Social norms, social distance, social approval and household electricity consumption: A field experiment in Cape Town.

Grant Smith (smtgra016)

Submitted for examination for a Masters by Dissertation in Economics, 2011

School of Economics, University of Cape Town.

Supervisor: Dr. Martine Visser

Declaration

1. I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own.
2. This paper entitled: "Social Norms, social distance, social approval and household electricity consumption: A field experiment in Cape Town" is my own work.
3. Each significant contribution to, and quotation in, this paper from the work(s) of other people has been attributed, and has been cited and referenced.
4. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

Signature ……(Grant Smith).

Acknowledgments.

I wish to thank the following people and organisations for their assistance in completing this work. While all errors remain my own, their efforts have been essential in running, completing and writing up this research. My supervisor, Dr. Martine Visser, for her patient supervision as well as help in securing the initial funding for this study. SIDA, for providing the initial funding. Don Early, Gary Ross, Angelo Carolus, Melanie Sonnenberg, Sarah Ward and Sarah Rushmere from the City of Cape Town municipality. The City of Cape Town for providing the funding to run this project for the second three quarters of 2010, allowing me to run this project for an entire year rather than the initially projected three months. Mike Little, Stefan Louw and Richard Bradstreet for assisting with quality control at the printers and data extraction. Oscar Whate, Craig Seyffert and Ikraam Tagodien at the printers - Lithotech - for going beyond the call of duty in order to solve the various logistical challenges that a project such as this presents. More than all these however, to my wife Rebecca for reading at least one chapter and the agreeable fact of your continued company; this is dedicated to you.

To Rebecca Smith.

For saying yes.

Contents.

Introduction.

South Africa wishes to consumes less electricity. There are primarily two reasons for this: the first is that environmental cost of generating electricity in South Africa is fairly high and the second is that the electricity grid is currently capacity constrained. South Africa's electricity consumption is exceptionally coal intensive, because of this the CO_2 (and other greenhouse gas) emissions per capita are amongst the highest in the world and higher than any other country with its level of development (Winkler & Marquard, 2009). As a result electricity consumption in South Africa currently comes at a large ecological cost. South Africa's electricity generation and distribution capacity have also been hard pressed to match the increased demand for electricity that has resulted from the general growth of the South African economy in the post-apartheid era and the country has been subject to large scale electricity black outs in recent years.

Given that South Africa's electricity generation is likely to remain coal intensive for a significant time, strategies that reduce the demand for electricity are ecologically desirable, since a reduction in demand would mean less electricity generation required overall. Strategies to reduce electricity demand are also desirable in light of the fact that it is also likely to be some time until sufficient new electricity generation capacity is installed to mitigate the peak demand on the electricity supply chain.

One of the strategies to manage consumption that has begun to attract attention around the world is to use social norms in reports to households about their own consumption. Initial tests of strategies using social norms in the USA have yielded promising results. At the same time, research into the way people are influenced by others has been growing as a field within economics, particularly with respect to social norms. In general it seems that we are influenced by social norms. Amongst other things, this influence varies with the number of people who are doing something (or not doing something), the social distance between us and the "normal" group and our degree of ignorance about the status quo. This influence is able to operate beyond situations of direct personal interaction with studies finding significant behaviour changes associated with receiving reports about normal levels of behaviour distributed by mail, the internet or over the airwaves.

There are several reasons that reporting norms may influence behaviour. Norms may signal important information which it would be difficult for a person to discover independently. Following norms may thus be a strategy that lowers the transactions cost of learning new important information for a person. The influence of social norms may itself be a function of the way in which

norms remain stable over time, which may be as a result of enforcement, internalisation or simply being an optimal coordinating mechanism. If a person has had prior experience of any of these or has imbibed a culture that generally defaults to conforming to norms, it is likely that they will have developed a heuristic that norms in general are to be conformed to, as much as possible or even as much as comfortable.

This research tested whether households in Cape Town would respond to reported social norms of electricity consumption in a manner consistent with theory and earlier studies.

Originally this research project was concerned with reactions to reported average levels of consumption (descriptive norms) that varied in terms of the social distance between the household and the average reported to it as well as implied social approval or disapproval (injunctive norms) with regard to where the household's own electricity consumption stood relative to the reported average level of consumption. In doing so, this experiment replicated elements of two interventions in the USA, Schultz et al (2007) and elements of OPower's home energy reports (Allcott, 2011; Violette et al, 2009). To test these elements four interventions were designed and households were allocated randomly to them.

However as fieldwork proceeded, interest grew to address a further field of inquiry: how the consumption of households receiving no mailed treatments compared to those that were. Constructing the appropriate control group, ex post was one of the challenges faced in analysis and which is dealt with in the methodology section.

There are two ways in which this study sought to extend the literature. First, this study sough to extend the field experimental literature on social norms and electricity consumption to a developing world context. As far as this researcher is aware, this study is the first instance of a norms based approach for mitigating electricity consumption that has been tested in a developing world context. Second, this study sought to assess the role played by social distance in households' reactions to reported social norms. This is an element that was not assessed in any of the USA studies. The results indicate that both injunctive norms and social distance were important influences upon how households responded to social norms.

This dissertation begins by examining relevant literature on social norms in chapter 1, we then move to discuss the rise and role of field experiments (such as this one) within economics in chapter 2.

Chapter 3 deals with the methodology of the field experiment itself as well as the analysis of data generated by the experiment. Chapter 4 reports and discusses the results of the data analysis.

Chapter 1.

Social Norms.

"No man is an island, entire of itself."

John Donne

Devotions upon Emergent Occasions

Meditation XVII

Introduction

We all, like John Donne, have probably felt at some time or another that our lives are inextricably bound up with and influenced by those around us. At the beginning of the 21st century research is showing that the influence of others upon us, in apparently non-direct and largely unconscious ways, appears to be truly stunning. This influence of others is already the subject of the fairly well established, century old[1] literature of an entire discipline, social psychology, but it has been receiving increasingly greater attention in within the literatures of economics over the last two decades. While acknowledgement of the social psychological shoulders upon which current economic research on this topic stands is not absent[2], this chapter concerns itself mainly with offering a brief review of what has been learnt about social influence and social norms within economics over the recent past.

Traditional economics, as George Akerlof noted (1997), "has been based upon methodological individualism." Within economics this has begun to change. Particularly, and of direct concern for this paper, one the directions in this new wave of research is examining the economic nature, role and determinants of social norms. Largely, this research has dealt with norms as conventions of behaviour and in particular how these conventions solve coordination problems in various markets. Electricity consumption in a city has definite elements of a coordination problem, particularly insofar as electricity provision is a large-scale commons problem, and so it is useful to examine this body of research. However it is the influence of reporting to people what the social norms is that we are most concerned with in this project.

This chapter will discuss research into various elements that are at play in the way others influence us through social norms. We begin by briefly discussing, section 1.1, two studies that have been

[1] Social psychology arguably begins with Norman Triplett's experiments (1898) with children which highlighted how the presence of another can increase physical effort. Although see Strube (2005) for a discussion of the veracity of Triplett's original findings.

[2] Section 1 pays particular homage to the discipline.

profoundly influential in developing the study of how others influence us, one of which has been reproduced at least 133 times.

From here we move on to discussing three elements that appear to determine the amount of influence exerted by reporting social norms. These are: the size of referent group, social distance and the presence of pluralistic ignorance. Section 1.2 deals with how influence appears to vary according to the size of the referent group. Section 1.3 examines the idea of social distance and how the influence of the referent group seems to change with the social distance between the individual/household and the referent group. Section 1.4 discusses pluralistic ignorance.

We then move on to discussing how social norms may help to reduce the transactions costs of learning (section 1.5) and a brief discussion of other general reasons why social norms can be stable over time and persist (section 1.6). The discussion in both of these sections deals with ways in which reporting social norms may be potent influences upon behaviour because of their evolutionary stability.

The final two sections cover research that bears on two operational elements of the natural field experiment at the heart of this. Section 1.7 examines the question whether social influence through reporting social norms can be exerted remotely, that is in situations that are not face-to-face, such as through inserts with mail. Section 1.8 then discusses descriptive and injunctive social norms. These are the two sorts of norms that are used in the inserts at the heart of the natural field experiment. Section 1.9 concludes.

1.10. Two early studies on the influence of others upon us generally.

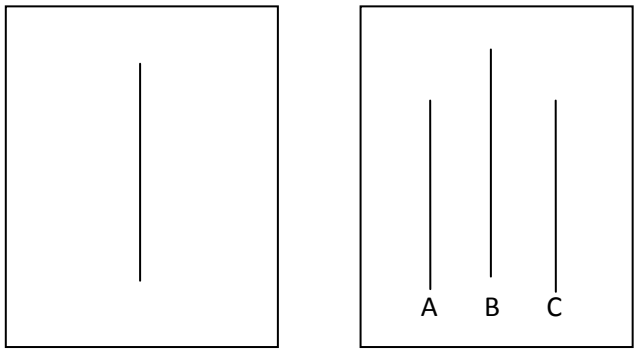
Two early experiments in particular within social psychology zeroed in on how groups might sway the decisions of individuals. Muzafar Sherif (1936) and Solomon Asch (1962) developed a line of enquiry to investigate the influence of majorities in small groups upon and individual.

Sherif's experiments wherein groups estimated by how much a point of light projected onto a wall moved during the experiment, demonstrated that a confident announcement by an undercover experimental confederate could sway a group's estimate of how far a point of light, that was in reality stationary, vacillated, could sway a group's estimate of the light's "movement" significantly toward their estimate. In one version of the experiment Sherif replaced half of the group each round so that each group from the second round onward was composed equally of new and old subjects and no person stayed for more than two rounds of experimentation. The bias introduced by the confederate in the initial trial persisted for several rounds in this situation.

The experiments by Solomon Asch (1962) are perhaps the most famous demonstrations of the fact that the opinion of others can affect our judgement. In these experiments a group is presented with two large cards. One card has a single vertical line drawn on it, as in card A below, while the other card has three vertical lines drawn on it, as in card B below. On the card with three lines drawn upon it, one line is of exactly the same length as the single line on the other card, while the two other lines are of markedly different lengths to the single line on the other card. Each of the three lines on the three-lined card are labelled. Respondents are asked to call out which amongst the three lines is the same length as the single line on the first card. This process is repeated a number of times for different cards, all cards being essentially alike in format to those shown below, with only line lengths varying. Respondents are asked to do this individually, one after the other. It is clear which line on the three-lined card is of the same length as the line on the sing-lined card. The lengths of the lines on the three-lined card differed by between a quarter of an inch to an inch and three quarters (Asch:1962, p.453). The experiment has one more twist though. All of the members of the group are conspirators in the experiment except for one person. That person is oblivious to the conspiracy and is the true subject of the experiment.

When asked to deliver their judgement about which two lines are equal in length the true subject is generally asked last out of all the group. In the initial rounds of the experiment all members pronounce what is obviously true, in terms of the above example that line B is the match of the single line in Card A. However, once the initial rounds are behind, the conspirators assert that lines which are obviously of different length to the single line on Card A as being equal in length to that line. This would be the equivalent of saying that line C or A were the same length as the line on card A. Conspirators are careful to agree about which line is the same length as the former line in their statements. Under these circumstances, when conspirators agree upon a falsehood, it has been found that between 20% to 40% of respondents apparently ignore their own senses and agree with the conspirators. In contrast, almost all respondents who do this test individually, apart from a group, match the correct lines all of the time.

What is amazing about this test and its results is that the task that is set is so simple. It should be clear what the correct answer is. Everyone should get it right, no matter what the groups says. Indeed, in interviews with the single critical subjects after each of their experiments Asch notes (Asch:1962, p.460) that subjects feel a conflict between their judgements and the rest of the group's, the subjects have apprehended with their senses what is clearly the case in terms of which lines are to match.



Card A. Card B.

However, what subjects talked about in the post experimental analysis was that, to varying degrees, the rest of the group's disagreement with them caused them to doubt the validity of their perception. In the words of one respondent "To me it seems I'm right, but my reason tells me I'm wrong, because I doubt that so many people can be wrong and I alone right" (Asch:1962, p.46 3-4) No doubt the specific reasons will vary person to person, but the thrust of this is pretty much the same. What causes subjects to pronounce a judgement that is at variance with their own perception is the fact of *all other people* claiming to see things another way.

Since its first publication in 1952, this experiment has been tried many times over all over the world. Bond and Smith conducted a meta-analysis of the experiment in 1996 and found 133 instances of Asch line experiments in the literature, conducted in seventeen different countries, with conformity of the sort described above found in all studies. One major result of Bond and Smith's meta-analysis was the finding that conformity was more intense in societies that also had a more collectivist norm operating in society at large, such as the former communist countries (Bond and Smith, 1996).

It appears that the influence upon the individual in the Asch experiments may be deeper than expected. Subsequent study (Berns *et al*, 2005) has shown that participants, when faced with this situation, actually see things as the rest of the group does. In fMRI studies Berns *et al* found that regions of activity in regions of the brain traditionally associated with perception were especially active when a subject conformed, whereas this was not the case with subjects who did not conform.

The literature on "groupthink", beginning with Janis (1971), provides a series of case studies (see Esser 1998 for an overview) which together offer suggestive evidence that decision making groups can foster these pressures to conform endogenously. Members of such groups, the literature contends, can together herd themselves toward otherwise inexplicably bad decisions by increasing degrees by a potent mix of increasing insularity and antipathy to outside groups (Esser, 1998). The

various suicide cult tragedies are potentially horrific examples of this as are some high profile policy blunders at government level, such as the Bay of Pigs, in the Kennedy administration, and the appeasement policy of the Chamberlain government (Esser, 1998).

1.11. The size of the referent group as a determinant of its influence.

You are sipping your coffee and checking your email over WiFi in a street-side branch of your preferred cafe when a man stops outside on the sidewalk and looks intently up at something; you notice because he pauses a good minute before seeming to give up looking for whatever it was he was looking for and moving on. You probably would not even have remembered the incident except that over the next forty five minutes first two men, then three, then five, then ten, then fifteen men pause at different times outside the same window to look up for, presumably, the same thing for a minute. What makes the situations more memorable is that each time more and more people actually stop to look up too, in fact not too long after the last crowd has dispersed and the fifteen men arrive to start looking up, a crowd grows rapidly larger than before and blocks the door of the cafe, you know because it stops you getting outside to see what could be so arresting.

The situation described above of course reflects the now famous experiment conducted in New York on two winter afternoons in 1968 by Stanley Milgram, Leonard Bickman and Lawrence Berkowitz[3] where larger initial crowds (confederates of the experimenters) draw ever larger crowds focused upon the same action or object, here, looking up.

The experiment is a demonstration of a fact about human nature that we perhaps intuitively realise to be true and which we will explore in this section: we are more likely to adopt a behaviour if more people adopt it and, all else being equal, less likely to adopt the behaviour if fewer people adopt it.

An experiment with informational cascades conducted by Salganik, Dodds and Watts (2006) demonstrates this effect operating without any face-to-face social interaction (as in Milgram et al's experiment above). One of the implications of non-face-to-face interaction is that a person's actions are not observable by others. This is reasonably obvious to each person, in this case, sitting in front of their computer. To carry out the experiment Salganik, Dodds and Watts created a website that hosted unreleased songs from unknown bands to which 14 341 participants were recruited from a teen interest site on the world wide web. Participants were not informed that they were now part of an experiment, they almost certainly thought that they had arrived at a music site much like many on the world wide web. Participants who landed on the site were invited to listen to the various songs, rate them and, if they chose to, download them.

[3] Milgram, Bickman & Berkowitz (1969).

In real-time, participants were randomly allocated to one of two treatments as they arrived on the website. The two treatments were: social influence and independent influence. In the social influence treatment participants were further randomly allocated into eight subgroups (or "worlds" in the language of the Salganik *et al* paper). In each of the social influence subgroups, participants could see how many times each song had been downloaded for their particular subgroup, while participants in the independent influence group saw no information beyond the song title and band name. Participants in the social influence groups thus had an indication of the preferences of others while participants in the independent influence group had no such indication. The results in the social influence subgroups present a fairly compelling story about the influence of norms as well as how cascades can develop when social norms are visible. Across the eight subgroups in the social influence group "the best songs rarely did poorly and the worst songs rarely did well" (Salganik *et al,* 2006) but otherwise the distribution of songs rated highly was pretty much random. It is the pattern of results within each subgroup that is most interesting.

Within each subgroup of the social influence group the highest rated songs were always those which had been downloaded more initially. The results demonstrated that with the passage of time those songs that were initially popular attracted proportionally more downloads than those songs that were less popular initially.

These results indicate that reporting the evolving social norm had a decisive influence upon the fate of each song's ultimate popularity within each subgroup. This is thus an example of a reported behaviour (as opposed to primarily observed behaviour) influencing people in proportion to the number of other people who had adopted the behaviour.

This relationship between the number of people adopting a behaviour, or a product, and the likelihood of others then adopting this behaviour may well be at work in all markets that are reasonably competitive. There is evidence in the marketing literature that the sort of exponential distributions observed in Salganik *et* al (2006) are also present in a diverse range of competitive markets. The distributions that seem to describe the download patterns of least to most popular songs in the social influence groups of Salganik *et al's* research have been noted in the marketing literature since McPhee (1963) as describing the distribution of market share and frequency of purchase of goods from coffee to syndicated cartoons (Ehrenberg *et al*, 1990). Within this literature the phenomenon is known as "double jeopardy" after McPhee (1963) who was the first to identify the phenomenon. Since Ehrenberg (1972) this sort of "double favouring" of market leaders has been noted in many markets across the world from breakfast cereals, to washing powder, coffee and toilet soap (Ehrenberg *et al*, 1990). That this "double jeopardy" pattern accords with the results of

Salganik *et al*'s (2006) experiment provides circumstantial evidence then that the sort of social norm-referencing process identified by Salganik *et al* may well be at work in consumer goods markets across the world. Of course this is, at this stage, merely suggestive and not proof of a hard and fast sort.

Several literatures within economics on public goods have also investigated the possible role that social comparison may play in influencing contributions to a public good[4] or other pro-social behaviour.

Fellner et al (2009) conduct a field experiment amongst potential TV licence evaders in Austria. They employ four mailing-based treatments amongst a random sample of potential evaders. The four treatments are a legal threat (which was designed to imply a high risk of detection if a consumer decided to evade payment of their TV licence), a moral appeal that was designed to appeal to an innate sense of fairness on the part of the consumer. Their third treatment was, in a similar vein to this experiment, a provision of "social information" that indicated to recipients what the level of compliance (in terms of paying their TV licences) was. The final treatment was a control treatment where a number of people were not mailed any communication whatsoever.

The results that Fellner et al (2009) observed have particular resonance with this study. Firstly it was found that, overall, the mailed groups exhibited higher levels of compliance than the non-mailed group. Within the mailed groups, it was only the legal threat that produced significant results overall though, the moral appeal and social information treatment did not seem to produce much of a significant effect overall. What is really interesting though is the pattern with which the results differ within the social information treatment. The results differ by whether the consumer lived in a community which believed[5] that the level of compliance was lower than what the social information treatment reported it to be or if the community believed it to be higher. Respondents who lived in communities which believed that TV licence compliance was below that reported were more likely to comply than those who lived in a community who believed the level of compliance to be higher than what was reported.

[4] Electricity may well be considered by many of its consumers to be what amounts to a public good, or a common pool resource, since their contemporaneous aggregate demand - if too great - may instigate a blackout that suspends supply of electricity to a neighbourhood, suburb or a city. Similarly, for those aware of the ecological impacts of electricity generation in South Africa, the impact of increased electricity generation in response to increased electricity consumption may link private to consumption public consequences.

[5] Fellner *et al* (2009) obtain figures for what communities believe to be the level of compliance from a survey of over 1000 households conducted in 2000. In the survey, respondents are asked whether they believed the frequency of licence fee evasion in Austria was between 0%-5%, 5%-10%, 10%-20%, or more than 30%. The survey was conducted by the Austrian Public broadcasting Company and also provides data for Traxler and Winter (2009, 2012).

This sort of result is consistent with a model of contribution behaviour that is truly conditional. Not only do we see that the likelihood of compliance increases amongst those who discover that the level of compliance was actually greater than what they at first thought, but we see the same effect at work but in opposite direction. Those who discover that the level of compliance is lower than what they at first expected become less likely to comply. If people were not conditionally cooperative or only partially so, we would expect to see either no effect in any group or an effect in only one group, rather than an effect in both groups.

This sort of effect has also been observed in the norms literature in social psychology, where it is known as the boomerang effect (Schultz *et al*, 2007). A large literature has developed within social psychology around the use of social norms to curb alcohol consumption at college in the form binge drinking (Prentice & Miller 1993, Perkins *et al* 2005). Studies and programs in this vein typically, as in the social information treatment in Fellner *et al* (2006), report the average number of drinks consumed by a typical college student to students, often freshmen. The goal is to demonstrate that a normal college student does not consume as much alcohol as a freshman may at first expect (Perkins, 2002). Results from these studies lend support to the view that people will converge to what they perceive to be the normal level of consumption. As with Fellner *et al* (2009) an unfortunate corollary of this is that in many cases reporting a normal level of drinking raises the consumption levels of those who were initially consuming less than the reported normal level of alcohol (Perkins *et al*, 2005; Weschler *et al*, 2003; Werch *et al*, 2000). Within the broader resource consumption literature in economics the data has not, so far, pointed to the presence of a boomerang effect. Ferraro and Price (2011) do not observe it amongst water users in Cobb County that were compared to norms on mailed interventions and, similarly, Allcott (2011) does not observe it amongst OPower households that receive communication which incorporates similar comparative social norms.

Within the tax compliance literature; effects such as those described above have been noted as far back as Spicer and Becker (1980) who found the same sort of comparison effects at work amongst there sample to whom the reported three different prevailing tax rates of others (a higher, lower and middle rate) and found that tax evasion was lower the lower the reported tax rate of others, suggesting that people used information about where the tax rate they were subject to stood relative to the average rate in order to decide whether to evade tax or not.[6] This result accords with

[6] The experiment was not designed to evaluate the effect of comparison per se but rather to investigate equity theory, whereby the individual alters their behaviour so as to place themselves in a position that resolves any perceived lack of equity. Thus a person who perceived themselves to be taxed more than the average may feel that they are being treated unfairly and so decide to under declare their income, while a person who

a particularly consistent finding from the tax compliance literature within economics, which is that citizens tend to be conditionally cooperative (Frey and Torgler, 2007 provide a helpful overview in their literature review section). What is meant by conditionally cooperative is that people are more likely to pay their taxes if they perceive that more people are also paying their taxes. In their paper Frey and Torgler (2007) estimate tax morale across a 30 country European survey[7]. They find very strong support in this data for tax behaviour and tax morale to be strongly determined by an individual's perception of the level of tax evasion in their country. The result is robust to several weighted probit estimation techniques and two stage least squares techniques.

Using a linear public good game design that employs the "strategy method" of Selten (1967), Fischbacher *et al* (2001) test how much each subject is prepared to give in a public goods game for varying indicated levels of contributions from other members of the group. Given the structure of the game and its payoffs, complete free riding was predicted. Fischbacher *et al* find in their experiment that 50% of the subjects contribute in a conditionally cooperative fashion, while a third free ride. Similarly Heldt (2005) finds that cross country skiers in two ski resorts make contributions to a ski track in a fashion in a fashion that is positively dependent upon the degree to which others contribute to the track.

Croson (2007) conducted a series of experiments that aimed to test the comparative static predictions of linear public goods games based on assumptions about how individuals would contribute to the public good. The results supported a model of giving that suggested reciprocity sensitive behaviour on the part of subjects. Subjects were significantly and positively influenced by their beliefs about how much others were giving, in particular, subjects overall seemed to try to match the median contribution of their group.

Within public goods games agents' behaviours may be motivated by conformity that is purely conformity or by some reciprocity based heuristic or motive. Conditional cooperation is obviously a potential candidate for explanation by reciprocity derived actions where agents contribute to a public good because they receive some benefit from the existence of the good. Purely conformist agents would contribute to a public good from which they received no benefit, purely because they perceived that others were doing so, where as a reciprocity based agent would not (Bardsley and Sausgruber, 2005)

perceives themselves to be under taxed compared to their peers may be more likely to fully comply out of a sense to restore equity (Adams, 1965, Homans 1961, Tyler and Smith, 1988). It is in this way that the result is relevant.

[7] The European Values Survey.

Frey and Meier (2004a, 2004b) find in a field experiment and wherein students at the University of Zurich are invited to contribute to two social funds that the University runs at the beginning of each year (2004a) and a subsequently deeper analysis in a companion paper (2004b) that conditional cooperation is mainly observable amongst students who were indifferent to making a contribution, with past behaviour serving to be the main predictor for the balance of the students.

In a similar application to charitable contributions one experiment, conducted via a public radio station by Croson and Shang (2009), was designed to test two major classes of theories that have been proposed to explain the relationship between one's own donation and the donations of others. One set of models[8] explains one's donations as being substitutes for the donations of others while the other models the individual's donations as complements for the donations of others. Of course what both models to do not presume to explain[9] is the relationship between the individual and others contributions *in* fact, but rather they seek to explain the influence of *information* about others donation/s upon really an individual's donation/s.

In many instances it seems that information, direct or mediated, about the number of people who are adopting a particular behaviour will positively influence the likelihood the behaviour being adopted.

1.12. Social distance and the influence of social norms.

Social psychology was the first discipline to really take investigation into social norms' influence and social influence seriously. Early research in social psychology seemed to suggest that an individual was influenced by another or by a group because that individual wanted, in some way, to be part of and like by that one other or that group (Schultz *et al*, 2008, Bandura, Ross & Ross, 1961, Deutsch and Gerard, 1955) this finds a definite resonance in the work of George Akerlof on social distance (1997) in economics. In this section we examine how this idea of social distance may influence responses to a norm. In particular we are trying to develop a hypothesis about how social distance between the respondent household and the referent group may influence the response of the respondent household.

[8] Croson and Shang understand these models as falling into 3 major categories: the pure altruism models after Becker (1974), Warr (1982) and Roberts (1984) which predict a dollar for dollar crowding out of individual giving by high normative giving, the impure altruism models such as those described by Andreoni (1989, 1990), Steinberg (1987), Ribar and Wilhelm (2002) which describe normative giving as crowding out less than one dollar of individual giving per dollar normatively given, and finally threshold models which differentiate effects of normative giving upon individual giving according to whether the threshold for the provision of the public good in question has been reached or not (e.g. Andreoni 1998 for public goods and Romano 1991 as well as Cornelli 1996 for private goods).

[9] And it would be almost mystical if they did.

Akerlof (1997) develops a simple model which captures the essence of how the idea of social distance may influence the direction and degree to which people conform to a reported referent group. The model serves to highlight the thought that a choice with social aspects may be subject to three influences in terms of a rational, social, actor's utility function. The first influence would be the utility derived from the good or action (or degree of consumption/participation) in itself, something common to all students of microeconomics. The second and third influences are antagonistic in Akerlof's formulation. These influences are the utility derived from being distinct and better than one's neighbours[10] and the utility from being similar to your neighbours. Akerlof's formulae for each influence are fairly simple and illustrate the concept clearly.

In terms of the utility derived from being distinct and better we have:

$$U = -d(x' - x) - ax^2 + bx + c$$

With x' denoting the average choice of one's neighbours and the right most quadratic form denoting the influence upon utility of the choice devoid of social considerations.

In terms of the utility derived from acting similarly to one's neighbours we have:

$$U = -p|x - x'| - ax^2 + bx + c$$

Here we see that utility is maximised the smaller the distance between oneself and one's neighbours becomes, whether one approaches from "below" or "above" one's neighbours. As Akerlof discusses (1997; p1009-1010) the status of marginal utility as social distance moves to zero depends upon the functional form of the utility model. Here we have expressed the idea that non-zero social distance decreases the utility from a choice in and of itself for a person. Again the quadratic function on the rightmost denotes influence upon utility of the choice devoid of social considerations.

Overall then, we have an actor who would choose some level or type of x so as to maximise a utility function of the form:

$$U = -d(x' - x) - p|x - x'| - ax^2 + bx + c$$

The key idea here is that the degree of discrepancy between one's own action and the actions of others drives one's actions.

The experimental literature has brought forth examples of where, it seems, the idea of social distance as guiding behaviour finds some support.

[10] Neighbours, for Akerlof, need not be geographically close but may just be close in social terms, such as peers or friends.

Sacerdote (2001) examines Grade Point Average (GPA) and fraternity data for freshmen students at Dartmouth, for evidence of social distance influencing GPA. Students at Dartmouth are randomly assigned to dormitories and rooms in their freshmen year, allowing for estimates of roommate influence upon GPA to be estimated.[11] Sacerdote finds a significant relationship between the GPA of roommates is supported by his data (although, a robust magnitude of this relationship is found to be somewhat difficult to determine). This experiment builds on similar findings in other research into the effects of peers upon GPA, albeit upon non-randomised data where researchers needed to deal with this characteristic when estimating the relationships, such as Harris (1998), the Coleman report of 1966 and to some extent the social background findings of Betts and Morell (1999).

Within the context of school attendance in poor villages in Mexico, Lalive and Catteneo (2009) find that the schooling decisions of children are affected by those of their peers. They identify a significant positive relationship between school attendance of one group of children upon another. In particular they find that if the some children are funded to come to school (in this case via the PROGRESSA program in Mexico) that not only do they attend more days of school than before, but so do their unfunded peers. In the Lalive and Catteneo (2009) analysis this secondary peer-influence effect is nearly as important as the direct funding effect.

The effect of those close to us appears to extend to our waistline too. Christakis and Fowler (2007) analyse 12 067 people in the Framingham Heart Study wherein people are assessed with repeated measurements over the 32 years from 1971 - 2003 and find that the number of obese persons (with a body mass of 30 or greater) were to be found in discernable clusters. The chance of a person being obese in the data increased by a considerable 57% if they had a friend that became obese. Similarly Leatherdale and Papadakis (2010) find after controlling for between-school random effects amongst 12 049 school going adolescents from 76 secondary schools in Ontario, Canada that that as obesity increases amongst students aged 16-18 years, the likelihood of a 14-16 year old student being obese increased significantly.[12] Similar effects have also been suggested in studies dealing with adolescent smoking (Cameron *et al*, 1999; Leatherdale *et al*, 2006).

In terms of morality; research by Ariely *et al* (2009) indicates that when people are confronted with a moral norm, such as for cheating, within their own group they are more likely to conform to that norm (be it honesty or cheating). However, when confronted with a moral norm, again say for

[11] Where roommates are selected by each other, self selection by students becomes a highly obscuring factor in subsequent analysis of peer effects, a difficulty that is usually solved by the use of instrumental variables such as in Gavira and Raphael (1999) and Borjas (1992).

[12] Leatherdale and Papadakis (2010) use a logistic regression approach. The log odds of a 14-16 year old being obese increase by 1.05 for every 1% increase in the number of 16-18 year old students who are obese in their secondary school.

cheating, in a group of which they are not a member, people are likely to be behave in the opposite fashion to the norm, that is, they will cheat less if they perceive cheating to be a norm in a group of which they are not a member. Thus it seems that when social distance is small, the tendency to conform to a social moral norm is strong but when social distance is larger that tendency is either weaker or operating in the opposite direction.

Moving to consider larger communities Analysis of these sort of effects has also been carried out within another sub genre of economics research, namely the neighbourhood effects literature. Here the idea is to look for effects that operate over groups of people that are socially close (as opposed to socially distant). These groups constitute the "neighbourhoods" in the neighbourhood effects literature. Neighbourhoods in this literature need not be groups whose members are necessarily geographically close to one another (Durlauf, 2004) but merely, in some way, socially close.[13]

Within the neighbourhood effects literature there are uncovered role model effects, where people are shown to imitate others as well as peer influence effects, where people contemporaneously influence each other (Durlauf, 2004).

Buchan *et al* (2006) have also conducted some lab experimental research in this area that suggests that social distance may not uniformly exercise an influence upon behaviour, but is rather modified by the cultural context of the respondents. Buchan *et al* (2006) were interested in assaying the effects of social distance upon other regarding preferences. To this end they conducted several laboratory experiments with students in the USA, China, Japan and South Korea. In each case the participants were paired off to play an investor game[14]. Before each experiment began the participants were randomly split into several groups each of which discussed a designated topic, that was unrelated[15] to the forthcoming experiment, for ten minutes. Once the ten minute discussion was concluded, half of the participants were paired with people from within their discussion group while the other half of participants were paired with people from outside of their discussion group.

[13] The exact measure of social distance differs from paper to paper within this literature as would be expected (Durlauf, 2004. Section 4.6).

[14] The investor game is a two player game in which both players are given an initial endowment of money. The first player has the option to transfer any positive amount of money to the second player or none at all. Any money transferred by the first player to the second at this stage is tripled. The second player can then decide how much money to return to the first player. Since the unique subgame perfect Nash equilibrium is for the second player to return no money to the first player (the second player is essentially playing a dictator game), the first player should send none in the first round.

[15] There were only two types of topics. In one groups participants were asked to introduce themselves and then to discuss "a personal question relating to their birthdays". This was deemed to be personal communication. In another group participants were instructed to answer questions from the World Almanac, with one participant noting down the answers. This was deemed impersonal communication. (Buchan *et al*, 2006; pp.379-80). These discussion topics are related to a further experimental question of whether the personal nature of group discussion would influence results of the investor game.

The researchers measured what proportion of their wealth was returned by the second player to the first player across the various treatments and countries. They find that the reciprocal behaviour by second players differs by whether the first player was part of the second player's pre-experimental discussion group, but in a fashion that is different between the USA and China groups (with the South Korean and Japanese groups not distinguishably different from the USA or China groups). With the experiments conducted in the USA, Croson *et al* find that second players give significantly greater proportions of their wealth, in the "return" round of the investor game, when they were paired with first players that were part of their pre-experiment discussion group compared to when they were paired with a first player from a different pre experiment discussion group[16]. There thus seems to be a negative relationship between social distance and the proportion of wealth given by the second player, which is to say that that decreasing social distance increases the proportion of wealth given by the second player in the "return" round of the investment game. Amongst the experimental groups from the more collectivist cultures of Japan and South Korea the amount returned by the second player seemed to be the same, regardless of whether the first player was part of their pre-experiment discussion group or not. In China (the most collective culture) the opposite was noted. Second players in these groups gave a greater proportion of their wealth back to the first player in cases where the first player was from a different pre-experiment discussion group and less to players who were from their own pre-experiment discussion group.

Buchan *et al* (2006) appeal to broader research in social psychology[17] in order to explain this differing behaviour. Specifically, they contend that their results are consistent with theories about how in groups and out groups are formed and how people relate to them in individualistic and collectivist cultures. Within individualist cultures (such as the USA) the needs of the individual come before the broader group, while the opposite is true the more collectivist a culture is (such as China). Within individualistic cultures groups are formed for a purpose, are formed more frequently and are more flexible and temporary. However, within collectivist cultures, groups are formed less frequently and are oriented around long term associations such as family.

Whatever the precise explanation; what should be noted is that the influence of social distance thus differed across cultures. In the USA and China, social distance exerted influence in opposite directions, while in South Korea and Japan it apparently exerted no influence at all. This sort of result

[16] Participants did not know beforehand that they would be playing an investor game, so pre-game agreements are not likely to have been reached (Buchan *et al*, 2006; p.380).
[17] Triandis *et al* (1990), Triandis *et al* (1999) and Han and Choe (1994) together provide an overview of the work they appeal to.

suggests that it is worthwhile investigating how social distance may influence household response in Cape Town.

1.13. Pluralistic ignorance and social norms.

Pluralistic ignorance is an important element to consider when discussing normative influence. Pluralistic ignorance is when most or all of a group are ignorant about what others in the group do or think. The concept of pluralistic ignorance has particular bearing upon the investigation of normative social influence in regard to college drinking in the USA, which has been a particularly active area of research. The literature in this regard is motivated by concern for college students over-drinking and how to combat this. The literature on normative influence and college drinking is predicated upon the thought that college students are, at least partly (but still), significantly motivated to drink too much because they think that overindulgence is the norm at college (Prentice& Miller, 1993, Perkins, 2002, Perkins *et al* 2005). Several studies have demonstrated that college students routinely overestimate how much it is normal to consume by way of alcohol as a student (Berkowitz 2004, Perkins, 2002 provides an overview of studies in this area) believing that college students consume, on average, more alcohol than they, in fact, do. Interventions in this area have thus been targeted at informing college students about what the normal level of alcohol consumption for their group really is and so reducing pluralistic ignorance on this issue.

The results of applying norms reporting to college drinking have been mixed. With over-consumers reporting the norm appears to have had the effect of moderating consumption (Agostinelli, Brown, & Miller, 1995; Haines & Spear, 1996; Neighbors et al., 2004) some of the time; for under-consumers particularly, reporting the normal level of consumption seems to have raised consumption toward the norm in a number of cases (Perkins *et al*, 2005; Wechsler et al.,2003; Werch et al., 2000). We will return to this topic when discussing descriptive and injunctive norms below, but the important point to note is that these studies indicate that normative influence can be exerted without social interaction.

It is easy to see how pluralistic ignorance could persist in situations where the actions of a significant share of the larger group were not easily observable or otherwise discoverable.

1.14. Thoughts on why social norms may be stable through time: social norms may reduce the transactions costs of learning important information.

Norms, however they may be maintained through time, can certainly serve the function of reducing transactions costs in conditions where multiple equilibria exist (Wärneryd, 1994). Mainly they can reduce transactions costs by indicating an desired or just efficient[18], action, thus allowing agents to find equilibrium action far quicker than they would otherwise.

It may be too difficult for a household or a single person to identify an equilibrium or, in our case, to know what a more efficient level of consumption is. Norms may signal to people what such a level could be. People may learn that adopting a social norm leads to a more optimal personal outcome being realised with lower transactions or search costs than if they had independently set out to gather the information required to inform optimal behaviour themselves. We will deal with this particular signalling function of norms when discussing pluralistic ignorance below.

One situation in which transactions cost of finding an equilibrium or a more efficient level of consumption may be if there is an over abundance of data. If there is an over abundance of data it may be difficult for people to tell relevant from irrelevant data or to derive a relevant summary of the data without incurring significant search costs, such as time foregone that could have been devoted to a paid work. Norms would signal what most others actors are doing and may be interpreted as best and so adopted.

In coordination games, the adoption of norms reduces the amount of costly mismatches in action before an equilibrium of coordination is reached. Consider here the example of which side of the road to drive on. Experimentation in an atmosphere where one is ignorant of the norm could lead to some very costly mismatches with other players. Norms reduce the transactions cost of discovering a coordinating equilibrium for the players involved by virtue of signalling to all what it is normal to do, Roth (1985) demonstrates this experimentally.

Of course merely reducing transactions costs does not mean that norms are necessarily enhancing of overall efficiency from an individual, household or group's point of view. Akerlof (1980, 1982) and Romer (1984) demonstrate, theoretically, that a norm may be present in a market and generate an equilibrium that is not efficient but persist nevertheless. Contracts that are fairly uniform, such as those in the Illinois agricultural market (Young and Burke, 2001), rather than capturing the idiosyncratic realities of agents different circumstances or more direct efficient contract concerns (such as monitoring costs, information asymmetries and risk appetites) may be real world evidence of a norm driving an inefficient market outcome.

[18] Given a particular set of circumstances.

Research by Potters *et al* (2005) demonstrate that there is an incentive for agents within public goods games as well as for exogenous third parties who are interested in maximising contributions to the public good (such as managers of a charity) to order their giving sequentially. In Potters *et al* (2005) subjects can be either informed about the public good and its attributes or uninformed.

In game formats when information has a low but positive cost of acquisition and subjects are allowed to endogenously determine the order of giving, sequential giving is the equilibrium that is arrived at for the most part rather than simultaneous giving. Sequential ordering of contributions is perhaps arrived at as an equilibrium in the public goods games of Potters *et al* (2005) when information has a low positive cost out of a coordination of motives. The first mover has an incentive to purchase information prior to making their contribution so as to select the best contribution (or best level of contribution) if there is uncertainty about this. However, subsequent players do not share this motive if they are aware of the first mover's incentive to acquire information and ability to choose. Subsequent players can move according to the signal sent by the first player, specifically, they can mimic the first player's contribution.

Anecdotally there is some evidence that this already happens with charities. Potters *et al* cite the initial contribution of Brook Astor to the New York library which precipitated three major subsequent gifts as such a real world case. Brook Astor was known to only give to charities that she had personally inspected.

When sequential ordering of contributing is imposed exogenously, leaders were shown to perform similarly to in the endogenous situation. They seemed to, correctly, anticipate that subsequent contributors would mimic their giving. Subsequent contributors did mimic the contributions by initial players.

Although total contributions were found to be highest when allowing players to arrive at the sequential equilibrium endogenously, total giving is also raised, compared to simultaneous giving, by imposing a sequential giving structure upon the game that requires the informed subjects to contribute first. In other words not only do players seem to prefer sequential ordering of contributions to public goods when there is uncertainty about the public good and when information is asymmetrically distributed, but they are more likely to contribute when they have a clear positive contribution signal from the initial mover.

What these experiments seem to demonstrate is that in public goods situations, sequential giving is preferred by players on what seem to be the grounds of reducing transactions costs as the sequential format allows for an efficient use of information, by mimicking contributions of the initial

mover. Such a study has relevance for this experiment since it suggests that people will follow the signalled actions of initial movers when there is a degree of uncertainty.

1.15. Further thoughts on why social norms may be stable through time.

There are several theories about why norms emerge and for why norms are influential (see Feldman & MacCoun, 2005 for an overview of some important ones). Within economics one of the more widely accepted explanations rests upon the notion that it is the ability of norms to coordinate expectations in situations where multiple equilibria are possible that renders them both powerful and stable (see Young, 2007)).

Within this framework it is this coordination of expectations that distinguishes a norm from an equilibrium in general. Consider (from Young, 2007) a game in which two players make simultaneous bids for a portion of a pot of money, with players receiving their bids so long as the total of their bids does not exceed the total amount of money in the pot, if they do exceed the pot value, both players receive nothing. In this situation there are several equilibria that lead to a no-zero return for each player, however a fifty-fifty division of the pot is the one that is most often selected (Roth, 1985) with the other possible variations almost never selected in such coordination games.

Here we see that a fifty-fifty division is a norm because it seems to coordinate both players' or parties' expectations to focus on a specific division out of many possible divisions. Roth (1985) dissolves this coordination by asking players to divide a pot that is denominated in chips of unequal monetary value (which are cashed in at the end of the game), thus creating potentially two different focal solutions to the game: divide the chips equally or the money equally. In this circumstance the incidence of the fifty-fifty norm drops dramatically since neither player is sure how the other will denominate their offer.

Coordination, apart from being a service offered by norms has also been identified by some as being a reason for people to conform to norms. One of the most salient examples in this regard is currency. There is exists a strong norm to carry the currency, or electronic access to the currency, of the country around with you in your wallet. It would be silly to carry rupees with you in South Africa simply because you would probably not be able to find someone willing to accept them in exchange for goods and services. With money the strong incentive to stick to the norm of exchange derives directly from the need to coordinate with others if you are to have your daily bread. Thus coordination itself may often be a factor contributing significantly to the stability of a norm through time.

The influence of others by way of approval, threat of disapproval or even direct enforcement is another manner in which a norm may be sustained through time. Even in games, such as the dictator game, where the split of the pot is the decision of one player alone, between 20 and 30 percent of games are fifty-fifty splits (Camerer, 1997). Beyond dictator and ultimatum games played in a laboratory setting, the fifty-fifty outcome is present in more general cases where there is something to be divided. It is observed in many real life instances such as joint ventures among corporations, share tenancy in agriculture, bequests to children, and splitting the difference is a common outcome of arbitration (Andreoni & Bernheim, 2009).

Where these divisions happen publicly, whether in studies in the laboratory or in real world contracts, agents may be motivated in large part by a concern to appear fair. Andreoni and Bernheim (2009) present a model of agent behaviour in dictator games that accommodates many of the results of dictator games where the division was fifty-fifty by incorporating a concern, on the agent's part, to appear fair to others in their explanatory models.

Norms may also persist because people are concerned about possible retribution from others if they are discovered to be breaking of have broken a norm. Retribution may take the form of social judgement possibly augmented with material action. Fehr, *et al* (2002) is one example of a growing literature that demonstrates that third parties are indeed often prepared to enforce norms by punishing those who transgress a norm, even if it costs them a considerable amount to enact this punishment. The expectation alone that departure from a norm will be punished by social disapproval or otherwise may be sufficient to sustain norms, as shown by Sugden (1986) and Coleman (1990).

Apart from the coordination motive or expectation of social disapproval or even punishment, people may internalise the norm and regulate their own compliance with the norm. To return to the fifty-fifty norm, it may be that agents consider a fifty-fifty division to be fair division of a pot of extra resources and that they wish be able to view themselves to be fair. Similar concerns may explain the tendency for people to comply with tax regimes (Wenzel, 2004). In these cases agents are motivated by a desire to feel good about themselves or at least not feel bad about themselves, and this is what drives their behaviour to be norm consistent even if they act knowing that they are unobserved.

Of course these motivations or channels of enforcement may be act together. Azar (2004) develops a model in which an agent with such internalisation as well as concerns for social approval can contribute to the persistence of a norm, even one that is somewhat costly for the agent, such as tipping for service in restaurants.

In terms of the labour market, which is the market that Akerlof (1980, 1982) and Romer (1984) apply their theoretical models to, Fehr *et al* (1998) find evidence, albeit in (laboratory) experimental labour markets, that norms of reciprocity, where present, seem to have both efficiency lowering and efficiency enhancing influences apart from any effects upon transactions costs. Fehr *et al* (1998) find that a norm of reciprocity holds wages higher than the market clearing level when there is an excess supply of labour and wages "that are persistently above the competitive level", however, reciprocity raises the effort of labour and hence the efficiency the effort-wage trade. Efficiency, beyond reducing transactions costs, notably search costs, is not however something that norms, by virtue of being norms, will necessarily increase (or decrease) that depends upon the nature of the norm, for instance, it seems fairly clear that a norm of revenge killing would keep a society from attaining a more welfare efficient equilibrium.

Many norms, or conventions, are often trifling behaviours to keep. For instance, unless you are on a sinking ship, letting women go first (through doors, sitting down at table) costs a man very little. Similarly breaking a norm by eating with your elbows on the table or wearing a tie longer or shorter than your beltline does not seem to do much to altar any person's well being. As Posner (2000) argues, however, the keeping of these more trifling norms may serve as a signal to those around you that you a preparedness to observe other, less trivial norms, of the same cultural set which may lead to a non-trivial outcomes such as a business deal or a second date.

1.16. Transmitting the influence of social norms in non-face-to-face circumstances.

Much of the early research on normative influence seemed to view normative social influence as operating only through social interaction, up until quite recently (see, Prentice & Miller, 1993). Indeed, the bulk of research in social psychology is still conducted within a setting rich in social interaction. However, a growing body of research, both in social psychology and economics, is demonstrating that normative influence need not only arise in a matrix of social interaction, but may be transmitted though more passive media. In terms of this non-social channel of possible influence, what is being transmitted is information. This information is of a particular sort. We turn now to a review of research which addresses this topic to evaluate how and when such information may be influential.

One manner in which normative influence may be propagated is by signs of many having passed a certain way or adopted a certain behaviour ahead of us. A fairly clear example of this is that of a path worn through the bush which suggest to us the way we should walk when hiking. In a study of student behaviour when picking up mail from a student mail room Cialdini, Reno and Kallgren filled

each student's mailbox with handbills; they found that when the floor was littered with handbills, students overwhelmingly dropped their handbills to the floor but when the floor of the mail room was swept and tidy, most students put the handbills into their bags and did not litter (Cialdini *et al*, 1990). Similarly, a pile of swept litter in a parking lot induces people in a parking lot to behave in a manner that is appropriate for a public space and not litter (Cialdini *et al*, 1990).

Situational norms are norms that are embedded in situations themselves and not in direct social interaction. Situational norms are almost certainly *learned* in social interaction. Once situation norms are learnt they are however not dependent on social interaction to be activated (Cialdini & Trost, 1998). A situational norm which displays this clearly is that of being silent in the library. Aarts and Dijksterhuis (2003) find, in a lab experiment, that questions about libraries influence respondents to more silent behaviour, as measured by respondents lowering their voices when asked questions about library attendance as a part of a more general questionnaire. A similar process may be at work to influence how people deal with money. An experiment by Ariely (2008) in MIT college dorm common rooms suggests that norms about honesty attach more closely to (or are activated more strongly by) money as opposed to other objects such as, in this case, coke cans. In the experiment six-packs of coke cans were left in common room fridges as well as a plate with six one dollar bills. Ariely reports that within 72 hours all of the coke cans in all of the common room fridges had gone, but that all of the dollar bills remained, untaken. While Ariely runs other experiments to demonstrate the fact that people seem to act more honestly when dealing with money than they do when dealing with objects other than money (Ariely & Mazar, 2006, Ariely *et al*, 2008) these all involve personal interaction with an experimenter in some fashion. In the case of the common rooms though it seems that being confronted with money or with a can of coke makes a difference to how honest you are, even if this means asking those present in the common room "whose cokes are these?" but balking at asking the same people "whose dollar bills are these?". It seems then that the norm for dealing honestly with money does not necessarily need social interaction in order to be activated.

Reporting norms, without social interaction, in order to mitigate behaviour, that seems to be driven by pluralistic ignorance, has also been applied to littering (Keep Tennessee Beautiful), recycling within hotel rooms (Goldstein *et al*, 2008) to get inoculations against influenza (Blanton *et al*, 2001) as well as energy consumption (Schultz et al, 2007, Alcott, 2011).

The experiment by Salganik, Dodds and Watts (2006) that measured popularity of downloads across various "social information" treatments is also an example of how people may use information about the frequency of prior behaviours to guide their choice of their own behaviours when there is

a significant degree of uncertainty about what a good choice would be (such as could very well characterise the market for music). The circumstantial evidence of the "double jeopardy" effect observed in the marketing literature may be circumstantial evidence of the same. One can imagine that if someone, such as a student living out of home and in digs for the first time, is uncertain about which washing powder to buy, they may well go for the washing powder with the most shelf space facing them partly, but significantly, on the basis that that brand is the most popular and hence perhaps the best washing powder.

1.17. Comparative Social Norms of two sorts: Descriptive and Injunctive.

As we close this chapter we turn to a brief description and discussion of the two sorts of norms that will be used in this study. These norms are: descriptive and injunctive norms. Both sorts of norms have mainly been used in studies in the field of social psychology but have recently been making their way into the field of economics (Alcott, 2011).

Descriptive norms describe the situation as it is. In the case of this experiment, descriptive norms describe to subject households what the average level of electricity consumption over the past month was for a household in the City of Cape Town or a household in their neighbourhood.

Reporting descriptive social norms has been used as a strategy in many programs at colleges in the USA in attempts to reduce college binge drinking. According to a survey conducted in 2002 by the Harvard School of public health, 746 colleges in the USA had adopted such, descriptive norm-reliant strategies (Weschler *et al*, 2003). Reporting social norms in a comparative fashion has, in these cases, been shown to influence consumption behaviour and a large literature has developed within social psychology around the use of social norms in to curb alcohol binge consumption (Prentice & Miller 1993, Perkins *et al* 2005). Studies and programs in this vein typically report the average number of drinks consumed by a typical college student to students, often freshmen, in order to demonstrate that a normal college student does not consume as much alcohol as a freshman may at first expect (Perkins, 2002). The norms reported by Croson and Shang (2009) in their public radio experiment described earlier were descriptive norms.

Results from these studies lend support to the view that people will converge to what they perceive to be the normal level of consumption. An unfortunate corollary of this is that in many cases reporting a normal level of drinking raises the consumption levels of those who were initially consuming less than the reported normal level of alcohol (Perkins *et al*, 2005; Weschler *et al*, 2003; Werch *et al*, 2000). This tendency for under consumers to increase their consumption toward the

norm has been termed the "boomerang" effect (Schultz *et al*, 2007). Where this study uses the term "boomerang effect", this is what is meant.

Injunctive norms are norms that embody a measure of social approval or disapproval and when reported thus convey some sense of social approval or disapproval for one, or one's household's, actions. Injunctive norms have been used in some trial cases to reduce the tendency for consumers reported as having consumed less than the norm to increase their consumption when they learn of the higher norm. This injunctive norm is an emoticon that is smiling ☺ if a household is reported as consuming less than the reported average and frowning ☹ if a household is reported as consuming more than the reported average.

By employing the two elements of an injunctive norm along with the descriptive norm one achieves a form of communication where, for an over-consumer, the two norms are harmonious and exercise influence in the same direction[19] , that is influence to consume less. However, for an under-consumer these two norms will exert influences that will be in conflict. The descriptive norm will exert a an influence upon the consumer to increase their consumption (in order to conform the normal level of consumption) while the approving injunctive norm will exercise an influence in the opposite direction (society approves of your actions, keep your consumption down below the norm).

According to the focus theory of social norms (Cialdini *et al*, 1991), a person's behaviour will orient towards the norm that they perceive (whether consciously or unconsciously) as more prominent at the time. The thought is that in all cases it will be the injunctive norm that will be most prominent in people's minds, with the descriptive norm providing information to calibrate any action on the part of the individual or household. Hence it is thought that when a person who is consuming less than the normal level of consumption receives an approving injunctive norm along with the descriptive feedback, they will focus upon the injunctive norm that, in their case, conveys social approval and remain under consuming relative to the descriptive norm, while over consumers reduce their consumption (since both norms suggest that this is what they should do) relative to the descriptive norm.

The effect of combining descriptive norms with injunctive norms in comparative normative feedback has been analysed in two natural field experiments: Schultz, Khazian and Zaleski, (2008) and Schultz, *et al* (2007) - which served as the initial inspiration for this dissertation. Schultz *et al* (2007) randomly assign 290 households in San Marcos, California, to one of two electricity feedback treatments: descriptive only or descriptive and injunctive. The results of their study seem to indicate that a

[19] Assuming that over consumption is socially undesirable and therefore socially disapproved of.

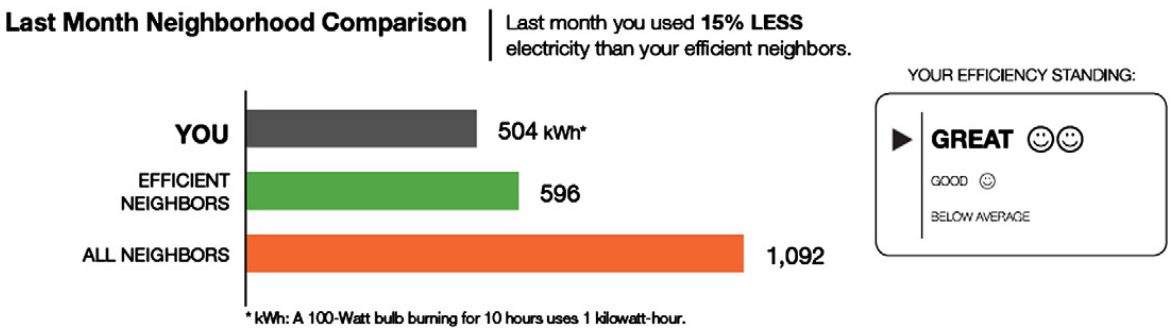
boomerang effect was operating amongst the households in the descriptive only treatment, while the boomerang effect was not evident amongst households in the descriptive and injunctive treatment. As a result, households in the descriptive and injunctive treatment reduced their electricity consumption as a group, while the consumption of the group of households in the descriptive only treatment did not change overall. Schultz, Khazian and Zaleski, (2008) extend this strategy to towel use amongst hotel guests. In this study Schultz, Khazian and Zaleski try to gauge the effect of descriptive and injunctive norms upon guests towel use. In particular, they wished to see which formulation of the normative message would encourage the guests to use their towels for longer periods (thus “recycling” them) before putting them into the washing basket for room service. By combining descriptive and injunctive norms they find that guests recycle their towels more than whether they are encouraged to do so by only a descriptive or an injunctive norm.

Alcott (2011) examines the energy saving program run by OPower with several energy utilities in the USA. The OPower program uses injunctive and descriptive norms in similar ways to which this study used them. OPower partners with utilities by sending households contracted to those utilities "home energy reports" that are several pages in length. These reports are not long essays but rather contain various modules.[20] One of the modules in the report compares households to their neighbours and their efficient neighbours descriptively and augments this with injunctive feedback as in fig. 1.8. (from Alcott, 2011) below. Households that were "great" in terms of their efficiency standing received to smiling emoticons and those that were "good" received one. Households that were below average in their efficiency standing initially received a frowning emoticon but Alcott records that customer complaints caused this practice to stop. The other modules provide tips about how to save electricity. When a new utility signed with OPower, households were randomly selected to receive home energy reports while others were selected to continue as normal. Comparison between these two groups allowed for the effect of the home energy report to be estimated.

Alcott (2011) estimates the that the home energy reports as a whole generate energy consumption that is between 1.4% to 3.3% less than the control group. Focusing his analysis in upon households that were close to each other in terms of energy consumption but on either side of the threshold for receiving the various injunctive norms he finds, using a regression discontinuity design, that the type of injunctive norm makes no difference to the energy reductions achieved.

Fig.1.8. Social comparison module in the OPower home energy report.

[20]More information can be found about the reports at http://opower.com/what-is-opower/reports/



1.9. Conclusion.

The last two decades in particular have seen economics move away from "methodological individualism" and increasingly investigate ways in which the group may influence the individual. One of the identifiable themes of this research concerns itself with social norms. From the literature it seems that people are conditionally cooperative and are thus more likely to adopt a behaviour if there is a signal that a majority have adopted the behaviour in question. A body of research is growing around how social distance may moderate this effect and it seems, from this, that the influence of a social norm upon an individual or household is generally more potent if the referent group is socially close to the individual or household. Given that norms provide information, it is not surprising to discover that the literature considers them to be effective modifiers of behaviour when there is a large degree of pluralistic ignorance about the topic concerned.

In terms of their stability through time, social norms may be stable since they may reduce the transactions costs of learning for an individual. Other reasons for their stability may be their utility for coordinating the actions of actors, social enforcement, people having internalised. These potential reasons for social norms being stable through time may further why norms have consistently found to be influential, even when there is no face to face interaction and they are reported.

Descriptive and injunctive norms are the two sorts of social norms that were used in the natural field experiment at the centre of this thesis. Descriptive norms were reported at two degrees of social distance, far (city) and close (neighbourhood), in order to judge the effect of social distance. Typically people adjust their behaviour towards the descriptive norms. When one is concerned with lowering consumption this produces desirable results amongst those consuming more than average, who lower their consumption, but can produce undesirable results amongst those consuming less than the average, who may increase their consumption as observed in some studies hoping to curb binge drinking by reporting descriptive norms. In order to combat this undesirable "boomerang effect" amongst people consuming less than the average a handful of experiments have augmented

descriptive norms with injunctive norms in feedback to consumers. Injunctive norms are meant to combat the tendency for people consuming lees than average to raise their consumption by communicating social approval of their actions. While some studies (Schultz, Khazian & Zaleski, 2008; and Schultz, *et al*, 2007) have found this strategy effective analysis of another large scale natural field experiment (Alcott, 2011) finds no evidence of injunctive norms having had an effect.

Chapter 2.

Field experiments in the literature of Economics.

Then Daniel said to the steward whom the chief of the eunuchs had assigned over Daniel, Hananiah, Mishael, and Azariah, "Test your servants for ten days; let us be given vegetables to eat and water to drink. Then let our appearance and the appearance of the youths who eat the king's food be observed by you, and deal with your servants according to what you see." So he listened to them in this matter, and tested them for ten days. At the end of ten days it was seen that they were better in appearance and fatter in flesh than all the youths who ate the king's food. So the steward took away their food and the wine they were to drink, and gave them vegetables.

Daniel 1 v11-16
The Bible
(English Standard Version)

The use of the experimental method to evaluate theory and help direct policy can be traced back at least as far as the passage from Daniel quoted above, recounting an event that occurred during first exile of the Jews from Jerusalem to Babylon under Nebuchadnezzer, at about 605BC[21]. The use of the experimental method in order to uncover truth in a more laboratory-like manner can arguably be traced as far back perhaps as the Muslim scholar Ibn al-Haytham writing somewhere between 965-1039 AD (Winter, 1953). Experiments were exceptionally important in the scientific renaissance being central to the work of Galileo Galilei[22]. His use of quantitative experiments to form and validate a theory of gravity is a model that science has used ever since the renaissance. Following his example the sciences have developed through the centuries largely on the back of experimental discoveries made as the method spread through the scientific disciplines to the point where physicists, biologists and chemists of various stripes for at least the last two centuries, have taken for granted the foundational role of the experimental method in validating their existing knowledge base and extending it.

[21] ESV Study Bible commentary on the bible: Daniel.
[22] Through whom such knowledge at least re entered the West (Fara, 2009).

For much of economics' history, economists considered that the only way open for them to move the discipline forward was through testing theory based on naturally occurring data[23] and not in controlled experiments as the quote from Samuelson and Nordhaus' text book (1985)[24] makes clear

> "The economic world is extremely complicated. There are millions of people and firms, thousands of prices and industries. One possible way of figuring out economic laws in such a setting is by controlled experiments. A controlled experiment takes place when everything else but the item under investigation is held constant. Thus a scientist trying to determine whether saccharine causes cancer in rats will hold "other things equal" and only vary the amount of saccharine. Same air, same light, same type of rat. Economists have no such luxury when testing economic laws. They cannot perform the controlled experiments of chemists or biologists because they cannot easily control other important factors. Like astronomers or meteorologists, they generally must be content largely to observe."

Economics, however, has only recently begun to adopt the method, beginning with laboratory experiments that were increasingly used from the 1960's with the work of Vernon Smith touching off growth in that field. Over the last three decades however, the experimental method has been increasingly employed by economists in order to investigate the likely cause of economic behaviour in various situations outside of the lab. It is with these field experiments that this chapter will be concerned.

This chapter begins with a discussion of early field experimental work, culminating in the important period of the 1930's when much of the theoretical foundations for field experiments were laid down, with the work of Fisher and Neyman in particular. From here we move on to discussing the large social experiments evaluating government policy that were conducted in the 1960's and 1970's. We then move to the more recent history of field experiments which has grown in leaps and bounds from the 1980's to the present. We describe the use of three major types of field experiment (artefactual, framed and natural) that have emerged during this period.

2.1.The early days: Field experiments up to the 1930's.

Two of the most important elements of field experiments are the control group and randomisation. In order to identify the effect of a treatment we philosophically wish to observe a subject at t before treatment and at $t+1$ after treatment and the same subject at $t+1$ without being treated.

[23] Recorded in national accounts, crop yields, sales figures, share prices and the like.
[24] pp.7-8

Observing the same subject with and without treatment at $t+1$ provides the perfect counterfactual. Practically this is of course impossible and so various methods of treatment effect estimation have had to devise methods that approximate this philosophical ideal. The random allocation of subjects to various treatments, including a control group is the experimental way of solving the counterfactual problem. Through randomly allocating subjects to various treatment and control groups, subjects face equal probability of being ascribed to any treatment. As a result, if a sufficiently large[25] number of subjects are chosen for the expected standard deviations treatment effects may be estimated at the desired level of precision by comparison of treatment and control groups.

Experiments in the fields of biology and medicine during the 19th century are perhaps the first body of experiments to use control groups as a necessary precondition for generating data that would provide the basis for scientifically distinguishable treatment results (Levitt and List, 2009). Levitt and List relate the story of Pasteur and Rossignol settling the question of the efficacy of Pasteur's vaccination methods (Rossignol had publicly questioned Pasteur's early work in immunization and Pasteur's results) using a control-group enabled experiment wherein the continued lives of fifty sheep randomly split into two groups (one treated with Pasteur's anthrax vaccine and then with anthrax, the other group just with anthrax) were the dependent variables of interest. The continued vitality of all the vaccinated sheep in the face of the apocalypse amongst the non-vaccinated control group (all had died within two days) was a dramatic vindication of Pasteur's thesis.

The earliest instances of the use of control groups spreading from medicine and biology to social sciences appear to have been educational studies conducted near the beginning of the 20th century. By 1901 at least one collaboration of field experimenters in education (Thorndike & Woodworth, 1901) had explicitly identified the need for and employed control groups to allow for such an analysis.

The case most often cited (See, for example: Stigler, 1992) as containing the first example of randomisation employed as a basis for designing experiments is that employed by C.S. Peirce and Joseph Jastrow in their sensitivity experiments conducted over a period from December 1883 to January 1884 and published in 1885. Anne Oakley contends that the early experimental researchers in education went so far in their efforts as to have "invented randomised assignment to experimental treatments...independently of, and considerably earlier than, R.A. Fisher's work at Rothamsted" (Oakley, 1998).

[25] see List, Sadoff and Wagner (2010).

Forsetlund, Chalmers and Bjørndal (2007) provide a helpful analysis of the studies cited by Oakley (1998) and find that none of the pre 1926 studies cited by Oakley amount to randomised field experiments, at least in the important sense that they did not use randomisation to select members of control groups and treatment groups. Mainly these studies seem to have relied upon matching according to prognostic variables in order to generate comparison groups. It seems then that while it is the case that these experiments in social science used what could be called elements of randomisation, it is not the case that they used randomisation in the thorough going fashion that Fisher used at Rothamsted in England or Neyman did at Bydgoszcz in Poland.[26]

For their part, Forsetlund, Chalmers and Bjørndal (2007) cite a 1928 study by H. H. Remmers done at Purdue university as the earliest field experiment in social science study to employ randomisation, along with 9 other studies published in the period 1932 to 1948 as the earliest examples of field experiments based upon randomisation in their design.

So, it would seem that experimental work done before 1930 seemed to be approaching a design based on randomisation, and that by the time the 1920's and 1930's came around "randomisation was in the air" (Rubin, 1990). Certainly theoretical work on randomisation as a condition to be applied in successful experiments in the field, was to a non-trivial degree, put into print before Fisher's work, notably by Charles Saunders Peirce. By way of illustration: a passage in C. S. Peirce's "Reasoning from Samples"[27] reads:

> "The truth is that induction is reasoning from a sample taken at random to the whole lot sampled. A sample is a random one, provided it is drawn by such machinery, artificial or physiological, that in the long run any one individual of the whole lot would get taken as often as any other. Therefore, judging of the statistical composition of a whole lot from a sample is judging by a method which will be right on the average in the long run, and, by the reasoning of the doctrine of chances, will be nearly right oftener than it will be far from right. That this does justify induction is a mathematical proposition beyond dispute..." (Peirce, 1931)

Work on the role of randomisation was thus approaching a point where it could inform the design of experiments such that a proper counterfactual could be constructed. However, before the 1920's

[26] As a result it would not be the case that their work could legitimately employ the body of statistical inference tools that are based upon randomisation such as are commonly used for hypothesis testing today.
[27] Section 16 "Reasoning from Samples" in Chapter 2 "Lessons from the history of science" in part I " General Historical Orientation "of "Principles of Philosophy". Available at http://www.textlog.de/charles_s_peirce.html. or http://www.textlog.de/peirce_principles.html

statistical knowledge was still very much fragmentary with many fields within the discipline that would be essential to any system of statistical inference characterised by confusion. In particular, there was no clear explanation of how to assign probabilities to various hypotheses – the inverse probability problem (Armitage, 2003). One of Fisher's contributions would be to develop the idea of randomisation as a basis for valid inference in a mathematical fashion such that clarified randomisation's role in validating tests of significance, especially as to how this related to assigning probabilities to various hypotheses – thus solving the inverse probability problem.[28]

At Rothamsted agricultural testing station, Sir John Russell had succeeded Adam Hall as director in 1912. Rothamsted manor had first seen experiments when in 1843, the owner of the manor, John Bennet Lawes, together with a young chemist, Joseph Henry Gilbert, began experiments to test various fertilisers and crop yield relationships. These were thus truly field experiments. Agricultural experiments continued at Rothamsted from 1843 without a stop and continue to this day at the manor.[29] As a result, when Russell assumed directorship of Rothamsted, he had sixty-nine years of experimental data to deal with. His sentiments on this front are captured in this excerpt from his "History of agricultural experiments in Great Britain:"

> "On taking charge at Rothamsted I found great files of records which I knew I could never deal with adequately...I knew that the Census Authorities had methods for extracting information from great masses of data, and in 1919 after the war I applied both to Oxford and to Cambridge universities for a young mathematician familiar with similar methods who would be prepared to examine our data and elicit further information that we had missed." (1966, p. 325 in Box, 1980).

The position of statistician at Rothamsted had thus been created with the view to organising, the now over seventy years worth of, experimental data and bringing modern statistical methods to bear upon it so as to extract "further information" that had been "missed.[30] It was this position that was offered to Fisher an which he took up in 1919.[31]

[28] . In the process also introducing the term "null hypothesis" into significance tests.
[29] Although now run under the auspices of "Rothamsted Research".
[30] Sir John also records that at the time he had only six months funding for the position. He hoped that this would be enough time to determine whether the position of statistician would be useful to Rothamsted and so justify making it permanent. One can only imagine the pressure that the newly married Fisher must have felt.
[31] At the end of World War I, Ronald Fisher was offered at least two jobs; one at the Galton laboratory under Karl Pearson, the other, by Sir John Russell at Rothamsted agricultural testing station, where he could be more of his own man. By 1919 Fisher and Pearson were emerging as the leading thinkers in statistics. It seems that they were not personalities that could easily work together; a rivalry that would grow to outright animosity had begun to develop between the two[31]. Almost certainly because of this growing rivalry, Fisher took up the position of statistician at Rothamsted in 1919 in preference to the Galton position.

Fisher's daughter, Joan Fisher Box, records that at the time Fisher began his work his view of the role of the statistician was very much that it would primarily be that of the estimator of information from data envisaged by Russell (Box, 1980). However, somewhere in the course of what must have been a gargantuan work, Fisher realised that all estimation can only generate as much information as the data generating process or processes allow. It was a realisation that had significant implications for experimental design as it shifted the "weight" of responsibility for generating good data and therefore good conclusions, on to experimental design. Fisher felt that the "moral balance" had shifted on to process which generated the data which was to be analysed by the statistician (Fisher, 1947).

The powerful argument which caused Fisher to declare a swing in the "moral balance" toward experimental design was, of course, his analysis for determining the maximum amount of information present in experimental data, otherwise known as Fisher information, fully published in 1925[32] but apparently hit upon by the man somewhere around 1922[33]. This theorem allowed Fisher to prove the maximal amount of information to be gained from any particular experiment with the result that, in his view, the role of the statistician changed:

> "The statistician is no longer an alchemist expected to produce gold from any worthless material offered him. He is more like a chemist capable of assaying exactly how much of value it contains, and capable also of extracting this amount, and no more. In these circumstances, it would be foolish to commend a statistician because his results are precise or to reprove because they are not. If he is competent in his craft, the value of the result follows solely from the value of the material given him. It contains so much information and no more. His job is only to produce what it contains." (R.A. Fisher quoted in Rao, 1952).

As a result, from some time in the period 1922 and 1925, Fisher began to focus his energies on informing the design of the experiments, primarily at Rothamsted, which generated the data and so to move his attention, naturally, to better experimental design.

In a 1923 paper with Mckenzie as well as an accompanying paper published in 1926, Fisher began to showcase how randomisation as well as blocking, replication and factorial design could be brought to bear upon experimental design such that the optimal amount of information could be produced in

32 Fisher, R.A. (1925). "Theory of statistical estimation" Proceedings of the Cambridge Philosophical society, 22: 700-725.

[33] Box, 1980; p1 "Thus by about 1922, Fisher recognized that a statistician, so long as he did his arithmetic right, had no responsibility for the value or worthlessness of his estimates; consequently, "the weight of his responsibility was thrown back on to the processes by which the data had come into existence"(1947, p. 435). Fisher accepted the design of experiments as his charge.

an experiment. This work culminated in his book "The Design of Experiments" published in 1935. The book was used for many years as a staple text on experimental design as well as statistical inference, it also perhaps served most effectively to increase the actual use of randomisation, blocking, replication and factorial design in the design of experiments in the field.

At the same time as Fisher was doing his work at Rothamsted, Jerzy Neyman was doing very similar work in Poland, also at an agricultural test station (in Bydgoszcz). Although his work was mainly concerned with surveys and sampling his early work, as evidenced in two long papers written in 1923, clearly took account of the foundational role randomisation played in validating hypothesis testing as we know it. For this, Neyman credited Fisher (Reid, 1982), however his work in this broader area did much to establish the experimental method in the field and his work on randomisation driven sampling, which produced his ground breaking 1934 paper on the subject, went on to inform much early work that dealt with sampling.

2.2. Social experiments in the mid-twentieth century.

The mid twentieth century saw many states using the randomised experimental methodology as part of the broader policy evaluation and formation process. These randomised social experiments[34] could span many neighbourhoods or even whole states. They typically had voluntary participation. Social experiments, in general, focused upon evaluations of local or national government policy, where the policy being evaluated is either existing or new, proposed, policy. In particular, social experiments seem to have focused upon elements of what could be called the welfare state (Levitt and List, 2009).

One the earliest social experiments (Levitt & List, 2009) directly examined electricity consumption. The study was conducted across six electricity board areas in Britain, between 1966 and 1972 and examined the response of consumption to various (randomly assigned) tariff structures. A Swedish study, conducted in 1975 in Eskilstuna, focused upon assaying the influence upon employment of extra staffing at employment offices. A randomly split sample saw one group making normal use of employment services (1.5 hours per person) while the other group received significantly more

[34] Again, what appears to count decisively, in the literature, in favour of a study being classed as a social experiment is that it employs randomisation. Levitt and List (2009) consider a Swedish social intervention to extend the years of compulsory schooling to nine years as falling outside the family of social experiments on the basis that assignment of the treatment was not done in a random fashion. Greenberg, Shroder and Onstott (1999) provide the following definition of a social experiment. "Social experiments are field studies of social programs in which individuals, households or (in rare instance) firms or organizations are randomly assigned to two or more alternative policy interventions, or "treatments". "The same authors quote the following criteria for the definition of social experiments from the "Digest of Social Experiments": "(i) random assignment (ii) policy intervention (iii) follow-up data collection and (iv) evaluation" (Greenberg and Shroder; 2004).

service (using 7.5 hours per person). The study found that greater use of services did in fact raise the chance of being employed, raise earnings and lower the period for which a person was unemployed. Since this experiment a larger literature has grown in Europe dealing with labour market policy[35] (Levitt & List, 2009).

Large Scale Social Experiments are recognised by most (see, for an overview, Greenberg[36] et al, 1999) to have had their genesis in the United States, a location of many subsequent social experiments, with the PhD dissertation of Heather Ross.[37] The project appears to itself have grown out of the debate in the USA around the welfare system that was generated by the 1966 Coleman Report. This debate centred upon the consequences in the labour market of existing welfare policies and the possibilities of different sorts of income supplementation.[38] It was Ross' suggestion that random assignment of several suggestions being made at the time be used to demonstrate the relative benefits of the various proposals in a social experiment. The title of the proposal document she wrote (while at the Brookings institution) is informative "A Proposal for Demonstration of New Techniques in Income Maintenance" in indicating the sort of epistemic value that early experimenters imagined could be delivered by these social experiments.

The actual experiment, when implemented[39], was an exercise in exploring the effects of negative income taxes and guaranteed levels of income. Around 1300 male-headed households were selected and treated with negative income tax rates (which varied between 30% to 70%) and guaranteed levels of income (which varied from 50% of the estimated poverty line for a family of four to 125% of the same) over three years. Every three months, during this period, they were asked to complete a questionnaire. The questionnaire explored issues such as family labour supply, consumption and expenditure patterns, general mobility, dependence on government, and social integration (Levitt and List, 2007).

Initial analysis of the results by Ross in 1970 indicated that the negative income tax intervention slightly raised the labour effort of households, a question of particular concern at the time. Since this analysis however, the data have been revisited several times and, it has been pointed out (by Moffit, 2001) that the interventions occurred in the context of a welfare system that, at the time, may have

[35] See Bjorklund and Regner (1996) as well as the various *Digests of social Experiments.*

[36] Greenberg is also one of the two editors of "Digests of Social Experiments", the other being Shroder.

[37] No doubt of at least a little fascination to academics is the price tag of Ross' research. Levitt and List indicate that the research cost more than US $5 million at the time, more than US $30 million in 2007, making it, in their estimation, "one of the most expensive dissertations in economics" (Levitt and List, 2007).

[38] A debate that current South Africa can readily identify with.

[39] Levitt and List (2007) indicate "that after the typical fiscal wrangling" Mathematica Inc. (based at Princeton, New Jersey) and the University of Wisconsin-Madison (where Heather Ross was based) were commissioned to conduct the study.

reduced labour effort (by means of large benefit-reduction rates) with the result that the negative income tax programmes performed better than existing programs even though the tax regime did reduce labour effort (which Levitt and list indicate is the majority view in the literature on this experiment), although this was not uniformly the case since in some areas it performed better than the status quo, while in others it performed worse. The fact that this sort of comparative analysis is possible is, however, illustrative of the virtues of the experimental method in general.

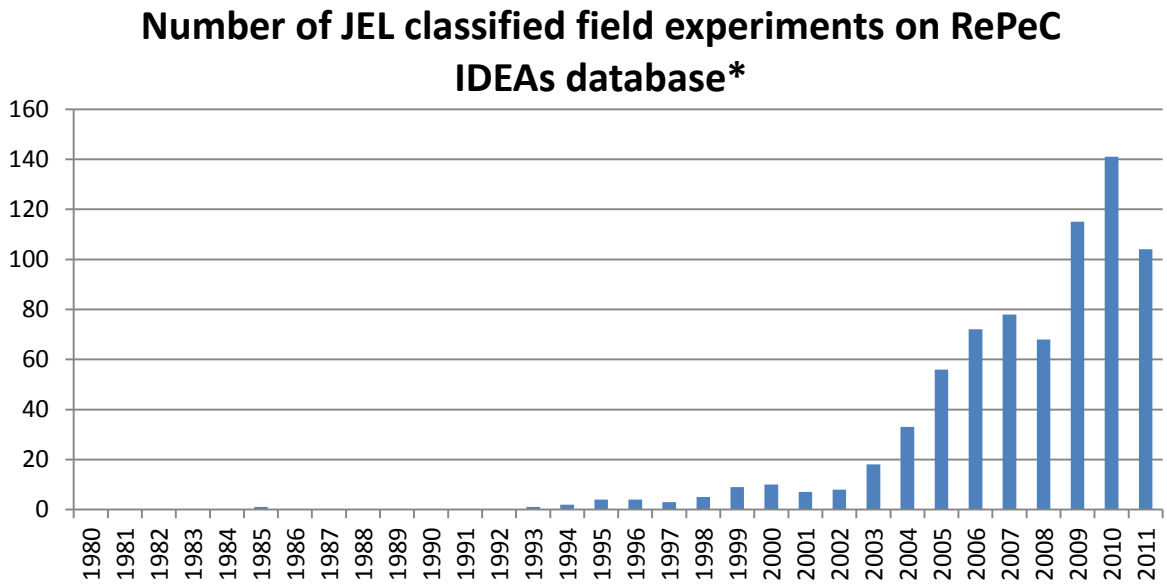
This experiment seems to have been catalytic in its influence upon the growth of social experiments in the United States of America and, thereby elsewhere in the world. Also influential were various treatises upon the virtues of social experiments such as that by Orcutt and Orcutt in 1968 (List & Rasul, 2010). During the 1980's the growth of the use of social experiments to evaluate policy provided a context for a largely constructive debate in economics between those favouring the use of social experiments and those in favour of structural econometrics (Manski and Garfinkel, 1992).

2.3. The current generation of field experiments: 1990's to the present.

Since the 1990's the number of field experiments appearing each year in various journals within the discipline of economics has increased steadily, as can be seen in figure 4.1. This section begins with a brief discussion of how the increased adoption of the field experimental method by academic economists differs from the earlier field experimental areas discussed in sections 2 and 3. The discussion then moves to discussing three sorts of field experiments, following Harrison and List's (2004) taxonomy, that are current in economics.

The most important way in which this current wave of field experiments differs from the earlier era of social experiments is in their orientation towards economic theory (Levitt & List, 2007). Field experiments appearing in the economic literature of the past two decades are concerned with testing and developing economic theory, particularly microeconomic theory. Although many field experiments (such as Alcott, 2010) address a practical policy issue directly, many do not (such as List & Lucking-Reiley, 2000) and some are a mixture (Duflo, Kremer & Robinson, 2009). Overwhelmingly most, however, address economic theory directly by testing the predictions of those theories in the field in an experimental fashion (including the two studies just cited). This is quite different to the social experiments described in the previous section which were largely concerned with testing policy measures.

Figure2.4.



*From http://ideas.repec.org/j/C93.html accessed 05 December 2011,8:40pm (SAST)

This is congruent with what some of the leading economists engaging in field experiments view as the methodologies role within economics. List and Levitt are advocates of field experiments being used in closer conjunction with other methodologies available to the economist. List, in particular has argued in several places[40] that field experiments have a particular *role* to play in taking economics forward. He views field experiments as providing researchers with a bridge between laboratory experiments and the study of naturally occurring data (such as that collected by labour surveys). List's view of how field experiments provide such a bridge may be summarised as a view that sees them providing a controlled test of the degrees to which conclusions drawn in the lab may be generalised while providing superior identification of influences (via the controls of an experiment – and the apposite employment of randomisation) to inform estimation on larger sets of "naturally occurring" data.

Field experiments have been used in this fashion over the last two decades in a growing number of field within economics. Some of the fields making noticeable use of the methodology are: labour economics, charity and investigations into the nature of the economic agent.

A large amount of the work has focused upon the nature of economic agents in development contexts, with topics such preferences, reciprocity, and norm enforcement being covered. Perhaps one of the most intriguing studies in this vein is the work done by Henrich, Bowles, Barr, et al (2001)

[40] See for instance List 2007, List 2008, List 2006, Levitt and List 2009.

amongst 15 small scale, fairly primitive societies around the world which seems to have produced some stunning data on the influence of culture upon the nature of "Economic Man".

In the field of labour economics, field experiments have been used increasingly to investigate a variety of topics across the lifecycle of a person (List and Rasul, 2010). Areas within labour economics investigated include the demand for education and human capital acquisition, wages, labour market programs, incentives and organisation in the workplace, marriage and the decision to retire.

The growing field of the economics of charity has also made particular use of the field experimental method, with notable contributions made by Peter Bohm, one the first researchers to meaningfully extend the experimental method beyond the laboratory and out to the field. Within this field researchers such as Rachel Croson have made several contributions to indentifying the influence of social comparison upon the propensity to contribute[41] that are of special relevance to the research presented in this thesis.

In 2004 Harrison and List presented a taxonomy of filed experiments that is fairly helpful in distinguishing the main ways in which experiments differ one from the other. Harrison and List describe a continuum along which the experimenter has varying control of the data generating process. We will discuss these briefly before concluding this chapter with an overview of each major category of field experiment.

At one end of the continuum are artefactual experiments where the researchers is very much in control of the entire data generating process. Artefactual field experiments may be very much like laboratory experiments in that subjects are to be found playing games or performing actions that have been explicitly defined by the experimenter. Importantly subjects are aware that they are in an experiment. Artefactual experiments differ from laboratory experiments in that they use subjects that are from the relevant real world setting rather than a convenience sample (such as students) and may often also be conducted outside of the laboratory but need not be.

In the middle of the continuum lie framed field experiments. Framed field experiments are always conducted outside of the laboratory in the "real world" setting appropriate for the theory being tested with subjects that are naturally to be found in that setting, not drawn from a convenience sample. Subjects in these experiments do not play tightly controlled games but rather are randomly assigned to a treatment group, where they may be exposed to something like a wage subsidy,

[41] See for instance Croson and Shang (2008) for the use of social comparison in "downward social information" and (2009) for the "upward" version.

counselling, given free de-worming medication, or a control group. Importantly, subjects are aware that they are being observed.

At the furthest end of the continuum lie natural field experiments. These are experiments that are designed by an experimenter in which subjects are randomly assigned to treatment or control groups. However, unlike the case of framed field experiments, subjects in natural field experiments are not aware of the fact that they are being observed and the treatment, such as the receipt of a flyer with their electricity bill, comes as part of their daily routine.

Before we move on further however, a word here on how field experiments stand distinct from another fast growing literature, natural experiments, is probably useful; since field and natural experiments may be conflated. Field experiments are distinguished from natural experiments mainly by consideration of how the randomisation is generated. In field experiments the researcher is responsible for random assignment of treatments, while with natural experiments the researcher looks for "naturally occurring" situations where various (potential or actual) influences (treatments) have been brought to bear upon people in a fashion that is, or is sufficiently close to, random.

It is important to note that by "naturally occurring" is meant that neither the situation (be it a market or some other setting) nor the "treatments" have been designed by the researcher although they may have been designed by somebody else. Thus assessing the effects of differential dorm assignment policies at various universities (who draw the same calibre and demographic spread of student) could be considered a natural experiment even though these policies would have been formulated and implemented by design on the part of the respective university authorities.

2.3.1. Artefactual field experiments.

Artefactual field experiments are fairly similar to laboratory experiments in that they are conducted under controlled conditions within a "lab" setting or using lab tools. The key difference between artefactual and lab experiments lie in their choice of subjects participating in the experiments. Artefactual experiments make a point of drawing their subjects from the general subject pool of interest (such as traders, farmers or rural tribes-people) whereas lab experiments do not place such a restriction upon the subjects recruited, often recruiting college students.

The benefit that is gained in this is to test the ability of theory to actually explain behaviour of people to whom it is imagined to apply. Of course in an artefactual experiment the subjects are drawn at random from the context of interest, allowing the exercise to be classed as an experiment.

Artefactual experiments perhaps began in earnest around the late 1960's and early 1970's with work such as that by Bernard Alpert (1967) and Peter Bohm (1972). Alpert's 1967 paper concerned itself with testing whether the results of the behavioural experiments, then entering use in business college courses, in fact extended to "real businessmen". His introduction to this paper explicitly highlights a concern about the external validity of conclusions drawn from experiments conducted with students or military personnel.

> "With the inclusion of behavioural courses in the business colleges, the results of behavioural experiments are being used as a basis for proposing resolutions to simulated business problems. But it is not known if the results of such experiments are valid for drawing conclusions about behaviour of businessmen. In many behavioural experiments, the samples have been drawn from the student population, ranging from nursery school to college, or else they have been drawn from military organisations, the next most accessible source of subjects." (Alpert, 1967).

His proposed epistemological remedy was to use businessmen as subjects in similar experiments and see whether the results hold amongst these business men to who it is supposed to apply.

The results of Alpert's 1967 paper indicated that reactions to various situations varied widely between student, military personnel and business people. The clear implication was that it may be dubious to extend results that obtain amongst students to beyond that universe of subjects.

Bohm's 1972 work on the difference in elicited value when the payoff and situation is hypothetical and when it is not. The paper has generated a significant amount of subsequent work in a similar vein in environmental economics (List & Gallet, 2001; Levitt and List, 2009). Peter Bohm's 1972 experiment involved examining the differences between respondents' stated willingness to pay for a sneak preview of a Swedish television program when one group stated their WTP when the preview and WTP were only hypothetical and the WTP's elicited when it was not hypothetical. Respondents were 20-70 year olds who were randomly sampled from Stockholm's population and thus likely customers in the market concerned.

Another major research project of recent years that indicates clearly the utility and utilisation of artefactual experiments to inform theoretical and developmental concerns is that of the seventeen person research team of Bowles, Henrich, Enslinger, Boyd, Smith Henrich, Hill, Camerer, Gil-White, Fehr, Gurven, Gintis, Marlowe, Mcelreath, Patton, Alvard, Tracer and Barr.

This unusually large team conducted experiments with fifteen small scale societies across the world. All societies were rural and were selected to a significant degree on the extent to which their

members lived their everyday lives separate from a market economy. The aim of the project was to test the extensibility of the canonical model of rational decision making as described by the rational agent, *Homo Economicus,* familiar to economics students since Alfred Marshall. The project's history and development is illustrative of the manner in which artefactual field experiments can be and are used to inform microeconomic theory and its development.

By the mid 1990's cross cultural experiments had uncovered meaningful deviations from microeconomic theory (Davis & Holt ,1993, Roth & Kagel, 1995) under a variety of conditions. In response to this accumulating evidence, microeconomic theorists began to propose modifications to microeconomic theory in order to explain these deviations and incorporate them into microeconomic theory (Camerer, 1997, Henrich et al, 2000). In proposing these modifications however, most theorists made explicit or implicit assumptions about human nature universally, with the necessary result that all humans everywhere were predicted to respond to various economic situations and problems in the same way (Henrich, 2000).

A large part of the reason for this may have been that these studies had, for the most part, employed that most prevalent of convenience samples, college student; the important point is that this is the reason Bowles, Henrich, et al most suspected (Bowles, Henrich et al, 2005). In 1996 Joseph Henrich had conducted ultimatum game experiments amongst the Machiguenga in the Peruvian Amazon. The Machiguenga behaved in the ultimatum games in a way that departed in a clear fashion from predictions of the canonical model of human behaviour used in economics and also differed significantly from results gained elsewhere in the West (Henrich 2000).[42] The use of a experimental procedure (the ultimatum game) that is consistent across target and control groups, but where respondents were aware that they were being observed,[43] in a replicable fashion renders the methodology artefactual in terms of Harrison and List's (2004) terminology.

This study inspired the subsequently larger study by Henrich, Bowles, et al (2005). This subsequent study was a systematic attempt to detect whether behaviour in well-established classes of economic games could be correlated to observed cultural differences and, importantly, whether these differences in play constituted departures from the predictions of the canonical model that could be systematically correlated with observed cultural differences.

[42] Henrich facilitated this comparison by means of two sorts of data for Western respondents. Most importantly, he ran a control group with college students in Los Angeles – which differed significantly to the Machiguenga. Henrich also used already-gathered data from similar experiments conducted by other researchers amongst western respondents in market-based societies (Henrich 2000).

[43] It is not clear that the Machiguenga understood that they were part of an experiment. Henrich records that the Machiguenga were gathered "under the auspices of "playing a fun game for money."”(Henrich 2000). This however, is not critical to considering the exercise an artefactual experiment.

In order to determine this, ultimatum games were played in all locations (with dictator and public goods games being played in addition to ultimatum games in some locations). The ultimatum game is highly artefactual in character.[44] It is also important to note that the ultimatum game is standard, like a game of singles tennis, its rules are the same everywhere and provide definite limits to the extent and direction of variation in playing conditions. The rules of the ultimatum game thus allow for a meaningful degree of comparability between results obtained at various places across the world. Of course, other factors may also vary across locations where the game is played; some of these are, of course factors the effect of whose variation the researchers were interested in ascertaining, while other factors may be confounding and these the researchers endeavoured to also standardise across locations. In terms of payoffs, instructions and how the game is played, the process is identical across all locations.

It was by following this procedure and utilising this methodology that the researchers in Henrich, Bowles, et al (2005) hoped to cast light upon the influence of cultural, for want of a more precise word, context upon the nature of agents involved in economic theory, especially the degree to which they could be described as selfish or unselfish. Their findings are compelling and indicate that variations in play in the ultimatum game can be correlated to variation in culture.

The systematic use of the methodology of games and the broader experimental method, the seeking out of variation in the variable of interest (culture separate from the influence of modern market economy societies), displays the sort of use researchers have been putting the artefactual field experimental method to.

Artefactual experiments have been used in an agent-examining sense in the developing world by many other researchers too, including Abigail Barr, Juan Camilo Cardenas and Jeffrey Carpenter in the past decade. Carpenter's research has included much by way of examining cooperation, competition and commons dilemmas. Often times Carpenter's work has been in development contexts but with the notable difference to the Henrich et al team (and in common with most of the rest of the literature) that their experiments involve subjects who would be considered to be very much a part of the modern market economy.

Artefactual experiments seem to have lent themselves to being able to explain non experimental outcomes apart from allowing closer examination of results generated within-lab. At least one of Carpenter's projects falls into this classification (Levitt and List, 2009), that being his work with Erika

[44] in the course of daily life one does not usually get given a sum of money which one can divide between oneself and another and keep, provided the other person agrees to the division, failing which neither party gets anything.

Seki amongst Japanese fishermen to investigate the likely influence of exposure to competition in various contexts on cooperation amongst workers across those various contexts.[45]

The inductive mechanisms of that agents themselves are often thought of as using have also been examined using artefactual field experiments. Sadoff, Levitt and List (2010) examine a cornerstone of many of game theory's important early theorems, namely that agents use backward induction to solve problems. They do so with an arresting artefactual experiment conducted with chess players. They find that chess players seem to often pass over the unique subgame perfect equilibrium that is predicted for backward inducting agents, in games such as the centipede game, in favour of a cooperative gain. This is despite demonstrating a high level of backward inducting ability in chess. The results suggest that other signals, independent of backward induction, exist by which such players can judge the probability of cooperation to be high. Especially suggestive is the fact that chess players reported that cooperative play (in terms of agreed draws near the end of a tournament) is something that they often do (Sadoff, Levitt, List, 2010)[46].

2.3.2.Framed field experiments.

Framed field experiments are defined by Harrison and List (2004) as being experiments that are very similar to artefactual field experiment. Researchers still define the parameters of the experiment very tightly and the subjects are aware that they are being observed. The major difference between framed and artefactual experiments is the degree to which natural context enters the experiment. Framed field experiments employ significant elements of the context, beyond using "natural" subjects as opposed to students or subjects drawn from some other merely convenient sample, in the experiment. As such framed field experiments are run outside of the laboratory setting and in the "real world" as part of subject's everyday lives. Framed field experiments have the attraction of generating data in a manner that is closer to being naturally occurring than data generated in the laboratory or under an artefactual field experimental framework.

The social experiments described in section 4.3 are a type of framed field experiment. Social Experiments differ from the sort of framed field experiments that we will describe below and which have been used more extensively over the past ten years in that the subjects in social experiments are not just aware that they have been selected but are also aware that they have been randomly selected. Most of the more recent framed field experiments that we will discuss have subjects who

[45] They interpret the data generated by their experiment as indicating that working in a more competitive environment decreases the probability of cooperation relative to what would be the case for working in a less competitive environment.

[46] See also Gränsmark (2010) as well as Bühren& Frank (2010) for some more recent work on the backward induction problem in a similar vein.

are not aware that they have been randomly selected. As a result any "randomisation bias" is conceivably dealt with. Social experiments were also typically conducted over a much longer term (often years) than the current wave of framed field experiments. Such long term studies ran a higher chance of suffering from attrition bias in their sample than the current wave which has seen studies typically conducted over a much shorter time.

Framed field experiments have been increasingly employed by development and development economics researchers over the last decade. One of the most active research groups in this field is the Abdul Latif Jameel Poverty Action Lab (J-PAL).It is indicative of the rise of this field experimental method that J-PAL, from one office at MIT in 2003 now has offices in Africa, Europe, Latin America and South East Asia (see for instance Duflo et al, 2009, Chattopadhyay & Duflo ,2004).

These projects are often mainly concerned with government policy or NGO program evaluation in a development context. As such these exercises may often be concerned with evaluating bundles of influences rather than in isolating the effect of each one by itself. As policies and programs usually have more than one element to them this approach makes good sense. However in some of the studies the resulting confounding influences does make it difficult to sharply identify the active economic ingredients.

Beyond the arena of policy and development economics, economic theory has also been the target of many researchers using this methodology over the past decade and more. Predictions of bidding patterns in various auction designs and the hypothetical valuation problem have been subjects of analysis by using framed field experiments. These sorts of studies have bearing, particularly in the USA upon how contingent valuation methods may be used to value public goods (List and Shogren, 1998b).

List and Shogren (1998b) in their baseball card experiment examined whether delegates and dealers at a sports card show[47] would submit bids for cards in hypothetical auctions that would be significantly inflated relative to bids submitted by those same people in actual auctions. Using different sorts of baseball cards and different numbers of those cards (1 card and 10 card auctions were employed) as good, dealers and non-dealers at the sports card fair participated first in a hypothetical second price auction and then an actual second price auction.

List and Shogren found that bidders did consistently inflate their bids in hypothetical auctions. However bidders did not inflate their hypothetical bids relative to their bids in an actual auction in a fashion that was uniform across different sorts of cards bid for, or uniform across auctions with

[47] In Denver Colorado, 1995. (List and Shogren, 1998b).

different numbers of cards bid for. When compared to results from other experiments testing valuation under hypothetical and real auction situations, such as radiated and irradiated meat, (Fox, *et al*, 1998) this result indicated that there probably did not exist a calibration function that was common across different goods or different types of market (or indeed, non-market auction) situations but rather such calibration functions would probably apply to clusters of goods and contexts (List and Shogren 1998a, 1998b).

Still with sports cards, List and Lucking-Reiley[48] (2000) test the validity of theory of demand reduction (Engelbrecht-Wiggans and Kahn's, 1998, theory) that predicts that first unit bids will be equivalent across uniform-price sealed-bid auctions and Vickrey sealed-bid auctions with Vickrey auctions still proving more allocatively efficient than the uniform price auction format since the format of the Vickrey auction (like the English Auction) elicits what is thought to be the true valuation of the good by a bidder through awarding the good to the highest bidder but only compelling the winning bidder to pay the second highest bid price[49]. To achieve a test of this theory, List and Lucking-Reiley randomly allocate subjects into Vickrey auction and uniform-price auctions and auction off a significant value of sports cards (328 sports cards were auctioned, valued at roughly ten thousand dollars in total) in two-unit, two person formats. While they found that the Vickrey auction format elicited a more allocatively efficient bid pattern, with participants underbidding in the uniform-bid format, they also found that, in contradiction to theory the first unit bids were not equal across the formats but, rather, were significantly higher in the uniform price auction format.

This particular experiment and the responses to it offers a nice example of the dialectic between testing and theoretical innovation that experimentation in general makes possible. In response to the experiment and subsequent lab experiments that confirmed List and Lucking-Reiley's anomalous result (see Engelmann and Grimm 2004, Porter and Vragov, 2003), Dan Levin (2005) proposed a theory that developed Engelbrecht-Wiggans and Kahn's (1998) theory of price reduction such that it could accommodate all of the results from List and Lucking-Reiley's (2000) experiment which generated an observation that was anomalous by the lights of Engelbrecht-Wiggans and Kahn's theory. Furthermore, Dan Levin constructed his theory in such a way that it produced testable predictions that future experiments can verify or falsify.

[48] Now David H. Reiley since Augsut 2002, when he changed his name from David Lucking-Reiley (http://www.u.arizona.edu/~dreiley).

[49] In a Vickrey Auction bidders submit sealed bids without knowing the amounts that other bidders are bidding. Although the highest bid wins, the price paid by the winner is the second highest bid. This bid format gives bidders a rational incentive to bid their highest actual valuation of the object being bid for.

2.3.3. Natural field experiments.

In the last twelve or so years, natural experiments have been used to address problems across almost every sub-field of microeconomics (List, 2011). In a natural field experiment, the researcher designs the treatments and randomly assigns subjects to the treatment and control groups. Importantly, the subjects are never aware that they are part of an experiment, with treatments being elements of everyday life. This also means that subjects are not able to select out of treatments. As a result, this methodology eliminates selection bias, and allows for parameter estimates that are more representative of the whole population of interest. It also eliminates bias that may arise from subjects acting differently to normal if they are aware that they are in an experiment. Parameter estimates are thus likely to be representative of not only the whole population but of that population as it would act under normal circumstances. We discuss two experiments to illustrate the typical value of a natural field experiment.

Since subjects cannot select themselves out of a study, the role of ethics committees is rendered much more important when a natural field experiment is undertaken. West and Gunn (1978) provide a helpful summary of the framework of ethical considerations of experiments in general involving human subjects from a psychological perspective, while natural experiments get around most of the ethical problems of deception discussed by Bonetti (1998). Certainly no subject should be made worse off than they would normally be in the course of their natural participation in a market or set of activities.

One experiment, conducted via a public radio station by Croson and Shang (2009), was designed to test two major classes theories that have been proposed to explain the relationship between one's own donation and the donations of others. One set of models[50] explains one's donations as being substitutes for the donations of others while the other models the individual's donations as complements for the donations of others. Of course what both models to do not presume to explain, in what would be an almost mystical sense if they did, the relationship between the individual and others contributions *in* fact, but rather the influence of *information* about others donations upon really an individual's donation/s.

[50] Croson and Sheng understand these models as falling into 3 major categories: the pure altruism models after Becker (1974), Warr (1982) and Roberts (1984) which predict a dollar for dollar crowding out of individual giving by high normative giving, the impure altruism models such as those described by Andreoni (1989, 1990), Steinberg (1987), Ribar and Wilhelm (2002) which describe normative giving as crowding out less than one dollar of individual giving per dollar normatively given, and finally threshold models which differentiate effects of normative giving upon individual giving according to whether the threshold for the provision of the public good in question has been reached or not (e.g. Andreoni, 1998, for public goods and Romano, 1991, as well as Cornelli, 1996, for private goods).

The manner in which this natural experiment is here used to collect data on the reaction of an individual's giving to information about the donations of others demonstrates nicely the utility of the natural experimental method for the development of theory in economics. They start by summarising the literature on models of generosity as being able to be split into two classes, where donations of others are substitutes or complements to an individual's donations. This allows for the description of a two competing and testable hypotheses: that information about the giving of others will be negatively correlated with an individual's giving if the relationship is of a substitutionary nature; competing against the hypothesis that information about the donations of others will be positively correlated with an individual's donations if the relationship is that of a complement.

The experiment also demonstrates how natural field experiments may be used with laboratory experiments in order to inform theory. Croson and Sheng use this natural field experiment as an opportunity to gather data for whether earlier laboratory experiments may generalise beyond the lab.

Armin Falk (2007) also pursued an interesting natural field experimental test of the importance of gift exchanges in soliciting donations via direct mail. As Falk notes at the outset of his paper, the theory and importance of gift exchange had already been discussed fairly widely in labour relations (Akerlof, 1982 and Bewley, 1999), customer relations (Kahneman, Knetsch & Thaler, 1986), bargaining (Camerer & Thaler, 1995), and price setting behaviour (Huck & Wallace, 2002). The experiment was constructed explicitly to test the degree to which hypothesised gift-exchange behaviour would hold in the everyday world. Falk randomly allocated 3 treatments (solicitation letters with no gift, solicitation letters small gifts and solicitation letters large gifts) across 10 000 potential donors to a charitable organisation.

He found, that a small gift increased contributions by 17% on average when compared to sending no gift, with this figure rising to 75% for a large gift. The results make a compelling argument that reciprocity in the form of gift-exchange extends beyond the only theoretical.

Natural experiments have been used in a similar fashion in many other areas of microeconomics too beyond these just discussed. In labour economics there are natural experiments addressing most areas of a person's life-cycle (see List and Rasul, 2010 for an overview), improving savings behaviour (Ashraf N., Karlan, D., and Yin W., 2006a, 2006b) institutional influence upon behaviour (Bertrand *et al*, 2006), as well as incentives (Azfar and Zinnes, 2006) to name some.

2.4. Conclusion.

Although the experimental method has long been a staple of the other sciences and the methodology to execute field experiments was largely described by Fisher and Neyman in the 1920's and 1930's it was not until fairly recently that economists adopted the methodology. Once economists began conducting experiments out in the field, the experimental literature fairly exploded. Today journals within economics will regularly carry articles based upon field experiments that may be artefactual, framed or natural field experiments. Of these, researchers particularly favour natural field experiments since the mixture of randomisation and "real-worldness" that it provides allows them to test theories free of selection bias and Hawthorne effect concerns.

Chapter 3.

Methodology: Experimental design, fieldwork and estimation.

This chapter describes the field experiment that was run between January to December, 2010 in the city of Cape Town. The experiment tested whether reporting the average level of electricity consumption, a form of social norm, would affect the electricity consumption of households. Norms were reported at two levels, city and neighbourhood, to different households in order to test whether social distance would be a factor that would significantly modify reaction to reported social norms. Injunctive norms were also included with half of the treatments in order to assess whether they would curb any "boomerang" effects as suggested by some of the literature on social norms.

The study was initially designed to test the different effects of varying social distance and including an injunctive norm. However soon after fieldwork began interest amongst several stakeholders about the effect of the treatments relative to a control group receiving no feedback was piqued. This raised several difficulties for analysis and these are discussed here too.

Also discussed are the methods used to estimate the relationship between household electricity consumption and the treatment elements that are used, in addition to climatic variables, how these variables are constructed and the expected nature of the relationships to household electricity consumption.

3.1. Treatments.

The treatments administered in this field experiment were designed to test the influence of two elements of normative feedback. These two elements were: the influence of injunctive norms and the influence of varying social distance between the reported norm and the household to which it is reported. Households in the various treatments received the treatment assigned to them as an insert in their electricity bill every time they received an electricity bill. In each treatment a household's own consumption for the past month was compared to the consumption of an average household for that same past month.

3.1.1. The social distance element of the treatments.

The element of social distance was included by incorporating two levels of social distance, distant and close, in the norms reported in the treatments. The socially distant norm reported to households was the consumption of an average household in the city last month. The socially close norm reported to households was the consumption of an average household in their neighbourhood last month. Both of these norms are descriptive in nature and so are termed descriptive norms.

It was expected at the outset that households would respond to reported descriptive consumption norms, and being compared to them, by changing their consumption towards that norm. Thus we expected that households who received feedback that they had consumed more than the average household in their neighbourhood or city would lower their consumption in the following month. Similarly, we expected that households who received feedback that they had consumed less than the average household in their neighbourhood or city would raise their consumption in the following month, the so-called boomerang effect. From an electricity conservation point of view a boomerang effect would be undesirable. It was hoped that the injunctive norm, discussed in section 1.2 would combat the boomerang effect.

In terms of varying social distance we would expect that a household would respond with a much stronger change in consumption when compared to an average household in its neighbourhood than an average household in the city. We thus expect that intensity of consumption response would be a negative function of social distance.

In terms of the monthly figures reported to treated households as average consumption figures for households in various neighbourhoods and the city as a whole, 2009 consumption figures were used. There were two reasons that "live", 2010 figures were not used.

First, calculating a live electricity consumption average for households in the city or a neighbourhood was untenable since the City's systems were not set up to calculate such averages. To set up an automated process to do this within the system would have entailed an expenditure that was well beyond the budget for this research project.

Second, it was likely that any live averages calculated and reported, particularly at the neighbourhood level would have been biased. This is likely to have been the case for two reasons. First, since the data on household consumption that is used to generate electricity bills to households is collected by a meter reading field force that may take more than one day to collect a neighbourhood's data, data may come in over two or more days. Second, all data collected by the field force goes through a checking process before its data is loaded onto the city's billing database and a bill is generated. Data from households from the same neighbourhood can leave this process at different times, which could lead to an artificially changing neighbourhood (and city) consumption average with each day, resulting in an average being reported to households that that represented only part of their neighbourhood, or city. The reasons that a household's electricity consumption data may leave this data checking process at a different time to data for another household from the same neighbourhood has to do with the manner in which the checking is conducted. Once data has

been collected by the meter reading field force and uploaded onto the city's electricity data base (as opposed to the billing database) it goes through an initial, automated, checking process. If any doubt is raised about the veracity of the consumption data for a household in this automated process, that household's data is passed on to a person who manually conducts the appropriate queries for this household. Depending on the reason for which a household's data needs to be manually checked it may spend a longer or shorter time in this process before being released onto the city's billing database.

Since reporting a live average to households was not tenable it was decided that the monthly average household consumption figures from 2009 for neighbourhoods and the city as a whole would be used instead. This data was provided with the kind assistance of the City of Cape Town municipality.

3.1.2. Injunctive norms.

In order to test the influence of an injunctive norm each of the socially distant and socially close treatments were split into two further categories of treatment. In one category only the descriptive norms were reported while in the other the descriptive norms were reported along with injunctive norms.

As a result, a total of four unique treatments were designed and administered to households in this field experiment: city level descriptive norms only (CD), city level descriptive norms augmented with injunctive norms (CDI), neighbourhood level descriptive norms only (ND), neighbourhood level descriptive norms augmented with injunctive norms (NDI). It was expected at the outset that the boomerang effect would be present amongst the descriptive-only treatments and absent amongst treatments incorporating an injunctive norm.

The injunctive norms used in this study attempted to convey social approval and social disapproval to households, conditional upon whether the household was consuming more or less than the reported average. In both cases the injunctive norm took the form of a stylised facial emoticon. If a household was consuming more than or equal to the average and was in a treatment that utilised injunctive norms the household received a socially disapproving injunctive norm. The socially disapproving injunctive norm was the following frowning emoticon ☹ in addition to the descriptive feedback. If a household was reported as consuming less than the reported average then it received the following smiling emoticon ☺ in addition to the descriptive feedback.

It was expected that conveying social approval to households, via an approving injunctive norm, for consuming less than the average reported to them would sustain them in this lower-than-average

consumption and mitigate against the boomerang effect. In terms of households reported as having consumed more than average it was expected that including an injunctive norm that conveyed social disapproval would work with the tendency of households to reduce their consumption towards the norm; it was uncertain to what degree the injunctive norm would strengthen this tendency.

3.1.3. Other information included with the feedback to households

In addition to comparing the household's consumption for the past month to either the average household in the city or their neighbourhood and the injunctive norms which half the treatment sample was selected to receive, households also received electricity savings tips with the treatments.

The same tips were given to all households and accompanied all of the treatments. These tips were designed by the City of Cape Town and targeted what were thought by them to generally be the largest areas of electricity use in households.[51]

The total number of treatments, along with the information provided in each one can be summarised in the table below.

Fig. 3.1. Information provided to households in each treatment

[51] The tips which appeared at the bottom of insert were:

Tips that may help you to use less electricity:

Use less hot water. If you can, take a shower rather than a bath. Use as little hot water as possible whether showering or bathing. Stack your dishes and wash them in one large load, rather than washing many smaller loads. If you go away for the weekend, switch your geyser off. The recommended temperature to set your geyser's thermostat to is 55^{o}C. Check what your thermostat is set to. Please ask a professional for help if you are not sure how to do this.

Use your big appliances less often. Hang up washing to dry as tumble dryers use a very large amount of electricity. Use stove plates and ovens for only as long as required to cook your food. Cook on as low a heat a possible. It is cheaper to boil water in a kettle than on a stove. Once water is boiling, only a low heat is required on a stove to keep it boiling. Boil only the water you need.

Conserve Heat: Closing the curtains helps greatly in keeping a room warm. Close the curtains in the early evening. Use heaters and air conditioners only when you are in the room and be sure to keep the room's doors and curtains closed as this helps to retain the heat. Whenever possible and convenient, use an extra pullover or a blanket instead of a heater.

	Social distance of descriptive norm		Injunctive norm	Electricity savings tips
Treatment	City	Neighbourhood		
CD	Yes			Yes
CDI	Yes		Yes	Yes
ND		Yes		Yes
NDI		Yes	Yes	Yes

Giving these tips to the treated households allows increased comparability with the growing literature on the use of social norms in managing energy consumption since most of the interventions used, so far, employ social norms in conjunction with information (Alcott, 2011). This does however introduce a confound if we are to try and identify the influence of norms, as a whole, versus the situation in the control group (only receiving electricity bills). Thus, when comparing our normative treatments to the control we will be unable to identify the effect of a norm separately to that of information.

With respect to identifying the effect of norms compared to a control group this is a methodological weakness, however, we decided that it was worthwhile to bundle the norm with information for two reasons. First, it allows greater comparability with similar studies in the first world who employ similar strategies. Second, within the fieldwork budget we had to decide which elements of the norms literature to focus upon. In this case the primary interest was in the role played by social distance and by injunctive norms. To this end we needed four treatments that differed systematically in these respects alone. Once the fieldwork was complete we could then compare these treatments to identify effects. Since an important and unobserved way in which households could differ from each other could be in their access to electricity saving information we decided that it would be prudent to provide all actively treated households with the information required to make the larger changes in consumption behaviour.

3.1.4. Format of the feedback given to households.

All households receiving one of the treatments are households that receive their electricity bills through the post (the reason for this is discussed in section 2 of this chapter). Since households receive their bills through the post, the treatments in this natural field experiment were

administered by means of a purpose designed insert included in the same envelope with the normal electricity bill and posted with it.

The inserts which constitute the treatments are all A4 size pages. Each insert is divided into two halves, a top and a bottom half. The bottom section of the insert contains the electricity saving tips just discussed and is constant across all treated households. The top half of the insert is the actively varying portion of the insert, containing information specific to each household's consumption and varying according to the specific treatment that the household has been assigned to. The inserts are printed in black and white only. Examples of full inserts are included in the appendix for this chapter.

Each insert is written in the language of the household. The language of the household is taken as the household language indicated on the City of Cape Town's database. The indicated household language is the same language in which the household receives their bills form the city. However, in the entire sample only English and Afrikaans were indicated as household languages. This casts some doubt upon the accuracy of this information in the City of Cape Town's contract account database, since a significant number of the households in the treatment samples were drawn from the townships of Gugulethu and Langa where the most common first languages are Xhosa and isiZulu. If households received feedback in a language that they struggled to understand it may have jeopardised the veracity of their reaction to the treatments.

3.2. Fieldwork and fieldwork sample.

This next section discusses how a sample for the experiment was selected and the process of executing the experimental treatment process. The sample aimed to reflect the socioeconomic breadth of the City of Cape Town and control for various potential sources of bias between treatments. The fieldwork process itself ran from early January 2010 to the end of December 2010. It involved matching the various treatment inserts to the electricity bills for the relevant households. This was done at the company responsible for printing the electricity bills for the city of Cape Town. A checking team of university students, employed by this project, checked the matched bills and inserts at the printers and identified errors where they occurred, so that they could be corrected. A major error in the matching process occurred in the middle of the year (August) which resulted in a large number of electricity accounts being matched to the incorrect inserts. Thankfully this matching error was detected by the students engaged in checking at the printers and the incorrectly matched accounts were not sent out. Upon investigation it was found that the mismatching arose from a change that had been made at the beginning of August to the computer code that matched the inserts with the electricity accounts. Since correcting the code would take some time, but

households still needed to receive their electricity accounts on time and could not be held back while the code was corrected, a large number of households did not receive their treatments during August. These households instead were mailed their electricity bills without any inserts while the computer code was corrected.

3.2.1. Selecting the fieldwork sample.

In selecting the fieldwork sample there were several concerns and challenges to be met. The first concern was that the study reflect, as much as possible, the socioeconomic breadth of the City of Cape Town. For this reason the likely relative income levels of households were estimated and this information used to guide the initial phase of sample selection. The next concern was to identify suburbs that had sufficient numbers of households consuming electricity on a credit basis. Once this had been achieved it remained to allocate households in an unbiased fashion to each treatment. The next three sub-sections elaborate on this process of selecting the fieldwork sample.

3.2.1.1. Income groups.

Different socioeconomic groups within the city of Cape Town were identified and included in the treatment and control sample in order to assess if they would react to the various treatments in different ways.

Since the City of Cape Town does not collect household level demographic data for its electricity customers we could not select households based upon their own level of income. Instead a proxy for income was used. This proxy was the general income of the suburb in which the household was found as determined by the most recent suburb-level income data for the city of Cape Town, the 2001 Census.

Using this Census' data, households were ranked from poorest to richest in order to create an income distribution of households within the city of Cape Town. Based on their position along the distribution each household was ascribed a percentile value between 1 and 100. Households were then organised into suburbs. Suburbs were then examined to see where the greatest proportion of their households' income lay in terms of their percentile position on the city's income distribution. If the greatest proportion of their households' income lay clearly in the lowest 33 percentiles then that suburb was considered a lower income suburb, if the greatest proportion of their households' income lay clearly in the middle 33 percentiles then that suburb was considered a middle income suburb and similarly if the greatest proportion of their households' income lay clearly in the upper 33 percentiles then that suburb was considered an upper income suburb.

3.2.1.2. Selecting suburbs in each income group that had sufficient households consuming electricity on a credit basis.

Once suburbs had been identified in each of the major income categories these suburbs were then examined in terms of how many of the households in them consumed their electricity on credit terms and how many consumed their electricity on prepaid terms. Since difference between these two sorts of households was an important basis for selecting households into the sample for this study it is necessary to explain, briefly the difference between the two.

Households consuming their electricity on the basis of credit are households who pay for their electricity once they have consumed it, they usually are billed every month by the city of Cape Town and receive these bills in the post. For the most part these households are billed for their electricity consumption, along with their water and other rates once a month. Electricity supply to these households would be cut off only after serial non-payment.

Households who consume electricity on prepaid basis will have a prepaid meter in their dwelling which governs whether electricity flows into the home or not installed in their dwelling dependant on whether the household has a positive balance of credits or not. Households must purchase encoded prepaid electricity credits in order to load them onto their prepaid meter. Purchasing of these credits may be done online at various websites, in stores or via cell-phones. There is enormous variation in the manner in which the encoded prepaid credits are ultimately presented on electronic or paper receipts to prepaid customers. Given these conditions, in order to present this study's treatments in a sensible and reasonably consistent fashion to households consuming electricity on a prepaid basis, a vast amount of programming work across several platforms would have been required. There is also an substantial variation in the frequency with which households purchase prepaid electricity, some households may purchase many times a month, some once a month, while others will purchase in bulk once every few months. This introduces significant practical difficulties into calculating either a household's monthly consumption or an average consumption over the period since the household's last purchase to allow a comparison to a norm.

Given the significant complexities that would need to have been overcome, and the cost of overcoming them, in order to administer this study's treatments to households consuming electricity on a prepaid basis, it was decided to only include households that consumed electricity on a credit basis in the sample.

As a result of this decision suburbs that had more households consuming electricity on a credit basis were preferred over suburbs with fewer such households on the basis that households with more

such households would most likely provide a sufficiently large sample for analysis. On this basis the following suburbs were selected in each income category.

3.2.1.3. Allocating households from each suburb to treatments in an unbiased fashion.

Once suburbs had been identified that had sufficient numbers of households consuming electricity on a credit basis it remained for those households to be allocated to each of the four treatments in a manner that was not likely to introduce bias. This was achieved in a two-step fashion.

This study primarily set out to compare the effects of the four different active treatments (CD, CDI, ND and NDI). Specifically it was hoped that the effect of social distance could be isolated (by taking differences between treatments in the following fashion {[ND+NDI] -[CD+CDI]}; ND - CD and NDI - CDI) and the effect of injunctive norms also isolated (by taking differences between treatments in the following fashion {[NDI + CDI] - [ND + CD]}; CDI - CD and NDI - ND) in the Capetonian context.

Comparison to the untreated households are also of obvious interest and are thus, where possible, made. Comparison to the untreated control group was not initially a goal of this study. Rather the control group was composed after households had been allocated to the printed insert treatments. As a result the control group was not selected in entirely the same random manner as the actively treated groups. This means that when comparing the untreated control group to the treated groups more conventional econometric techniques are preferred to simpler measures such as a difference of means.

The untreated control group was selected from households which had not been allocated to receive the printed inserts but were from the same suburbs as the households selected to be actively treated were the households which came to constitute the untreated control group. This led to a bias in pre treatment consumption between the untreated control and treated groups. There were two very practical reasons for this. Some suburbs, particularly the lower income suburbs, did not have sufficient households to allow for five treatments in meaningful numbers, some, in fact were totally exhausted by the allocation to the printed treatments. As a result suburbs are represented in different proportions in the untreated control group compared to the untreated group. When households were selected from a suburb into a treatment, this was done according to a consumption band within that suburb (to maintain balance in terms of pre-treatment consumption .

When the untreated control was selected, all households in the suburb were selected[52] into it in order to maintain a sufficient sample size. However since the untreated group would come from different consumption bands in different portions, this led to bias in levels of pre treatment consumption. Households were allocated to each of the four active treatment types in a random manner that was not likely to introduce bias. This was achieved in a two-step fashion. First, households were ranked according to their average daily electricity consumption within their neighbourhoods[53]; only households with true readings[54] were used and outliers were excluded from the ranking. Second, within each neighbourhood ranking, households were grouped into groups of four in a contiguous fashion, that is, the four households with the highest consumption would be in one group of four, the four households with the next four highest levels of consumption would be in the second group of four, and so on. Various groups of four were selected at random and each household within them uniquely assigned to one of the study's treatments. In this manner we attempted to avoid bias in the treatment sample.

It was important that a sufficiently large number of households be ascribed to separate treatments in a manner that minimised the difference in pre-treatment levels of consumption across treatments for two reasons. First, pre-treatment levels of consumption may be a proxy for a number of household characteristics that are important determinants of future consumption but are unobservable; these could include income, efficiency of energy use and household size. Allocating across the four treatments a large number of households who are close in terms of consumption would help minimise the chance of bias between treatments in terms of these unobserved characteristics. Second, previous consumption levels may be predictors of future levels of consumption (due to the just-mentioned unobserved characteristics) and as such a bias in terms of initial pre-treatment consumption levels.

The sampling procedure just described led to a sample across the four treatments at the beginning of the study that looked as in table 2.1.3 below

[52] Although households that are defined as outliers according to the criteria in the "Bias" section below are excluded from analysis (as was the case with households in the treated group).
[53] The definition of a neighbourhood for this study was what is known in the City of Cape Town as a Meter Reading Unit (MRU). An MRU is a small area of households that is read by one of the meter-reading teams within the city of Cape Town in a day. A suburb may contain several meter reading units.
[54] As will be discussed in the estimation sample section, households may have their electricity consumption estimated, and be billed on that basis, as opposed to actually read off of their electricity meters.

Table3.2.3.1. Sample selected for treatment: numbers of households in each treatment by suburb.							
	CD	CDI	ND	NDI	Control	All mailed households	Total households
Total	**1423**	**1378**	**1411**	**1415**	**5510**	**5627**	**11137**
Upper Income	***603***	***577***	***569***	***590***	***3060***	***2339***	***5399***
Constantia	157	161	156	162	203	636	839
Durbanville	182	176	174	173	2105	705	2810
Pinelands	161	154	149	162	292	626	918
Tamboerskloof	103	86	90	93	460	372	832
Middle Income	***513***	***505***	***539***	***519***	***1769***	**2076**	***3845***
Athlone	139	149	148	138	662	574	1236
Bonthewul	88	77	104	102	200	371	571
Grassy Park	156	151	159	152	287	618	905
Heideveld	60	64	56	59	136	239	375
Salt River	70	64	72	68	484	274	758
Lower Income	***307***	***296***	***303***	***306***	***681***	**1212**	***1893***
Gugulethu	163	150	154	160	295	627	922
Langa	54	52	55	54	158	215	373
Mannenberg	90	94	94	92	228	370	598

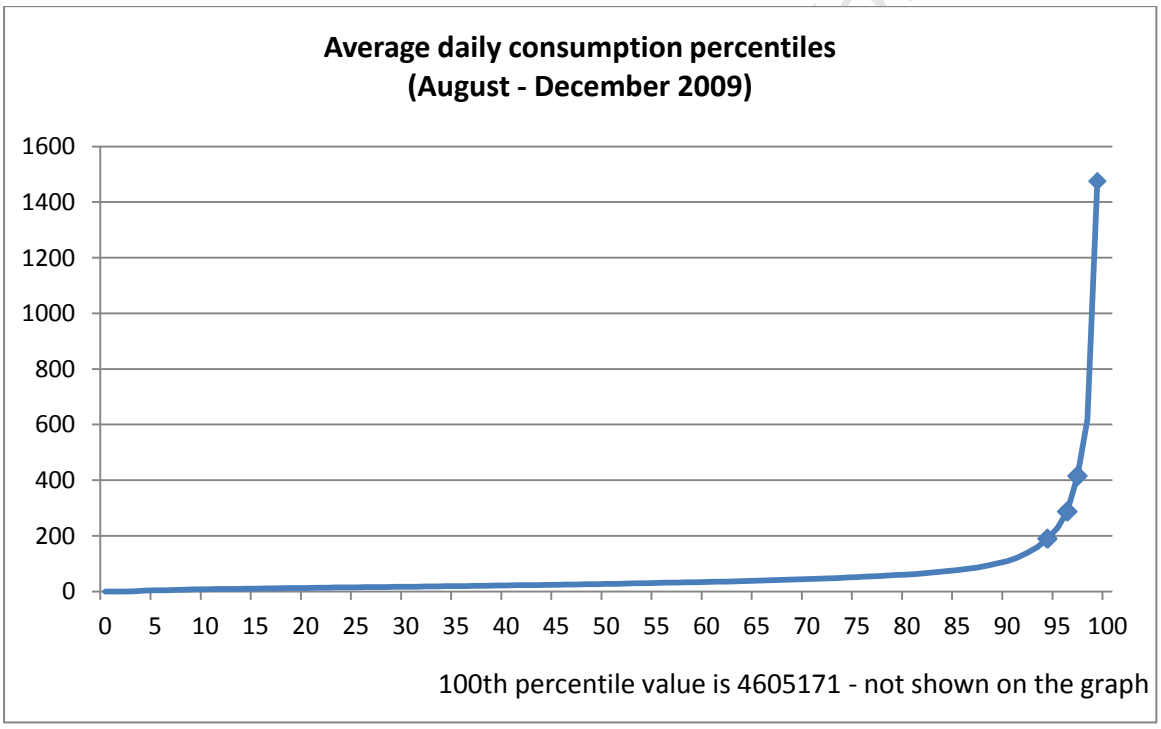
Bias:

In any experimental design it is important to demonstrate that randomisation that resulted in an allocation of households that is unbiased across treatments. Using the data which was available on electricity consumption, August - December 2009, the average daily electricity consumption for the period 1 August 2009 to 31 December 2009 was calculated. The median value and mean values for

this measure of consumption is shown for the pre-treatment sample in the table below. The means were compared between treatments (pairwise and using a t-test) and no significant differences, at the 5% level, were found between the various treatment cells.

Two things stand out from the table however. One is the large degree of heterogeneity found in the data - all standard deviations are quite large with respect to the means. The second is that the median consumption for all groups is very similar while the means seem to vary substantially, but apparently not significantly.

Looking at the average consumption percentiles we see that the top quintile appears to be made up of households with consumption that is significantly different to the rest of the households. A significant contributor to the observed heterogeneity appears to be these households with abnormally large consumption. At the lower end of the consumption values are a significant number



of households with consumption at zero or very low. In order to remove the influence of very low and abnormally high consumption we also consider households that recorded average daily consumption that was below 200kwh (excluding roughly the top quintile) and those that consumed less than 3kwh on average each day (excluding roughly the bottom quintile). For this constrained sample we obtain a sample with a much reduced, although still large, level of heterogeneity. The means of pre-treatment consumption for each treatment cell (including the control) obtained for this constrained sample are tested against each other in a pair-wise fashion and no significant differences are found between them. This remains true when we test the pre-treatment consumption for each treatment within each broad income level.

Table 3.2.3.2 Pre-treatment mean and median average daily consumption levels (Aug- Dec, 2009).						
		Treatments				
		CD	CDI	ND	NDI	Control
Unconstrained Total Sample						
	Median	27.86	29.29	27.35	27.64	26.04
	Mean	65.38	67.87	59.22	60.30	1726.95
	(sd)	(277.95)	(281.58)	(194.17)	(218.94)	(70935.09)
	(n)	1423	1378	1411	1415	5510
Excluding lowest and upper quintile of the sample (Constrained sample)						
Total sample	Median	26.60	28.41	26.47	26.77	26.31
	Mean	36.45	38.73	36.48	35.80	37.82
	(sd)	(31.27)	(34.28)	(31.06)	(29.31)	(35.35)
	(n)	1357	1324	1356	1355	4627
Upper Income	Median	42.66	45.93	42.73	42.95	29.83
	Mean	52.74	56.55	53.61	51.50	38.83
	(sd)	(38.27)	(40.86)	(37.92)	(34.27)	(32.19)
	(n)	562	546	530	548	2736
Middle income	Median	23.55	23.47	23.42	23.28	24.05
	Mean	28.17	29.05	27.86	27.709	38.92
	(sd)	(18.28)	(20.95)	(18.62)	(18.22)	(40.96)

	(n)	498	491	531	508	1349
Lower Income	Median	16.43	15.76	16.62	15.94	18.75
	Mean	19.48	21.41	21.21	20.78	29.95
	(sd)	(15.36)	(20.85)	(18.35)	(19.80)	(34.64)
	(n)	297	287	295	299	542

3.2.1.4. Power calculations.

Power calculations were made in order to determine the level of statistical power that would be achieved at the conventional 5% confidence levels. Since all of the results would essentially be comparisons made between two groups (either different active treatments or an active treatment to the untreated control group)[55] this calculation was based on the formula for the comparison of two means set out below.

$$n_1 = \frac{\left(\sigma_1 + \frac{\sigma_2}{r}\right)\left(z_{1-\alpha/2} + z_{1-\beta}\right)^2}{(\mu_1 - \mu_2)^2}$$

which we can rearrange to give us a calculation in terms of power in the following way

$$\sqrt{\frac{n_1(\mu_1 - \mu_2)}{(\sigma_1 + \frac{\sigma_2}{r})}} - z_{1-\alpha/2} = z_{1-\beta}$$

n_2 is the control group's sample size while n_1 is the treated group's sample size which we are solving for. Similarly σ_2 denotes the estimated likely variance for the control sample and σ_1 for the treated sample, while μ_2 denotes the mean proportional consumption for the control group and μ_1 for the treated group. As will be discussed later, consumption is measured as the household's daily average consumption for a particular month as a proportion of their daily average consumption over August - December 2009. $z_{1-\alpha/2}$ is the confidence level for a conventional confidence level "α" of 5% when large numbers are involved (1.96) and $z_{1-\beta}$ is the critical statistic for the level of power desired, here 80%, when large numbers are involved (1.29).

A challenge in calculating power was that while data would be generated monthly over a full year, 2010, data upon which power calculations could be made was only available from August 2009 to December 2009 at the time of the calculations. Although not ideal since electricity consumption is expected to be seasonal and symmetric, drawing our power calculation from data for the last half of the year, does probably capture the likely variation over an entire year. It is certainly better than no power calculation.

[55] Although households were to be measured each month, since they were to be mailed each month also, this rendered each successive after a mailing month a different treatment to the one before it since we would be observing the effect, potentially, of two different amounts of mailings received.

Estimates for the mean and variance in consumption for each treatment were drawn from this period). Power calculations are conducted for the typical variance of a typical month. Power calculations are conducted for several different effect sizes including the expected treatment effect of 2.5% observed in the OPower data by Allcott (2011).

In terms of comparison at the overall sample level between a treatment and the control group this experiment is not expected to achieve a power of 80%, the conventional level of statistical power for these exercises (Ferraro and Price, 2011), at the 5% level to detect most reasonably expected effect sizes. This is particularly true for comparisons amongst the lower income groups and between treatments. As a result this study will be cautious in its interpretation of results for these comparisons.

Power calculations.

Groups being compared to the control group			Effect size					
Control group (sample size)	Compared to (assumed individual mailed treatments within the...	minimum sample size of one mailed treatment cell)	1%	2%	2.5%	3%	4%	5%
(n=5510)	Total sample	(n=1397)	0.06	0.13	0.15	0.21	0.32	0.44
(n=3060)	Upper income	(n=569)	0.06	0.07	0.09	0.11	0.15	0.20
(n=1769)	Middle income	(n=505)	0.06	0.09	0.12	0.15	0.21	0.29
(n=681)	Lower income	(n=296)	0.05	0.06	0.07	0.07	0.09	0.13
Comparisons between individual cells of treated groups								
	Total sample	(n=1397)	0.06	0.09	0.12	0.14	0.21	0.30
	Upper income	(n=569)	0.06	0.08	0.09	0.11	0.15	0.20
	Middle income	(n=505)	0.06	0.08	0.10	0.12	0.16	0.22

Lower income	(n=296)	0.05	0.06	0.06	0.07	0.08	0.10

3.2.2. Fieldwork.

Households were mailed as part of this field experiment from the beginning of January 2010 until the end of December 2010.

During the fieldwork period one of the primary challenges was to get the correct feedback to the correct households as determined by the sample framework. In order to do this a number of possibilities were explored but, in the end, it was found that to effect this process at the company who would print the inserts (as well as the city's electricity and rates bills) was likely to work best. As a result, over the fieldwork period, the inserts were printed out at the printing company responsible for printing and mailing the electricity and rates bills for the city of Cape Town[56]. The printing company was able to automatically match the inserts and bills so that the insert bearing a particular household's information was mailed with that household's electricity bill.

Once the inserts had been printed, matched to their bills and placed together with those bills into unsealed envelopes ready for sealing and mailing a team of four university students checked a sample of the batches that had been prepared for mailing. At the printers, the sum total of envelopes to be mailed were split into several batches, with each batch being normally 50 envelopes or less. For every batch, the two top and the two bottom envelopes were checked as well as two further envelopes pulled at random from the rest of the batch. In this way generally at least 10% of each batch was checked. If a mismatched insert was discovered in a sample, then the whole batch from which it was taken was checked to see if mismatches had occurred systematically. Where systematic mismatches were discovered, The printing company re-matched the bills and inserts, which were then checked again. Mismatches were discovered only very seldom, four times in the whole year, which speaks strongly about the accuracy of the matching process.

Unfortunately not all households in the sample were mailed each month, there were a variety of reasons for this, sometimes a household was not mailed its insert if the city wished to include another insert with its bill, such as when rates rebates are due or some other communication specific to the household needed to be made. In these circumstances many such households' bills would be diverted to relevant city insert procedures and would skip this study's insert procedure,

[56] Lithotech.

ultimately being mailed that month without its treatment insert. On another occasion, in the middle of the year, systematic printing and matching errors, unanticipated at the outset, were discovered when the bills and their matched inserts entered the checking process[57]. Since correcting these errors required roughly a week, it was decided to mail households their electricity bills without their intended inserts so that these households could receive their bills in reasonable time, rather than suffer complications arising from electricity bills unduly delayed while a solution was sought!

Households could also drop out of the treatment sample if, during the fieldwork period they changed from consuming electricity on a credit basis to consuming electricity on a prepaid basis.

3.3. Estimation sample.

Once the fieldwork was complete and data was available for estimation, it needed to be decided who to include in the estimation sample. There were a number of criteria for inclusion in the estimation sample which sought to ensure that the consumption figures measured were true readings and not estimations, that households and not businesses were included, that only households with consistent membership were likely to be included and that there would be as little bias as possible between the control and treatment samples in terms of initial consumption. We will discuss these now.

3.3.1. Consumption outliers.

Households with consumption that was very different to the consumption found in the rest of the sample, were excluded from analysis. These were households typically in the top or bottom 5 percentiles of the consumption distribution. They were excluded since their large consumption values severely biased estimates of the average treatment effect.

Specifically this meant that households that in any month consumed more than 5 times their daily average for the pre-treatment period were excluded. Households that consumed more than 200 kWh or less than 3 kWh per day on average in the pre-treatment period or the treatment period fell firmly into the tails of the consumption distribution and were excluded. Another bias that needed to be controlled for was households that were unusually distant from the norm reported to the in terms of consumption. The great differences between these households and the norm bias the estimates of the average treatment effects estimated upon the interaction between the specific treatment and the difference between the household's consumption and the reported norm.

[57] The consumption figures on the electricity bills were suddenly quite different from that on the inserts.

The estimation sample for each month in the treatment period is expressed in terms of the pre-treatment consumption in table 3.3.2 below. The pre-treatment sample displays decreased bias in terms of the untreated control group that was selected after fieldwork began although the bias is still significant. The unmailed control group has a significantly higher pre-treatment consumption than all the other groups in August and October and significantly higher than the NDI treatment in July. In February the unmailed control group has a significantly lower pre-treatment consumption than all treatments. In August the NDI treatment is slightly but significantly above the CD treatment in terms of pre-treatment daily average consumption of the households in each respective treatment.

Table 3.3.1. Pre-treatment average daily consumption for the each months estimation sample					
	CD	CDI	ND	NDI	UN
Feb *mean*	37.90272	37.75656	37.78975	37.26859	32.95459
(sd)	(36.75156)	(33.9276)	(36.0155)	(35.72448)	(31.20395)
n	375	359	382	384	3407
March	32.58764	32.84359	32.08038	31.59595	32.9908
	(28.68919)	(27.23796)	(27.09159)	(26.84557)	(31.22723)
	922	907	885	945	3399
April	32.82002	33.48453	32.25187	32.17043	32.95218
	(29.57751)	(27.89212)	(28.1171)	(27.83278)	(31.20472)
	796	783	792	822	3414
May	33.74475	34.02719	33.15854	33.53999	33.47724
	(29.10667)	(29.01151)	(29.20035)	(29.24205)	(31.39605)
	709	689	689	725	3328
June	33.71587	33.9394	33.21606	32.29029	34.40461
	(31.92697)	(29.28075)	(30.54632)	(27.93835)	(32.08966)
	629	625	623	678	3105
July	33.10217	33.46353	32.18952	31.93702	34.35337
	(29.2645)	(27.43801)	(28.00637)	(27.48587)	(31.90128)
	828	804	838	880	3185
August	29.80986	28.78837	28.02435	26.60141	33.83449
	(23.84846)	(19.37722)	(20.48719)	(16.79725)	(31.65186)
	360	367	381	412	3256
September	34.2562	33.60344	32.6289	32.28966	33.87135
	(29.59548)	(26.82833)	(28.61119)	(27.34659)	(31.86076)
	803	807	811	852	3231
October	27.36757	27.18888	26.24531	25.42278	33.25476
	(22.89298)	(18.89134)	(19.44101)	(16.73259)	(31.48297)
	557	561	565	600	3369
November	33.52249	33.84031	33.731	33.52731	33.30471
	(29.28754)	(27.56606)	(29.06618)	(28.15285)	(31.4974)
	818	810	718	756	3365
December	31.47055	31.68648	31.19115	30.92846	32.74262
	(28.3757)	(26.17953)	(27.091)	(26.24784)	(31.22831)
	960	929	943	981	3457

3.4. Estimation of treatment effects: Variables of interest and their expected influence upon consumption.

The behaviour that was targeted by the treatments in this study was the consumption of electricity by households. Consumption, in particular differences in consumption between households in various treatments, is thus the measure used to assess the impact of treatments in this study. Measures of consumption will thus be the dependant variable in the analysis.

In terms of explaining the behaviour of household electricity consumption there are various variables of interest. These are the treatment-related variables, whose explanatory power with respect to consumption we wish to analyse, and the climatic variables which are likely to predict electricity consumption generally.

Before we move on to discussing these variables, a word first about what is meant by "month" in the estimation phase and about the subscript notation used.

As described in the fieldwork section, consumption for households consuming electricity on a credit basis are read at different times of the month. This affects what is meant by the term "month" in the estimation process. The term "month" does not refer here necessarily (or even most often) to a calendar month but rather to the period of time between two meter readings with the month named after the calendar month in which the second reading occurred. Thus the period that this study refers to as the January 2010 pre-treatment period may in fact contain more days from December 2009 if a household's meter is read in the first half of January 2010 than if it is read in the second half of January 2010 and yet still be called January 2010 consumption. A similar situation holds for each month's consumption.

Since the data is panel data we use subscript i $\{i|\, i \in \mathbb{N}, i \geq 1\}$ to indicate individual households[58] and the subscript t $\{t|\, t \in \mathbb{N}, 1 \leq t \leq 12\}$ to indicate month in the conventional fashion.

3.4.1. Measuring consumption.

The measure of household electricity consumption used for analysis in chapter 4 will be two similar proportional measures. The reason for this is the bias that exists between the mailed treatment cells and the control group. These two measures are termed "annual proportional consumption" and 'monthly proportional consumption". This section will briefly describe how these two measures are calculated and describe how they should improve comparison.

[58] Or put more strictly, electricity accounts.

We begin with the most simple measure of consumption which is just the absolute amount of electricity consumed by a household in a month period. We express consumption simply as: $consumption_{it}$ $\{consumption|\ consumption \in \mathbb{R}\}$. Consumption is the aggregate amount of electricity consumed by a household since it's last meter reading and is measured in kilowatt-hours (kWh).

We now calculate the households' average daily consumption on a monthly basis as $avg\_cons_{it} = \frac{consumption_{it}}{number\ of\ days_{it}}$

$number\ of\ days_{it}$ $\{number\ of\ days|numberof\ days\ \in \mathbb{Z}\}$ is the number of days since the last meter reading.[59]

Monthly proportional consumption is simply the average consumption in a month during the treatment period avg_cons_{it} divided by its average daily consumption over the August 2009 to January 2010 pre-treatment period, or

$$prop\_month\_cons_{it} = \frac{avg\_cons_{it}|_{t=February\ 2010}^{t=December\ 2010}}{\left\{\frac{\sum_{t=August\ 2009}^{t=January\ 2010} consumption_{it}}{\sum_{t=August\ 2009}^{t=January\ 2010}(number\ of\ days_{it})}\right\}}$$

Similarly annual proportional consumption is calculated as

$$prop\_annual\_cons_{it} = \frac{\left\{\frac{\sum_{t=February\ 2010}^{t=December\ 2010} consumption_{it}}{\sum_{t=February\ 2010}^{t=December\ 2010}(number\ of\ days_{it})}\right\}}{\left\{\frac{\sum_{t=August\ 2009}^{t=January\ 2010} consumption_{it}}{\sum_{t=August\ 2009}^{t=January\ 2010}(number\ of\ days_{it})}\right\}}$$

The denominator of $prop\_annual\_cons_{it}$ and $prop\_month\_cons_{it}$ is calculated in a similar method to the denominators of the other proportional measures, of climatic variables, whose calculations will be described later in this chapter. In all of these cases the denominator calculates the pre-treatment daily average of some variable of interest. In other words, the calculations are the same in all respects except for the climatic aspect being measured. In the interests of brevity then, this sort of calculation of the pre-treatment average is denoted $avg_{X_{ipretreatment}}$.

[59] The negative or undefined values of avg_cons_{it} resulting from differently signed $consumption_{it}$ and $number\ of\ days_{it}$ or a zero value for $number\ of\ days_{it}$ are possible (when meter readings are estimated) but excluded from the estimation sample by the estimation sample criteria for average household consumption described in section 3.3.2 above.

The reason for adopting a proportional measure of consumption as the basis of comparison between the treatment and control cells rather than just absolute consumption is that the control cell had a pre-treatment level of absolute consumption that was significantly different to the mailed treatments. This could well signal that households were likely to naturally continue consuming significantly different amounts of electricity (Bruhn & Mckenzie, 2009). [60] In order to eliminate this difference, it was decided to calculate a households consumption during the treatment period as a proportion of its historic pre-treatment consumption. In this we assume that households are likely to follow similar patterns of consumption through the year. This would lead to their consumption being similar, as a proportion of pre-treatment consumption, at the same time of the year if there was no other systematically different influence at work, such as an experimental treatment. Assuming this to be true, then differences in proportional consumption (annual and monthly in this paper's terminology) would indicate differences due to receiving the various experimental treatments or not.

3.4.2. Treatment variables.

Each household was assigned a treatment variable, that was constant across all months for that household, based upon which treatment it had been ascribed to. Each treatment variable assumed a value of 1 if the household was in the treatment and a value of zero if it was not. There were thus four treatment variables as below and an untreated control group.

$$CD_i = \begin{cases} 1, & if\ household\ is\ in\ city\ level\ descriptive\ norm\ only\ treatment \\ 0, & otherwise \end{cases}$$

$$CDI_i = \begin{cases} 1, & if\ household\ is\ in\ city\ level\ descriptive\ norm\ treatment\ wih\ injunctive \\ 0, & otherwise \end{cases}$$

$$ND_i = \begin{cases} 1, & if\ household\ is\ in\ neighbourhood\ level\ descriptive\ norm\ only\ treatment \\ 0, & otherwise \end{cases}$$

$$NDI_i = \begin{cases} 1, & if\ household\ is\ in\ neighbourhood\ level\ descriptive\ norm\ treatment\ with\ injunctive \\ 0, & otherwise \end{cases}$$

$$UN_i = \begin{cases} 1, & if\ household\ is\ in\ the\ control\ group \\ 0, & otherwise \end{cases}$$

From previous research, discussed in the chapter on social norms, we suspect that in addition to matters of comfort - such as being warm enough, fed enough and entertained enough - there is a further determinant of a household's electricity consumption in situations when the household is

[60] It could thus signal that these households were different in terms of their size and stock of electric appliances, which would be key determinants of electricity consumption.

explicitly compared to the consumption of an average household μ_{it} . In such circumstances we expect households to maximise utility, all else being equal, by minimising the difference between their own consumption and the reported level of average consumption. In other words

$$U = f(consumption_{it-1} - \mu_{it-1}, comfort_{it})$$

specifically

$$max(U) = min\left(consumption_{it-1} - \mu_{it-1}\right) + max\ (comfort)$$

Thus we expect a household to increase its consumption, and thus its proportional consumption, more than it otherwise would have, if it is reported as having consumed less than the average household in the city or it neighbourhood. Similarly, we expect a household to decrease its proportional consumption if it is reported as having consumed more than the average in proportion to the percentage it is above the average household in the city or it neighbourhood. In order to test this we construct variables that capture how far each household's reported consumption was above or below the average as a percentage of the household's own) consumption. Each of these variables interacts with the specific treatment type.

We thus have

$$above_{i\,t} = \begin{cases} 1, & if\ reported\ in\ insert\ as\ having\ consumed\ above\ the\ reported\ average\ in\ t \\ 0, & otherwise\ (including\ control\ households, which\ receive\ no\ report\ in\ insert \end{cases}$$

$$below_{i\,t} = \begin{cases} 1, & if\ reported\ in\ insert\ as\ having\ consumed\ below\ the\ reported\ average\ in\ t \\ 0, & otherwise\ (including\ control\ households, which\ receive\ no\ report\ in\ insert) \end{cases}$$

and

$$pdiff_{i\,t} = \left| \frac{HHconsumption_{it-1} - \mu_{it-1}}{HHconsumption_{it-1}} \right| \times 100$$

which are combined as below

$$(\ below_{i\,t} \times\ pdiff_{i\,t}) = pdiff\_below_{i\,t} \quad \text{and} \quad (\ above_{i\,t} \times\ pdiff_{i\,t}) = pdiff\_above_{i\,t}$$

which interact with the treatment variables in the estimation phase.

The likely relationship of these treatment variables to the consumption variable $prop\_month\_cons_{it}$ can be thought of as follows.

If households maximise utility in the fashion described above then, where $X_{i\,t}$ is a vector of other explanatory variables that a household responds to in order to optimise comfort and e_{it} is the error term, we have

$$prop\_month\_cons_{it} = \alpha + \beta_A(pdiff\_above_{it-1}) + \beta_B(pdiff\_below_{it-1}) + \beta_1 X_{it} + e_{it}$$

wherein we expect $\beta_A < 0$ and $\beta_B > 0 \;\; \forall \;\; \mu_{it} \neq consumption_{it}$.

In order to factor in the effect of the injunctive norm (I_{it}) we must expand the utility function to include perceived social approval. If we continue to assume an additive utility function, incorporating that a household wishes to maximise perceived social approval subject to comfort and distance from the norm we have the following utility function[61]

$$max(U) = min\left(HHconsumption_{it-1} - \mu_{it-1}\right) + \max(perceived\ social\ approval) + max\ (comfort)$$

Now if $perceived\ social\ approval = f\left(HHconsumption_{it-1} - \mu_{it-1}\right)$ such that

$$I_{it} \equiv perceived\ social\ approval_{it} = \begin{matrix} 1\ if & (\mu_{it-1} - consumption_{it-1}) > 0 \\ 0\ \ if & (\mu_{it-1} - consumption_{it-1}) \leq 0 \end{matrix}$$

then when some households are receiving injunctive norms with their feedback and some are not, we have:

$$prop\_month\_cons_{it} = \alpha + \beta_A(pdiff\_above_{it-1}) + \beta_{AI}(pdiff\_above_{it-1})I_{it} + \beta_B(pdiff\_below_{it-1}) + \beta_{BI}(pdiff\_below_{it-1})I_{it} + \beta_1 X_{it} + e_{it}$$

$\beta_{BI} < \beta_B$ for households reported below the norm. For households reported as having consumed above the norm, the expected relationship between β_A and β_{AI} is unclear beyond β_{AI} , $\beta_A < 0$.

In order to assess the impact of varying social distance (city vs neighbourhood effects) upon a household's reaction to this feedback we need to expand the utility function further to include social distance as an element determining utility. Since we suspect that households are likely to respond in

[61] An additive function is used for the sake of tractability, the form does not jeopardise the generality of the point though.

direct proportion to the social distance between the reported norm[62] and the household's own consumption, we augment the utility function as follows:

$$max(U) = min[city_{it}.(\mu_{it} - consumption_{it})] + max[city_{it}.(perceived\ social\ approval)] + min[neighbour_{it}.(\mu_{it} - consumption_{it})] + max[neighbour_{it}.(perceived\ social\ approval)] + max\ (comfort)$$

where

$$neighbour_{it} = \begin{cases} 0\ if & CD\ or\ CDI = 1 \\ 1\ if & ND\ or\ NDI = 1 \end{cases}$$

and

$$city_{it} = \begin{cases} 0\ if & ND\ or\ NDI = 1 \\ 1\ if & CD\ or\ CDI = 1 \end{cases}$$

since each injunctive and non-injunctive treatment is combined separately with a neighbourhood and a city level norm we have

$$prop\_month\_cons_{it} = \alpha + \beta_{CDA}(pdiff\_above_{it-1}).CD_{it} + \beta_{CDIA}(pdiff\_above_{it-1}).CDI_{it} + \beta_{NDA}(pdiff\_above_{it-1}).ND_{it} + \beta_{NDIA}(pdiff\_above_{it-1}).NDI_{it} + \beta_{CDB}(pdiff\_below_{it-1}).CD_{it} + \beta_{CDIB}(pdiff\_below_{it-1}).CDI_{it} + \beta_{NDB}(pdiff\_below_{it-1}).ND_{it} + \beta_{NDIB}(pdiff\_below_{it-1}).NDI_{it} + \beta_1 X_{it} + e_{it}$$

$$prop\_month\_cons_{it} = \alpha_{it} + (\beta_{CDA}.CD_{it} + \beta_{CDIA}.CDI_{it} + \beta_{NDA}.ND_{it} + \beta_{NDIA}.NDI_{it})(pdiff_{above\,it-1}) + (\beta_{CDB}.CD_{it} + \beta_{CDIB}.CDI_{it} + \beta_{NDB}.ND_{it} + \beta_{NDIB}.NDI_{it})(pdiff_{below\,it-1}) + \beta_x X_{it} + e_{it}$$

From which we expect $\beta_{NDA} < \beta_{CDA} < 0$; $\beta_{NDIA} < \beta_{CDIA} < 0$ for households reported as having consumed more than the norm. For households reported as having consumed less than the norm we expect $\beta_{NDB} > \beta_{CDB} > 0$ and $\beta_{NDIB} < \beta_{CDIB}$.

In order to avoid inappropriately large amounts of multicollinearity when estimating these relationships, separate regressions are run comparing each treatment to the same control sample in the estimation phase.

3.4.3. Cumulative number of times that a household has been mailed a treatment.

[62] Both injunctive and descriptive norms.

Previous studies of how a household's energy consumption reacts to feedback about how the household's energy consumption compares to an average level of consumption found that the treatment effect did not decline as a function of how many times a household had been mailed (Alcott, 2011). This make sense for two reasons. First, the value of the reported average is new every month. Second, the behaviourally active component is the comparison itself, so long as there is a difference between household and average, we can expect the household to consume differently to what it would otherwise have done. In terms of the comparison itself, the households are not consuming more and more of the same good with each mailing but rather most are consuming different comparisons with each mailing.

There are however, for each treatment, elements beyond the actively changing comparison that remain steady. These include the electricity saving tips and the general format of the treatment. It might be that the marginal effectiveness of these unchanging elements will decline as people tire of receiving the inserts. Considering that electricity savings tips take up half of the space on each of the inserts, and are clearly labelled "Tips that may help you to use less electricity", it is very likely that households will have perceived this as the unchanging element of the treatments. These tips are also generic to each treatment. As a result, when comparisons with the control group are made, it is necessary to include a variable, in addition to the treatment-specific variables, that will be able to capture this element that is shared between treatments.

It will be interesting to see if households react to cumulative mailings in a manner consistent with declining effectiveness of successive mailings or the opposite, or if they react at all.

In order to capture this effect we thus have, interacting with each treatment for each household, a variable, $cumu\_mail_{it}$, that measures how many times a household has been mailed at any time t

$$\{cumu\_mail_{it} \mid cumu\_mail_{it} \in \mathbb{Z},\ 0 \leq cumu\_mail_{it} \leq 11\}$$

Consistent with the above thought, we assume an interaction of the following nature[63] in terms of our utility function

$$max(U) = min(\mu_{it} - consumption_{it}) + max(perceived\ social\ approval) + max(comfort) \mp cumu\_mail_{it}$$

Thus, if we gather the various other treatment-specific terms into T_{it} [64]we can express the electricity consumption function as

[63] Again the broader point that the marginal utility of acting to close the gap reported on the treatment inserts is declining with the number of inserts received is dependent upon the function capturing this in a declining, monotonic fashion rather than the specific way it does this.

$$prop\_month\_cons_{it} = \alpha + \beta_T T_{it} + \beta_{cumu}(cumu\_mail_{it}) + \beta_x X_{it} + e_{it}$$

It is difficult to say *a priori* what the nature of the difference in the effect of cumulative mailing of the various treatments will be, if any.

3.4.4. Climatic variables.

Climate will likely have a significant influence upon the electricity consumption of households. There are three particular climatic elements which are likely to influence household electricity consumption, these are: the amount of daylight minutes each day, the amount of rainfall and temperature. In order to preserve scaling with the proportional consumption variable, the climactic variables are also all average daily values for each month expressed as proportions of January 2010 pre-treatment levels. Information for all climatic variables is only available at a city level. Values for climatic variables at the household level are determined on the basis of climatic measurements done between the two meter reading dates that delineate a month's consumption for a household.

Temperature is measured in terms of heating degree days (HDD) and cooling degree days (CDD). Heating and cooling degree days are calculated as the number of degrees Fahrenheit that the temperature is below or above balance point temperature. Balance point temperature is defined as the temperature at which temperature-sensitive energy demand is minimised (Ruth and Lin, 2006). Deviations from balance point temperature require increased energy use in order to bring the household temperature back to balance point. Every degree Fahrenheit below balance point temperature requires one HDD of energy to raise the temperature of the home back to balance point temperature. Similarly every degree Fahrenheit above balance point temperature requires a CDD of energy to cool the home back to balance point temperature. Most studies use a balance point temperature of 65°F (Ruth & Lin, 2006). In order to maintain comparability with these studies, 65°F is also used as the balance point temperature here.[65]

[64] $T_{it} = \{ (pdiff\_above_{it-1}).CD_{it}, (pdiff\_above_{it-1}).CDI_{it}, (pdiff\_above_{it-1}).ND_{it}, (pdiff\_above_{it-1}).NDI_{it}, (pdiff\_below_{it-1}).CD_{it}, (pdiff\_below_{it-1}).CDI_{it}, (pdiff\_below_{it-1}).ND_{it}, (pdiff\_below_{it-1}).NDI_{it} \}$

[65] Although in many studies balance point temperature is assumed to be 65°F (Ruth & Lin, 2006) this is a somewhat arbitrary value since the actual balance point temperature would depend upon many factors which would vary between households. These would be factors such as: size of house, insulation, efficiency of refrigerators, geographic situation with respect to aspect. There is also the consideration of variation in the temperatures which household members find comfortable. Apart from varying between individual people, there is some evidence that these "comfort" temperatures vary between countries (Nicol & Humphreys, 2002). Ruth and Lin (2006) also identify variation across states within the USA. There are however no clear indications of what might constitute a reasonable balance point temperature in Cape Town, which is another reason why 65°F is used.

As mentioned, the specific variable used as a measure of temperature over a particular month in this study is the average daily heating/cooling degree days for a household that month as a proportion of the average daily heating/cooling degree days for that household in the January pre-treatment period. The calculations for average daily heating/cooling degree days and their expression as a proportion of January pre-treatment levels are as below.

$$avg\_HDD_{it} = \frac{HDD_{it}}{number\ of\ days_{it}} \quad ; \qquad avg\_CDD_{it} = \frac{CDD_{it}}{number\ of\ days_{it}}$$

thus

$$prop_{month_{it}} = \frac{avg_{HDD\,it}}{avg_{HDD\,ipretreatment}} \quad ;$$

$$prop\_month\_CDD_{it} = \frac{avg\_CDD_{it}}{avg\_CDD_{ipretreatment}}$$

with $\{prop\_jan\_HDD | prop\_month\_HDD \in \mathbb{R},\ 0 \leq prop\_jan\_HDD\}$

and $\{prop\_jan\_CDD | prop\_month\_CDD \in \mathbb{R},\ 0 \leq prop\_jan\_CDD\}$

For rainfall, a similar procedure converts the average daily millimetres of rainfall for a month to that average expressed as a proportion of the average daily millimetres of rainfall in January 2010.

We thus have:

$$avg\_rain_{it} = \frac{mm\ of\ rainfall_{it}}{number\ of\ days_{it}}$$

and from this

$$prop\_month\_rain_{it} = \frac{avg\_rain_{it}}{avg\_rain_{ipretreatment}}$$

with $\{prop\_jan\_rain | prop\_jan\_rain \in \mathbb{R},\ 0 \leq prop\_jan\_rain\ \}$

The amount of daylight in a day was measured as the minutes of civil twilight[66] in a day. Again, for each household in each month, we calculate the average daily minutes of daylight for a month and

[66] Civil Twilight is the period when the sun is no more than 6 degrees below the horizon at either sunrise or sunset. The horizon should be clearly defined and the brightest stars should be visible under good atmospheric

from these averages then express each month's average daily minutes of daylight as a proportion of January 2010's average daily minutes of daylight as below.

$$avg\_daylight_{it} = \frac{daylight_{it}}{number\ of\ days_{it}}$$

and from this

$$prop\_month\_daylight_{it} = \frac{daylight_{it}}{avg\_rain_{ipretreatment}}$$

with $\{prop\_jan\_daylight | prop\_jan\_daylight \in \mathbb{R},\ 0 \leq prop\_jan\_daylight\ \}$.

In terms of our assumed utility function:

$$max(U) = min(\mu_{it} - consumption_{it}) + \max(perceived\ social\ approval) + max(comfort) - cumu\_mail_{it}$$

the $comfort$ term captures the sense of general material well-being that a person or a household wishes to feel, that is, the amount of entertainment, food, warmth, washed and ready clothing which maximises their utility, all else being equal. As such $comfort$ is a function, directly and indirectly, of the climatic variables as indicated below. Again, for the sake of tractability, the relationship between comfort and the climatic variables is expressed additively. Nevertheless, so long as the parameters express a monotonic relationship in the indicated direction between the parameter's climatic variable and $comfort$, other functional forms are equivalent.

$$comfort_{it} = \omega_{HDD} HDD_{it} + \omega_{CDD} CDD_{it} + \omega_{daylight} daylight_{it} + \omega_{rainfall} rainfall_{it}$$

with ω_{HDD} , ω_{CDD} , $\omega_{rainfall} < 0$ and $\omega_{daylight} > 0$.

we expand the $\beta_1 X_{it}$ term in our consumption function to capture this relationship

$$prop\_month\_cons_{it} = \alpha + \beta_T T_{it} + \beta_{cumu}(cumu\_mail_{it}) + \beta_1 X_{it} + e_{it}$$

thus becomes

$$prop\_jan\_cons_{it} = \alpha + \beta_T T_{it} + \beta_{cumu}(cumu\_mail_{it}) + \beta_{HDD} prop\_jan\_HDD_{it} + \beta_{CDD} prop\_jan\_CDD_{it} + \beta_{daylight} prop\_jan\_daylight_{it} + \beta_{rainfall} prop\_jan\_rainfall_{it} + e_{it}$$

wherein we expect β_{HDD} , β_{CDD} , $\beta_{rainfall} < 0$ and $\beta_{daylight} > 0$.

---

conditions (i.e. no moonlight, or other lights). One still should be able to continue with ordinary outdoor activities.

3.5. Estimation of the treatment effects: Estimation techniques used.

Two methods were used in order to assess which, if any, of the various treatments were likely to have affected households' consumption. The two methods used were the sign test and random effect regressions. The sign test is a non-parametric measure and the random effect regressions are parametric measures. In both cases the treatment effect was assessed by comparing the treated sample to the control sample.

In all cases we wish to see whether the relative socioeconomic level of a household's suburb would modify any treatment effect and so analysis is split by these socioeconomic groups. Given that we expect households to respond to being compared to an average level of consumption by changing their consumption towards that average, analysis always takes account of whether a household was reported as above or below the average.

3.5.1 Fixed effect versus Random effect regressions.

Since households were mailed treatments on a monthly basis and we have data on consumption for those months, it is appropriate that we estimate treatment effects using a panel data method. .

Table 3.5.1.Within and Between variation for variables of interest.

Variable		Mean	Std. Dev.	Min	Max	Observations
prop_precons_da	overall	1.118844	1.140136	1.13E-05	4.999924	N = 71010
	between		1.063034	0.00166	4.967099	n = 7582
	within		0.535276	-2.96572	5.336827	T = 9.3656
CD_pdiff_norm_above	overall	0.820741	6.774877	0	99.79395	N = 71010
	between		5.174521	0	67.703	n = 7582
	within		4.619062	-66.8823	90.02182	T = 9.3656
CD_pdiff_norm_below	overall	6.295817	34.90928	0	399.6921	N = 71010
	between		29.99908	0	247.6796	n = 7582
	within		21.98751	-241.384	334.338	T = 9.3656
CDI_pdiff_norm_above	overall	0.88319	6.969426	0	99.92434	N = 71010
	between		5.368581	0	60.75448	n = 7582
	within		4.746172	-59.8713	90.12448	T = 9.3656
CDI_pdiff_norm_below	overall	6.199174	35.09752	0	399.6443	N = 71010
	between		30.18811	0	279.8243	n = 7582
	within		22.03555	-273.625	297.9643	T = 9.3656
ND_pdiff_norm_above	overall	0.977679	7.727965	0	395.7784	N = 71010
	between		5.296978	0	72.40186	n = 7582
	within		5.830982	-71.4242	328.2254	T = 9.3656
ND_pdiff_norm_below	overall	4.333964	27.41358	0	399.5742	N = 71010
	between		21.23906	0	244.2898	n = 7582
	within		19.95784	-239.956	346.9059	T = 9.3656
NDI_pdiff_norm_above	overall	0.958433	7.292432	0	163.1431	N = 71010
	between		5.124216	0	70.03453	n = 7582
	within		5.421478	-69.0761	142.266	T = 9.3656
NDI_pdiff_norm_below	overall	4.819013	29.29336	0	399.8933	N = 71010
	between		23.18255	0	257.2778	n = 7582
	within		21.13244	-252.459	348.5557	T = 9.3656
cumu_mail	overall	2.368455	3.144713	0	10	N = 71010
	between		2.717849	0	10	n = 7582
	within		1.566738	-3.38155	7.618455	T = 9.3656
prop_precdd	overall	1.00498	1.281126	0	4.771705	N = 71010
	between		0.318455	0	4.09616	n = 7582

	within		1.255616	-1.46485	4.852499	T = 9.3656

Table 3.5.1.1.Within and Between variation for variables of interest (continued)

prop_prehdd_da	overall	1.524098	3.857352	0	117.7955	N = 70986
	between		2.362019	0.005265	46.17566	n = 7580
	within		2.798987	-44.423	74.33734	T = 9.36491
prop_premin_da	overall	0.900414	0.113042	0.669142	1.093525	N = 71010
	between		0.029023	0.746886	1.09007	n = 7582
	within		0.110426	0.693266	1.120699	T = 9.3656
prop_prerain_da	overall	0.693179	0.938012	0	33.91765	N = 71010
	between		0.538017	0.010795	12.63801	n = 7582
	within		0.720746	-11.4226	24.80776	T = 9.3656
N = Total sample size aggregated across all periods						
n= Total number of unique households						
T= Mean number of months that each household's consumption appears in the data						

In order to decide which approach to adopt to estimating treatment effects in the panel, the data should first be examined, in particular the nature of the within and between variation of the dependent monthly consumption variable and the various potential explanatory variables. Table 3.5.1 above reports data for the nature of variation of these variables of interest. Suburb indicators are not reported since they are constant for our sample through the year and thus exhibit no within variation.[67] As observed in the table 3.5.1, there is a large degree of both within and between variation in the consumption, treatment and climatic variables. Accordingly, we should use an approach that captures this.

The Random Effects estimator (RE) is able to control for both within and between variation through calculating a weight (θ) in a feasible general least squares manner. The weight θ serves to weight the between or FE estimator in the RE approach, essentially rendering the RE approach a weighted mixture of BE and FE. In the case of all variation being within variation $\theta = 1$, and in the case of all variation being between variation $\theta = 0$. As such the RE approach seems a logical candidate for estimating the treatment effect in data such as that described above. Indeed looking at the values of θ for RE regressions over this data suggested that there would be significant difference between the

[67] Households that moved neighbourhoods (almost none) over the year were omitted from the analysed sample.

RE and FE approaches, although with theta usually at around 0.815, this difference would not be very great.

However the RE estimator is valid only under more stringent assumptions than the FE estimators. In terms of the generic individual effects model for panel data where

$$y_{it} = \alpha_i + \beta x_{it} + \varepsilon_{it}$$

we have a time invariant error component, α_i , and a time variant component, ε_{it} . Within the FE framework may be correlated with the regressors x_{it} since it is dealt with by the FE estimation anyway. The time varying error component ε_{it} may not be correlated with regressors x_{it}. However, within the RE framework it is necessary that $E(\varepsilon_{it}|\alpha_{it}, x_{it}) = E(\varepsilon_{it}|, x_{it}) = 0$, in other words that the time invariant error component also be distributed independently of the regressors x_{it}.

To test whether these assumptions hold, a Hausman type test is conducted. Since not all accounts in the actively treated sample were mailed every month the panel in this case is unbalanced. This means that the usual approach of Arellano (1993) and Wooldridge (2002, pp. 290-91), using a Wald test, needs to be extended since it is valid only in balanced panels. Schaffer and Stillman (2010) generalise from the approach of Arellano (1993) and Wooldridge (2002) supply such an extension by estimating θ for every observation.

The result of these Hausman tests, for every regression reported in chapter 4 provides strong evidence that the additional independence assumptions of the RE estimator are not met. As such, the FE estimators have priority in interpretation.

Nevertheless both RE and FE estimates are reported. Although the statistical evidence weighs in favour of the FE approach for this dataset, it does have one drawback that should cause pause. In this dataset no household changes suburb over the year period. This renders the suburb time-invariant for each household. Since the households suburb cannot be included in FE regression estimating the household level treatment effect.

Bruhn and Mckenzie (2009) make a compelling case for the inclusion of dummy variables that explicitly signal systematic variation in an experimentally generated dataset. Excluding these, where they can be included explicitly may introduce bias into coefficient estimates, such as those estimating the treatment effect. The suburb indicators perform that role in this dataset. Across all non-FE regressions they explain a significant amount of variation. This is to be expected in light of the discussion in chapter 1. Since similar people and households that are similar in many respects may cluster together in neighbourhoods (and hence suburbs) and may just as significantly be then

subject to very similar suburb level influences (such as services, crime, quality of education, transport costs) we would expect there to be significant differences between suburbs. Since the suburbs in this dataset are non-contiguous with each other, it is likely that these differences would quite marked. Finally, we know that the households in these suburbs have different income characteristics, since they were selected upon that basis.

As a result, while the coefficients calculated from the FE approach are preferred in terms of interpreting the analysis, both RE and FE results are reported since the RE approach allows suburbs to be included explicitly.

Since the standard errors around the parameter estimates are likely to vary substantially robust estimates of the standard errors are used. Since these are also likely to vary in some fashion at the neighbourhood level due to unobserved neighbourhood level characteristics, the standard errors are allowed to cluster at the neighbourhood level.

3.5.1.1. Comparing treated households to untreated households: Regression specifications.

The four treatments are compared to the control group in four separate regressions for each income group. In each case the dependant variable is $prop\_month\_cons_{it}$. Thus for treatment CD, in an estimation sample consisting of control group households and households treated with CD we have:

$$\begin{aligned} prop\_month\_cons_{it} \; &= \beta_{CDA}(CD_i)(pdiff\_above_{it-1}) + \beta_{CDB}(CD_i)(pdiff\_below_{it-1}) \\ &+ \beta_{CDIA}(CDI_i)(pdiff\_above_{it-1}) + \beta_{CDIB}(CDI_i)(pdiff\_below_{it-1}) \\ &+ \beta_{NDA}(ND_i)(pdiff\_above_{it-1}) + \beta_{NDB}(ND_i)(pdiff\_below_{it-1}) \\ &+ \beta_{NDIA}(NDI_i)(pdiff\_above_{it-1}) + \beta_{NDIB}(NDI_i)(pdiff\_below_{it-1}) + \beta_{cumu} cumu\_mail_{it-1} \\ &+ \beta_x \boldsymbol{X}_{it} + v_{it} \end{aligned}$$

where $\beta_x \boldsymbol{X}_{it}$ is the vector of climatic variables and v_{it} is the composite error term.

This equation is estimated separately in each income bracket to produce estimates of treatment, effects amongst upper, middle and lower income suburbs of households.

3.5.1.2. Expected results of comparing treated households to untreated households.

Given theory and previous studies discussed in chapter one, we would expect that households reported as having consumed above the norm will lower their consumption towards that norm in the following month, or a at least increase their consumption by less than the control households in the following month to leave them consuming more than control households in the following month in terms of their average daily consumption as a proportion of their January 2010 consumption.

Similarly we would expect that households reported as having consumed below the norm will increase their consumption in the following month or at least decrease their consumption by less than control group households, driven in both cases by a behavioural tendency for people to conform to reported norms.

Given this, we expect that for the two descriptive-only norm treatments that all else being equal,

$$\beta_{CDA} < 0 \text{ and } \beta_{CDB} > 0$$

$$\text{and } \beta_{NDA} < 0 \text{ and } \beta_{NDB} > 0$$

We also expect that for households receiving feedback that they had consumed less than average a positive injunctive norm (☺) will mitigate against the tendency for these households to increase their consumption in the following month or at least decrease their consumption by less than control group households, by reinforcing their pro-social behaviour.

When injunctive norms are introduced we thus expect that

$$\beta_{CDIB} < \beta_{CDB} \text{ and } \beta_{NDIB} < \beta_{NDB}$$

Introducing a socially closer norm, such as a neighbourhood norm rather than a city level norm, should intensify any reaction to the reported norm. We thus expect, for all income groups that:

$$\beta_{NDIB} < \beta_{CDIB} \text{ and } \beta_{NDIA} < \beta_{CDIA} \text{ amongst injunctive norms, and}$$

$$\beta_{NDB} > \beta_{CDB} \text{ and } \beta_{NDA} < \beta_{CDA} \text{ amongst descriptive norms.}$$

We also expect $\beta_{cumu} > 0$ for each type of treatment.

.

3.5.1.3. Comparing various treatments: Regression specifications.

When we compare the various mailed treatments to each other, we will mainly be interested in comparing the influence of varying social distance and the influence of an injunctive norm.

To analyse both questions we make use of the following framework. As discussed earlier, we are assuming that if we report to households how their consumption compares to a average household's consumption then their consumption in the following month will be a function of the difference between their own consumption and the average household's in the following fashion.

$$prop\_month\_cons_{it} = \beta_{\mu}(\text{consumption}_{it-1} - \mu_{it-1}) + \beta_{x}X_{it}$$

Where HH_{it-1} indicates the household's consumption, μ_{it-1}indicates the average household's consumption and X_{it} is a vector of other consumption-influencing variables, such as the climatic variables discussed in chapter 3. This implies that if households tend towards the norm, that $\beta_{\mu} > 0$.

When we introduce varying social distance and the varying presence of an injunctive norm we assume that the tendency for household's to move their own consumption towards the norm may be augmented or curtailed by the introduction of a socially closer average and/or the introduction of an injunctive norm in the following fashion.

$$prop\_month\_cons_{it} = [\beta_{\mu} + \beta_{I}I_{it-1} + \beta_{N}N_{it-1} + \beta_{NI}(N_{it-1}I_{it-1})][HH_{it-1} - \mu_{it-1}] + \beta_{x}X_{it}$$

With I_{it-1} a dummy indicating whether the household was mailed an injunctive norm or not, N_{it-1} a dummy indicating whether the household was mailed a neighbourhood level average (1) or a city level average (0) and $N_{it-1}I_{it-1}$ an interaction of those two terms. For the sake of space only the results of the Fixed Effect regressions reported (in table 4.3) since the FE estimator is the most appropriate for this data. As before, these regressions are run over the whole year's data as well as over the two halves of the year in order to take account for potentially different variation in treatment effect.

3.6. Summary.

In order to test whether reporting social norms would alter households' electricity consumption a natural field experiment was designed and executed in the city of Cape Town from January to December 2010. Originally the experiment was designed to test the effect of varying social distance and of including an injunctive norm, making for four treatments. However soon after fieldwork began interest in the effect of the various treatments relative to households was piqued. This meant that while treatments were randomised amongst the four mailed treatments, the control was not randomly allocated. In order to account for the bias that arose as a result of this incomplete randomisation with respect to the unmailed control, consumption is measured as a proportion of pre-treatment consumption levels

In order to test whether social distance would modify the influence of reporting social norms, two sorts of descriptive norms were used: city level norms (for the socially distant treatments) and a neighbourhood level norm (for the socially close treatments). For each of these, half of the households receiving inserts received feedback that was augmented with an injunctive norm. In total there were 4 treatment cells (CD, CDI, ND, NDI) with similar numbers of households in each. Households who were not allocated to any insert treatment but which were from the same suburbs as the treated households formed the control sample.

The experiment was conducted amongst a sample of households that included suburbs representative of the lower, middle and upper thirds of the city's income distribution. The norms were reported to households upon A4 inserts included with their electricity bills. There was a brief problem in the middle of fieldwork that necessitated a change in the background calculation methodology of the printing company in order to be solved. However mailing largely continued through the fieldwork period. Data will be examined for any break around this period.

We expect to find that the neighbourhood level treatments elicit stronger reactions than the city level treatments. We expect that households who are reported as having consumed above the norm will reduce their consumption toward the reported norm in the subsequent month. We expect to observe that the injunctive norm will curb any boomerang effects that are present in the data. Thus we expect that households who are reported as having consumed below the average will increase their consumption in the subsequent month but that this effect will either be moderated or entirely reversed by the augmenting the descriptive norm with an injunctive norm.

Chapter 3 Appendix: Treatment Inserts.

CD (Below)

The average household in Cape Town used **925 KwH** of electricity last month.

Your household used **128.000 KwH** of electricity last month.

Your household consumed **less** than the average.

Tips that may help you to use less electricity:

Use less hot water. If you can, take a shower rather than a bath. Use as little hot water as possible whether showering or bathing. Stack your dishes and wash them in one large load, rather than washing many smaller loads. If you go away for the weekend, switch your geyser off. The recommended temperature to set your geyser's thermostat to is 55^{0}C. Check what your thermostat is set to. Please ask a professional for help if you are not sure how to do this.

Use your big appliances less often. Hang up washing to dry as tumble dryers use a very large amount of electricity. Use stove plates and ovens for only as long as required to cook your food. Cook on as low a heat a possible. It is cheaper to boil water in a kettle than on a stove. Once water is boiling, only a low heat is required on a stove to keep it boiling. Boil only the water you need.

Conserve Heat: Closing the curtains helps greatly in keeping a room warm. Close the curtains in the early evening. Use heaters and air conditioners only when you are in the room and be sure to keep the room's doors and curtains closed as this helps to retain the heat. Whenever possible and convenient, use an extra pullover or a blanket instead of a heater.

CITY OF CAPE TOWN | ISIXEKO SASEKAPA | STAD KAAPSTAD

CD (Above)

The average household in Cape Town used **925 KwH** of electricity last month.

Your household used **961.000 KwH** of electricity last month.

Your household consumed **more** than the average.

Tips that may help you to use less electricity:

Use less hot water. If you can, take a shower rather than a bath. Use as little hot water as possible whether showering or bathing. Stack your dishes and wash them in one large load, rather than washing many smaller loads. If you go away for the weekend, switch your geyser off. The recommended temperature to set your geyser's thermostat to is 55°C. Check what your thermostat is set to. Please ask a professional for help if you are not sure how to do this.

Use your big appliances less often. Hang up washing to dry as tumble dryers use a very large amount of electricity. Use stove plates and ovens for only as long as required to cook your food. Cook on as low a heat a possible. It is cheaper to boil water in a kettle than on a stove. Once water is boiling, only a low heat is required on a stove to keep it boiling. Boil only the water you need.

Conserve Heat: Closing the curtains helps greatly in keeping a room warm. Close the curtains in the early evening. Use heaters and air conditioners only when you are in the room and be sure to keep the room's doors and curtains closed as this helps to retain the heat. Whenever possible and convenient, use an extra pullover or a blanket instead of a heater.

CITY OF CAPE TOWN | ISIXEKO SASEKAPA | STAD KAAPSTAD

CDI (Below)

The average household in Cape Town used **925 KwH** of electricity last month.

Your household used **128.000 KwH** of electricity last month

Your household consumed **less** than the average.

Tips that may help you to use less electricity:

Use less hot water. If you can, take a shower rather than a bath. Use as little hot water as possible whether showering or bathing. Stack your dishes and wash them in one large load, rather than washing many smaller loads. If you go away for the weekend, switch your geyser off. The recommended temperature to set your geyser's thermostat to is 55°C. Check what your thermostat is set to. Please ask a professional for help if you are not sure how to do this.

Use your big appliances less often. Hang up washing to dry as tumble dryers use a very large amount of electricity. Use stove plates and ovens for only as long as required to cook your food. Cook on as low a heat a possible. It is cheaper to boil water in a kettle than on a stove. Once water is boiling, only a low heat is required on a stove to keep it boiling. Boil only the water you need.

Conserve Heat: Closing the curtains helps greatly in keeping a room warm. Close the curtains in the early evening. Use heaters and air conditioners only when you are in the room and be sure to keep the room's doors and curtains closed as this helps to retain the heat. Whenever possible and convenient, use an extra pullover or a blanket instead of a heater.

CDI (Above)

The average household in Cape Town used **925 KwH** of electricity last month.

Your household used **961.000 KwH** of electricity last month.

Your household consumed **more** than the average.

Tips that may help you to use less electricity:

Use less hot water. If you can, take a shower rather than a bath. Use as little hot water as possible whether showering or bathing. Stack your dishes and wash them in one large load, rather than washing many smaller loads. If you go away for the weekend, switch your geyser off. The recommended temperature to set your geyser's thermostat to is 55°C. Check what your thermostat is set to. Please ask a professional for help if you are not sure how to do this.

Use your big appliances less often. Hang up washing to dry as tumble dryers use a very large amount of electricity. Use stove plates and ovens for only as long as required to cook your food. Cook on as low a heat a possible. It is cheaper to boil water in a kettle than on a stove. Once water is boiling, only a low heat is required on a stove to keep it boiling. Boil only the water you need.

Conserve Heat: Closing the curtains helps greatly in keeping a room warm. Close the curtains in the early evening. Use heaters and air conditioners only when you are in the room and be sure to keep the room's doors and curtains closed as this helps to retain the heat. Whenever possible and convenient, use an extra pullover or a blanket instead of a heater.

ND (Below)

The average household in your neighbourhood used **1630.752 Kwh** of electricity last month.

Your household used **1628.000 KwH** of electricity last month.

Your household consumed **less** than the average.

Tips that may help you to use less electricity:

Use less hot water. If you can, take a shower rather than a bath. Use as little hot water as possible whether showering or bathing. Stack your dishes and wash them in one large load, rather than washing many smaller loads. If you go away for the weekend, switch your geyser off. The recommended temperature to set your geyser's thermostat to is 55°C. Check what your thermostat is set to. Please ask a professional for help if you are not sure how to do this.

Use your big appliances less often. Hang up washing to dry as tumble dryers use a very large amount of electricity. Use stove plates and ovens for only as long as required to cook your food. Cook on as low a heat a possible. It is cheaper to boil water in a kettle than on a stove. Once water is boiling, only a low heat is required on a stove to keep it boiling. Boil only the water you need.

Conserve Heat: Closing the curtains helps greatly in keeping a room warm. Close the curtains in the early evening. Use heaters and air conditioners only when you are in the room and be sure to keep the room's doors and curtains closed as this helps to retain the heat. Whenever possible and convenient, use an extra pullover or a blanket instead of a heater.

ND (Above)

The average household in your neighbourhood used **749.087 Kwh** of electricity last month.

Your household used **1012.000 KwH** of electricity last month.

Your household consumed **more** than the average.

Tips that may help you to use less electricity:

Use less hot water. If you can, take a shower rather than a bath. Use as little hot water as possible whether showering or bathing. Stack your dishes and wash them in one large load, rather than washing many smaller loads. If you go away for the weekend, switch your geyser off. The recommended temperature to set your geyser's thermostat to is 55°C. Check what your thermostat is set to. Please ask a professional for help if you are not sure how to do this.

Use your big appliances less often. Hang up washing to dry as tumble dryers use a very large amount of electricity. Use stove plates and ovens for only as long as required to cook your food. Cook on as low a heat a possible. It is cheaper to boil water in a kettle than on a stove. Once water is boiling, only a low heat is required on a stove to keep it boiling. Boil only the water you need.

Conserve Heat: Closing the curtains helps greatly in keeping a room warm. Close the curtains in the early evening. Use heaters and air conditioners only when you are in the room and be sure to keep the room's doors and curtains closed as this helps to retain the heat. Whenever possible and convenient, use an extra pullover or a blanket instead of a heater.

CITY OF CAPE TOWN | ISIXEKO SASEKAPA | STAD KAAPSTAD

NDI (Below)

The average household in your neighbourhood used **1630.752 Kwh** of electricity last month.

Your household used **1628.000 KwH** of electricity last month.

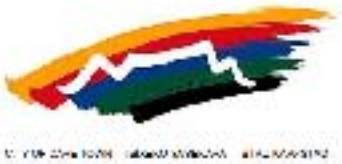
Your household consumed **less** than the average.

Tips that may help you to use less electricity:

Use less hot water. If you can, take a shower rather than a bath. Use as little hot water as possible whether showering or bathing. Stack your dishes and wash them in one large load, rather than washing many smaller loads. If you go away for the weekend, switch your geyser off. The recommended temperature to set your geyser's thermostat to is 55°C. Check what your thermostat is set to. Please ask a professional for help if you are not sure how to do this.

Use your big appliances less often. Hang up washing to dry as tumble dryers use a very large amount of electricity. Use stove plates and ovens for only as long as required to cook your food. Cook on as low a heat as possible. It is cheaper to boil water in a kettle than on a stove. Once water is boiling, only a low heat is required on a stove to keep it boiling. Boil only the water you need.

Conserve Heat: Closing the curtains helps greatly in keeping a room warm. Close the curtains in the early evening. Use heaters and air conditioners only when you are in the room and be sure to keep the room's doors and curtains closed as this helps to retain the heat. Whenever possible and convenient, use an extra pullover or a blanket instead of a heater.



NDI (Above)

The average household in your neighbourhood used **749.087 Kwh** of electricity last month.

Your household used **1012.000 KwH** of electricity last month.

Your household consumed **more** than the average.

Tips that may help you to use less electricity:

Use less hot water. If you can, take a shower rather than a bath. Use as little hot water as possible whether showering or bathing. Stack your dishes and wash them in one large load, rather than washing many smaller loads. If you go away for the weekend, switch your geyser off. The recommended temperature to set your geyser's thermostat to is 55°C. Check what your thermostat is set to. Please ask a professional for help if you are not sure how to do this.

Use your big appliances less often. Hang up washing to dry as tumble dryers use a very large amount of electricity. Use stove plates and ovens for only as long as required to cook your food. Cook on as low a heat as possible. It is cheaper to boil water in a kettle than on a stove. Once water is boiling, only a low heat is required on a stove to keep it boiling. Boil only the water you need.

Conserve Heat: Closing the curtains helps greatly in keeping a room warm. Close the curtains in the early evening. Use heaters and air conditioners only when you are in the room and be sure to keep the room's doors and curtains closed as this helps to retain the heat. Whenever possible and convenient, use an extra pullover or a blanket instead of a heater.

Chapter 4

Analysis and Results.

This chapter presents the results of the analysis outlined in chapter 3. It has two major sections. First we will look at whether there was a difference in consumption between the control group and the various mailed treatments. In doing this the apparently different effect of mailings amongst the various income groups is noted as well as an indication that reactions to the mailed treatments may have changed from the sixth mailing. From here we move on to an analysis of whether the various mailed treatments differed from each other. In particular we are interested to see whether there was any significant difference in consumption that could be due to varying social distance or the presence of an injunctive norm.

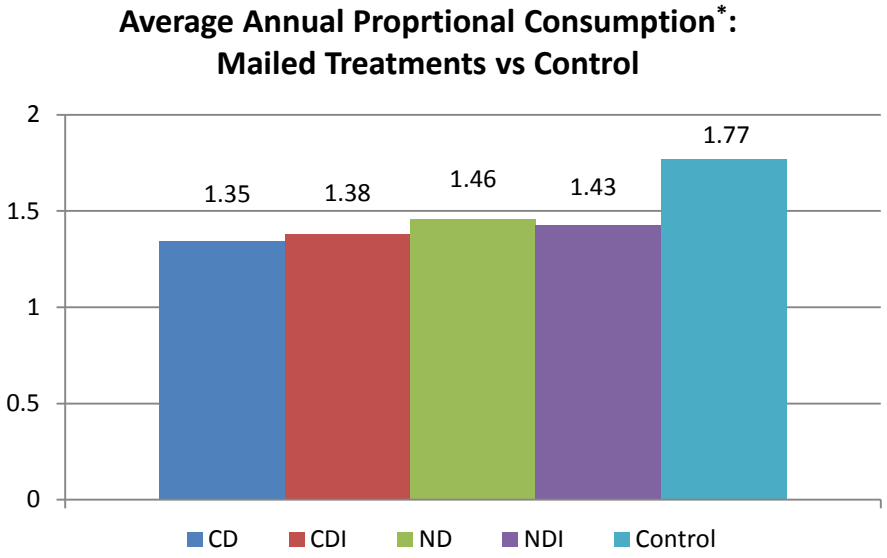
4.1 Treatment effects for the whole year compared to the control group.

As discussed in the methodology section, the preferred metric for comparison of consumption between the various mailed treatments and control group is that of proportional consumption. To recapitulate further: in this section, as in the rest of this chapter, there are two measures of proportional consumption that are used. The first is average daily consumption during the treatment period as a proportion of consumption during the entire pre-treatment period, which will be referred to as "annual proportional consumption" . The second is average daily consumption in a treatment-period month as a proportion of the average daily consumption in the entire pre-treatment period, which will be referred to as "monthly proportional consumption".

Examining the difference between the mean annual proportional consumption values for the various household treatment groups, shown in the histogram below, we find there is a significant and large difference in consumption between control group and the mailed groups. The difference between the control group and the mailed treatments is significant at the 1% level and smaller.

Examining the consumption of the various mailed treatment groups reveals only one difference in consumption that is significant, that is between CD and ND. Although we analyse this more in section 4.3, it seems from these figures that, contrary to expectation, the more socially distant City Descriptive norm (CD) achieved a lower level of consumption than did the Neighbourhood Descriptive norm (ND). The significance of this difference is somewhat sensitive to whether one calculates an unweighted average, in which case it is significant at the 5% level, or weights suburbs equally (as above), where the t-stat for this difference is 1.775, indicating significance at the 10% level.

Figure 4.1.1



* Average daily consumption over treatment period as a proportion of average daily consumption in the pre-treatment period

A simple OLS cross sectional regression (reported in table 4.1.1) controls for variation specific to suburbs in a more conventional regression fashion, (not the equal-weighting approach in figure 4.1.1). These regressions confirm that there is a significant difference in consumption between the control sample when and all of the treated samples at the overall year level.

Table 4.1.1: Estimated treatment effects compared to the control group in terms of the estimated effect upon annual average proportional consumption
(full regression results in the appendix)

	OLS	OLS (Robust)
CD	-0.4603754	-0.4603754
	(0.0606044)***	(0.0570777)***
CDI	-0.4167257	-0.4167257
	(0.0611908)***	(0.0596710)***
ND	-0.3500760	-0.3500760
	(0.0610709)***	(0.0605511)***
NDI	-0.3946388	-0.3946388
	(0.0599847)***	(0.0626385)***

Standard errors shown in brackets underneath coefficient estimates

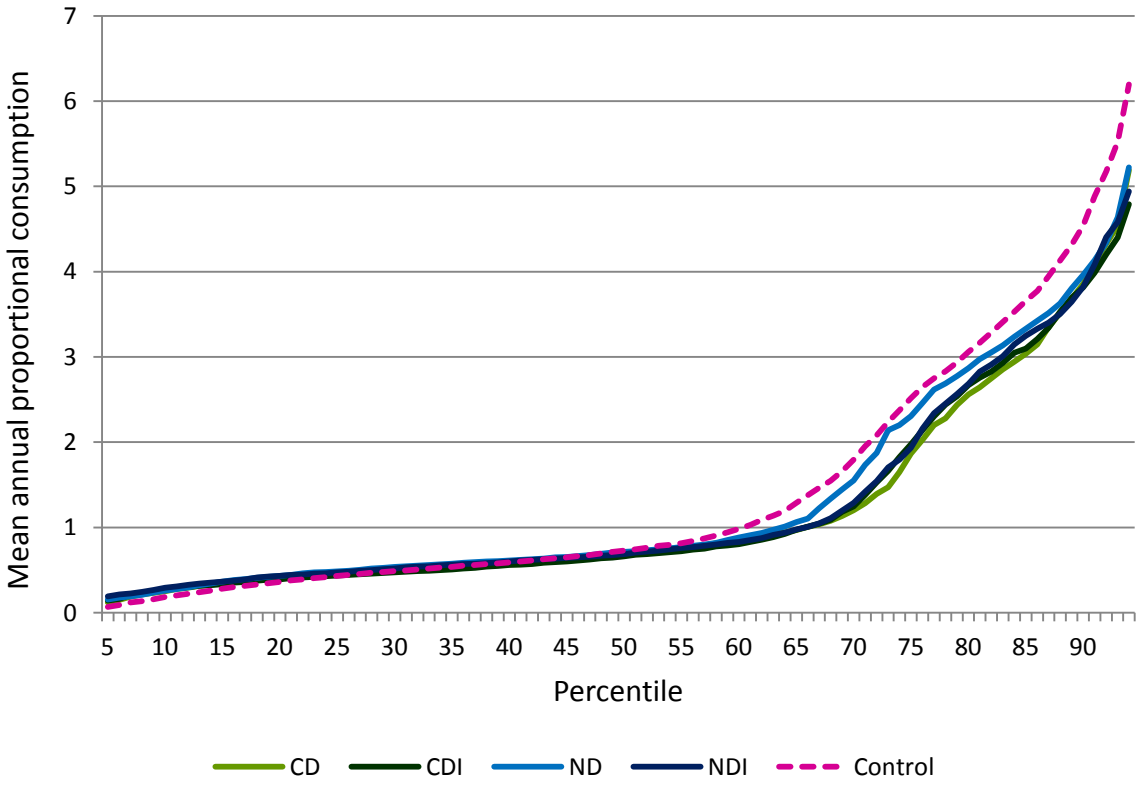
* $p<0.1$; ** $p<0.05$; *** $p<0.01$

Figure 4.1.2 below shows the mean annual proportional consumption values for the 5th to 95th percentiles for each treatment group. The relative percentile distributions suggest that much of the difference between the mailed treatment groups and the control group derives

from differences observed in the upper percentiles, especially percentiles 60 and above. Similarly, the difference between the ND and CD treatment appears to be driven by differences in consumption amongst households in the upper half of percentiles. The relative distribution of the ND and CD treatments diverges between roughly the 65th and 85th percentile, with ND showing slightly higher consumption on average in that range.

Figure 4.1.2

Distribution of mean Annual Proportional Consumption* for Treatment and Control Groups: comparison of percentile values

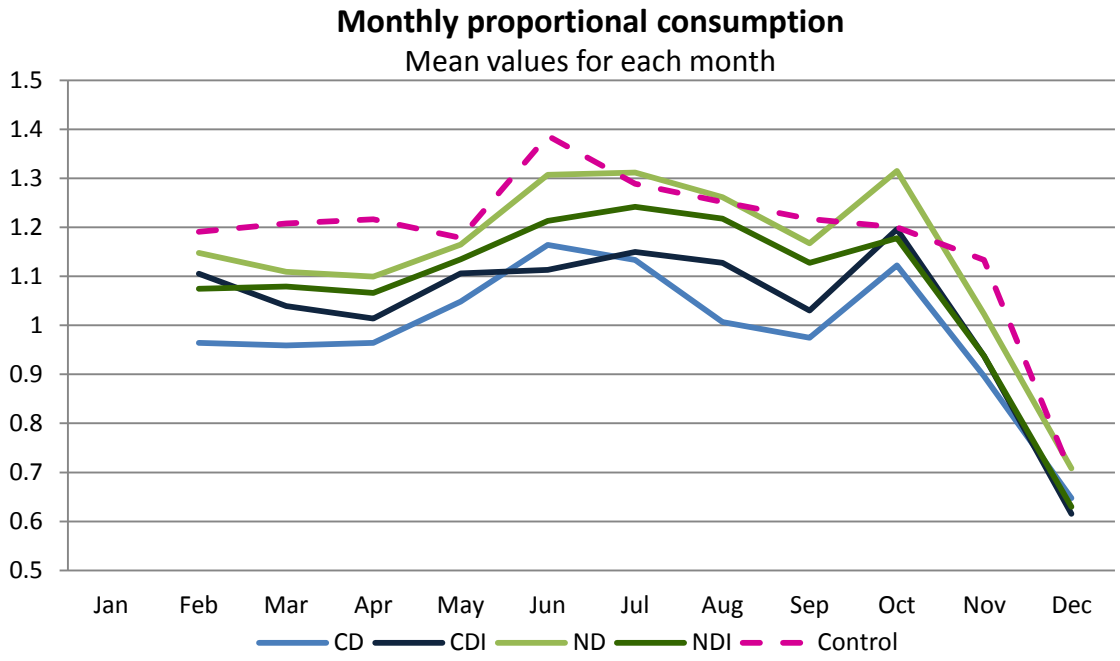


*Average daily consumption for the treatment period as a proportion of average daily consumption in the pre-treatment period

In terms of the variance of the treatment effect over time, the data is somewhat noisy for various treatments. Monthly proportional consumption (shown in figure 4.1.2) figures differed between treatments through different periods of the year. The period in which there seems to have been the clearest effect is during the four months after the initial mailing. After this period it seems that the city level treatments become increasingly more effective than the neighbourhood level treatments. Figure 4.1.3 suggests that the households receiving the city level treatments achieved lower monthly proportional

consumption than the control group for much of the year. In contrast, and unexpectedly, the neighbourhood level treatments did not always consume less than the control, for most of the July - October period they seem to have consumed either the same or, sometimes, in excess of the control.

Figure 4.1.3



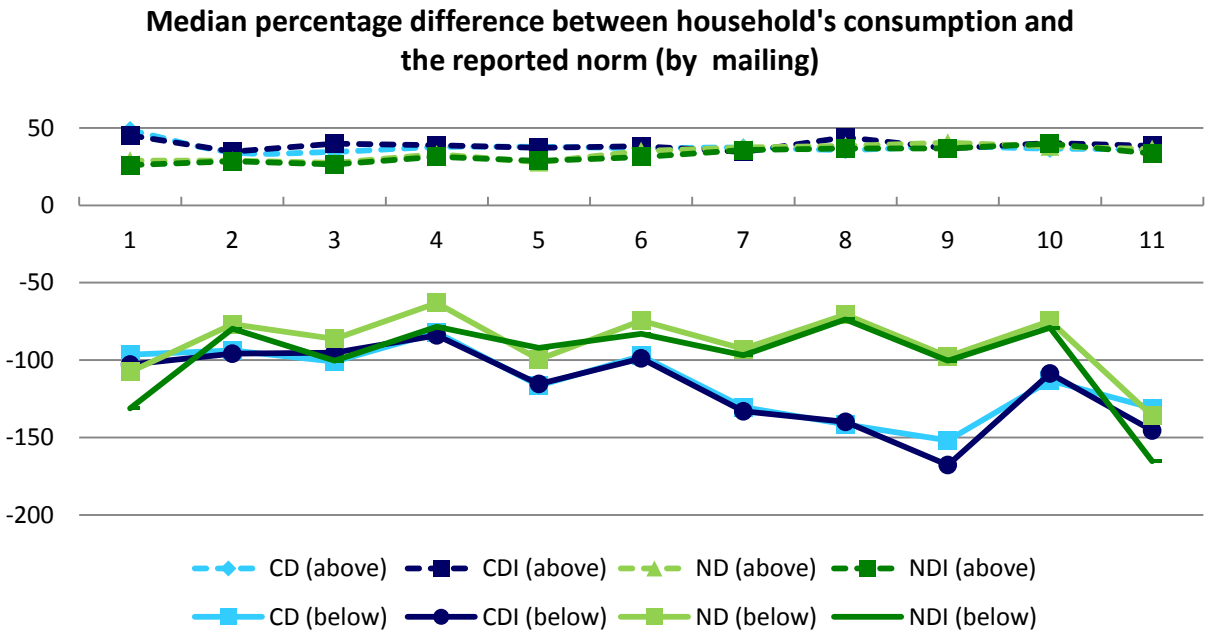
Regression analysis was used to analyse overall treatment effects at the month to month level and to take advantage of the panel nature of the data. As discussed in the methodology, these regressions attempt to control for various confounds that may be present in the data.

One of the more important potential confounds is that which may exist between social distance and the reported difference between a household's consumption and the norm to which they are compared. If we expect, as described in the literature review of social norms and the methodology chapter, that households will change their consumption in proportion to the degree to which their consumption differs from the reported normal consumption, then we may have a situation where declining social distance is positively correlated with declining differences between household and reported normal consumption and the two elements confound each other. As discussed in the methodology section it is for this reason that we include a measure of the difference between the household's reported consumption and the reported normal consumption, specifically the difference as a percentage of the household's own reported consumption.

Since data for the norms was drawn from the real world, it is plausible that the reported difference between household consumption and the norm will on average be greater when the norm is a city norm as opposed to a neighbourhood norm since households are likely to

be more similar in many respects to other households in the neighbourhood than to other households in the city. Figure 4.1.4 below suggests that this was the case mainly for households that were reported as being below the norm, with the differences between households receiving city norms noticeably greater than that for neighbourhood level norms between mailing number 6 and 10. Although somewhat hidden by the line graph's scale, there is also a relatively large difference in percentage difference form mailings 1 to 5 between city and neighbourhood treatment households when they were reported as having consumed more than the norm. Given this data it seems that we should control for this confound.

Figure 4.1.4



The regression results in table 4.1.2 reports two model specifications. In one specification monthly proportional consumption is modelled as a function of being mailed any treatment, while the other specification models monthly proportional consumption as a function of treatment-specific variables. Both are run over the analysis sample (control and treated) for the entire year. Since the "mailed" variable is "1" if the household was mailed last month and "0" if the household was in the control group, the mailed variable is essentially time invariant for each household; as such it cannot be estimated by the FE estimator. For the FE estimator results it is thus plausible to interpret the estimated relationship between cumulative mailings received and consumption as incorporating an effect beyond just the average effect of receiving an additional mailing, but as encompassing something of the difference between being mailed at all or being in the control group too.

The regression results in table 4.1.2 indicate two major types of treatment effect. The most dominant effect seems to be cumulative in nature. The relationship between additional treatment mailings received and consumption relative to the control is significant and

Table 4.1.2 Estimated treatment effects compared to the control group amongst all income groups

	RE	RE (robust)	FE	FE (robust)
Specific treatments				
L.CD_pdiff_norm_above	-0.0016533	-0.0016533	-0.0006902	-0.0006902
	(3.33)***	(4.59)***	(1.31)	(1.85)*
L.CD_pdiff_norm_below	0.0000001	0.0000001	-0.0000036	-0.0000036
	(0.01)	(0.03)	(0.73)	(1.72)*
L.CDI_pdiff_norm_above	-0.0018652	-0.0018652	-0.0007213	-0.0007213
	(3.77)***	(4.88)***	(1.37)	(1.90)*
L.CDI_pdiff_norm_below	0.0000133	0.0000133	-0.0000309	-0.0000309
	(0.41)	(0.25)	(0.92)	(0.48)
L.ND_pdiff_norm_above	0.0004388	0.0004388	0.0006418	0.0006418
	(2.14)**	(1.01)	(3.11)***	(1.50)
L.ND_pdiff_norm_below	-0.0000004	-0.0000004	-0.0000005	-0.0000005
	(0.20)	(0.29)	(0.26)	(0.38)
L.NDI_pdiff_norm_above	0.0000250	0.0000250	0.0003552	0.0003552
	(0.10)	(0.05)	(1.39)	(0.90)
L.NDI_pdiff_norm_below	-0.0000012	-0.0000012	-0.0000015	-0.0000015
	(0.42)	(0.23)	(0.51)	(0.28)
L.cumu_mail	-0.0178553	-0.0178553	-0.0144496	-0.0144496
	(12.04)***	(9.84)***	(9.39)***	(7.23)***
θ_5	0.75	0.75		
θ_{50}	0.81	0.81		
θ_{95}	0.83	0.83		
Generic Mailing				
L.mailed	-0.2142120	-0.2142120		
	(8.31)***	(7.30)***		
L.cumu_mail	-0.0138290	-0.0138290	-0.0140950	-0.0140950
	(9.01)***	(7.00)***	(9.21)***	(7.11)***
θ_5	0.76	0.76		
θ_{50}	0.82	0.82		
θ_{95}	0.83	0.83		

Standard errors shown in brackets underneath coefficient estimates
* $p<0.1$; ** $p<0.05$; *** $p<0.01$

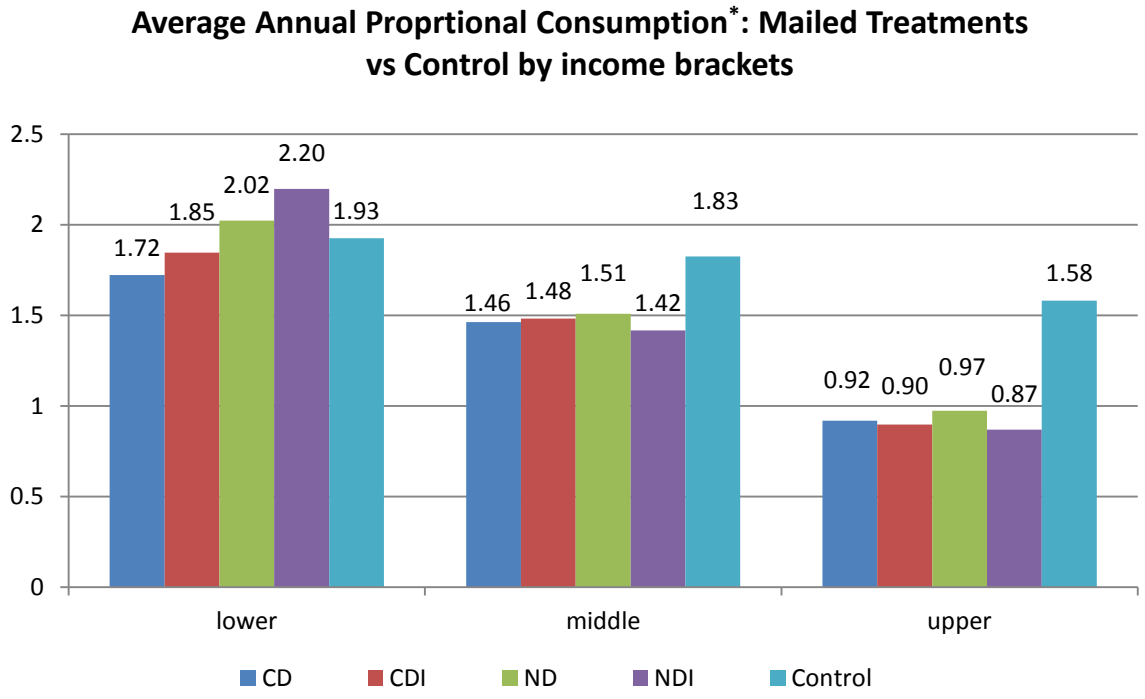
negative for all estimators. Put simply, the more treatment mailings a households received over the year, the less it would consume relative to the control group. This suggests that one of the most important elements of these treatments may have been cumulative.

The other result in the regressions in table 4.1.2 is somewhat unexpected. In the sample as a whole, it is the City level treatments which seem to have proved the more effective, even when controlling for difference between a household's consumption and the value of the average reported to it. In particular it is households who received feedback that they had consumed more than an average household in the City that responded with a significant decline in consumption; households who received feedback that they had consumed less than an average city level household did not consume significantly more or less than the control group. Neither of the neighbourhood level treatments were associated with significantly lower levels of consumption than the control group. Given the indications of the prior literature this is somewhat unexpected.

4.2. Overall treatment effects compared to the control group differ by income.

So far we have considered effects relative to the control group at the level of the total sample. However, the results also seem to have varied with income level. Figure 4.2.1 shows the relative results for each treatment among the three major income groups.

Figure 4.2.1



* Average daily consumption over treatment period as a proportion of average daily consumption in the pre-treatment period

From figure 4.2.1 we can see that it is amongst the upper and middle income suburbs where the difference between treated and control households appears to have been the greatest over the entire year. In both these suburban income groups the difference between each treatment and the control is significant at the 1% level. While there appears to be a strong treatment effect amongst households in the middle and upper income suburbs there is no

significant difference between the various mailed treatments themselves. This suggests, for households in the upper and middle income suburbs at least, that the treatment effect is something which is associated with an element that is generic to the various treatments rather than any of the differences between them.

Unlike households in the upper and middle income suburbs, mailed households in the lower income suburbs did not consume significantly less than the control group. It seems that households in this income bracket receiving the neighbourhood level treatments consumed more than those receiving the city level treatments. In particular both neighbourhood level treatments are associated with significantly higher consumption (at the 5% level) than is the case with the CD (City Descriptive only) treatment (but not the CDI treatment). Given the close similarity amongst treatments in the other income groups, this variation explains why neighbourhood treatments appeared to underperform the city level treatments in the overall sample.

Judging by the significant differences in consumption, portrayed in figure 4.2.1 above, it seems clear that there was a significant average treatment effect in terms of the overall year period. This treatment effect, for the overall year, was not observed amongst all income groups but was only found amongst the middle and upper income groups.

We now turn to look at the treatment effect at the month level. Here we are able to take advantage of the panel nature of this dataset and incorporate much richer data than was tractable in the cross sectional analysis of the overall year. This enables us to control for the potential confound between social distance and the difference between a household's consumption and the value of the reported average discussed in section 4.1. As discussed in chapter 3 (section 3.5.1), although Random effects and Fixed Effect estimators are reported (both estimated with and without robust clustered standard errors) the Fixed Effect estimators with robust standard errors are probably the most appropriate estimators for this data. Regressions were run over the whole year as well as the two halves of the year for each income group. Analysis for the two periods in the year are reported in addition to the whole year since it seems, from roughly month 6, that the variance in consumption of various groups seems to change. The change in fieldwork strategy (slight though it was) that occurred around this period is another reason that it is sensible to do that. In addition to the treatment-specific variables, the regression models also control for the climatic variables discussed in chapter 3.

As we examine the regression results for each income group the identification of specific treatment effects becomes a lot less clear. What we observe seems to be an untidy weaving of complex relationships that in their aggregate produce the clear results of section 4.1. This is most likely a combination of a large degree of natural variance in the consumption data and what appears to be a high degree of similarity in effect between the various treatments. Overall however, the regressions seem to be consistent with a description of the treatment

effect as being significantly cumulative in nature and being strongest amongst the upper and middle income suburbs. However the detail is not emphatic in this regard.

Making use of the month to month data to make comparisons between treatments it seems that for the upper income group (table 4.2.1) the CD treatment was the most significant treatment influence upon consumption. This is unexpected. Looking regression run on data for the overall year obscures another important influence upon consumption which was the cumulative effect of receiving any treatments. For the upper income group receiving more and more mailings initially lowered consumption, on average, relative to the control group. However in the second half of the year it seems that this effect began to go the other way, with additional mailings increasing consumption, on average, relative to the control group. This pattern implies that, at least upper income households, may tire, after some time, of receiving inserts with their electricity bill and begin to act in opposition to them.

Middle income households were also significantly influenced by the mailed treatments in a cumulative fashion when we consider the whole year period. The coefficient of -0.013 upon the cumulative mailing repressor indicates that on average, each additional mailing reduced consumption below the control group 1.3 percentage points. However as table 4.2.2 suggests, this effect seems to reverse sometime in the second half of the year. This is much like the pattern observed in the upper income households and may suggest similar forces at work in the middle income households. This might have been something that was especially influenced by the two city level treatments and the neighbourhood descriptive treatment, ND. These three treatments seem to decline in effectiveness from the first to the second half of the year amongst middle income households.

Unlike the other treatments amongst the middle income suburbs, the NDI treatment seems to have become a stronger moderating influence upon consumption in the latter half of the year. The reaction of households to this format of being reported as having consumed more than the norm was especially strong in the second half of the year. This result suggests that the NDI format is the most desirable format of the four treatments for continuous norms-based consumption reporting, at least amongst middle income households.

Households in the lower income suburb sample appear to have been subject to two sorts of treatment influences: the moderating cumulative effect of receiving any sort of treatment and the declining effectiveness of the city level treatments. The significant coefficient on the cumulative mailing regressor in table 4.2.3 suggests the familiar cumulative effect. The City

Table 4.2.1. Estimated treatment effects compared to the control group amongst Upper income suburbs
Influence upon Proportional Monthly Consumption

Specific treatments	RE	RE (robust)	FE	FE (robust)	FE (robust) Mailings 1-5	FE (robust) Mailings 6 - 11
L.CD_pdiff_norm_above	-0.0022624	-0.0022624	-0.0008931	-0.0008931	-0.0007879	-0.0010019
	(3.83)***	(4.99)***	(1.38)	(1.71)*	(0.0012316)	(0.0006232)
L.CD_pdiff_norm_below	-0.0000007	-0.0000007	-0.0000046	-0.0000046	0.0001262	-0.0000058
	(0.15)	(0.88)	(0.93)	(5.00)***	(0.0001707)	(0.0000227)
L.CDI_pdiff_norm_above	-0.0020209	-0.0020209	-0.0005681	-0.0005681	-0.0002470	-0.0002413
	(3.67)***	(5.73)***	(0.95)	(1.59)	(0.0007166)	(0.0004257)
L.CDI_pdiff_norm_below	-0.0000600	-0.0000600	-0.0000839	-0.0000839	-0.0001167	-0.0001203
	(1.35)	(0.70)	(1.87)*	(0.87)	(0.0000867)	(0.0001307)
L.ND_pdiff_norm_above	-0.0002791	-0.0002791	0.0003241	0.0003241	-0.0006301	-0.0001075
	(0.59)	(0.79)	(0.66)	(0.94)	(0.0006002)	(0.0005142)
L.ND_pdiff_norm_below	-0.0000003	-0.0000003	-0.0000003	-0.0000003	-0.0000130	-0.0000005
	(0.18)	(0.25)	(0.15)	(0.21)	(0.0000138)	(0.0000022)
L.NDI_pdiff_norm_above	-0.0003468	-0.0003468	0.0002705	0.0002705	-0.0005407	-0.0003540
	(0.77)	(0.84)	(0.59)	(0.65)	(0.0006388)	(0.0004850)
L.NDI_pdiff_norm_below	0.0000004	0.0000004	0.0000005	0.0000005	-0.0000057	0.0000017
	(0.16)	(0.08)	(0.18)	(0.09)	(0.0000554)	(0.0000048)
L.cumu_mail	-0.0100302	-0.0100302	-0.0025020	-0.0025020	-0.0513893	0.0201796
	(4.86)***	(4.57)***	(1.16)	(1.05)	(0.0055049)***	(0.007051)***
θ_{50}	0.81	0.81				
Generic mailing						
L.mailed	-0.4277227	-0.4277227				
	(12.67)***	(11.60)***				
L.cumu_mail	-0.0019171	-0.0019171	-0.0021672	-0.0021672	-0.0510011	0.0206217
	(0.90)	(0.81)	(1.02)	(0.92)	(0.0054775)***	(0.0069725)***
θ_{50}	0.81	0.81				

Standard errors shown in brackets underneath coefficient estimates
* $p<0.1$; ** $p<0.05$; *** $p<0.01$

Table 4.2.2. Estimated treatment effects compared to the control group amongst middle income suburbs
Influence upon Proportional Monthly Consumption

Specific Treatments	RE	RE (robust)	FE	FE (robust)	FE (robust) Mailings 1-5	FE (robust) Mailings 6 - 11
L.CD_pdiff_norm_above	-0.0009893	-0.0009893	-0.0001081	-0.0001081	0.0008386	0.0003457
	(1.05)	(1.57)	(0.11)	(0.18)	(0.0007228)	(0.0008743)
L.CD_pdiff_norm_below	0.0000062	0.0000062	0.0000047	0.0000047	0.0000002	0.0000067
	(0.49)	(1.03)	(0.38)	(0.74)	(0.0000062)	(0.0000013)***
L.CDI_pdiff_norm_above	-0.0020137	-0.0020137	-0.0007177	-0.0007177	-0.0008846	-0.0000264
	(1.87)*	(2.32)**	(0.64)	(0.89)	(0.0009301)	(0.0017331)
L.CDI_pdiff_norm_below	0.0000937	0.0000937	0.0000540	0.0000540	-0.0012571	0.0000235
	(1.74)*	(2.47)**	(0.98)	(0.90)	(0.0003296)***	(0.0000723)
L.ND_pdiff_norm_above	-0.0014228	-0.0014228	-0.0008124	-0.0008124	-0.0018099	-0.0011725
	(1.89)*	(2.17)**	(1.06)	(1.20)	(0.0010924)*	(0.0008868)
L.ND_pdiff_norm_below	-0.0000021	-0.0000021	-0.0000128	-0.0000128	0.0000034	-0.0001160
	(0.10)	(0.12)	(0.57)	(0.68)	(0.0000413)	(0.0000772)
L.NDI_pdiff_norm_above	-0.0023537	-0.0023537	-0.0014233	-0.0014233	-0.0015025	-0.0013936
	(3.04)***	(4.34)***	(1.79)*	(2.46)**	(0.0012281)	(0.0007025)**
L.NDI_pdiff_norm_below	-0.0000368	-0.0000368	-0.0000453	-0.0000453	0.0000422	-0.0003024
	(1.64)	(1.28)	(2.02)**	(1.42)	(0.0000112)***	(0.0001076)***
L.cumu_mail	-0.0165133	-0.0165133	-0.0131154	-0.0131154	0.0060221	0.0170021
	(6.16)***	(5.21)***	(4.72)***	(3.68)***	(0.0092222)	(0.0097211)*
θ_{50}	0.82	0.82				
Generic mailing						
L.mailed	-0.1838915	-0.1838915				
	(4.10)***	(3.51)***				
L.cumu_mail	-0.0133031	-0.0133031	-0.0132245	-0.0132245	0.0052152	0.0127483
	(4.79)***	(3.74)***	(4.78)***	(3.71)***	(0.0091355)	(0.0097458)
θ_{50}	0.82	0.82				

Standard errors shown in brackets underneath coefficient estimates
* $p<0.1$; ** $p<0.05$; *** $p<0.01$

Table 4.2.3 . Estimated treatment effects compared to the control group amongst lower income suburbs

Influence upon Proportional Monthly Consumption

Specific Treatments	RE	RE (robust)	FE	FE (robust)	FE (robust) Mailings 1-5	FE (robust) Mailings 6 - 11
L.CD_pdiff_norm_above	0.0007860	0.0007860	0.0009797	0.0009797	-0.0023178	0.0049517
	(0.49)	(0.85)	(0.61)	(1.15)	(0.0007938)***	(0.0015662)***
L.CD_pdiff_norm_below	0.0000218	0.0000218	-0.0000264	-0.0000264	0.0000332	-0.0000448
	(0.38)	(0.51)	(0.46)	(0.87)	(0.0001128)	(0.0000324)
L.CDI_pdiff_norm_above	0.0004186	0.0004186	0.0010823	0.0010823	-0.0000022	0.0029269
	(0.25)	(0.26)	(0.65)	(0.73)	(0.0027416)	(0.0015426)*
L.CDI_pdiff_norm_below	0.0000083	0.0000083	-0.0000950	-0.0000950	-0.0000863	-0.0001264
	(0.09)	(0.12)	(0.97)	(1.21)	(0.0000425)**	(0.0001765)
L.ND_pdiff_norm_above	0.0006798	0.0006798	0.0007589	0.0007589	0.0014679	0.0006797
	(2.41)**	(1.26)	(2.69)***	(1.41)	(0.0016712)	(0.0004873)
L.ND_pdiff_norm_below	-0.0000203	-0.0000203	-0.0000509	-0.0000509	0.0000158	-0.0000095
	(0.44)	(0.34)	(1.10)	(0.82)	(0.0000686)	(0.0000804)
L.NDI_pdiff_norm_above	0.0004872	0.0004872	0.0005923	0.0005923	-0.0027034	0.0007274
	(1.27)	(0.88)	(1.54)	(1.17)	(0.0020851)	(0.0007746)
L.NDI_pdiff_norm_below	-0.0000234	-0.0000234	-0.0000315	-0.0000315	0.0000152	0.0000413
	(0.81)	(0.49)	(1.07)	(0.62)	(0.0000523)	(0.0000710)
L.cumu_mail	-0.0259588	-0.0259588	-0.0284823	-0.0284823	0.0012247	0.0067839
	(6.90)***	(5.72)***	(7.32)***	(5.70)***	(0.0140139)	(0.0154729)
θ_{50}	0.81	0.81				
Generic mailing						
L.mailed	0.1740466	0.1740466				
	(2.48)**	(2.24)**				
L.cumu_mail	-0.0281271	-0.0281271	-0.0282909	-0.0282909	0.0001791	0.0069812
	(7.24)***	(5.67)***	(7.31)***	(5.69)***	(0.0138224)	(0.0154877)
θ_{50}	0.81	0.81				

Standard errors shown in brackets underneath coefficient estimates

* $p<0.1$; ** $p<0.05$; *** $p<0.01$

level treatments seem to have initially been a moderating influence upon consumption, particularly for households above the norm in the CD format and below the norm in the CDI format. However, in the second half of the year the influence of the city level treatments seems to have reversed from one of moderating consumption to one that promoted consumption, particularly if households were reported as having consumed more than average.

Within each income group we see hints of the busy complexity by which reference to social norms may likely influence consumption. However, when we step back we observe a pattern that suggests that there was a lot more similarity than difference between the various treatments over most of the treatment period. The most important treatment effect seems to have been cumulative in nature and generic to all treatments. Nevertheless we do see hints that the neighbourhood level treatments began to perform a stronger moderating role than the city level treatments in the latter months of the treatment period.

4.3.Specific differences in effects between mailed treatments

One of the questions this project originally sought to address was whether there would be a relatively different effect between varying the social distance between the household and the norm it was compared to and if the presence of an injunctive norm would reduce the tendency of those households reported as having consumed less than the norm to raise their consumption toward the norm. In these final two sections we examine those questions.

To analyse both questions we make use of the following framework. As discussed in chapter 3 we are assuming that if we report to households how their consumption compares to a average household's consumption then their consumption in the following month will be a function of the difference between their own consumption and the average household's in the following fashion.

$$consumption_{it} = \beta_{\mu}(HH_{it-1} - \mu_{it-1}) + \beta_x \boldsymbol{X}_{it}$$

Where HH_{it-1} indicates the household's consumption, μ_{it-1}indicates the average household's consumption and $\boldsymbol{X}_{it}$ is a vector of other consumption-influencing variables, such as the climatic variables discussed in chapter 3. This implies that if households tend towards the norm, that $\beta_{\mu} > 0$.

When we introduce varying social distance and the varying presence of an injunctive norm we assume that the tendency for household's to move their own consumption towards the norm may be augmented or curtailed by the introduction of a socially closer average and/or the introduction of an injunctive norm in the following fashion.

$$consumption_{it} = [\beta_{\mu} + \beta_I I_{it-1} + \beta_N N_{it-1} + \beta_{NI}(N_{it-1} I_{it-1})\,][HH_{it-1} - \mu_{it-1}] + \beta_x \boldsymbol{X}_{it}$$

With I_{it-1} a dummy indicating whether the household was mailed an injunctive norm or not, N_{it-1} a dummy indicating whether the household was mailed a neighbourhood level average (1) or a city level average (0) and $N_{it-1}I_{it-1}$ an interaction of those two terms. For the sake of space only the results of the Fixed Effect regressions reported (in table 4.3) since the FE estimator is the most appropriate for this data. As before, these regressions are run over the whole year's data as well as over the two halves of the year in order to take account if potentially different variation in treatment effect.

Table 4.3 presents no evidence of any consistently significant difference between treatments along the vectors of social distance or the presence of injunctive norms. The pattern of influence of these treatment elements seems inconsistent across time and income groups and is not significant at the overall year level for any income group.

Neighbourhood and injunctive treatments seem to have been somewhat significant influences upon consumption in the first half of the year and only amongst the middle and lower income groups. It seems that these specific influences may have worn off as the year went on, since they are no longer significant in the second half of the year for households in these income groups.

When the relationship to consumption was significant, both injunctive norms and socially closer averages appear to have moderated consumption, which is the expected direction of influence. In the case of varying social distance this provides a weak indication that when social distance significantly influences the effect a social norm has upon consumption it is more influential if the norm is socially closer to the household than if it is not. Of course, this data gives only the weakest indication of this (one half of the year for one demographic group). Whether this social distance effect is one that is just very small or is hidden by the large variance in the sample is difficult to tell. However the indication for future research in this vein is clear: a larger sample will be needed to analyse the effect of varying social distance.

It is worth considering how the injunctive norm behaved when it did appear to significantly influence consumption. The fact that β_I <0 provides some evidence that the injunctive norm is behaving in the expected fashion. Having accounted for the direct influence of $HH_{it-1} - \mu_{it-1}$ by itself, through β_μ, a β_I <0 is consistent with the theory that the presence of an injunctive norm is likely to work against the tendency for households that consumed below the reported average to raise their consumption. At the same time a β_I <0 is also consistent with the theory that injunctive norms will not alter the direction of the household's change in consumption as it alters its consumption back towards the norm it was reported as having consumed more than. The overriding conclusion from this data though is that the injunctive norm did not significantly influence household consumption generally. This is either because it just simply does not normally act in this way, or (more likely) that the effect is too small to be detected in a sample this size given the large amount of natural variation in consumption.

Table 4.3. Specific differences between mailed treatments by income group - (FE, robust standard errors).

Influence upon Proportional Monthly Consumption

Whole year	All income	Upper income	Middle income	Lower income
β_μ	-0.0007213	-0.0003458	-0.0006839	-0.0008805
	(0.0001177)***	(0.0001682)**	(0.0002275)***	(0.0001956)***
β_N	0.0001091	-0.0000125	0.0003626	-0.0004675
	(0.0001593)	(0.0002305)	(0.0002767)	(0.0003158)
β_I	-0.0002542	-0.0003421	-0.0003331	-0.0001434
	(0.0001747)	(0.0003262)	(0.0003002)	(0.0002895)
β_{NI}	0.0000486	0.0003218	-0.0001548	-0.0000294
	(0.0002317)	(0.0003787)	(0.0003967)	(0.0004685)
Mailings 1-5	All income	Upper income	Middle income	Lower income
β_μ	-0.0001201	-0.0002655	0.0004901	0.0000392
	(0.0001962)	(0.0002073)	(0.0004580)	(0.0003434)
β_N	-0.0002216	0.0001459	-0.0004438	-0.0007852
	(0.0002453)	(0.0003449)	(0.0005053)	(0.0004319)*
β_I	-0.0003956	0.0004359	-0.0009603	-0.0006609
	(0.0002813)	(0.0004020)	(0.0005731)*	(0.0004710)
β_{NI}	0.0005834	-0.0004652	0.0012244	0.0010231
	(0.0003426)*	(0.0005143)	(0.0006480)*	(0.0006219)
Mailings 6-11	All income	Upper income	Middle income	Lower income
β_μ	-0.0004158	-0.0001872	-0.0004637	-0.0003243
	(0.0001654)**	(0.0002501)	(0.0003030)	(0.0002760)
β_N	0.0000150	0.0000859	0.0002187	-0.0007428
	(0.0002305)	(0.0003336)	(0.0003727)	(0.0005198)
β_I	-0.0001526	-0.0005702	-0.0000718	0.0000222
	(0.0002454)	(0.0005024)	(0.0004256)	(0.0003733)
β_{NI}	0.0000038	0.0006461	-0.0002815	-0.0002689
	(0.0003349)	(0.0005709)	(0.0005662)	(0.0007301)

Standard errors shown in brackets underneath coefficient estimates

* $p<0.1$; ** $p<0.05$; *** $p<0.01$

The interaction between the injunctive norm and the social distance regressor appears to have only mattered significantly once, for middle income households in the first half of the year. In that instance the interaction appears to have had the result of raising consumption above what it would normally have been. Why this was the case is difficult to say. We would expect households to react more strongly to injunctive feedback from people more like them. However, perhaps the nature of those others in the community plays some role in how injunctive comparisons with them are received. The middle income suburbs in Cape Town are mostly in the Cape-flats area. The Cape flats is an area that has a long standing and grave problem with inter-personal crime, specifically gangs. It could be that households in this area, when they react to injunctive comparisons to their neighbours, may be more likely to project this anti-social description upon the referenced neighbour and react anti-socially themselves. Again, it is difficult to verify this given our data and the inconsistency of the response over time.

The strongest and most significant influence upon consumption amongst the treatments over the whole year appears to have been the difference between the household's consumption and the reported norm $(HH_{it-1} - \mu_{it-1})$, captured in β_μ. The regression results for the overall year find that $\beta_\mu < 0$ for all income groups. This implies that, as expected, households will tend to move their consumption towards the reported norm; raising it if they are below and reducing it if they are above.

This is an interesting result since it strongly suggests that households do respond to being compared to averages. It implies that moderations in consumption are more likely to be induced by reporting to households that they consumed more than the average, that that they consumed less than the average.

The general pattern of the regression results suggest that this tendency is something that is found in all households. The effect is mainly only observable when considering it over the whole year. This may be a function of a larger sample size that is achieved if one considers the whole year's worth of data rather than just one half of it.

Since this is an element that is common to all of the treatments this is also consistent with the analysis that compared the mailed treatments to the control group. There the most important treatment factor seemed to be the fact that one was merely mailed one of the treatments - it did not matter very much which one.

4.4. Conclusion.

Households in eleven different Cape Town suburbs, spanning the major income brackets found in the city, were randomly allocated to receive one of four different inserts with their electricity bill for 2010. Each of these treatments compared the household's consumption to that of an average household for the same period. These treatments made use of two different degrees of social distance (city vs. neighbourhood) with half including injunctive norms and half not.

Analysis of the consumption of the households in the various treatments did not reveal any significant difference between treatments in terms of the different degrees of social distance, nor did injunctive norms seem to exert any detectable influence upon consumption.

The element that did emerge as a significant influence upon consumption for households in the various treatments was the difference in consumption between the household and the average reported to it. Specifically, households were found to significantly adjust their consumption towards the reported norm. This is consistent with theory and prior work on the topic of social norms generally. It is interesting that this element alone seemed to influence consumption rather than whether the average was for the neighbourhood or the city.

Assessing the difference in consumption between mailed households and households in the control was hampered by the fact that households were not allocated to the control group in the same, random, fashion as households were allocated to the various mailed treatments. Nevertheless some feasible analysis was possible considering that the control households were all drawn from the same suburbs, that consumption was measured in a proportional fashion and that various other covariates such as suburb and climate were controlled for.

Comparisons between the mailed treatments and the control group reveal that there seems to have been a difference in consumption between the mailed and control households. This difference appears to mainly derive from some element that is generic to the mailed treatments and to be exerted in a cumulative fashion. There are two elements that could be said to be generic to all mailed treatments: the tips and the fact that each household is being compared to an average. It is impossible to identify one of these elements as the source of influence to the exclusion of the other in comparisons to the control group. Given the significance of the difference between the household's consumption and the average reported to it amongst mailed households, it is almost certainly a significant factor accounting for differences in proportional consumption of household in the control and mailed groups.

Indications from the second half of the year that city level treatments may have been declining in effectiveness while neighbourhood level treatments were increasing in effectiveness are intriguing. However, analysis of that question is beyond the scope of this project and perhaps beyond the ability of this data to answer.

The most important conclusion that this analysis argues for is the fact that norms appear to be effective influences upon consumption.

Appendix: Chapter 4.

Fixed Effects regressions amongst all income groups for the whole year: Influence upon Annual proportional consumption

	FE	FE (Robust)
CD	-0.4603754	-0.4603754
	(0.0606044)***	(0.0570777)***
CDI	-0.4167257	-0.4167257
	(0.0611908)***	(0.0596710)***
ND	-0.3500760	-0.3500760
	(0.0610709)***	(0.0605511)***
NDI	-0.3946388	-0.3946388
	(0.0599847)***	(0.0626385)***
treated_da_prop_hdd	0.0148949	0.0148949
	(0.0183528)	(0.0082104)*
treated_da_prop_cdd	0.2485411	0.2485411
	(0.7668215)	(0.6522683)
treated_da_prop_rain	-0.0659630	-0.0659630
	(0.0768890)	(0.0336942)*
treated_da_prop_min	6.3695910	6.3695910
	(6.9301890)	(5.3690677)
Athlone	-0.5585746	-0.5585746
	(0.1160115)***	(0.1386295)***
Bonthewul	-0.2313747	-0.2313747
	(0.1098185)**	(0.1273887)*
Constantia	-1.3369619	-1.3369619
	(0.0926212)***	(0.0906847)***
Durbanville	-1.1643265	-1.1643265
	(0.0936875)***	(0.0980095)***
Grassy_Park	-0.6036786	-0.6036786
	(0.1016696)***	(0.1094482)***
Heideveld	-0.0930520	-0.0930520
	(0.1472195)	(0.1801948)
Langa	0.1603058	0.1603058
	(0.1436152)	(0.1715265)
Mannenberg	-0.1751117	-0.1751117
	(0.1268271)	(0.1430885)
Pinelands	-0.5906657	-0.5906657
	(0.1314260)***	(0.1349402)***
Salt_River	-0.4705786	-0.4705786
	(0.1259327)***	(0.1378155)***
Tamboerskloof	-0.1771541	-0.1771541
	(0.1192532)	(0.1303329)
_cons	-3.5927228	-3.5927228
	(5.5116770)	(4.2430277)
R^2	0.07	0.07
N	7,586	7,586

Estimated treatment effects compared to the control group amongst all income groups.

Proportional monthly consumption	RE	RE (Robust)	FE	FE (Robust)
L.CD_pdiff_norm_above	-0.0016533	-0.0016533	-0.0006902	-0.0006902
	(0.0004970)***	(0.0003599)***	(0.0005259)	(0.0003730)*
L.CD_pdiff_norm_below	0.0000001	0.0000001	-0.0000036	-0.0000036
	(0.0000047)	(0.0000017)	(0.0000050)	(0.0000021)*
L.CDI_pdiff_norm_above	-0.0018652	-0.0018652	-0.0007213	-0.0007213
	(0.0004945)***	(0.0003822)***	(0.0005256)	(0.0003789)*
L.CDI_pdiff_norm_below	0.0000133	0.0000133	-0.0000309	-0.0000309
	(0.0000328)	(0.0000534)	(0.0000334)	(0.0000647)
L.ND_pdiff_norm_above	0.0004388	0.0004388	0.0006418	0.0006418
	(0.0002052)**	(0.0004336)	(0.0002063)***	(0.0004291)
L.ND_pdiff_norm_below	-0.0000004	-0.0000004	-0.0000005	-0.0000005
	(0.0000021)	(0.0000014)	(0.0000021)	(0.0000015)
L.NDI_pdiff_norm_above	0.0000250	0.0000250	0.0003552	0.0003552
	(0.0002535)	(0.0004622)	(0.0002559)	(0.0003937)
L.NDI_pdiff_norm_below	-0.0000012	-0.0000012	-0.0000015	-0.0000015
	(0.0000029)	(0.0000052)	(0.0000029)	(0.0000053)
L.cumu_mail	-0.0178553	-0.0178553	-0.0144496	-0.0144496
	(0.0014828)***	(0.0018150)***	(0.0015383)***	(0.0019985)***
prop_precdd_da	0.0126592	0.0126592	0.0152851	0.0152851
	(0.0024852)***	(0.0026176)***	(0.0025040)***	(0.0027057)***
prop_prehdd_da	0.0051427	0.0051427	0.0053493	0.0053493
	(0.0009179)***	(0.0011665)***	(0.0009224)***	(0.0011698)***
prop_premin_da	-1.5308173	-1.5308173	-1.5714521	-1.5714521
	(0.0292724)***	(0.0364089)***	(0.0294771)***	(0.0374904)***
prop_prerain_da	-0.0185216	-0.0185216	-0.0180758	-0.0180758
	(0.0038070)***	(0.0046238)***	(0.0038129)***	(0.0045583)***
Athlone	-0.6771813	-0.6771813		
	(0.0518695)***	(0.0611158)***		
Bonthewul	-0.2429845	-0.2429845		
	(0.0639940)***	(0.0784372)***		
Constantia	-1.1035623	-1.1035623		
	(0.0553768)***	(0.0564456)***		
Durbanville	-0.9320356	-0.9320356		
	(0.0450054)***	(0.0524888)***		
Grassy_Park	-0.6021489	-0.6021489		
	(0.0539989)***	(0.0632107)***		
Heideveld	-0.4014742	-0.4014742		
	(0.0715722)***	(0.0803843)***		
Langa	-0.1328928	-0.1328928		
	(0.0706041)*	(0.0856409)		
Mannenberg	-0.3224151	-0.3224151		
	(0.0604190)***	(0.0732561)***		
Pinelands	-0.6857512	-0.6857512		
	(0.0558561)***	(0.0635942)***		
Salt_River	-0.5299525	-0.5299525		
	(0.0577589)***	(0.0699582)***		
Tamboerskloof	-0.2522311	-0.2522311		
	(0.0557898)***	(0.0678436)***		
_cons	3.1817001	3.1817001	2.5391980	2.5391980
	(0.0454512)***	(0.0606398)***	(0.0246024)***	(0.0321652)***
Chi2	7,043.59	3,721.72		
F			486.92	254.37
P	0.00	0.00	0.00	0.00.
R^2	.	.	0.10	0.10
R2_O	0.09	0.09	0.02	0.02
R2_B	0.09	0.09	0.01	0.01
R2_W	0.10	0.10	0.10	0.10
Rho	0.76	0.76	0.80	0.80
θ_50	0.81	0.81		
N	67,534	67,534	67,534	67,534
N_Clusters	.	7,575	.	7,575

* $p<0.1$; ** $p<0.05$; *** $p<0.01$

Estimated treatment effects compared to the control group amongst all income groups.

Proportional monthly consumption	RE	RE (Robust)	FE	FE (Robust)
L.mailed	-0.2142120	-0.2142120		
	(0.0257760)***	(0.0293448)***		
L.cumu_mail	-0.0138290	-0.0138290	-0.0140950	-0.0140950
	(0.0015345)***	(0.0019748)***	(0.0015309)***	(0.0019814)***
prop_precdd_da	0.0157378	0.0157378	0.0152936	0.0152936
	(0.0025113)***	(0.0027076)***	(0.0025042)***	(0.0027053)***
prop_prehdd_da	0.0050194	0.0050194	0.0053652	0.0053652
	(0.0009169)***	(0.0011587)***	(0.0009224)***	(0.0011711)***
prop_premin_da	-1.5634955	-1.5634955	-1.5723880	-1.5723880
	(0.0295143)***	(0.0374749)***	(0.0294667)***	(0.0374860)***
prop_prerain_da	-0.0190173	-0.0190173	-0.0181527	-0.0181527
	(0.0038014)***	(0.0046410)***	(0.0038126)***	(0.0045603)***
Athlone	-0.7229919	-0.7229919		
	(0.0525622)***	(0.0614346)***		
Bonthewul	-0.2470183	-0.2470183		
	(0.0645343)***	(0.0786556)***		
Constantia	-1.1162290	-1.1162290		
	(0.0556239)***	(0.0567772)***		
Durbanville	-1.0297873	-1.0297873		
	(0.0467672)***	(0.0537226)***		
Grassy_Park	-0.6014925	-0.6014925		
	(0.0544487)***	(0.0635175)***		
Heideveld	-0.4181269	-0.4181269		
	(0.0721934)***	(0.0809271)***		
Langa	-0.1434983	-0.1434983		
	(0.0712065)**	(0.0860030)*		
Mannenberg	-0.3378661	-0.3378661		
	(0.0609513)***	(0.0736847)***		
Pinelands	-0.6857022	-0.6857022		
	(0.0563175)***	(0.0639346)***		
Salt_River	-0.5931134	-0.5931134		
	(0.0587302)***	(0.0710141)***		
Tamboerskloof	-0.3065262	-0.3065262		
	(0.0566214)***	(0.0675173)***		
_cons	3.3480668	3.3480668	2.5383583	2.5383583
	(0.0500257)***	(0.0661902)***	(0.0245265)***	(0.0321000)***
Chi2	7,088.78	3,731.16		
F			1,262.45	659.49
P	0.00	0.00	0.00	0.00
R^2	.	.	0.10	0.10
R2_O	0.09	0.09	0.01	0.01
R2_B	0.09	0.09	0.01	0.01
R2_W	0.10	0.10	0.10	0.10
Rho	0.76	0.76	0.80	0.80
θ_50	0.82	0.82		
N	67,534	67,534	67,534	67,534
N_Clusters	.	7,575	.	7,575

* $p<0.1$; ** $p<0.05$; *** $p<0.01$

Estimated treatment effects compared to the control group amongst upper income suburbs

Proportional monthly consumption	RE	RE (Robust)	FE	FE (Robust)	FE (Robust) mailings 1-5	FE (Robust) mailings 6-11
L.CD_pdiff_norm_above	-0.0022624	-0.0022624	-0.0008931	-0.0008931	-0.0007879	-0.0010019
	(0.0005907)***	(0.0004532)***	(0.0006486)	(0.0005228)*	(0.0012316)	(0.0006232)
L.CD_pdiff_norm_below	-0.0000007	-0.0000007	-0.0000046	-0.0000046	0.0001262	-0.0000058
	(0.0000046)	(0.0000008)	(0.0000049)	(0.0000009)***	(0.0001707)	(0.0000227)
L.CDI_pdiff_norm_above	-0.0020209	-0.0020209	-0.0005681	-0.0005681	-0.0002470	-0.0002413
	(0.0005506)***	(0.0003528)***	(0.0005999)	(0.0003573)	(0.0007166)	(0.0004257)
L.CDI_pdiff_norm_below	-0.0000600	-0.0000600	-0.0000839	-0.0000839	-0.0001167	-0.0001203
	(0.0000445)	(0.0000860)	(0.0000449)*	(0.0000961)	(0.0000867)	(0.0001307)
L.ND_pdiff_norm_above	-0.0002791	-0.0002791	0.0003241	0.0003241	-0.0006301	-0.0001075
	(0.0004758)	(0.0003546)	(0.0004879)	(0.0003444)	(0.0006002)	(0.0005142)
L.ND_pdiff_norm_below	-0.0000003	-0.0000003	-0.0000003	-0.0000003	-0.0000130	-0.0000005
	(0.0000019)	(0.0000014)	(0.0000019)	(0.0000014)	(0.0000138)	(0.0000022)
L.NDI_pdiff_norm_above	-0.0003468	-0.0003468	0.0002705	0.0002705	-0.0005407	-0.0003540
	(0.0004500)	(0.0004135)	(0.0004609)	(0.0004146)	(0.0006388)	(0.0004850)
L.NDI_pdiff_norm_below	0.0000004	0.0000004	0.0000005	0.0000005	-0.0000057	0.0000017
	(0.0000026)	(0.0000052)	(0.0000026)	(0.0000052)	(0.0000554)	(0.0000048)
L.cumu_mail	-0.0100302	-0.0100302	-0.0025020	-0.0025020	-0.0513893	0.0201796
	(0.0020647)***	(0.0021929)***	(0.0021519)	(0.0023882)	(0.0055049)***	(0.0070510)***
prop_precdd_da	0.0149075	0.0149075	0.0198539	0.0198539	0.0230263	-0.2719996
	(0.0029808)***	(0.0031095)***	(0.0029992)***	(0.0031801)***	(0.0029208)***	(0.0145685)***
prop_prehdd_da	0.0047694	0.0047694	0.0048200	0.0048200	0.0075588	0.0029340
	(0.0010211)***	(0.0013059)***	(0.0010270)***	(0.0012824)***	(0.0026462)***	(0.0014075)**
prop_premin_da	-1.2993180	-1.2993180	-1.3552365	-1.3552365	-1.3095370	-0.4598688
	(0.0357028)***	(0.0451237)***	(0.0358932)***	(0.0462127)***	(0.0720607)***	(0.0632938)***
prop_prerain_da	-0.0090408	-0.0090408	-0.0085695	-0.0085695	-0.0047403	0.0475790
	(0.0044348)**	(0.0055237)	(0.0044520)*	(0.0053847)	(0.0075723)	(0.0127698)***
Constantia	-0.8667151	-0.8667151				
	(0.0495228)***	(0.0557263)***				
Durbanville	-0.6737763	-0.6737763				
	(0.0401451)***	(0.0506204)***				
Pinelands	-0.4396939	-0.4396939				
	(0.0497141)***	(0.0630584)***				
_cons	2.7025315	2.7025315	2.1303159	2.1303159	2.1256906	1.3409126
	(0.0462329)***	(0.0674031)***	(0.0303199)***	(0.0392344)***	(0.0660903)***	(0.0521599)***
Chi2	3,028.81	1,631.52				
F			202.28	114.87	42.30	96.54
P	0.00	0.00	0.00	0.00	0.00	0.00
R^2	.	.	0.08	0.08	0.06	0.15
R2_O	0.10	0.10	0.02	0.02	0.04	0.01
R2_B	0.10	0.10	0.00	0.00	0.03	0.05
R2_W	0.08	0.08	0.08	0.08	0.06	0.15
Rho	0.74	0.74	0.79	0.79	0.85	0.79
θ_50	0.81	0.81				

N	33,422	33,422	33,422	33,422	15,118	18,304
N_Clusters	.	3,621	.	3,621	3,603	3,588

* $p<0.1$; ** $p<0.05$; *** $p<0.01$

Estimated treatment effects compared to the control group amongst upper income suburbs

Proportional monthly consumption	RE	RE (Robust)	FE	FE (Robust)	FE (Robust) mailings 1-5	FE (Robust) mailings 6-11
L.mailed	-0.4277227	-0.4277227				
	(0.0337651)***	(0.0368638)***				
L.cumu_mail	-0.0019171	-0.0019171	-0.0021672	-0.0021672	-0.0510011	0.0206217
	(0.0021347)	(0.0023536)	(0.0021326)	(0.0023591)	(0.0054775)***	(0.0069725)***
prop_precdd_da	0.0200855	0.0200855	0.0198609	0.0198609	0.0229607	-0.2717924
	(0.0030038)***	(0.0031842)***	(0.0029990)***	(0.0031794)***	(0.0029034)***	(0.0145442)***
prop_prehdd_da	0.0045458	0.0045458	0.0048320	0.0048320	0.0075784	0.0029270
	(0.0010185)***	(0.0012878)***	(0.0010267)***	(0.0012829)***	(0.0026464)***	(0.0014072)**
prop_premin_da	-1.3522176	-1.3522176	-1.3547150	-1.3547150	-1.3090147	-0.4610135
	(0.0358793)***	(0.0461924)***	(0.0358666)***	(0.0461465)***	(0.0719617)***	(0.0632585)***
prop_prerain_da	-0.0100475	-0.0100475	-0.0086151	-0.0086151	-0.0048201	0.0477620
	(0.0044227)**	(0.0055648)*	(0.0044503)*	(0.0053825)	(0.0075661)	(0.0127521)***
Constantia	-0.7647010	-0.7647010				
	(0.0503748)***	(0.0544563)***				
Durbanville	-0.7559277	-0.7559277				
	(0.0407431)***	(0.0501872)***				
Pinelands	-0.3307456	-0.3307456				
	(0.0506671)***	(0.0620568)***				
oL.mailed			0.0000000	0.0000000	0.0000000	0.0000000
			(0.0000000)	(0.0000000)	(0.0000000)	(0.0000000)
o.Constantia			0.0000000	0.0000000	0.0000000	0.0000000
			(0.0000000)	(0.0000000)	(0.0000000)	(0.0000000)
o.Durbanville			0.0000000	0.0000000	0.0000000	0.0000000
			(0.0000000)	(0.0000000)	(0.0000000)	(0.0000000)
o.Pinelands			0.0000000	0.0000000	0.0000000	0.0000000
			(0.0000000)	(0.0000000)	(0.0000000)	(0.0000000)
_cons	2.9139515	2.9139515	2.1264414	2.1264414	2.1213535	1.3366799
	(0.0494326)***	(0.0720858)***	(0.0301223)***	(0.0390111)***	(0.0659265)***	(0.0519655)***
Chi2	3,171.18	1,672.50			.	.
F			524.39	292.45		
P	0.00	0.00	0.00	0.00	.	.
R^2	.	.	0.08	0.08	0.06	0.15
R2_O	0.12	0.12	0.02	0.02	0.03	0.01
R2_B	0.13	0.13	0.00	0.00	0.03	0.05
R2_W	0.08	0.08	0.08	0.08	0.06	0.15
Rho	0.74	0.74	0.79	0.79	0.85	0.79
θ_50	0.81	0.81			.	.
N	33,422	33,422	33,422	33,422	15,118	18,304
N_Clust	.	3,621	.	3,621	3,603	3,588

* $p<0.1$; ** $p<0.05$; *** $p<0.01$

Estimated treatment effects compared to the control group amongst middle income suburbs

Proportional monthly consumption	RE	RE (Robust)	FE	FE (Robust)	FE (Robust) mailings 1-5	FE (Robust) mailings 6-11
L.CD_pdiff_norm_above	-0.0009893	-0.0009893	-0.0001081	-0.0001081	0.0008386	0.0003457
	(0.0009410)	(0.0006284)	(0.0009598)	(0.0006161)	(0.0007228)	(0.0008743)
L.CD_pdiff_norm_below	0.0000062	0.0000062	0.0000047	0.0000047	0.0000002	0.0000067
	(0.0000125)	(0.0000060)	(0.0000125)	(0.0000063)	(0.0000062)	(0.0000013)***
L.CDI_pdiff_norm_above	-0.0020137	-0.0020137	-0.0007177	-0.0007177	-0.0008846	-0.0000264
	(0.0010787)*	(0.0008670)**	(0.0011155)	(0.0008035)	(0.0009301)	(0.0017331)
L.CDI_pdiff_norm_below	0.0000937	0.0000937	0.0000540	0.0000540	-0.0012571	0.0000235
	(0.0000540)*	(0.0000379)**	(0.0000552)	(0.0000601)	(0.0003296)***	(0.0000723)
L.ND_pdiff_norm_above	-0.0014228	-0.0014228	-0.0008124	-0.0008124	-0.0018099	-0.0011725
	(0.0007516)*	(0.0006545)**	(0.0007671)	(0.0006772)	(0.0010924)*	(0.0008868)
L.ND_pdiff_norm_below	-0.0000021	-0.0000021	-0.0000128	-0.0000128	0.0000034	-0.0001160
	(0.0000220)	(0.0000177)	(0.0000223)	(0.0000187)	(0.0000413)	(0.0000772)
L.NDI_pdiff_norm_above	-0.0023537	-0.0023537	-0.0014233	-0.0014233	-0.0015025	-0.0013936
	(0.0007754)***	(0.0005422)***	(0.0007949)*	(0.0005779)**	(0.0012281)	(0.0007025)**
L.NDI_pdiff_norm_below	-0.0000368	-0.0000368	-0.0000453	-0.0000453	0.0000422	-0.0003024
	(0.0000224)	(0.0000288)	(0.0000224)**	(0.0000318)	(0.0000112)***	(0.0001076)***
L.cumu_mail	-0.0165133	-0.0165133	-0.0131154	-0.0131154	0.0060221	0.0170021
	(0.0026807)***	(0.0031683)***	(0.0027814)***	(0.0035639)***	(0.0092222)	(0.0097211)*
prop_precdd_da	0.0138803	0.0138803	0.0164152	0.0164152	-0.0002787	-0.2961602
	(0.0046780)***	(0.0053886)***	(0.0047139)***	(0.0056007)***	(0.0044889)	(0.0178393)***
prop_prehdd_da	0.0066059	0.0066059	0.0069268	0.0069268	-0.0127866	-0.0010382
	(0.0020218)***	(0.0032358)**	(0.0020291)***	(0.0032707)**	(0.0048857)***	(0.0028680)
prop_premin_da	-1.6649417	-1.6649417	-1.7100497	-1.7100497	-0.3061280	-0.8600320
	(0.0547515)***	(0.0747868)***	(0.0552035)***	(0.0775084)***	(0.1190236)**	(0.1033607)***
prop_prerain_da	-0.0285492	-0.0285492	-0.0283367	-0.0283367	0.0867706	0.0577350
	(0.0078399)***	(0.0115654)**	(0.0078142)***	(0.0114093)**	(0.0158960)***	(0.0209151)***
Athlone	-0.0798457	-0.0798457				
	(0.0541838)	(0.0544947)				
Bonthewul	0.3604601	0.3604601				
	(0.0671586)***	(0.0736561)***				
Heideveld	0.1961986	0.1961986				
	(0.0752316)***	(0.0755637)***				
Salt_River	0.0699531	0.0699531				
	(0.0605137)	(0.0642853)				
_cons	2.7066898	2.7066898	2.7179556	2.7179556	1.4719583	1.9194155
	(0.0606263)***	(0.0862220)***	(0.0463344)***	(0.0706073)***	(0.1132939)***	(0.0838148)***
Chi2	2,308.21	1,211.43				
F			178.62	91.80	18.93	90.88
P	0.00	0.00	0.00	0.00.	0.00.	0.00.
R^2	.	.	0.10	0.10	0.05	0.20
R2_O	0.03	0.03	0.02	0.02	0.00	0.03
R2_B	0.01	0.01	0.01	0.01	0.02	0.02
R2_W	0.10	0.10	0.10	0.10	0.05	0.20
Rho	0.77	0.77	0.80	0.80	0.89	0.79

θ_50	0.82	0.82				
N	22,435	22,435	22,435	22,435	10,039	12,396
N_Clust	.	2,607	.	2,607	2,590	2,582

* $p<0.1$; ** $p<0.05$; *** $p<0.01$

Estimated treatment effects compared to the control group amongst middle income suburbs

Proportional monthly consumption	RE	RE (Robust)	FE	FE (Robust)	FE (Robust) mailings 1-5	FE (Robust) mailings 6-11
L.mailed	-0.1838915	-0.1838915				
	(0.0448708)***	(0.0524441)***				
L.cumu_mail	-0.0133031	-0.0133031	-0.0132245	-0.0132245	0.0052152	0.0127483
	(0.0027753)***	(0.0035528)***	(0.0027658)***	(0.0035639)***	(0.0091355)	(0.0097458)
prop_precdd_da	0.0168302	0.0168302	0.0163909	0.0163909	-0.0003961	-0.2999446
	(0.0047309)***	(0.0056085)***	(0.0047135)***	(0.0056017)***	(0.0044699)	(0.0178241)***
prop_prehdd_da	0.0064486	0.0064486	0.0069483	0.0069483	-0.0127999	-0.0010133
	(0.0020186)***	(0.0032131)**	(0.0020288)***	(0.0032715)**	(0.0049090)***	(0.0028699)
prop_premin_da	-1.6982161	-1.6982161	-1.7118769	-1.7118769	-0.3036414	-0.8434660
	(0.0552855)***	(0.0774578)***	(0.0551583)***	(0.0774133)***	(0.1194088)**	(0.1034701)***
prop_prerain_da	-0.0289310	-0.0289310	-0.0285797	-0.0285797	0.0873070	0.0570990
	(0.0078172)***	(0.0115698)**	(0.0078080)***	(0.0114284)**	(0.0160400)***	(0.0208859)***
Athlone	-0.1172701	-0.1172701				
	(0.0558915)**	(0.0556590)**				
Bonthewul	0.3549936	0.3549936				
	(0.0682087)***	(0.0735949)***				
Heideveld	0.1829671	0.1829671				
	(0.0764816)**	(0.0763123)**				
Salt_River	0.0166464	0.0166464				
	(0.0629916)	(0.0674485)				
_cons	2.8491350	2.8491350	2.7163321	2.7163321	1.4561097	1.9093002
	(0.0713461)***	(0.1037314)***	(0.0462502)***	(0.0705094)***	(0.1136898)***	(0.0837353)***
Chi2	2,312.04	1,199.12				
F			462.43	237.12	42.22	232.59
P	0.00	0.00	.	.	.	.
R^2	.	.	0.10	0.10	0.05	0.20
R2_O	0.03	0.03	0.01	0.01	0.00	0.03
R2_B	0.01	0.01	0.02	0.02	0.01	0.03
R2_W	0.10	0.10	0.10	0.10	0.05	0.20
Rho	0.77	0.77	0.80	0.80	0.89	0.79
Thta_50	0.82	0.82				
N	22,435	22,435	22,435	22,435	10,039	12,396
N_Clusters	.	2,607	.	2,607	2,590	2,582

* $p<0.1$; ** $p<0.05$; *** $p<0.01$

Estimated treatment effects compared to the control group amongst Lower income suburbs

Proportional monthly consumption	RE	RE (Robust)	FE	FE (Robust)	FE (Robust) mailings 1-5	FE (Robust) mailings 6-11
L.CD_pdiff_norm_above	0.0007860	0.0007860	0.0009797	0.0009797	-0.0023178	0.0049517
	(0.0016018)	(0.0009206)	(0.0016191)	(0.0008507)	(0.0007938)***	(0.0015662)***
L.CD_pdiff_norm_below	0.0000218	0.0000218	-0.0000264	-0.0000264	0.0000332	-0.0000448
	(0.0000567)	(0.0000427)	(0.0000574)	(0.0000303)	(0.0001128)	(0.0000324)
L.CDI_pdiff_norm_above	0.0004186	0.0004186	0.0010823	0.0010823	-0.0000022	0.0029269
	(0.0016443)	(0.0016099)	(0.0016690)	(0.0014840)	(0.0027416)	(0.0015426)*
L.CDI_pdiff_norm_below	0.0000083	0.0000083	-0.0000950	-0.0000950	-0.0000863	-0.0001264
	(0.0000942)	(0.0000704)	(0.0000980)	(0.0000783)	(0.0000425)**	(0.0001765)
L.ND_pdiff_norm_above	0.0006798	0.0006798	0.0007589	0.0007589	0.0014679	0.0006797
	(0.0002824)**	(0.0005378)	(0.0002816)***	(0.0005369)	(0.0016712)	(0.0004873)
L.ND_pdiff_norm_below	-0.0000203	-0.0000203	-0.0000509	-0.0000509	0.0000158	-0.0000095
	(0.0000461)	(0.0000602)	(0.0000461)	(0.0000624)	(0.0000686)	(0.0000804)
L.NDI_pdiff_norm_above	0.0004872	0.0004872	0.0005923	0.0005923	-0.0027034	0.0007274
	(0.0003848)	(0.0005545)	(0.0003845)	(0.0005075)	(0.0020851)	(0.0007746)
L.NDI_pdiff_norm_below	-0.0000234	-0.0000234	-0.0000315	-0.0000315	0.0000152	0.0000413
	(0.0000290)	(0.0000476)	(0.0000294)	(0.0000511)	(0.0000523)	(0.0000710)
L.cumu_mail	-0.0259588	-0.0259588	-0.0284823	-0.0284823	0.0012247	0.0067839
	(0.0037637)***	(0.0045352)***	(0.0038915)***	(0.0049931)***	(0.0140139)	(0.0154729)
prop_precdd_da	0.0095754	0.0095754	0.0055699	0.0055699	-0.0216481	-0.5672197
	(0.0080194)	(0.0082182)	(0.0081028)	(0.0085833)	(0.0074708)***	(0.0376847)***
prop_prehdd_da	0.0058847	0.0058847	0.0066982	0.0066982	-0.0225455	-0.0046222
	(0.0031092)*	(0.0034737)*	(0.0031141)**	(0.0037265)*	(0.0075816)***	(0.0028205)
prop_premin_da	-2.0531916	-2.0531916	-2.0383826	-2.0383826	-0.3969800	-0.5688061
	(0.0921299)***	(0.1088284)***	(0.0929763)***	(0.1130395)***	(0.2025962)*	(0.1932608)***
prop_prerain_da	-0.0486775	-0.0486775	-0.0490525	-0.0490525	0.0929419	0.0239576
	(0.0116877)***	(0.0149951)***	(0.0116717)***	(0.0153748)***	(0.0292552)***	(0.0238352)
Langa	0.1936195	0.1936195				
	(0.0871726)**	(0.0899362)**				
Gugulethu	0.3435957	0.3435957				
	(0.0701713)***	(0.0738062)***				
_cons	3.3657466	3.3657466	3.4662086	3.4662086	1.9715822	2.1392295
	(0.0919955)***	(0.1152972)***	(0.0751569)***	(0.0958568)***	(0.1976485)***	(0.1382125)***
Chi2	1,392.87	712.09				
F			112.00	55.54	13.55	56.00
P	0.00	0.00	0.00	0.00	0.00	0.00
R^2	.	.	0.12	0.12	0.06	0.24
R2_O	0.02	0.02	0.01	0.01	0.00	0.04
R2_B	0.00	0.00	0.04	0.04	0.02	0.02
R2_W	0.12	0.12	0.12	0.12	0.06	0.24
Rho	0.74	0.74	0.78	0.78	0.87	0.77
θ_50	0.81	0.81				
N	11,677	11,677	11,677	11,677	5,185	6,492

N_Clusters	.	1,347	.	1,347	1,328	1,341

* $p<0.1$; ** $p<0.05$; *** $p<0.01$

Estimated treatment effects compared to the control group amongst Lower income suburbs

Proportional monthly consumption						
L.mailed	0.1740466	0.1740466				
	(0.0703160)**	(0.0777762)**				
L.cumu_mail	-0.0281271	-0.0281271	-0.0282909	-0.0282909	0.0001791	0.0069812
	(0.0038845)***	(0.0049597)***	(0.0038685)***	(0.0049747)***	(0.0138224)	(0.0154877)
prop_precdd_da	0.0056971	0.0056971	0.0052487	0.0052487	-0.0220768	-0.5620274
	(0.0081380)	(0.0085678)	(0.0081029)	(0.0085743)	(0.0074699)***	(0.0368574)***
prop_prehdd_da	0.0060833	0.0060833	0.0066057	0.0066057	-0.0225196	-0.0051837
	(0.0031022)**	(0.0035250)*	(0.0031144)**	(0.0037084)*	(0.0075761)***	(0.0028237)*
prop_premin_da	-2.0305087	-2.0305087	-2.0567594	-2.0567594	-0.3890042	-0.5832869
	(0.0928654)***	(0.1115578)***	(0.0925966)***	(0.1120951)***	(0.2024317)*	(0.1913621)***
prop_prerain_da	-0.0484422	-0.0484422	-0.0488224	-0.0488224	0.0932498	0.0341084
	(0.0116456)***	(0.0150112)***	(0.0116607)***	(0.0152896)***	(0.0293164)***	(0.0241611)
Langa	0.1874689	0.1874689				
	(0.0886757)**	(0.0900323)**				
Gugulethu	0.3333762	0.3333762				
	(0.0715525)***	(0.0736620)***				
oL.mailed			0.0000000	0.0000000	0.0000000	0.0000000
			(0.0000000)	(0.0000000)	(0.0000000)	(0.0000000)
o.Langa			0.0000000	0.0000000	0.0000000	0.0000000
			(0.0000000)	(0.0000000)	(0.0000000)	(0.0000000)
o.Gugulethu			0.0000000	0.0000000	0.0000000	0.0000000
			(0.0000000)	(0.0000000)	(0.0000000)	(0.0000000)
_cons	3.2468374	3.2468374	3.4810916	3.4810916	1.9642564	2.1466147
	(0.1069089)***	(0.1323655)***	(0.0747584)***	(0.0952519)***	(0.1977468)***	(0.1362860)***
Chi2	1,399.02	710.66				
P	0.00	0.00	.	.	.	.
R^2	.	.	0.12	0.12	0.06	0.24
R2_O	0.02	0.02	0.01	0.01	0.00	0.04
R2_B	0.00	0.00	0.03	0.03	0.02	0.02
R2_W	0.12	0.12	0.12	0.12	0.06	0.24
Sigma_U	1.12	1.12	1.21	1.21	1.26	1.23
Sigma_E	0.65	0.65	0.65	0.65	0.49	0.68
Rho	0.75	0.75	0.78	0.78	0.87	0.77
Thta_5	0.72	0.72				
Thta_50	0.81	0.81				
Thta_95	0.83	0.83				
N	11,677	11,677	11,677	11,677	5,185	6,492

N_Clust	.	1,347	.	1,347	1,328	1,341
F			288.09	143.35	31.18	144.60
Tcon			.	.	.	.

* $p<0.1$; ** $p<0.05$; *** $p<0.01$

Specific differences between mailed treatments: All income groups - (FE, robust standard errors). Influence upon Proportional Monthly Consumption

	Whole year	Mailings 1-5	Mailings 6-11
pdiff_norm	-0.0007213	-0.0001201	-0.0004158
	(0.0001177)***	(0.0001962)	(0.0001654)**
mru_pdiff_norm	0.0001091	-0.0002216	0.0000150
	(0.0001593)	(0.0002453)	(0.0002305)
injunctive_pdiff_norm	-0.0002542	-0.0003956	-0.0001526
	(0.0001747)	(0.0002813)	(0.0002454)
mru_inj_pdiff_norm	0.0000486	0.0005834	0.0000038
	(0.0002317)	(0.0003426)*	(0.0003349)
prop_precdd_da	0.0333289	-0.0000907	-0.3191215
	(0.0031592)***	(0.0035133)	(0.0225250)***
prop_prehdd_da	0.0430438	-0.0123299	-0.1428050
	(0.0050491)***	(0.0117132)	(0.0474347)***
prop_premin_da	-1.4821961	0.0318719	-1.8440373
	(0.0578706)***	(0.0938789)	(0.3837425)***
prop_prerain_da	-0.0496510	0.1389961	0.0787651
	(0.0077874)***	(0.0186980)***	(0.0195040)***
_cons	2.2823019	0.9677151	2.9839587
	(0.0563045)***	(0.0846892)***	(0.4360729)***
F	205.65	46.01	206.51
P	0.00.	0.00	0.00.
R^2	0.12	0.05	0.20
R2_O	0.05	0.03	0.05
R2_B	0.04	0.04	0.00
R2_W	0.12	0.05	0.20
Rho	0.80	0.89	0.78
N	31,040	13,699	17,341
N_Clusters	4,047	4,020	3,999

Specific differences between mailed treatments: Upper income suburbs

(FE, robust standard errors).

Influence upon Proportional Monthly Consumption

	Whole year	Mailings 1-5	Mailings 6-11
pdiff_norm	-0.0003458	-0.0002655	-0.0001872
	(0.0001682)**	(0.0002073)	(0.0002501)
mru_pdiff_norm	-0.0000125	0.0001459	0.0000859
	(0.0002305)	(0.0003449)	(0.0003336)
injunctive_pdiff_norm	-0.0003421	0.0004359	-0.0005702
	(0.0003262)	(0.0004020)	(0.0005024)
mru_inj_pdiff_norm	0.0003218	-0.0004652	0.0006461
	(0.0003787)	(0.0005143)	(0.0005709)
prop_precdd_da	0.0318981	0.0072799	-0.1710699
	(0.0033479)***	(0.0041550)*	(0.0244153)***
prop_prehdd_da	0.0665614	0.0905841	0.1146552
	(0.0059742)***	(0.0204663)***	(0.0419288)***
prop_premin_da	-0.8396733	0.0398571	0.5177546
	(0.0672913)***	(0.1100682)	(0.3310459)
prop_prerain_da	-0.0630082	-0.0465951	0.1049686
	(0.0100631)***	(0.0293215)	(0.0264549)***
_cons	1.3799058	0.6218178	0.0275437
	(0.0636631)***	(0.0985123)***	(0.3782728)
F	70.32	10.99	69.61
P	0.00.	0.00	0.00.
R^2	0.09	0.03	0.17
R2_O	0.05	0.01	0.04
R2_B	0.05	0.01	0.01
R2_W	0.09	0.03	0.17
Rho	0.78	0.89	0.76
N	12,460	5,566	6,894
N_Clusters	1,615	1,607	1,591

Specific differences between mailed treatments: Middle income suburbs -

(FE, robust standard errors).

Influence upon Proportional Monthly Consumption

	Whole year	Mailings 1-5	Mailings 6-11
pdiff_norm	-0.0006839	0.0004901	-0.0004637
	(0.0002275)***	(0.0004580)	(0.0003030)
mru_pdiff_norm	0.0003626	-0.0004438	0.0002187
	(0.0002767)	(0.0005053)	(0.0003727)
injunctive_pdiff_norm	-0.0003331	-0.0009603	-0.0000718
	(0.0003002)	(0.0005731)*	(0.0004256)
mru_inj_pdiff_norm	-0.0001548	0.0012244	-0.0002815
	(0.0003967)	(0.0006480)*	(0.0005662)
prop_precdd_da	0.0301443	0.0052113	-0.2703780
	(0.0058707)***	(0.0063611)	(0.0294287)***
prop_prehdd_da	0.0254833	-0.0228505	-0.0984974
	(0.0082048)***	(0.0147671)	(0.0504901)*
prop_premin_da	-1.6625714	0.1447758	-1.6498869
	(0.0979623)***	(0.1513562)	(0.4215292)***
prop_prerain_da	-0.0207817	0.1968309	0.1113945
	(0.0122762)*	(0.0269313)***	(0.0317221)***
_cons	2.4886109	0.9014110	2.7183182
	(0.0960158)***	(0.1377686)***	(0.4764602)***
F	87.07	24.01	87.76
P	0.00.	0.00	0.00.
R^2	0.13	0.07	0.22
R2_O	0.04	0.00	0.05
R2_B	0.02	0.00	0.00
R2_W	0.13	0.07	0.22
Rho	0.79	0.89	0.78
N	11,208	4,892	6,316
N_Clusters	1,510	1,503	1,491

Specific differences between mailed treatments: Lower income suburbs -

(FE, robust standard errors).

Influence upon Proportional Monthly Consumption

	Whole year	Mailings 1-5	Mailings 6-11
pdiff_norm	-0.0008805	0.0000392	-0.0003243
	(0.0001956)***	(0.0003434)	(0.0002760)
mru_pdiff_norm	-0.0004675	-0.0007852	-0.0007428
	(0.0003158)	(0.0004319)*	(0.0005198)
injunctive_pdiff_norm	-0.0001434	-0.0006609	0.0000222
	(0.0002895)	(0.0004710)	(0.0003733)
mru_inj_pdiff_norm	-0.0000294	0.0010231	-0.0002689
	(0.0004685)	(0.0006219)	(0.0007301)
prop_precdd_da	0.0328168	-0.0050944	-0.6126694
	(0.0086580)***	(0.0097516)	(0.0458215)***
prop_prehdd_da	0.0416970	-0.1172891	-0.5492301
	(0.0130525)***	(0.0304106)***	(0.0791477)***
prop_premin_da	-2.3161956	0.1113460	-5.3821122
	(0.1538370)***	(0.2254244)	(0.7284144)***
prop_prerain_da	-0.1401341	0.3730304	-0.0246419
	(0.0224943)***	(0.0574557)***	(0.0423513)
_cons	3.5503699	1.3723761	7.5172112
	(0.1524145)***	(0.2087046)***	(0.8028071)***
F	63.64	21.15	61.57
P	0.00.	0.00	0.00.
R^2	0.14	0.09	0.25
R2_O	0.05	0.02	0.06
R2_B	0.00	0.00	0.00
R2_W	0.14	0.09	0.25
Rho	0.78	0.87	0.77
N	7,372	3,241	4,131
N_Clusters	922	910	917

Chapter 5.

Conclusion.

Most of the electricity consumed by households in South Africa is consumed by households that are in one of the major municipalities (such as Johannesburg, Pretoria, Port Elizabeth, Durban, Bloemfontein, Pietermaritzburg and Cape Town). Many of these households consume electricity that is bought on a credit basis, receiving a bill for the electricity consumed. This bill represents an opportunity for Eskom and the municipalities (who often buy electricity from Eskom and sell it on to households) to manage household electricity consumption down. One of the methods for doing this that is beginning to be investigated around the world, but most notably in the USA, is to report to households the average level of electricity of consumption for a household and then to compare the households to that average. This approach has been found to be successful in the USA (Alcott, 2011 & Schultz *et al*, 2007).

Originally this study also sought to examine two possible determinants of the effectiveness of reporting social norms and whether a similar strategy would prove successful in a major South African municipality - Cape Town - where many households are likely to be much poorer than the households in the USA studies. These two possible determinants were the social distance between the household and the norm to which they were compared and the communication of social approval or dissaproval - via injunctive norms.

However during the course of the fieldwork interest grew in what the difference in consumption between households receiving the mail and those not receiving any mails at all would be. The ecological cost of generating electricity in South Africa is great. Given the long lead time to changing from a coal-intensive energy mix to one dominated by cleaner sources of energy, this is likely to remain the case for a long time. At the same time, the electricity generation and distribution infrastructure in South Africa is capacity constrained, with the threat of black outs often looming. For these two reasons assessing the degree to which household electricity consumption could be reduced though this sort of reporting is desirable.

Comparisons between the various mailed treatments do not provide evidence that either the injunctive norm or the two degrees of social distance employed in this study made any significant difference to consumption. However, they do provide evidence that households respond in a significant fashion, when made aware of their own consumption via social comparisons. This is consistent with findings from earlier work in the field. This finding suggests that one of the most

important elements of norms feedback that households respond to is this difference itself, generic to all treatments in this project, rather than other elements that surround it.

In comparisons with the unmailed control group this study noted what seemed to feasibly be a treatment effect. Households in the unmailed control group consumed more (as a proportion of their pre-treatment consumption) than treated households. Lower consumption amongst treated households seemed to be a function of some element that was generic to all of the mailed treatments (either tips, or difference in consumption compared to a norm, or both) and which worked cumulatively.

What remains the most astounding result of this research, as well as other research like it (Allcott, 2011, Schultz 2007, Violette et al, 2009), is that the electricity consumption of households can be influenced by merely reporting normative levels of consumption and comparing households to those norms.

That this is found to be effective in the case of electricity accords with a growing literature in behavioural economics around the influence of social norms as well as findings from social psychology. Taken together with the broader literature, this research suggests that reporting social norms may be an effective way for Eskom and municipalities to mitigate household electricity consumption in South Africa.

References.

Aarts, H., & Dijksterhuis, A. (2003). The silence of the library: Environment, situational norm, and social behavior. Journal of Personality and Social Psychology, 84, 18–28.

Abrahamse W., (2005). "A review of intervention studies aimed at household energy conservation." *Journal of Environmental Psychology.* Volume 25, Issue 3, September 2005, Pages 273-291

Adams, J. Stacey, 1965. Inequity in social exchange. In: Berkowitz, Leonard (Ed.), Advances in Experimental Social Psychology. Academic Press, New York, pp. 167-299.

Adato M., Carter, Michael R. and May, J. (2006) 'Exploring poverty traps and social exclusion in South Africa using qualitative and quantitative data', Journal of Development Studies, 42: 2, 226 — 247

Agostinelli, G., Brown, J., & Miller, W. (1995). Effects of normative feedback on consumption among heavy drinking college students. Journal of Drug Education, 25, 31–40.

Akerlof G. (1980). "A Theory of Social Custom, of Which Unemployment May be One Consequence." *The Quarterly Journal of Economics*, Vol. 94, No. 4 (Jun., 1980), pp. 749-775

Akerlof G. (1982) "Labor Contracts as Partial Gift Exchange." Quarterly Journal of Economics 97 (November 1982): 543-69.

Akerlof, George. A: (1997) "Social distance and social decisions" in *Econometrica*, (65) pp.1005-1027.

Akerlof, G., Kranton, R.E., 2000. Economics and identity. Quarterly Journal of Economics 115, 715–753.

Akerlof, G., Kranton, R.E., (2002). "Identity and schooling: some lessons for the economics of education." *Journal of Economic Literature* 40, 1167–1201.

Allcott H. (2009). "Social Norms and Energy Conservation." Working paper. MIT-CEEPR (Series); 2009-014, 2009.

Allcott H. (2011). "Social norms and energy conservation" J. Public Econ. Volume 95, Issues 9-10, October 2011, Pages 1082-1095

Alpert, B. (1967). " Non-Businessmen as Surrogates for Businessmen in Behavioral Experiments." The Journal of Business . Vol. 40, No. 2, Apr., 1967.

Andreoni, J. (1989). _Giving with impure altruism: applications to charity and Ricardian equivalence, Journal of Political Economy, vol. 97 (6) (January), pp. 1447–58.

Andreoni, J. (1990). _Impure altruism and donations to public goods: a theory of warm glow giving, Economic Journal, vol. 100 (401) (June), pp. 464–77.

Andreoni, J. (1998). "Toward a theory of charitable fundraising", Journal of Political Economy, vol. 106 (6) (December), pp. 1186–213.

Andreoni J. and Bernheim D. (2009). “Social image and the 50–50 norm: a theoretical and experimental analysis of audience effects.” *Econometrica*, Vol. 77, No. 5 (September, 2009), 1607–1636.

Arbuthnott, J. (1710). "An argument for divine providence, taken from the constant regularity observed in the births of both sexes." *Philosophical Transaction of the Royal Society of London* 27: 186–190.

Armitage, P. (2003). “Fisher, Bradford Hill and randomization”. *International Journal of Epidemiology* 32: 925–928

Ariely, D. (2008). “Predictably Irrational: The Hidden Forces that Shape our Decisions.” Harper Collins.

Ariely D. and Mazar N. (2006). “Dishonesty in Everyday Life and its Policy Implications.” Journal of Public Policy and Marketing. Vol. 25(1) spring 2006, pp. 117-26.

Ariely D., Amir O. and Mazar N. (2008). “The Dishonesty of Honest People: A Theory of Self-Concept Maintenance.” Journal of Marketing Research. vol. XLV (December, 2008) pp. 633-44.

Ariely D., Ayal s., & Gino F. (2009). “Contagion and Differentiation in Unethical Behavior : The Effect of One Bad Apple on the Barrel.” Psychological Science (20) 3. Pp. 393-8.

Arellano, M. (1993)."On the testing of correlated effects with panel data. Journal of Econometrics". Vol. 59, Nos. 1-2, pp. 87-97.

Asch, Solomon. (1962). “Group forces in the modification and distortion of judgements”, Chapter 16 (p.450-501) in “Social Psychology” (7th printing) by Solomon Asch. Prentice Hall, New Jersey. 7th printing: June, 1962. First printing: September 1952.

Ashraf N., Karlan, D., and Yin W. (2006a) "Deposit Collectors," *Advances in Economic Analysis & Policy*: Vol. 6: Iss. 2, Article 5. Available at: http://www.bepress.com/bejeap/advances/vol6/iss2/art5

Ashraf, N., D. Karlan and W. Yin (2006b). "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines." Quarterly Journal of Economics 121(2): 673-697.

Azar, O.H. (2004). " What sustains social norms and how they evolve? The case of tipping." Journal of Economic Behavior & Organization, Vol. 54 (2004) 49–64.

Azfar O. and Zinnes C. (2006). “Which incentives work? An experimental analysis of incentives for trainers.” The IRIS Center, Department of Economics, University of Maryland University at College Park.

Bandura, A.;Ross, Dorothea;Ross, Sheila A. (1961). "Transmission of aggression through imitation of aggressive models." The Journal of Abnormal and Social Psychology, Vol 63(3), Nov 1961, 575-582. doi: 10.1037/h0045925

Bardsley, N., & Sausgruber, R. (2005). Conformism and reciprocity in public good provision. *Journal of Economic Psychology*, *26*(5), 664–681.

Becker, G. (1974). "A theory of social interactions", Journal of Political Economy, vol. 82 (6) (November-December), pp. 1063–93.

Bertrand M., Djankov S., Hanna R., and Mullainathan S. (2006). " Does Corruption Produce Unsafe Drivers?" NBER Working Paper No. 12274. Issued in June 2006. NBER Program(s): CF LS

Berns G. S, Chappelow J, Zink C. F., Pagnoni G., Martin-Skurski M. E., Richards J. (2005) "Neurobiological Correlates of Social Conformity and Independence During Mental Rotation", Biological Psychiatry, Volume 58, Issue 3, 1 August 2005, Pages 245-253

The Holy Bible. English Standard Version (2001). Crossway Bibles.

Björklund and Régner, (1996) "Experimental evaluation of European Labour Market Policy." In: G. Schmid, J. O'Reilly and K. Schömann, Editors, *International Handbook of labour market policy and Evaluation*, Edward Elgar, Cheltenham (1996), pp. 89–114.

Blanton H., Stuart A.E., and Van den Eijnden R.J.J.M (2001). "An Introduction to Deviance-Regulation Theory: The Effect of Behavioral Norms on Message Framing." Personality and Social Psychology Bulletin (27) pp. 848-58.

Berkowitz, A.D. (2004). An overview of the social norms approach. In L. Lederman & L. Stewart (Eds.), Changing the culture of college drinking (pp. 193–214). Cresskill, NJ: Hampton Press.

Betts, Julian R., and Darlene Morell, (1999) "The Determinants of Undergraduate GPA: The Relative Importance of Family Background, High School Resources, and Peer Group Effects," Journal of Human Resources, XXXIV (1999), 268-293.

Bewley, T. F. (1999): Why Wages Don't Fall During a Recession. Cambridge, MA: Harvard University Press.

Bohm P. (1972). "Estimating Demand for Public Goods: An Experiment." European Economic Review 3.

Bohnet, Iris and Frey, Bruno S. Frey (1999) "Social Distance and other-regarding Behavior in Dictator Games." The American Economic Review, Vol. 89, No. 1 (Mar., 1999), pp. 335-339.

Bond R. & Smith P.B. (1996). "Culture and Conformity: A Meta-Analysis of Studies Using Asch's (1952b, 1956) Line Judgment Task." Psychological Bulletin, 1996, Vol. 119, No. 1, pp. 111-137.

Bonetti, S. (1998). "Experimental economics and deception." *Journal of Economic Psychology*. Volume 19, Issue 3, June 1998, Pages 377-395

Box J. F. (1980). "R. A. Fisher and the Design of Experiments, 1922-1926." The American Statistician, Vol. 34, No. 1 (Feb., 1980), pp. 1-7

Brock, W. and S. Durlauf, (2001), "Interactions-Based Models," in *Handbook of Econometrics*, *Vol. 5*, J. Heckman and E. Leamer, eds., Amsterdam: North Holland.

Brock, W. and S. Durlauf, (2003), "A Multinomial Choice Model with Social Interactions," *The Economy as an Evolving Complex System III*, L. Blume and S.Durlauf, eds., Oxford University Press.

Buchan N, Johnson E & Croson R, (2006) "Let's get personal: An international examination of the influence of communication, culture and social distance on other regarding preferences, Journal of Economic Behavior & Organization, Volume 60, Issue 3, July 2006, Pages 373-398

Bruhn, M., & McKenzie, D. (2009). In Pursuit of Balance: Randomization in Practice in Development Field Experiments. *American Economic Journal: Applied Economics*, *1*(4), 200-232.

Bühren C.,Frank B. (2010). "Chess players' performance beyond 64 squares: A case study on the limitations of cognitive abilities transfer." MAGKS Joint Discussion Paper Series in Economics. No. 19-2010.

Camerer, Colin F. (1997) "Progress in Behavioral Game Theory." Journal of Economic Perspectives, Autumn 1997, 11(3), pp. 167-88.

Camerer, C., and Thaler R. (1995): "Ultimatums, Dictators, and Manners," *Journal of Economic Perspectives*, 9, 209–219.

Cameron R., Brown K. S., Best J. A., Pelkman C. L., Madill C. L., Manske S. R., and Payne M. E. (1999). "Effectiveness of a social influences smoking prevention program as a function of provider type, training method, and school risk." *American Journal of Public Health*, 89, 1827–1831.

Case A. C., and Katz L.F. (1991). "The Company You Keep: The Effect of Family and Neighbourhood on Disadvantaged Youths," National Bureau of Economic Research Working Paper No. 3705, 1991.

Charness G, Haruvy E. and Sonsino D. (2007). "Social distance and reciprocity: An Internet experiment. " Journal of Economic Behavior & Organisation. 63 (2007) 88–103.

Chattopadhyay, R., & Duflo, E. (2004). "Women as policy makers: Evidence from a randomized policy experiment in India." *Econometrica*, *72*(5), 1409-1443.

Christakis N. A. and Fowler J.H. (2007). "The Spread of Obesity in a Large Social Network over 32 Years." New England Journal of Medicine. Volume 357 (4) pp. 370-379.

Cialdini R., Reno R. and Kallgren C. (1990). "A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places." Journal of Personality and Social Psychology 1990, Vol. 58, No. 6, pp. 1015-1026.

Cialdini, R.B., Kallgren, C.A., & Reno, R.R. (1991). "A focus theory of normative conduct." Advances in Experimental Social Psychology, 24, 201–234.

Cialdini R B, Trost M R. (1998)." Social influence: social norms, conformity, and compliance." In *The Handbook of Social Psychology*, ed. DT Gilbert, ST Fiske, G Lindzey, 2:151–92. Boston: McGraw-Hill. 4th ed.

Coakley C.W., and Heise M. A. (1996). "Versions of the sign test in the presence of ties." *Biometrics*. Vol. 52. No.4 (Dec., 1996), pp.1242-1251.

Coleman, James Samuel, et al., (1966) "Equality of Educational Opportunity" (Washing-ton, DC: U. S. Government Printing Office, 1966).

Coleman, James S. (1990). Foundations of Social Theory. Cambridge MA: Harvard University Press.

Cornelli, F. (1996). "Optimal selling procedures with fixed costs", Journal of Economic Theory, vol. 71 (1) (October), pp. 1–30.

Croson R. (2007). "Theories of commitment, altruism and reciprocity: evidence from linear public goods games." *Economic Inquiry*. 2007 (45) no. 2. Pp. 199-216.

Croson R., Shang J. (2008). "The impact of downward social information on contribution decisions." Experimental Economics (11) pp. 221–233

Croson R., Shang J. (2009). "A Field Experiment in charitable contribution: the impact of social information on the voluntary provision of public goods." *The Economic Journal.* 119 (October) pp.1422-39.

Croson, R., Handy, F., & Shang, J. (2006). "The relationship between norms, social information and subsequent giving: results from a donor survey and a lab experiment ." (Working paper). University of Texas at Dallas.

Davis, D.D., Holt, C.A. (1993). "Experimental Economics." Princeton Univ. Press, Princeton.

Deutsch, Morton; Gerard, Harold B. (1955). "A study of normative and informational social influences upon individual judgment." The Journal of Abnormal and Social Psychology, Vol 51(3), Nov 1955, 629-636.

Dixon, W. J. and Mood, A. M. (1946). "The statistical sign test." *Journal of the American Statistical Association* 41, 557-566.

Duflo E, Kremer M and Robinson (2009). "Nudging Farmers to use fertilizer: theory and experimental evidence from Kenya." NBER working paper 15131.

Durlauf, S.E. (2004), "Neighborhood effects," In: J.V. Henderson and J-F. Thisse (Eds.), Handbook of Regional and Urban Economics Vol. 4, Amsterdam: Elsevier Science, pp. 2173-2242.

Ehrenberg A.S. (1972) "Repeat-Buying: Theory and Applications", 2nd ed. London: Charles Griffin; New York: Oxford University Press.

Ehrenberg A.S., Barwise T.P. & Goodhardt G.J. (1990), "Double Jeopardy revisited". The Journal of Marketing, Vol. 54, No. 3 (Jul., 1990), pp. 82-91.

The English Standard Version (ESV) Study Bible (2008). Crossway Bibles.

Engelbrecht-Wiggans R., and Kahn C. M. (1998). Multi-Unit Auctions with Uniform Prices. *Economic Theory*, Vol. 12, No. 2 (Sep., 1998), pp. 227-258.

Engelmann D., Grimm V., (2004). "Bidding behaviour in multi-unit auctions—an experimental investigation and some theoretical insights." Working Paper, Royal Holloway.

Esser, J. (1998). "Alive and well after 25 years: a review of groupthink research." *Organisational Behaviour and Human Decision Processes* (73) February/March, pp. 116-41.

Falk A. (2007). "Gift exchange in the field." *Econometrica*, Vol. 75, No. 5 (Sep., 2007), pp. 1501-1511.

Fara R. (2009) "Science: a four thousand year history." Oxford

Fehr E, Kirchsteiger G, & Riedl A. (1998). "Gift exchange and reciprocity in competitive experimental markets." European Econ. Rev., 42 (1998), pp. 1‐34

Fehr E., Fischbacher U. and Gächter S. (2002). "Strong reciprocity, human cooperation, and the enforcement of social norms." Human Nature, 13, 1-25.

Festinger, L. (1954). "A theory of social comparison processes." *Human Relations*, *7*, 117–140.

Feldman Y, MacCoun R (2005). "Some Well-Aged Wines for the "New Norm" Bottles, Implications of Social Psychology to Law and Economics." In: Parisi F. & Smith V (eds) The Law and Economics of Irrational Behavior, pp. 358–396. Stanford University Press.

Fellner G., Sausgruber R., & Traxler C. (2009) "Testing Enforcement Strategies in the Field: Legal Threat, Moral Appeal and Social Information. " (September 1, 2009). MPI Collective Goods Preprint, No. 2009/31; CESifo Working Paper Series No. 2787. Available at SSRN: http://ssrn.com/abstract=1468344 or http://dx.doi.org/10.2139/ssrn.1468344

Ferraro P & Price M (2011)."Using non-pecuniary strategies to influence behavior: evidence from a large scale field experiment." NBER working paper 17189, July 2011.

Fischbacher, U., Gaechter, S., & Fehr, E. (2001). "Are people conditionally cooperative? Evidence from a public goods experiment." *Economics Letters*, *71*, 397–404.

Fisher, R.A. (1925). "Theory of Statistical Estimation" Proceedings of the Cambridge Philosophical society, 22: 700-725.

Fisher, R.A. (1926). "The arrangement of field trials". Journal of the Ministry of Agriculture of Great Britain (33) pp. 503-513.

Fisher, R. A. (1947). "The development of the theory of experimental design" Proceedings of the international statistical conference. 3, 434-439. And Collected papers, 4, no 212.

Fisher, R. A., McKenzie W.A. (1923). "Studies in crop variation. II. The manurial response of different potato varieties." The Journal of Agricultural Science (1923), 13: 311-320 Cambridge University Press.

Frey B., Meier S. (2004a). Pro-social behaviour in a natural setting. *Journal of Economic Behavior and Organization* 54, 65-88.

Frey, B., & Meier, S. (2004b). "Social comparisons and pro-social behavior: testing 'conditional cooperation' in a field experiment." *American Economic Review*, *94*, 1717–1722.

Frey B., & Torgler B. (2007). "Tax morale and conditional cooperation." *Journal of Comparative Economics.* Vol. 35 pp. 136-159.

Forsetlund,L., Chalmers, I., and Bjørndal, A.(2007). "When was random allocation first used to generate comparison groups in experiments to assess the effects of social interventions?" Econ. Innov. New Techn. Vol 16 (5) pp.371-384.

Fourcade, Marion (2009) *Economists and Societies: Discipline and Profession in the United States, Britain and France, 1890s-1990*s (Princeton University Press).

Fox J.A., Shogren J.F., Hayes D.J., and Kliebenstein J.B. (1998). "CVM-X: Calibrating Contingent Values with Experimental Auction Markets." *American Journal of Agricultural Economics*, Vol. 80, No. 3 (Aug., 1998), pp. 455-465

Goldstein N. J., Cialdini R.B. and Vladas Griskevicius V. (2008). "A Room with a Viewpoint: Using Social Norms to Motivate Environmental Conservation in Hotels." *Journal of Consumer Research* , Vol. 35, No. 3 (October 2008), pp. 472-482

Gränsmark P. (2010). "Social Screening and Cooperation Among Expert Chess Players." Preliminary version for the EALE conference, London, 2010.

Greenberg D. and Shroder M. (2004). The Digest of Social Experiments, The Urban Institute Press, Washington (2004).

Greenberg D., Shroder M. and Onstott M. (1999). "The Social Experiment Market." Journal of Economic Perspectives (13) 3 pp. 157-172. *American Economic Association.*

Haines, M., & Spear, S. (1996). Changing the perception of the norm: A strategy to decrease binge drinking among college students. Journal of American College Health, 45, 134–140.

Haines. M. A., Vaughan. G.M. (1979). "Was 1898 a 'great date' in the history of experimental social psychology?" Volume 15 Issue 4, Pages 323 – 332.

Harris J.R. (1998). "The Nurture Assumption: Why Children Turn Out The Way They Do." (New York, NY: Free Press, 1998).

Harrison G., & List J. A. (2004) "Field experiments." *Journal of Economic Literature*, *42*(4), 1009–1055.

Han G. & Choe S. (1994). "Effects of family, region, and school network ties on interpersonal intentions and the analysis of network activities in Korea." In: U. Kim, H.C. Triandis, C. Kagitcibai, S.C. Choi and G. Yoon, Editors, *Individualism and Collectivism: Theory, Method, and Applications*, Sage Publications, Thousand Oaks, CA (1994), pp. 189–199.

Heldt, T. (2005). "Conditional Cooperation in the Field: Cross-Country Skiers? Behavior in Sweden", unpublished manuscript, Uppsala University.

Henrich, J. (2000) "Does culture matter in economic behavior? Ultimatum game bargaining among the Machiguenga of the Peruvian Amazon." *The American Economic Review* 90.4 (2000): 973-979.

Henrich J., Boyd B., Bowles S., Camerer C., Fehr E., Gintis H., McElreath R. (2001) "In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies." *American Economic Review.* 91(2) May, 2001. Pp. 73-8.

Henrich, J., et al.(2005) " Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies." *Behavioral and brain sciences* 28.6 (2005): 795-814.

Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, Richard McElreath et al.(2005) " Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies." *Behavioral and brain sciences* 28, no. 6 (2005): 795-814.

Hoffman, Elizabeth; McCabe, Kevin; Smith, Vernon L. (1996) "Social Distance and Other-Regarding Behavior in Dictator Games". The American Economic Review, Vol. 86, No. 3 (Jun., 1996), pp. 653-660.

Hoffman, Elizabeth; McCabe, Kevin; Smith, Vernon L. 1999. "Social Distance and other-regarding Behavior in Dictator Games." The American Economic Review, Vol. 89, No. 1 (Mar., 1999), pp. 340-341.

Homans, George C. 1961, Social Behavior: Its Elementary Form. Harcourt, New York.

Huck, S., and Wallace B. (2002): "Reciprocal Strategies and Aspiration Levels in a Cournot–Stackelberg Experiment," *Economics Bulletin*, 3, 1–7.

Janis, I. (1971). "Groupthink." *Psychology Today,* November, pp.43–46, 74–76.

Jager J. (1983). "Climate and energy systems: a review of their interactions." Wiley, New York.

Keep Tennessee Beautiful. http://www.ktnb.org/

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

Lalive R. & Catteneo A. (2009). "Social interactions and schooling decisions." *The Review of Economics and Statistics.* 91(3) pp. 457-77.

Leatherdale S. T., Cameron R., Brown K. S., Jolin M., & Kroeker C. (2006). "The influence of friends, family and older peers on smoking among elementary school students: Low risk students in high risk schools." *Preventive Medicine*, 42, 218–222.

Leatherdale S. and Papadakis S. (2010). "A Multi-Level Examination of the Association Between Older Social Models in the School Environment and Overweight and Obesity Among Younger Students." *Journal of Youth and Adolescence*. DOI 10.1007/s10964-009-9491-z.

Levin, D. (2005). "Demand Reduction in Multi-Unit Auctions: Evidence from a Sportscard Field Experiment: Comment." *The American Economic Review*, Vol. 95, No. 1 (Mar., 2005), pp. 467-471

Levitt, S., & List, J. (2007). "What do laboratory experiments measuring social preferences tell us about the real world?" *Journal of Economic Perspectives*, *21*(2), 153–174.

Levitt, S., & List, J. (2009)." Field experiments in economics: The past, the present, and the future." *European Economic Review* 53 (2009) pp.1–18.

List J.A. (2006). "The Behavioralist Meets the Market: Measuring Social Preferences and Reputation Effects in Actual Transactions." *Journal of Political Economy,* 2006, vol. 114, no. 1.

List J.A. (2007). “Field Experiments: A Bridge Between Lab and Naturally-Occurring Data.” *NBER Working Paper No. W12992*, March 2007.

List J.A. (2008). “Introduction to field experiments in economics with applications to the economics of charity.” *Experimental Economics* (11) pp. 203-212.

List J. A., (2009). “An introduction to field experiments in economics.” *Journal of Economic Behavior and Organization*, 70 (2009), pp.439-442.

List J. A., (2011). "Why economists should conduct field experiments and 14 tips for pulling one off." *Journal of Economic Perspectives*, Vol.25 (3) pp. 3-16, summer 2011.

List J A., & Levitt S., (2009). “ Field experiments in economics: the past, the present, and the future.” *European Economic Review*, 53 (2009) pp.1-18.

List J A., Sadoff S., Wagner M. (2010). “So you want to run an experiment, now what? Some Simple Rules of Thumb for Optimal Experimental Design.” NBER Working Paper No. 15701, January 2010.

List J. A., Shogren J.F., Margolis, M. (1998a). “Hypothetical–actual bid calibration of a multi-good auction”. Economics Letters. 60 (1998) pp. 263–268

List J. A., Shogren J.F. (1998b). "Calibration of the difference between actual and hypothetical valuations in a field experiment." *Journal of Economic Behaviour & Organisation*. 37 pp.193-205

List J.A. & Gallet C.A. (2001). " What Experimental Protocol Influence Disparities Between Actual and Hypothetical Stated Values?" *Environmental and Resource Economics* **20:** 241–254, 2001.

List J. A., and Lucking-Reiley, D. (2000). Demand Reduction in Multiunit Auctions: Evidence from a Sportscard Field Experiment. *The American Economic Review*, Vol. 90, No. 4 (Sep., 2000), pp. 961-972

List J. A., Lucking-Reiley, D and Engelbrecht-Wiggans R. (2005). Demand Reduction in Multi-Unit Auctions: Evidence from a Sportscard Field Experiment: Reply. *The American Economic Review*, Vol. 95, No. 1 (Mar., 2005), pp. 472-476

List J. A., and Rasul I. (2010). “Field Experiments in Labour Economics.” NBER Working Paper 16062. June, 2010.

C.J.H. Midden, J.E. Meter, M.H. Weenig and H.J.A. Zieverink. (1983). “ Using feedback, reinforcement and information to reduce energy consumption in households: A field-experiment.” *Journal of Economic Psychology* **3** (1983) (1), pp. 65–86.

Manski C.F., Garfinkel I.(1992). In: "Evaluating Welfare and Training Programs." Harvard University Press, Cambridge, pp.1–22.

Manski, C.F. 1993). "Identification of Endogenous Social Effects: The Reflection Problem." Review of Economic Studies. 60, pp. 531-42.

Manski C.F. (2000). “Economic analysis of social interactions.” *Journal of Economic Perspectives.* 14(3) pp. 115-136.

McPhee, W. N. (1963), Formal Theories of Mass Behavior. New York: The Free Press.

Milgram S., Bickman L., Berkowitz L. (1969). “Note on the drawing power crowds of different size.” Journal of Personality and Social Psychology. 1969, Vol. 13, No. 2, 79-82

Moffitt, R., (2001), "Policy Interventions, Low-Level Equilibria, and Social Interactions," in Social Dynamics, S. Durlauf and H. P. Young, eds., Cambridge: MIT Press.

Neighbors, C., Larimer, M., & Lewis, M. (2004). "Targeting misperceptions of descriptive drinking norms: Efficacy of a computer delivered personalized normative feedback intervention." Journal of Consulting and Clinical Psychology, 73, 434–447.

Nicol J.F. and Humphreys M.A. (2002). “Adaptive thermal comfort and sustainable thermal standards for buildings.” *Energy and Buildings* (34) pp. 563-72.

Oakley A. (1998). "Experimentation and social interventions: a forgotten but important history." British Medical Journal (BMJ) 1998;317:1239–42.

Orcutt G.H. & Orcutt A.G. (1968) “Incentive and Disincentive Experimentation for Income Maintenance Policy Purposes”, American Economic Review 58: 754-73.

Peirce C.S. (1931). "Principles of Philosophy." http://www.textlog.de/peirce_principles.html

Perkins H. W (2002). " Social Norms and the Prevention of Alcohol Misuse in Collegiate Contexts." JOURNAL OF STUDIES ON ALCOHOL / SUPPLEMENT NO. 14,2002.

Perkins H. W., Haines P.M. & Rice R. (2005). “Misperceiving the College Drinking Norm and Related Problems: A Nationwide Study of Exposure to Prevention Information, Perceived Norms and Student Alcohol Misuse.” Journal of Studies on Alcohol and Drugs. Vol. 66 (4), July 2005.

Peyton Young H. (2008). “Social Norms”. New Palgrave Dictionary of Economics. Palgrave Macmillan.

Porter D., Vragov, R. (2006). “An experimental examination of demand reductions in multi-unit versions of the uniform-price, Vickrey, and English auctions.” Managerial and Decision Economics Volume 27 Issue 6, Pages 445 – 458

Posner E. A. (2000). "Law and Social Norms: The Case of Tax Compliance." Virginia Law Review , Vol. 86, No. 8, Symposium: The Legal Construction of Norms (Nov., 2000), pp. 1781-1819

Potters J., Sefton M., Vesterlund L. (2005). “After you—endogenous sequencing in voluntary contribution games.” *Journal of Public Economics.* (89) pp. 1399-1419.

Prentice, D.A., & Miller, D.T. (1993). "Pluralistic ignorance and alcohol use on campus: Some consequences of misperceiving the social norm." Journal of Personality and Social Psychology, 64, 243–256.

RAO, C. R.(1952). "Advanced Statistical Methods in Biometric Research." (New York: John Wiley & Sons, 1952), 165-172.

Reid C. (1982). “Neyman – From Life”. Springer, New York.

Reno, R., Cialdini, R., & Kallgren, C.A. (1993). “The transsituational influence of social norms.” *Journal of Personality and Social Psychology*, 64, 104–112.

Ribar, D.C. and Wilhelm, M.O. (2002). "Altruistic and joy of giving motivations in charitable behavior." Journal of Political Economy, vol. 110 (2) (April), pp. 425–57.

Roberts, R.D. (1984). _A positive model of private charity and public transfers, Journal of Political Economy, vol. 92 (1) (February), pp. 136–48.

Romano, R.E. (1991). "When excessive consumption is rational", American Economic Review, vol. 81 (3) (June), pp. 553–64.

Romer, D. (1984). "The Theory of Social Custom: A Modification and Some Extensions. " *The Quarterly Journal of Economics (1984) 99(4): 717-727*

Roth, Alvin E. (1985). Toward a focal point theory of bargaining. In Game-Theoretic Models of Bargaining, Alvin E. Roth, ed. Cambridge UK and New York: Cambridge University Press.

Roth, A. (1993) “On the early history of experimental economics”. Journal of the History of Economic Thought, 15, Fall, 1993, 184-209. Adapted version accessed on 21 April 2010 at http://kuznets.harvard.edu/~aroth/history.html

Roth, A E., and Kagel J H K. (1995) *The handbook of experimental economics*. Vol. 1. Princeton: Princeton University Press, 1995.

Rubin D.B. (1990). “Comment: Neyman (1923) and Causal Inference in Experiments and Observational Studies.” *Statistical Science*. Volume 5, Number 4 (1990), 472-480.

Ruth M. and Lin A-C. (2006). “Regional energy demand and adaptions to climate change: methodology and applications to the state of Maryland, USA.” *Energy Policy*. 34. pp.2820-33.

Sacerdote B. (2001). “Peer Effects with Random Assignment: Results for Dartmouth Roommates.” The Quarterly Journal of Economics, Vol. 116, No. 2 (May, 2001), pp. 681-704

Sadoff S., Levitt S., List J. A. (2010). “Checkmate: Exploring Backward Induction Among Chess Players.” American Economic review. *Forthcoming.*

Salganik M. J., Dodds P. J., & Watts D.J. (2006). "Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market." SCIENCE. Vol(311) 10 February 2006.

Samuelson P. A. & Nordhaus W. D. (1985). "Economics", 12th ed. New York: McGraw-Hill.

Schaffer, M.E., & Stillman, S. (2010). " xtoverid: Stata module to calculate tests of overidentifying restrictions after xtreg, xtivreg, xtivreg2 and xthtaylor."
http://ideas.repec.org/c/boc/bocode/s456779.html

Schultz P.W., Nolan J., Cialdini R., Goldstein N. and V. Griskevicius (2007). “The Constructive, Destructive and Reconstructive Power of Social Norms.” Psychological Science Vol.18(5). pp.429-34.

Schultz, P.W., Khazian, A. M., & Zaleski, A. C. (2008). Using normative social influence to promote conservation among hotel guests. *Social Influence*, 3, 4–23.

Schultz, P.W., Tabanico, J., & Rendón, T. (2008). Normative beliefs as agents of influence: Basic processes and real-world applications. In R. Prislin & W. Crano (Eds.), *Attitudes and attitude change* (pp. 385-409). New York: Psychology Press.

Selten, R. (1967) , "Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments," pp. 136-168 in: H. Sauermann (ed.), Beiträge zur experimentellen Wirtschaftsforschung, Vol. I, J.C.B. Mohr (Paul Siebeck), Tübingen. (Quoted in Fischbacher, 2001).

Sherif, M. (1936). "The psychology of social norms." Oxford, England: Harper.

Spicer, Micheal W., Becker, Lee A., (1980). "Fiscal inequity and tax evasion: An experimental approach." National Tax Journal 33, 171-175.

Staats H., Harland P. and Wilke H.A.M. (2004) "Effecting durable change. A team approach to improve environmental behavior in the household", *Environment and Behavior* **36** (2004) (3), pp. 341–367.

Steinberg, R.S. (1987). "Voluntary donations and public expenditures in a federal system", American Economic Review, vol. 77 (1) (March), pp. 24–36.

Stigler, Stephen. M. (1992). "A Historical View of Statistical Concepts in Psychology and Education Research". *American Journal of Education.* 101, November.

Strube. M. J. (2005). "What Did Triplett Really Find? A Contemporary Analysis of the First Experiment in Social Psychology." *The American Journal of Psychology*, Vol. 118, No. 2 (Summer, 2005), pp. 271-286.

Sugden R. (1986). "The Economics of Rights, Cooperation and Welfare." Oxford: Basil Blackwell.

Thaler R. and Sunstein C. (2008) "Nudge". Penguin

Thorndike EL, Woodworth RS. (1901). "The influence of improvement in one mental function upon the efficiency of other functions." *Psychol Rev* 8:247-61, 384-95, 553-64.

Traxler C. (2010). "Social norms and conditional cooperative taxpayers." *European Journal of Political Economy* (26) pp. 89-103.

Traxler, C., and J. Winter (2009): "Survey Evidence on Conditional Norm Enforcement," Working Paper 2009-03, Max Planck Institute for Research on Collective Goods, Bonn.

Traxler C and Winter J. (2012). "Survey evidence on conditional norm enforcement". *European Journal of Political Economy* vol. 28 pp.390-8.

Triandis H.C, Bontempo R., Villareal M.J., Asai M. and N. Lucca. (1988). "Individualism and collectivism: cross-cultural perspectives on self-ingroup relationships." *Journal of Personality and Social Psychology* 54, pp. 323–338.

Triandis H.C., McCusker C. and Hui C.H. (1990). "Multimethod probes of individualism and collectivism." *Journal of Personality and Social Psychology.* 59, p. 1006.

Triplett, N. (1898). "The Dynamogenic Factors in Pacemaking and Competition." *The American Journal of Psychology*, Vol. 9, No. 4 (Jul., 1898), pp. 507-533.

Tyler, Tom R. and Smith, Heather J. 1998. Social justice and social movements. In: D. T. Gilbert, S. T.Fiske and G. Lindzey (eds.), The Handbook of Social Psychology, Vol. 3. Boston: McGraw-Hill: pp. 595-629.

Vesterlund L. (2003). "The informational value of sequential fundraising." Journal of Public Economics, Volume 87, Issues 3-4, pp. 627-657.

Violette D., Provencher B. & Klos M. (2009). "Impact evaluation of Positive Energy SMUD pilot study." Boulder, Colorado. Summit Blue Consulting.

Visser M., Martinsson P., Kocher M. (2009). "Social background, cooperative behaviour and norm enforcement." University of Gothenburg, School of Business, Economics and Law. Working papers in Economics. Working paper No. 385. October, 2009.

Wärneryd, K. (1994). "Transaction cost, institutions, and evolution." Journal of Economic Behaviour and Organization 25, 219-239.

Warr, P.G. (1982). Pareto optimal redistribution and private charity, Journal of Public Economics, vol. 19 (1) (October), pp. 131–8.

Webley, Paul, Robben, Henry, Morris, Ira, (1988). "Social comparison, attitudes and tax evasion in a shop simulation". Social Behaviour 3, 219-228

Wechsler, H., Nelson, T., Lee, J.E., Seiberg, M., Lewis, C., & Keeling,R. (2003). Perception and reality: A national evaluation of social norms marketing interventions to reduce college students' heavy alcohol use. Quarterly Journal of Studies on Alcohol, 64, 484–494.

Wenzel M. (2004). " An analysis of norm processes in tax compliance." Journal of Economic Psychology. Volume 25, Issue 2, April 2004, Pages 213-228

Werch, C.E., Pappas, D.M., Carlson, J.M., DiClemente, C.C., Chally, P.M., & Sinder, J.S. (2000). Results of a social norm intervention to prevent binge drinking among first-year residential college students. Journal of American College Health, 49, 85–92.

West S. G., Gunn S. P. (1978)."Some issues of ethics and social psychology." *American Psychologist*. (January, 1978), pp. 30-8.

Winkler, H and Marquard, A. 2009. 'Changing development paths: From an energy-intensive to low-carbon economy in South Africa'. *Climate and Development*. 1(2009): 47 – 65.

Winett, R.A.; Leckliter, I.N.; Chinn, D.E.; Stahl, B. and Love, S.Q. (1985) "Effects of television modeling on residential energy conservation, *Journal of Applied Behavior Analysis* **18** (1985), pp. 33–44.

Winter, H. J. J. (1953), "THE OPTICAL RESEARCHES OF IBN AL-HAITHAM." Centaurus, 3: 190–210.

Wooldridge, J.M. (2002). Econometric Analysis of Cross Section and Panel Data. Cambridge, MA: MIT Press.

Young H.P. (2007). "Social Norms." University of Oxford Discussion Paper series Number 307. January 2007.

Young, H. P. and Burke M.A. (2001). “Competition and custom in economic contracts: a case study of Illinois agriculture.” American Economic Review 91, 559-573.