Een deel van hen kreeg daarvoor een kleine beloning, een ander deel een flinke beloning. Daarna werden ze gevraagd te zeggen wat ze er zelf van vonden. Diegenen met een hoge beloning waardeerden het seminar laag, terwijl diegenen met een lage beloning het hoog waardeerden.

Overzicht van de hoofdstukken

Het voorgaande is een beknopt overzicht van de verschillende thema's uit dit proefschrift. De thema's worden dieper uitgewerkt in de hoofdstukken zelf. Het proefschrift begint met een algemene inleiding in het onderwerp psychologie en economie. De overige hoofdstukken zijn ruwweg ingedeeld naar de twee thema's. Hoofdstukken 2 tot en met 5 gaan in op de vraag waarom mensen geven. De nadruk in hoofdstukken 3, 6 en 7 ligt op de effecten van beloningen. Dit onderscheid is uiteraard enigszins kunstmatig, want in brede zin zijn beloningen soms ook giften, en andersom. Hieronder volgt een overzicht van de inhoud van de individuele hoofdstukken, zonder diep in te gaan op de modellen en verklaringen. Het is vooral bedoeld als leidraad.

Hoofdstuk 2. In dit hoofdstuk worden redenen uiteengezet waarom mensen geven. Er zijn veel redenen denkbaar: altruïsme, ruil, rechtvaardigheid, signalen geven en sociale waardering komen allen aan bod. Echter, niet allen zijn even

geloofwaardig als verklaringen. Zoals gezegd zijn er twee bijzonderheden aan geefgedrag: bijna altijd wordt iets teruggegeven (reciprociteit) en bijna nooit in geld (inadequaatheid). Met deze twee kenmerken in gedachten wordt de verklaringskracht van iedere theorie beoordeeld.

Een eerste voor de hand liggende verklaring is dat mensen een ruil beogen met het geven van cadeaus aan elkaar. Dit kan inderdaad een goede verklaring zijn mits mensen geduldig genoeg zijn, want ze krijgen bijna nooit meteen een cadeau terug. Altruïsme lijkt aannemelijk, maar kan niet verklaren waarom mensen zo zelden geld geven in plaats van cadeaus. Het verklaart dus maar een klein deel van alle cadeau's. Een andere verklaring is rechtvaardigheid. Maar er wordt aannemelijk gemaakt dat dit niet alles kan verklaren, want in sommige experimenten handelen mensen in strijd met de veronderstelde rechtvaardigheidstheorieën. De vraag naar sociale waardering kan deze experimenten wel verklaren, en ook waarom giften inadequaat zijn en bovendien waarom er reciprociteit bestaat (zie vooral hoofdstuk 3). Een andere, wellicht minder voor de hand liggende verklaring, is dat mensen iets willen signaleren met hun geefgedrag, bijvoorbeeld hoe rijk ze zijn of hoe goed ze de ontvanger kennen. Het kan zelfs zo zijn dat iemand iets aan zichzelf laat zien, bijvoorbeeld dat hij rechtvaardig is omdat hij fooien achterlaat terwijl hij net zo goed het restaurant had kunnen verlaten zonder fooi.

Het hoofdstuk eindigt met een discussie waarin ik pleit voor een hybride verklaring. Bijvoorbeeld: mensen geven niet omdat ze eerlijk zijn maar omdat ze eerlijk willen lijken en daar sociale waardering voor krijgen. Verder beargumenteer ik dat het belangrijk is om te weten met welk doel mensen geven voordat een institutie ontworpen wordt. Zo kan het bijvoorbeeld averechts werken om mensen te belonen voor hun gift als ze dit doen om sociale waardering te krijgen.

Hoofdstuk 3 borduurt voor op de gedachte dat mensen geven teneinde sociale waardering te krijgen. Allereerst probeer ik aan te tonen dat mensen om sociale waardering geven en dat hier een element van status in zit: mensen willen vooral meer gewaardeerd worden dan hun buren. Op basis hiervan verklaar ik reciprociteit en adequaatheid. Reciprociteit volgt uit het status effect: Wanneer iemand geeft krijgt hij daarvoor waardering en dit motiveert de ontvanger om ook te geven teneinde ook die waardering te krijgen. De eerste gever krijgt het liefst meer waardering dan de ontvanger. Door geen geld te geven maar cadeaus, maakt hij het duurder voor de ontvanger om iets terug te geven en blijft hij voorop lopen in de race om status. Bovendien verklaart deze theorie waarom mensen

meer geven als ze hiervoor publiekelijk bedankt worden, zoals vaak het geval is bij liefdadigheid. Door de publiekelijke bekendheid die eraan gegeven wordt is het mogelijk om sociale waardering te krijgen, zelfs als de directe ontvanger anoniem blijft.

Er valt een interessante relatie te leggen tussen de theorie uit dit hoofdstuk en die van averechtse effecten van beloningen. Er wordt sociale waardering gegeven voor een gift omdat een gift een opoffering is voor de gever. Het belonen van giften maakt de opoffering kleiner, en bijgevolg ook de sociale waardering. Een compensatie voor geven kan op deze manier averechts werken.

Vervolgens wordt in *hoofdstuk 4* geefgedrag bekeken vanuit een meer macro-economisch perspectief. Wanneer men veronderstelt dat het geven en wedergeven een ruil ten doel heeft, dan lijkt het aannemelijk dat op termijn het marktmechanisme al het geefgedrag zal verdringen. Naarmate de markt groeit in omvang wordt deze efficiënter en wordt geven een slechter alternatief.

Echter, in dit hoofdstuk wordt beargumenteerd dat geefgedrag niet alleen als ruilmechanisme dient, maar ook symbolische waarde heeft. Hoofdstuk 3 ging eerder al in op de gedachte dat een gift sociale waardering teweegbrengt. Deze symbolische waarde komt niet tot stand via het marktmechanisme omdat deze een betrekkelijk anoniem karakter heeft. Hierdoor wordt geefgedrag niet in zijn geheel verdrongen. Het model laat tevens zien dat geefgedrag kan blijven bestaan ondanks dat de markt efficiënter zou zijn geweest. Bovendien kan het voorkomen dat de markt al het geven verdringt terwijl geven meer efficiënt is.

In de voorgaande hoofdstukken is er beargumenteerd dat er om verschillende redenen nut ontleend wordt aan geefgedrag. In *hoofdstuk 5* worden de consequenties hiervan nader beschouwd. Er wordt verondersteld dat mensen graag een meerprijs willen betalen voor goederen die geproduceerd zijn met behulp van technieken die sociale externaliteiten verminderen. Met sociale externaliteiten worden bijvoorbeeld productiemethoden bedoeld die schadelijke effecten op het milieu hebben, maar ook die gebruik maken van kinderarbeid of een onrechtvaardig laag loon betalen. De meerprijs voor producten met minder sociale externaliteiten kan als een gift aan de werknemers of aan de maatschappij opgevat worden.

Problematisch is dat consumenten niet kunnen beoordelen welke techniek de producent gebruikt heeft. Ze kunnen bijvoorbeeld niet zien of er gebruik is gemaakt van kinderarbeid. Daarom zullen ze niet bereid zijn om een meerprijs te betalen.

Op hun beurt zijn producenten niet bereid te investeren in productiemethoden met minder sociale externaliteiten.

Er worden twee manieren bekeken om dit probleem op te lossen: een standaard invoeren, of labels invoeren. Wanneer de overheid een standaard op de productiemethode invoert, dan betekent dat dat alle consumenten gedwongen zijn om het product tegen die standaard te kopen, of om niets te kopen. Wanneer labels of certificaten worden ingevoerd dan krijgt de producent (en daarmee de consument) de vrijheid om te produceren volgens hun eigen gekozen productietechniek of volgens de techniek waarvoor ze een certificaat krijgen. De consumenten met een hoge betalingsbereidheid zullen het gecertificeerde goed kopen, die met een lagere betalingsbereidheid het ongecertificeerde goed. Vervolgens wordt bekeken welke van de twee mogelijkheden sociaal optimaal is. Er wordt gevonden dat een certificaat beter is voor een interval van consumentenheterogeniteit. Wanneer consumenten of relatief homogeen zijn of relatief heterogeen, dan is een standaard welvaartsmaximaliserend.

Hoofdstuk 6 bestudeert de effecten van beloningen op zelfvertrouwen in een principaal-agent model. Het model is een uitbreiding op de literatuur die laat zien dat een bonus een signaal kan zijn dat het een moeilijke taak betreft of dat de principaal de talenten van de agent laag inschat. Dit verklaart waarom beloningen negatieve consequenties kunnen hebben. In die literatuur is de focus gericht op beloningen zoals vastgelegd in een contract. In dit hoofdstuk wordt de aanname gemaakt dat de uitkomst alleen geobserveerd kan worden door de principaal. Dit maakt een contract onmogelijk. Er wordt dus gekeken naar het effect van onverwachte beloningen.

Dit hoofdstuk biedt een mogelijke verklaring waarom onverwachte beloningen in eerste instantie gegeven worden. Een cruciale veronderstelling hierbij is dat de principaal over meer informatie beschikt dan de agent wat betreft succesvolheid. De principaal weet of de opdracht een succes is of niet, terwijl de agent slecht een idee hierover kan vormen. De theorie heeft dus betrekking op situaties waar de agent nog in een leerfase zit: een kind die piano leert spelen of een werknemer die pas begonnen is aan zijn nieuwe baan. Een andere mogelijkheid is dat de agent geen overzicht op het geheel van activiteiten omdat hij zich specialiseert op een klein onderdeel ervan, terwijl de principaal het totale proces kan overzien.

In deze opzet kan een onverwachte beloning een signaal zijn dat de agent het tot een succes heeft weten brengen. Een beloning is dus goed nieuws. Dit schroeft

het zelfvertrouwen van de agent omhoog, die op zijn beurt de volgende periode meer gemotiveerd is. We contrasteren dit met de literatuur die laat zien hoe een bonus slecht nieuws kan geven. We geven een conditie aan wanneer een beloning goed nieuws is. Deze conditie sluit goed aan bij de vele experimentele resultaten.

In het laatste hoofdstuk, *hoofdstuk 7*, wordt er gekeken naar veranderingen in preferenties. Terwijl de prikkels om gedragsveranderingen teweeg te brengen uitgebreid bestudeerd zijn, is er weinig aandacht geweest binnen economie naar veranderingen in voorkeuren. Sociaal psychologen hebben daarentegen veel aandacht besteed aan de vorming van preferenties, waarbij ze hebben proberen aan te tonen dat voorkeuren niet stabiel zijn. Voorkeuren lijken inderdaad niet stabiel te zijn, maar in veel gevallen wel voorspelbaar. In dit hoofdstuk worden prikkels en veranderingen in voorkeuren onderzocht. Veel van de veranderingen kunnen verklaard worden door aan te nemen dat mensen hun keuze achteraf rationaliseren.

Het basisidee is dat mensen er een onaangenaam gevoel bij krijgen (cognitieve dissonantie) wanneer hun geloof en gedrag niet consistent zijn met elkaar (bijvoorbeeld, je rookt terwijl dat je gelooft dat het slecht is). Om van dit onaangename gevoel af te komen kun je je gedrag proberen te rationaliseren, dus jezelf ervan te overtuigen dat je de goede keuze hebt gemaakt. Dit kan bijvoorbeeld door informatie te zoeken die overeenkomt met je gedrag, en informatie te vermijden die er niet mee in overeenstemming is.

Als toepassing wordt in dit hoofdstuk speciale aandacht besteed aan consumptiegoederen met minder sociale externaliteiten, zoals milieuvriendelijke producten (zie ook hoofdstuk 5). Het belangrijkste resultaat is dat een lage subsidie op dit soort goederen een positieve verandering in de voorkeur ervoor teweegbrengt, maar een hoge subsidie niet. Dit komt overeen met resultaten uit experimenten. De conclusie is dan ook dat hoge subsidies teveel van het goede zijn: ze beïnvloeden het gedrag op dat moment maar niet dat van in de toekomst.

References

Ahlbrecht, M., and M. Weber (1997), Preference for gradual resolution of uncertainty, *Theory and Decision* 43, 167-185.

Ainslie, G. (1991), Deviation of "Rational" Economic Behavior from Hyperbolic Discount Curves, *American Economic Review* 81 (2), 334-340.

Akerlof, G. (1970), The Market for "Lemons": Quality Uncertainty and the Market Mechanism, *Quarterly Journal of Economics* 84 (3), 488-500.

Akerlof, G. (1982), Labor Contracts as Partial Gift Exchange, *Quarterly Journal of Economics* 97, 543–569.

Akerlof, G. (1991), Procrastination and Obedience, *American Economic Review* 81 (2), 1-19.

Akerlof, G. (1997), Social Distance and Social Decisions, *Econometrica* 65, 1005–1027.

Akerlof, G. and W. Dickens (1982), The economic consequences of cognitive dissonance, *American Economic Review* 72 (3), 307-319.

Andreoni, J. (1989), Giving with Impure Altruism: Applications to charity and Ricardian Equivalence, *Journal of Political Economy* 97 (6), 1447-1458.

Andreoni, J. (1990), Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving, *The Economic Journal* 100, 464–477.

Andreoni, J. (1993), An Experimental Test of the Public-Goods Crowding-out Hypothesis, *American Economic Review* 83 (5), 1317-1327.

Andreoni, J. (2001), The Economics of Philantropy, in N. Smeltser and P. Baltes (eds.), *International Encyclopedia of Social and Behavioural Sciences,* Elsevier Pergamon, Oxford.

Andreoni, J., and R. Petrie (2000), Social Motives to Giving: Can these Explain Fund Raising Institutions?, Mimeo, University of Wisconsin Madison.

Aristotle (1994), *The Nicomachean Ethics*, translated by H. Rackham, Harvard University Press: Harvard.

Arnott, R., and J. Stiglitz (1991), "Moral Hazard and Nonmarket Institutions: Dysfunctional Crowding Out or Peer Monitoring?" *American Economic Review* 81, 179–190.

Arnsperger, C. (2001), *Gift-giving Practices and Non-contextual Habitus*, in *Gifts and Interests*, A. Vandevelde (ed.), Peeters.

Aronson, E. (1988), *The Social Animal,* W.H. Freeman and Company, New York, 5th Edition.

Aronson, E., Turner, J. and M. Carlsmith (1963), Communicator credibility and communicator discrepancy as determinants of opinion change, *Journal of Abnormal and Social Psychology* 67, 31-36.

Arrow, K. (1972), Gifts and Exchanges, *Philosophy and Public Affairs* 1 (4), 343-362.

Arrow, K. (1997), Invaluable Goods, *Journal of Economic Literature* XXXV, 757-765.

Axelrod, R. (1984), *The Evolution of Cooperation*, Basic Books.

Baker, G., R. Gibbons, and K. Murphy (1994), Subjective performance measures in optimal incentive contracts, *Quarterly Journal of Economics* CIX (4), 1125-1156.

Bansal, S. (2002), *Environmental Regulation in the Presence of Environmentally Conscious Consumers*, Ph.D. thesis Indian Statistical Institute, Delhi.

Barkow, J. (1989), Darwin, Sex, and Status, University of Toronto Pess, Toronto.

Baumeister, R. (1998), *The Self*, in: D. Gilbert, S. Fiske, and G. Lindzey (eds.), The Handbook of Social Psychology, McGraw-Hill, New York.

Baumol, W. and W. Oates (1988), *The Theory of Environmental Policy*, Cambridge University Press, Cambridge, 2nd ed.

BBC News (2002), Diamonds: A rebel's best friend, http://news.bbc.co.uk/1/hi/world/africa/745194.stm

Becker, G. (1974), A Theory of Social Interactions, *Journal of Political Economy,* 82, 1063–1093.

Becker, G. (1996), *Accounting for Tastes*, Harvard University Press.

Bem, D. (1972), *Self-Perception Theory*, in L. Berkowitz (ed.), Advances in experimental social psychology vol. 6, academic Press, New York.

Bénabou, R., and J. Tirole (2001), Willpower and Personal Rules, Mimeo November.

Bénabou, R., and J. Tirole (2002a), Intrinsic and Extrinsic Motivation, forthcoming in *Review of Economic Studies.*

Bénabou, R., and J. Tirole (2002b), Self-Confidence and Personal Motivation, *Quarterly Journal of Economics* 117(3), 871-915.

Berkowitz, E., R. Kerim, S. Hartley and W. Rudelius (1994), *Marketing*, Irwin (4th ed.).

Bernheim, B., and O. Stark (1988), Altruism within the family reconsidered: do nice guys finish last?, *American Economic Review* 78 (5), 1034-1045.

Binmore, K. (1998), *Game Theory and the Social Contract*, vol. I, MIT Press.

Bjørner, T., L. Hansen, C. Russell, and T. Olsen (2002), The Effect of the Nordic Swan Label on Consumers' Choice, AKF Forlaget.

Bolle, F. (2000), Why to buy your darling flowers: on cooperation and exploitation, *Theory and Decision*, forthcoming.

Bolton, G. and E. Katok (1998), An experimental test of the crowding out hypothesis: The nature of beneficent behavior, *Journal of Economic Behavior & Organization* 37, 315-221.

Bolton, G. and A. Ockenfels (1998), Strategy and Equity: An ERC-analysis of the Güth-van Damme game, *Journal of Mathematical Psychology* 42, 215-226.

Bolton, G., and A. Ockenfels (2000), ERC: A Theory of Equity, Reciprocity, and Competition, *American Economic Review,* 90, 166–193.

Bowles, S. (1998), Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions, *Journal of Economic Literature*, XXXVI, 75–111.

Brehm, J. (1956), Postdecision changes in the desirability of alternatives, *Journal of Abnormal and social Psychology* 52, 384-389.

Brigham, J. (1991), *Social Psychology*, Harper Collins Publishers, 2nd ed.

Camerer, C. (1988), Gifts as Economic Signals and Social Symbols, *American Journal of Sociology,* 94, S180–S214.

Camerer, C., L. Babcock, G. Loewenstein, R. Thaler (1997), Labor Supply of New York City Cab Drivers: One Day at a Time, *Quarterly Journal of Economics* 112 (2), 407-441.

Caplow, T. (1982), Christmas Gifts and Kin Networks, *American Sociological Review,* 47, 383–392.

Carmichael, H. and W. MacLeod (1997), Gift giving and the evolution of cooperation, *International Economic Review* 38 (3), 485-509.

Carrillo, J., and T. Mariotti (2000), Strategic ignorance as a self-disciplining device, *Review of Economic Studies* 67, 529-544.

Charness, G. and M. Rabin (2002), Understanding Social Preferences with Simple Tests, *Quarterly Journal of Economics*, forthcoming.

Charness, G., G. Frechette, and J. Kagel (2001), How Robust is Laboratory Gift Exchange, Mimeo.

Cho, I., and D. Kreps (1987), Signalling games and stable equilibria, *Quarterly Journal of Economics* 102, 179-221.

Codere, H. (1950), *Fighting with property*, Unversity of Washington Press.

Coleman, J. (1990), *Foundations of Social Theory*, Belknap Press: Cambridge, MA.

Cooper, R., D. de Jong, R. Forsythe, and T. Ross (1996), Cooperation Without Reputation, Experimental Evidence from Prisoner's Dilemma Games, *Games and Economic Behavior,* 12, 187–218.

Cooper, B., C. Garia-Penalosa, and P. Funk (2001), Status Effects and Negative Utility Growth, *The Economic Journal* 111 (473),.642-665.

Corneo, G., and O. Jeanne (1997) Conspicuous consumption, snobbism and conformism, *Journal of Public Economics* 66 (1), 55-71.

Croyle, R. and J. Cooper (1983), Dissonance Arousal: Physiological Evidence, *Journal of Personality and Social Psychology* 45, 782-791.

Dasgupta, P. (2000), Economic Progress and the Idea of Social Capital,325–424 in: P. Dasgupta and I. Serageldin (eds.), *Social Capital: A Multifaceted Perspective*, The World Bank: Washington.

Deci, E., and R. Ryan (1985), *Intrinsic Motivation and Self-Determination in Human Behavior*, Plenum Press, New York.

Deci, E., R. Koestner, and R. Ryan (1999), A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation, *Psychological Bulletin* 125 (6), 627-668.

Diamond, P. (1982), Aggregate-Demand Management in Search Equilibrium, *Journal of Political Economy* 90 (5), 881-894.

Diamond, P. (1984), Money in Search Equilibrium, *Econometrica* 52 (1), 1-20.

Dibb, S., L. Sinhin, W. Pride and O. Ferrell (1997), *Marketing: Concepts and Strategies*, Houghton Mifflin, Boston (3rd European ed.).

Dickens, W. (1986), Crime and punishment again: the economic approach with a psychological twist, *Journal of Public Economics* 30, 97-107.

Diecidue, E., and J. van de Ven (2003), Aspiration Levels in Decision Making, Mimeo INSEAD/Tilburg University.

Diecidue, E., and P. Wakker (2001), On the Intuition of Rank-Dependent Utility, *The Journal of Risk and Uncertainty* 23 (3), 281-298.

Douglas, M. and B. Isherwoord (1978), *The world of goods*, Allen lane (ed. 1979).

Drucker, P. and R. Heizer (1967), *To make my name good*, University of California Press.

Economist, The (2003), Regulating the diamond trade: A crook's best friend, January 4th.

Etzioni, A. (1988), *The Moral Dimension: toward a new economics,* the Free Press.

Falk, A., E. Fehr, and U. Fischbacher (1999), On the Nature of Fair Behavior, Working Paper no. 17 University of Zürich.

Fehr, E. and S. Gächter (1999), Reciprocity and Economics, Implications of Reciprocal behaviour for labour markets, Mimeo University of Zurich.

Fehr, E., and S. Gächter (2000), Fairness and Retaliation: The Economics of Reciprocity, *Journal of Economic Perspectives,* 14, 159–181.

Fehr, E., and K. Schmidt (1999), A Theory of Fairness, Competition, and Cooperation, *Quarterly Journal of Economics*, 114, 817–868.

Festinger, L. (1957), *A theory of cognitive dissonance*, Row, Peterson and Company.

Festinger, L., and J. Carlsmith (1959), Cognitive consequences of forced compliance, *Journal of Abnormal and Social Psychology* 58, 203-210.

Forge, A. (1972), The Golden Fleece, *Man* 15 (4), 527-540.

Frank, R. (1985), *Choosing the Right Pond: Human Behavior and the Quest for Status*, Oxford University Press: Oxford.

Frank, R. (1988), *Passions within Reason: the strategic role of emotions,* W.W. Norton & Company, New York.

Frank, R. (1997), The Frame of Reference as a Public Good, *The Economic Journal* 107 (445), 1832-1847

Freedman, J. (1963), Long-term behavioral effects of cognitive dissonance, *Journal of Experimental Social Psychology* 1, 145-155.

Frey, B. (1997a), A Constitution for Knaves Crowds Out Civic Virtues, *The Economic Journal,* 107, 1043–1053.

Frey, B. (1997b), *Not Just for the Money: An Economic Theory of Personal Motivation*, Edward Elgar: Brookfield.

Frey, B., and R. Jegen (2002), Motivation Crowding Theory: A Survey of Empricial Evidence, *Journal of Economic Surveys*, forthcoming.

Friedman, M., and L. Savage (1948), The Utility Analysis of Choices Involving Risk, *Journal of Political Economy* 56 (4), 279-304.

Fudenberg, D., and J. Tirole (1991), *Game Theory*, MIT Press, Cambridge MA..

Fukuyama, F. (1999), *The great disruption; Human nature and the reconstruction of social order*, the Free Press, New York.

Gächter, S., and E. Fehr (1999), Collective Action as a Social Exchange, *Journal of Economic Behavior & Organization* 39, 341–369.

Gächter, S., and A. Falk (1999), Reputation or Reciprocity,Working Paper 19, University of Zurich, Zurich.

Galor, O., and J. Zeira (1993), Income Distribution and Macroeconomics, *Review of Economic Studies,* 60, 35–52.

Glazer, A., and K. Konrad (1996), A Signaling Explanation for Charity, *American Economic Review,* 86, 1019–1028.

Gneezy, U., and A. Rustichini (2000), A Fine is a Price, *The Journal of Legal Studies,* 29, 1–18.

Goeree, J., C. Holt, and S. Laury (2002), Private Costs and Public Benefits: Unraveling the Effects of Altruism and Noisy Behavior, *Journal of Public Economics*, 83, 255-276.

Goldman, S. (1978), Comment: gift equilibria and Pareto optimality, *Journal of Economic Theory* 18, 368-370.

Grant, S., A. Kajii, and B. Polak (1998), Intrinsic Preference for Information, *Journal of Economic Theory* 83, 233-259.

Grant, S., A. Kajii, and B. Polak (2000), Preference for Information and Dynamic Consistency, *Theory and Decision* 48, 263-286.

Gregory, C. (1989), Gifts,109–118 in: J. Eatwell et al. (eds.), *The New Palgrave Social Economics*, Macmillan: London.

Güth, W. (1995), On Ultimatum Bargaining Experiments: A personal review, *Journal of Economic Behaviour and Organization* 27, 329-344.

Güth, W., and H. Kliemt (1994), Competition or Co-Operation: On the Evolutionary Economics of Trust, Exploitation and Moral Attitudes, *Metroeconomica,* 45, 155–187.

Güth, W., and H. Kliemt (1998), The Indirect Evolutionary Approach, *Rationality and Society,* 10, 377–399.

Güth, W. and E. van Damme (1998), Information, Strategic Behavior, and Fairness in ultimatum bargaining: an experimental study, *Journal of Mathematical Psychology* 42, 227-247.

Harbaugh, T. (1998), The Prestige Motive for Making Charitable Transfers, *American Economic Review,* 88, 277–282.

Harsanyi, J. (1969), Rational-Choice Models of Political Behavior vs. Functionalist and Conformist Theories, *World Politics,* 22, 513–538.

Henry, O. (1906) The Gift of the Magi and Other Short Stories, Dover Thrift, New York (ed. 1992).

Hirsch, F. (1976), *Social Limits to Growth*, Harvard University Press: Cambridge, MA.

Holländer, H. (1990), A Social Exchange Approach to Voluntary Cooperation, *American Economic Review,* 80, 1157–1167.

Homans, G. (1958),Social Behavior as Exchange, *American Journal of Sociology,* 63, 597–606.

Homans, G. (1961), *Social Behaviour: Its Elementary Forms*, Routlegde & Kegan Paul: London.

Hume, D. (1886), *A Treatise of Human Nature*, Book II, in T. Green and T. Grose (eds.), *Hume: The Philosophical Works*, Scienta Verlag Aalen.

Iannaccone, L. (1992), Sacrifice and Stigma: reducing free-riding in cults, communes, and other collectives, *Journal of Political Economy* 100 (2), 271-291.

James, J., and E. Gutkind (1985), Attitude change revisited: Cognitive dissonance theory and development policy, *World Development* 13 (10/11), 1139-1149.

Jones, E., and R. Kohler (1959), The effects of plausibility on the learning of controversial statements, *Journal of Abnormal and Social Psychology* 57, 315-320.

Kahneman, D. (1994), New challenges to the rationality assumption, *Journal of Institutional and Theoretical Economics* 150, 18-36.

Kahneman, D. (2000), *Experienced Utility and Objective Happiness: A moment-based approach*, in: Kahneman, D. and A. Tversky (eds.), Choices, Values, and Frames, Cambridge University Press, New York.

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

Kahneman, D., P. Wakker, and R. Sarin (1997) Back to Bentham? Explorations of Experienced Utility, *Quarterly Journal of Economics* 112 (2), 375-405.

Kenrick, D., S. Neuberg, and R. Cialdini (1999), *Social Psychology*, Allyn and Bacon: Boston.

Keyzer, M. (2002), Labeling and the realization of cultural values, *De Economist* 150 (4), 487-511.

Khalil, E. (1997), Etzioni versus Becker: Do moral sentiments differ from ordinary tastes?, *De Economist* 145 (4), 491-520.

Kirchhoff, S. (2000), Green Business and Blue Angels, *Environmental and Resource Economics* 15, 403-420.

Kohn, A. (1993), *Punished by Rewards*, Houghton-Mifflin, Boston.

Kranich, L. (1994), Gift Equilibria and Pareto Optimality Reconsidered, *Journal of Economic Theory,* 64, 298–300.

Kranton, R. (1996), Reciprocal Exchange: A Self-Sustaining System, *American Economic Review,* 86, 830–851.

Kranton, R. (1996b), The formation of cooperative relationships, *Journal of Law, Economics, & Organization*, 214-233.

Kropotkin, P. (1904), *Mutual Aid; A factor of evolution*, William Heinemann, London.

Laibson, D. (1997), Golden Eggs and Hyberbolic Discounting, *Quarterly Journal of Economics,* 443-477.

Lancaster, K. (1966), A New Approach to Consumer Theory, *Journal of Political Economy,* 74, 132–157.

Leibenstein, H. (1950), Bandwagon, snob, and Veblen effects in the theory of consumer's demand, *Quarterly Journal of Economics* 64 (2), 183-207.

Lepper, M., D. Greene, and R. Nisbett (1973), Undermining Children's Interest with Ectrinsic Reward: A test of the "overjustification" hypothesis, Journal of Personality and social Psychology 28 (1), 129-173.

Loewenstein, G., and D. Adler (1995), A bias in the prediction of tastes, *The Economic Journal* 105, 929-937.

Loomes, G. (1998), Probabilities vs Money: A test of some fundamental assumptions about rational decision making, *The Economic Journal,* 108, 477-489.

Lopes, L. (1984), Risk and Distributional Inequality, *Journal of Experimental Psychology: Human Perception and Performance* 10 (4), 465-485.

Magat, W. and W. Viscusi (1992), *Informational Approaches to Regulation*, MIT Press, Cambridge, Mass.

Mandeville, B. (1714) *The fable of the bees: or, private vices, public benefits,* Oxford University Press, London (ed. 1957).

Mankiw, G. (1998), *Principles of Economics*, Dryden Press.

Mas-Colell, A., M. Whinston, and J. Green (1995), *Microeconomic Theory*, Oxford University Press, New York.

Mason, C. (2002), On the Economics of Eco-Labeling, Mimeo University of Wyoming.

Mauss, M. (1925/1980), *The Gift*, translated from *Essai sur le don,* Routledge and Kegan Paul: London.

Milgrom, P. and J. Roberts (1986), Price and Advertising Signals of Product Quality, *Journal of Political Economy* 94 (4), 796-821.

Mullanaithan, S., and R. Thaler (2000), Behavioral Economics, *NBER Working Paper* 7948.

Nagel, R. (1995), Unraveling in Guesssing Games, An experimental Study, *American Economic Review* 85(5), 1313-1326.

Ng, Y. (1997), A Case for Happiness, Cardinalism, and Interpersonal Comparability, *The Economic Journal* 107 (445), 1848-1858

Payne, J., D. Laughhunn, and R. Crum (1980), Translation of gambles and aspiration level effects in risky choice behavior, *Management Science* 26 (10), 1039-1060.

Payne, J., D. Laughhunn, and R. Crum (1981), Further tests of aspiration level effects in risky choice behavior, *Management Science* 27 (8), 953-958.

Polanyi, K. (1957), *The Great Transformation*, Beacon Press: Boston.

Prendergast, C., and L. Stole (2001), The non-monetary nature of gifts, *European Economic Review* 45 (10), 793-1810.

Pruitt, D. (1968), Reciprocity and Credit Building in a Laboratory Dyad, *Journal of Personality and Social Psychology,* 8, 143–147.

Rabin, M. (1993), Incorporating fairness into game theory and economics, *American Economic Review* 83 (5), 1281-1301.

Rabin, M. (1994), Cognitive dissonance and social change, *Journal of Economic Behavior and Organization* 23, 177-194

Rabin, M. (1995), Moral preferences, moral constraints, and self-serving biases, Berkeley Department of Economics Working Paper No 95-241.

Rabin, M. (1998), Psychology and Economics, *Journal of Economic Literature* XXXVI, 11-46.

Rabin, M. (2000), *Diminishing Marginal Utility of Wealth Cannot Explain Risk Aversion*, in Kahneman, D., and A. Tversky (eds.), *Choices, Values, and Frames*, Cambridge University Press, NeW York.

Rabin, M. (2002), A Perspective on Psychology and Ecpnomics, *European Economic Review* 46, 657-685.

Robben, H., and T. Verhallen (1994), Behavioral Costs as Determinants of Cost Perception and Preference Formation for Gifts to Receive and Gifts to Give, *Journal of Economic Psychology,* 15, 333–350.

Roth, A (1995), *Introduction to Experimental Economics*, in:Kagel, J. and A. Roth (eds.), *Handbook of Experimental Economics*, Princeton University Press, Princeton.

Rubinstein, A. (2001), A theorist's view of experiments, *European Economic Review* 45, 651-218.

Ruffle, B. (1999), Gift-giving with emotions, *Journal of Economic Behavior and Organization* 39, 399-420.

Sapolsky, R. (1998), *Why zebras don't get ulcers*, Freeman, New York.

Satow, K. (1975), Social Approval and Helping, *Journal of Experimental Social Psychology,* 11, 501–509.

Schwartz, B. (1967), The Social Psychology of the Gift, *The American Journal of Sociology,* 73, 1–11.

Slovic, P. (1995), The construction of preference, *American Psychologist* 50 (5), 364-371.

Smith, V. (1962), An Experimental Study of Competitive Behavior, *Journal of Political Economy* 70 (2), 111-137.

Solnick, S., and D. Hemenway (1996), The Deadweight Loss of Christmas: Comment, *American Economic Review,* 86, 1299–1305.

Souvorov, A. (2003), Addiction to Rewards, Mimeo GREMAQ, Toulouse.

Spence, M. (1973), Job market signaling, *Quarterly Journal of Economics* 87, 355-374.

Starmer, C. (1998), Developments in Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk, *Journal of Economic Literature* XXXVIII, 332-382.

Sugden, R. (1989), Spontaneous Order, *Journal of Economic Perspectives,* 3, 85–97.

Tirole, J. (1988), *The Theory of Industrial Organisation*, MIT Press, Cambridge, Mass.

Tirole, J. (2002), Rational irrationality: Some economics of self-management, *European Economic Review* 46, 633-655.

Thaler, R. (1980), Toward a positive Theory of Consumer Choice, *Journal of Economic Behavior and Organization* 1, 39-60.

Thaler, R. (1999), Mental Accounting Matters, *Journal of Behavioral Decision Making* 12, 183-206.

Thaler, R. (2000), From Homo Economicus to Homo Sapiens, *Journal of Economic Perspectives* 14 (1), 133-141.

Titmuss, R. (1970), *The Gift Relationship*, George Allen & Unwin Ltd.: London.

Trivers, R. (1971), The evolution of reciprocal altruism, *Quarterly Review of Biology* 46, 35-57.

Tversky, A., and D. Kahneman (1992), Advances in prospect theory, cumulative representation of uncertainty, *Journal of Risk and Uncertainty 5, 297-323.*

Van Damme, E. (1989), Stable equilibria and forward induction, *Journal of Economic Theory* 48 (2), 476-496.

VandeVelde, A. (2000), Towards a Conceptual Map of Gift Practices, in *Gifts and Interests*, A. Vandevelde (ed.), Peeters.

Van Kempen, L. (2003), Fooling the eye of the Beholder: Deceptive Status Signalling among the Poor in Developing Countries, *Journal of International Development*, 15, 157-177.

Varian, H. (1992), *Microeconomic Analysis*, W.W. Norton & Company, New York (3rd ed.).

Veblen, T. (1899), *The theory of the leisure class : an economic study of institutions*, The New American Library, New York (ed. 1953).

Wakker, P. (1988), Nonexpected Utility as Aversion of Information, *Journal of Behavioral Decision Making* 1, 169-175.

Waldfogel, J. (1993), The Deadweight Loss of Christmas, *American Economic Review,* 83, 1329–1336.

Waldfogel, J. (1996), The Deadweight Loss of Christmas: Reply, *American Economic Review,* 86, 1306–1308.

Webley, P., and R. Wilson (1989), Social Relationships and the Unacceptability of Money as a Gift, *Journal of Social Psychology,* 129, 85–91.

Williamsion, O. (1985), *The Economic Institutions of Capitalism*, Free Press, New York.

Wright, R. (1994), *The Moral Animal*, Brown & Company: London.

Yellen, J. (1990), The Transformation of the Kalahari !Kung Tribe, *Scientific American,* 262, April, 96–105.

Zwick, R., I. Erev, and D. Budescu (1999), The Psychological and Economical Perspectives on Human Decisions in Social and Interactive Contexts, in Budescu, D., I. Erev, and R. Zwick (eds.), *Games and Human Behavior: Essays in honor of Amnon Rapoport*, Lawrence Erlbaum Associates, New Jersey.

Center for Economic Research, Tilburg University, The Netherlands
Dissertation Series

No.	Author	Title	Published
1	P.J.J. Herings	Static and Dynamic Aspects of General Disequilibrium Theory; ISBN 90 5668 001 3	June 1995
2[1]	Erwin van der Krabben	Urban Dynamics: A Real Estate Perspective - An institutional analysis of the production of the built environment; ISBN 90 5170 390 2	August 1995
3	Arjan Lejour	Integrating or Desintegrating Welfare States? - a qualitative study to the consequences of economic integration on social insurance; ISBN 90 5668 003 X	September 1995
4	Bas J.M. Werker	Statistical Methods in Financial Econometrics; ISBN 90 5668 002 1	September 1995
5	Rudy Douven	Policy Coordination and Convergence in the EU; ISBN 90 5668 004 8	September 1995
6	Arie J.T.M. Weeren	Coordination in Hierarchical Control; ISBN 90 5668 006 4	September 1995
7	Herbert Hamers	Sequencing and Delivery Situations: a Game Theoretic Approach; ISBN 90 5668 005 6	September 1995
8	Annemarie ter Veer	Strategic Decision Making in Politics; ISBN 90 5668 007 2	October 1995
9	Zaifu Yang	Simplicial Fixed Point Algorithms and Applications; ISBN 90 5668 008 0	January 1996
10	William Verkooijen	Neural Networks in Economic Modelling - An Empirical Study; ISBN 90 5668 010 2	February 1996
11	Henny Romijn	Acquisition of Technological Capability in Small Firms in Developing Countries; ISBN 90 5668 009 9	March 1996
12	W.B. van den Hout	The Power-Series Algorithm - A Numerical Approach to Markov Processes; ISBN 90 5668 011 0	March 1996
13	Paul W.J. de Bijl	Essays in Industrial Organization and Management Strategy; ISBN 90 5668 012 9	April 1996

No.	Author	Title	Published
14	Martijn van de Ven	Intergenerational Redistribution in Representative Democracies; ISBN 90 5668 013 7	May 1996
15	Eline van der Heijden	Altruism, Fairness and Public Pensions: An Investigation of Survey and Experimental Data; ISBN 90 5668 014 5	May 1996
16	H.M. Webers	Competition in Spatial Location Models; ISBN 90 5668 015 3	June 1996
17	Jan Bouckaert	Essays in Competition with Product Differentiation and Bargaining in Markets; ISBN 90 5668 016 1	June 1996
18	Zafar Iqbal	Three-Gap Analysis of Structural Adjustment in Pakistan; ISBN 90 5668 017 X	September 1996
19	Jimmy Miller	A Treatise on Labour: A Matching-Model Analysis of Labour-Market Programmes; ISBN 90 5668 018 8	September 1996
20	Edwin van Dam	Graphs with Few Eigenvalues - An interplay between combinatorics and algebra; ISBN 90 5668 019 6	October 1996
21	Henk Oosterhout	Takeover Barriers: the good, the bad, and the ugly; ISBN 90 5668 020 X	November 1996
22	Jan Lemmen	Financial Integration in the European Union: Measurement and Determination; ISBN 90 5668 021 8	December 1996
23	Chris van Raalte	Market Formation and Market Selection; ISBN 90 5668 022 6	December 1996
24	Bas van Aarle	Essays on Monetary and Fiscal Policy Interaction: Applications to EMU and Eastern Europe; ISBN 90 5668 023 4	December 1996
25	Francis Y. Kumah	Common Stochastic Trends and Policy Shocks in the Open Economy: Empirical Essays in International Finance and Monetary Policy; ISBN 90 5668 024 2	May 1997
26	Erik Canton	Economic Growth and Business Cycles; ISBN 90 5668 025 0	September 1997

No.	Author	Title	Published
27	Jeroen Hoppenbrouwers	Conceptual Modeling and the Lexicon; ISBN 90 5668 027 7	October 1997
28	Paul Smit	Numerical Analysis of Eigenvalue Algorithms Based on Subspace Iterations; ISBN 90 5668 026 9	October 1997
29	Uri Gneezy	Essays in Behavioral Economics; ISBN 90 5668 028 5	October 1997
30	Erwin Charlier	Limited Dependent Variable Models for Panel Data; ISBN 90 5668 029 3	November 1997
31	Rob Euwals	Empirical Studies on Individual Labour Market Behaviour; ISBN 90 5668 030 7	December 1997
32	Anurag N. Banerjee	The Sensitivity of Estimates, Inferences, and Forecasts of Linear Models; ISBN 90 5668 031 5	December 1997
33	Frans A. de Roon	Essays on Testing for Spanning and on Modeling Futures Risk Premia; ISBN 90 5668 032 3	December 1997
34	Xiangzhu Han	Product Differentiation, Collusion and Standardization; ISBN 90 5668 033 1	January 1998
35	Marcel Das	On Income Expectations and Other Subjective Data: A Micro-Econometric Analysis; ISBN 90 5668 034 X	January 1998
36	Jeroen Suijs	Cooperative Decision Making in a Stochastic Environment; ISBN 90 5668 035 8	March 1998
37	Talitha Feenstra	Environmental Policy Instruments and International Rivalry: A Dynamic Analysis; ISBN 90 5668 036 6	May 1998
38	Jan Bouwens	The Use of Management Accounting Systems in Functionally Differentiated Organizations; ISBN 90 5668 037 4	June 1998
39	Stefan Hochguertel	Households' Portfolio Choices; ISBN 90 5668 038 2	June 1998
40	Henk van Houtum	The Development of Cross-Border Economic Relations; ISBN 90 5668 039 0	July 1998
41	Jorg Jansen	Service and Inventory Models subject to a Delay-Limit; ISBN 90 5668 040 4	September 1998

No.	Author	Title	Published
42	F.B.S.L.P. Janssen	Inventory Management Systems: control and information issues; ISBN 90 5668 041 2	September 1998
43	Henri L.F. de Groot	Economic Growth, Sectoral Structure and Unemployment; ISBN 90 5668 042 0	October 1998
44	Jenke R. ter Horst	Longitudinal Analysis of Mutual Fund Performance; ISBN 90 5668 043 9	November 1998
45	Marco Hoeberichts	The Design of Monetary Institutions; ISBN 90 5668 044 7	December 1998
46	Adriaan Kalwij	Household Consumption, Female Employment and Fertility Decisions: A microeconometric analysis; ISBN 90 5668 045 5	February 1999
47	Ursula Glunk	Realizing High Performance on Multiple Stakeholder Domains: A Resource-based Analysis of Professional Service Firms in the Netherlands and Germany; ISBN 90 5668 046 3	April 1999
48	Freek Vermeulen	Shifting Ground: Studies on the Intersection of Organizational Expansion, Internationalization, and Learning; ISBN 90 5668 047 1	February 1999
49	Haoran Pan	Competitive Pressures on Income Distribution in China; ISBN 90 5668 048 X	March 1999
50	Laurence van Lent	Incomplete Contracting Theory in Empirical Accounting Research; ISBN 90 5668 049 8	April 1999
51	Rob Aalbers	On the Implications of Thresholds for Economic Science and Environmental Policy; ISBN 90 5668 050 1	May 1999
52	Abe de Jong	An Empirical Analysis of Capital Structure Decisions in Dutch Firms; ISBN 90 5668 051 X	June 1999
53	Trea Aldershof	Female Labor Supply and Housing Decisions; ISBN 90 5668 052 8	June 1999
54	Jan Fidrmuc	The Political Economy of Reforms in Central and Eastern Europe; ISBN 90 5668 053 6	June 1999
55	Michael Kosfeld	Individual Decision-Making and Social Interaction; ISBN 90 5668 054 4	June 1999

No.	Author	Title	Published
56	Aldo de Moor	Empowering Communities - A method for the legitimate user-driven specification of network information systems; ISBN 90 5668 055 2	October 1999
57	Yohane A. Khamfula	Essays on Exchange Rate Policy in Developing Countries; ISBN 90 5668 056 0	September 1999
58	Maurice Koster	Cost Sharing in Production Situations and Network Exploitation; ISBN 90 5668 057 9	December 1999
59	Miguel Rosellón Cifuentes	Essays on Financial Policy, Liquidation Values and Product Markets; ISBN 90 5668 058 7	December 1999
60	Sharon Schalk	Equilibrium Theory: A salient approach; ISBN 90 5668 059 5	December 1999
61	Mark Voorneveld	Potential Games and Interactive Decisions with Multiple Criteria; ISBN 90 5668 060 9	December 1999
62	Edward Droste	Adaptive Behavior in Economic and Social Environments; ISBN 90 5668 061 7	December 1999
63	Jos Jansen	Essays on Incentives in Regulation and Innovation; ISBN 90 5668 062 5	January 2000
64	Franc J.G.M. Klaassen	Exchange Rates and their Effects on International Trade; ISBN 90 5668 063 3	January 2000
65	Radislav Semenov	Cross-country Differences in Economic Governance: Culture as a major explanatory factor; ISBN 90 5668 065 X	February 2000
66	Alexandre Possajennikov	Learning and Evolution in Games and Oligopoly Models; ISBN 90 5668 064 1	March 2000
67	Eric Gaury	Designing Pull Production Control Systems: Customization and Robustness; ISBN 90 5668 066 8	March 2000
68	Marco Slikker	Decision Making and Cooperation Restrictions; ISBN 90 5668 067 6	April 2000
69	Ruud Brekelmans	Stochastic Models in Risk Theory and Management Accounting; ISBN 90 5668 068 4	June 2000
70	Flip Klijn	A Game Theoretic Approach to Assignment Problems; ISBN 90 5668 069 2	June 2000

No.	Author	Title	Published
71	Maria Montero	Endogenous Coalition Formation and Bargaining; ISBN 90 5668 070 6	June 2000
72	Bas Donkers	Subjective Information in Economic Decision Making; ISBN 90 5668 071 4	June 2000
73	Richard Nahuis	Knowledge and Economic Growth; ISBN 90 5668 072 2	October 2000
74	Kuno Huisman	Technology Investment: A Game Theoretic Real Options Approach; ISBN 90 5668 073 0	October 2000
75	Gijsbert van Lomwel	Essays on Labour Economics; ISBN 90 5668 074 9	November 2000
76	Tina Girndt	Cultural Diversity and Work-Group Performance: Detecting the Rules; ISBN 90 5668 075 7	December 2000
77	Teye A. Marra	The Influence of Proprietary Disclosure Costs on the Decision to Go Public: Theory and Evidence; ISBN 90 5668 076 5	January 2001
78	Xiaodong Gong	Empirical Studies on the Labour Market and on Consumer Demand; ISBN 90 5668 078 1	February 2001
79	Kanat Çamlibel	Complementarity Methods in the Analysis of Piecewise Linear Dynamical Systems; ISBN 90 5668 079 X	May 2001
80	Joost Driessen	Empirical Studies on the Pricing of Bonds and Interest Rate Derivatives; ISBN 90 5668 080 3	June 2001
81	Bram van den Broek	Uncertainty in Differential Games; ISBN 90 5668 081 1	June 2001
82	Ricardo Recht	Organisational Culture and Privatisation: A Case Study of the Argentinean Railway Sector; ISBN 90 5668 082 X	June 2001
83	Willem Verhagen	Inflation Targeting and Interest Rate Policy; ISBN 90 5668 083 8	June 2001
84	Vincent Verouden	Essays in Antitrust Economics; ISBN: 90 5668 086 2	June 2001
85	Rosalia Vazquez Alvarez	A Nonparametric Approach to the Sample Selection Problem in Survey Data; ISBN 90 5668 087 0	June 2001
86	Theo Leers	Public Pensions and Population Ageing: An Economic Analysis of Fertility, Migration, and Social-Security Policy; ISBN 90 5668 084 6	June 2001

No.	Author	Title	Published
87	Bernard Conlon	Consumer Rationality in Choice; ISBN 90 5668 085 4	June 2001
88	Judith Timmer	Cooperative behaviour, uncertainty and operations research; ISBN 90 5668 088 9	September 2001
89	Enrico Diecidue	Nonexpected Utility and Coherence; ISBN 90 5668 089 7	November 2001
90	Alexander Konovalov	Essays in General Equilibrium Theory; ISBN 90 5668 091 9	November 2001
91	Stefan Stremersch	Essays on Marketing Strategy in Technology-Intensive Markets; ISBN 90 5668 090 0	December 2001
92	Rian Drogendijk	Expansion Patterns of Dutch Firms in Central and Eastern Europe: Learning to Internationalize; ISBN 90 5668 092 7	December 2001
93	Han Donker	Takeovers. An Empirical Study of Ownership Structure and the Wealth to Shareholders in Dutch Takeovers; ISBN 90 5668 093 5	December 2001
94	Erica van Herpen	Perceptions and Evaluations of Assortment Variety; ISBN 90 5668 094 3	December 2001
95	Jana Vyrastekova	Cheap Talk and Spiteful Preferences in Ultimatum Games: Experiments and Evolutionary Rationale; ISBN 90 5668 095 1	March 2002
96	Tao Lin	Statistics of Extremes in the Space of Continuous Functions; ISBN 90 5668 096 X	May 2002
97	Sander Onderstal	Papers in Auction Theory; ISBN 90 5668 097 8	May 2002
98	Wolf Wagner	Risk Sharing under Incentive Constraints; ISBN 90 5668 098 6	June 2002
99	Willem-Jan van den Heuvel	Integrating Modern Business Applications with Objectified Legacy Systems; ISBN 90 5668 099 4	June 2002
100	Jacco Wielhouwer	Optimal Tax Depreciation and its Effects on Optimal Firm Investments; ISBN 90 5668 100 1	June 2002
101	Peter Roosenboom	Corporate Governance Mechanisms in IPO Firms; ISBN 90 5668 101 X	September 2002

No.	Author	Title	Published
102	Victoria Shestalova	Essays in Productivity and Efficiency; ISBN 90 5668 103 6	September 2002
103	Paula van Veen-Dirks	Flexibility and Control: An Empirical Study Relating Production Flexibility to the Design of Performance Management Systems; ISBN 90 5668 102 8	September 2002
104	Michal Matějka	Management Accounting in Organizational Design: Three Essays; ISBN 90 5668 105 2	November 2002
105	Maarten Pronk	Market Liquidity around Earnings Announcements; ISBN 90 5668 106 0	November 2002
106	Nancy Kamp-Roelands	Towards a framework for auditing environmental reports; ISBN 90 5668 108 7	November 2002
107	Ana Ferreira	Statistics of extremes: estimation and optimality; ISBN 90 5668 107 9	November 2002
108	Alexei Goriaev	On the behaviour of Mutual Fund Investors and Managers; ISBN 90 5668 110 9	November 2002
109	Paul Ingenbleek	Money for Value Pricing from a Resource-Advantage Perspective; ISBN 90 5668 109 5	December 2002
110	Frank Wijen	Stakeholder influence and organizational learning in environmental management; ISBN 90 5668 111 7	December 2002
111	Michele Belot	Labor market institutions in OECD countries: Origins and consequences; ISBN 90 5668 113 3	January 2003
112	Federica Teppa	Estimating preference parameters in individual financial decision making; ISBN 90 5668 112 5	January 2003
113	Dmitry Danilov	The effects of pretesting in econometrics with applications in finance; ISBN 90 5668 104 4	February 2003
114	Jacco Thijssen	Investment under uncertainty, market evolution and coalition spillovers in a game theoretic perspective; ISBN 90 5668 114 1	May 2003

No.	Author	Title	Published
115	Laura Spierdijk	Empirical Studies of Market Microstructure; ISBN 90 5668 115 X	June 2003
116	Nicole van Beeck	A New Decision Support Method for Local Energy Planning in Developing Countries; ISBN 90 5668 116 8	June 2003
117	Grzegorz Pawlina	Corporate Investment under Uncertainty and Competition: A Real Options Approach; ISBN 90 5668 118 4	June 2003
118	Jana P. Fidrmuc	The Impact of Blockholders on Information Signalling, Productivity, and Managerial Disciplining; ISBN 90 5668 117 6	June 2003
119	Sjoerd Beugelsdijk	Culture and Economic Development in Europe; ISBN 90 5668 121 4	September 2003
120	Bas van Groezen	The Wealth of Generations: Ageing, Pensions and Welfare from a Macroeconomic Perspective; ISBN 90 5668 120 6	October 2003
121	Pham Do Kim Hang	Essays in Game Theory and Natural Resource Management; ISBN 90 5668 122 2	September 2003
122	Jeroen van de Ven	Psychological Sentiments and Economic Behaviour; ISBN 90 5668 123 0	October 2003

[1] Copies can be ordered from Thela Thesis, Prinseneiland 305, 1013 LP Amsterdam, The Netherlands, phone: + 31 20 6255429; fax: +31 20 6203395; e-mail: office@thelathesis.nl