# THE INEFFICIENCY OF SPLITTING THE BILL*

### *Uri Gneezy, Ernan Haruvy and Hadas Yafe*

When agents are ascribed selfish motives, economic theory points to grave inefficiencies resulting from externalities. We study a restaurant setting in which groups of diners are faced with different ways of paying the bill. The two main manipulations are splitting the bill between the diners and having each pay individually. We find that subjects consume more when the cost is split, resulting in a substantial loss of efficiency. Diners prefer the individual pay to the inefficient split-bill method. When forced to play according to a less preferred set of rules, they minimise their individual losses by taking advantage of others.

Economic theory is unambiguous in its prediction that if externalities exist, outcomes are likely to be inefficient when agents selfishly maximise. The literature on externalities, as well as its derivatives in public goods, tragedy of the commons and moral hazard studies, has shown that externalities lead to inefficient levels of production and consumption. This result depends crucially on the general assumption taken by such studies that human agents maximise selfish payoffs without regard for others.

With the emergence of behavioural economics, economists have come to question whether people actually ignore costs imposed on others when reaching economic decisions. If altruism is common, the various proposals in the literature to solve externality problems may be unnecessary or even harmful. For example, the government in a public good setting may actually reduce voluntary contributions by interfering with the provision of a public good (Andreoni, 1993). Similarly, increased government monitoring for corruption may backfire by reducing intrinsic other-regarding behaviour (Bohnet *et al.*, 2001; Schulze and Frank, 2003), and the mere sanctioning of an activity may be counterproductive (Gneezy and Rustichini, 2001).

Experimental studies, with few exceptions, find evidence against theories based purely on selfish motives. The studies find that people free ride but not to the extent economic theory predicts (Dawes and Thaler, 1988). Hence, despite the strong predictions generated by classical theory in externality settings, social scientists often question the truths provided by it.

To test economic predictions, we investigate a familiar environment. The *unscrupulous diner's dilemma* is a problem faced frequently in social settings. When a group of diners jointly enjoys a meal at a restaurant, often an unspoken agreement exists to divide the check equally. A selfish diner could thereby enjoy exceptional dinners at bargain prices. Whereas a naive approach would appear to suggest that this problem is not likely to be severe, it appears that even the best of friends can sometimes find it rather severe.[1] Furthermore, this dilemma typifies a class of

---

[1] Ross:… plus tip, divided by six. Ok, everyone owes 28 bucks. Phoebe: No, uh uh, no way, I'm sorry, not gonna happen. (Friends, Season 2, Episode 5).

serious social problems from environmental protection and resource conservation to eliciting charity donations and slowing arms races (Glance and Huberman, 1994).

We observe and manipulate conditions for several groups of six diners at a popular dining establishment. In one treatment the diners pay individually; in a second treatment they split the bill evenly between the six group members. In yet a third treatment, the meal is paid for entirely by the experimenter. Economic theory prescribes that consumption will be smallest when the payment is individually made, and largest when the meal is free, with the even split treatment in-between the other two. The restaurant findings are consistent with these predictions. A fourth treatment, in which each participant pays only 1/6 of her own consumption costs and the experimenter pays the remainder, is introduced to control for possible unselfish and social considerations. The marginal cost imposed on the participants in this treatment is the same as in the even split treatment. However, the externalities are removed: in the even split case, increasing an individual's consumption by $1 increases the individual's cost, as well as the cost of each of the other participants, by $1/6. In the fourth treatment, this will increase only the individual's cost by $1/6 but will have no effect on the payment of the other participants. In other words, the negative externality present in the even-split treatment is completely eliminated. If participants are completely selfish, the fourth treatment should not affect their consumption relative to the second treatment of the even split. On the other hand, if they care also for the well being of the other participants (or for social efficiency), they can be expected to consume more in the last treatment than in the even split treatment.

The efficiency implication of the different payment methods is straightforward. When splitting the bill, diners consume such that the marginal social cost they impose is larger than their own marginal utility and, as a result, they over-consume relative the social optimum. In fact, it is easy to show that the only efficient payment rule is the individual one. It turns out that subjects' preferences are consistent with increasing efficiency. When asked to choose, prior to ordering, whether to split the bill or pay individually, 80% choose the latter. That is, they prefer the environment without the externalities. However, in the presence of externalities, they nevertheless take advantage of others.

One example of an environment in which the selfishness hypothesis was studied is public good games; see Davis and Holt (1993) and Ledyard (1995) for comprehensive reviews. Public Goods experiments in which non-contribution is a dominant strategy typically find that subjects are sensitive to free riding incentives but nonetheless cooperate at a level that cannot be fully explained by mainstream economic theory.[2] However, as the typical public goods game is repeated (regardless of whether opponents are the same or different), contributions fall substantially (Kim and Walker, 1984; Isaac and Walker, 1988; Andreoni, 1988; Asch *et al.*, 1993; Weimann, 1994). In all these studies, subjects contribute less and less

---

[2] In fact, even in public goods games where some positive contribution is best-response, subject tend to substantially over-contribute relative to their best-response (Keser, 1996).

the longer they play. In other words, it seems that subjects may be contributing in part due to inexperience or confusion under lab conditions. Kim and Walker (1984) reviewed previous experiments that found little or no free riding. They raised serious concerns about lab experiments among which were misunderstanding and vagueness as well as insufficient economic incentives. In an experiment designed to overcome the criticisms raised, they indeed found that selfish behaviour was in fact prevalent after only a few repetitions. Andreoni (1995) raised similar criticisms, which he labelled collectively 'confusion'. In order to explore this issue, he designed a zero-sum version of the public goods game, in which the sets of strategies and corresponding token payoffs were the same as the public goods game but where token payoffs were mapped to monetary payoffs by the earning ranks of the subjects. This mapping eliminated the monetary incentive to cooperate; indeed cooperation dropped significantly but not entirely. Andreoni (1995, p. 900) concluded that 'on average about 75 percent of the subjects are cooperative, and about half of these are confused about incentives, while about half understand free-riding but choose to cooperate out of some form of kindness'.

A traditional lab environment could present some limitations when extrapolating to real life settings. Such limitations may result from participants' lack of familiarity with the lab setting. It could be argued that subjects should be observed in settings with which they are familiar and experienced. For example, in a field experiment conducted during orange picking season in Israel (Erev *et al.*, 1993), with different groups of four workers facing different payment schemes, it was found that, in line with the theoretical prediction, a collective payment resulted in substantial free riding and 30% loss in production.

The current study proposes the restaurant setting as a setting with which subjects are expected to be familiar, thereby reducing the possibility of confusion. The idea of studying human economic behaviour in a restaurant setting is not new. In a study discussed in Thaler (1980) costumers at an all-you-can-eat pizza restaurant were randomly given free lunches. These consumers ate less than the control group who paid the $2.50 normal bill. The main conclusion of that study was that, unlike the prescription of economic theory, people do not ignore sunk costs.

The paper is organised as follows: Section 1 sketches the theory as it pertains to the unscrupulous diner's dilemma and derives the appropriate hypotheses implied by the theory. Section 2 details the design and procedures for the restaurant setting. Section 3 lists and explains the results and investigates possible implications of gender issues. Section 4 presents a related laboratory experiment. Section 5 concludes.

## 1. Theory

In this Section we first introduce the mainstream assumptions and the resulting social inefficiency under the even-split and free-meal treatments. We then posit the hypotheses implied by the theory.

## 1.1. *Mainstream Assumptions*

According to standard economic assumptions, consumers will find it optimal to increase consumption when marginal benefit exceeds marginal cost and to lower consumption when the opposite holds. Therefore, at the utility maximising consumption level, marginal cost must equal marginal benefit. It is also a standard assumption that the marginal utility is decreasing (clearly the marginal utility reaches zero at some point, else consumers in the free meal treatment would consume at a level of infinity). Given these standard assumptions, economic theory predicts a negative relation between the marginal cost of the food and its consumption.

If the individuals do not internalise the negative externalities they impose on others, they will over-consume relative to the social best in all but the individual pay treatment. In particular, if the six diners elected a social planner to 'dictate' the allocations in the even split treatment, this planner would be able to increase the value received by each diner. Similarly, under the free-meal treatment, the party financing the dinner could pay the diners to consume at individual-pay levels, such that all diners as well as the paying party would be better off. It follows that the individual-pay outcome is a Pareto improvement relative the other treatments.

## 1.2. *Hypotheses*

Several hypotheses emerge from the theory of selfish utility-maximising consumers. Since marginal benefit must equal marginal cost, it will take greater and greater consumption to equate marginal benefit to marginal cost, as we move from individual-pay to even-split and from even-split to free-meal. The following three hypotheses emerge:

> *Hypothesis* 1: Diners will eat more in the even-split treatment relative to the individual-pay treatment.
> *Hypothesis* 2: Diners will eat more in the free-meal treatment relative to the even-split treatment.
> *Hypothesis* 3: Diners will eat more in the free meal treatment relative to the individual pay treatment.

## 2. Design of the Restaurant Setting

Subjects were recruited through signs posted around the Technion campus, which promised a large amount for a one-hour experiment and invited them to call for information about the experiment. Upon calling, they were informed that the experiment would be conducted at a popular restaurant near the Technion campus. They were asked to show up at a specific time (during lunchtime). Six subjects, three males and three females, were invited for each time slot. A conscious effort was made not to invite students who were familiar with each other to the same treatment. Upon arrival, subjects received a show up payment of 80 NIS (roughly $20 at the time of the study) and short instructions (see Appendix A for the translation of the instructions from Hebrew). They were cautioned to maintain absolute silence for 10 minutes, during which all participants were asked to

complete the questionnaires in front of them. The questionnaires requested subjects to rate themselves on a wide range of emotions. They were told to expect the same questionnaire at the end of the meal.

In the instructions for the questionnaire, subjects were informed they would be able to order from the restaurant menu following the completion of the questionnaire. They were asked to indicate their orders on a designated sheet of paper. Subjects wrote down their orders individually and separately, without any ability to communicate or coordinate with other participants. The intent of the questionnaire was to maintain the independence between observations, and the questionnaire was effective in keeping subjects silent.

Treatments differed only in the payment mechanism specified in the instructions: in the individual payment treatment, subjects were told that they would pay for their own meal. In the even split treatment, subjects were told that the bill would be evenly split between the six of them. In the free meal treatment, subjects were told that the meal would be fully paid for by the experimenter. Four groups of six subjects participated in the individual payment treatment and four groups participated in the even split treatment. Two groups participated in the free lunch treatment. Two of the groups in each of the first two treatments were asked how they would prefer to pay – individually or by splitting the bill – prior to being informed of the actual payment mechanism.

The menu included a variety of renowned international cuisine, with numerous delectable categories to encompass a wide range of tastes. Waiters were instructed not to communicate or otherwise interact with subjects before picking up the order sheet. That is, subjects had contact only with the experimenters before they ordered. The same two experimenters attended all treatments.

## 3. Results

### 3.1. *The Three Main Treatments*

Table 1 below summarises the results of the field study. For each of the three treatments discussed in Section 1, the first column reports the gender of the subject, the second column reports the number of items that subject ordered, and the third column reports the cost of the subject's meal. Subjects are ordered by the cost of their meals from highest to lowest.

Note the variability in subjects' costs for any given treatment: For example, the difference in cost between the least expensive subject and the most expensive subject in the free meal treatment was NIS 119. Normally, such heterogeneity could pose somewhat of a problem for hypothesis testing. However, despite this enormous variability there is a fairly small overlap in meal costs between treatments. Treatment 3, for example, has only two observations out of 12 that fall below the highest observation of 24 observations in treatment 1. This surprisingly small overlap is clearly depicted in Figure 1.

The x-axis lists three values, corresponding to the three treatments. The y-axis represents the cost of the meals. Each point in the plot represents the meal cost in

Table 1

*Summary of the Results*

| | Individual pay | | | Even split | | | Free meal | | |
|---|---|---|---|---|---|---|---|---|---|
| Subject | Sex | Items | Cost | Sex | Items | Cost | Sex | Items | Cost |
| 1 | F | 2 | 59 | F | 2 | 81 | M | 4 | 168 |
| 2 | M | 2 | 54 | M | 2 | 73 | M | 5 | 123 |
| 3 | M | 2 | 50 | M | 2 | 71 | M | 3 | 101 |
| 4 | M | 2 | 49 | F | 1 | 66 | F | 4 | 94 |
| 5 | M | 2 | 47 | F | 2 | 64 | M | 3 | 81 |
| 6 | F | 1 | 46 | F | 3 | 62 | M | 3 | 75 |
| 7 | F | 2 | 45 | F | 2 | 60 | F | 3 | 69 |
| 8 | M | 2 | 45 | M | 2 | 59 | F | 2 | 61 |
| 9 | F | 2 | 43 | F | 2 | 59 | F | 2 | 57 |
| 10 | M | 2 | 43 | M | 3 | 56 | M | 3 | 59 |
| 11 | M | 2 | 40 | M | 2 | 52 | F | 2 | 51 |
| 12 | F | 2 | 40 | F | 2 | 47 | F | 2 | 49 |
| 13 | F | 2 | 39 | M | 2 | 46 | | | |
| 14 | M | 2 | 39 | F | 2 | 46 | | | |
| 15 | F | 1 | 35 | M | 2 | 45 | | | |
| 16 | F | 1 | 35 | M | 2 | 45 | | | |
| 17 | M | 1 | 35 | M | 2 | 44 | | | |
| 18 | F | 2 | 31 | M | 2 | 40 | | | |
| 19 | F | 2 | 31 | M | 2 | 40 | | | |
| 20 | F | 2 | 30 | F | 1 | 39 | | | |
| 21 | M | 1 | 16 | F | 1 | 37 | | | |
| 22 | M | 1 | 16 | F | 1 | 35 | | | |
| 23 | F | 1 | 15 | M | 2 | 33 | | | |
| 24 | M | 1 | 12 | F | 1 | 22 | | | |
| Avg. | | 1.67 | 37.3 | | 1.87 | 50.9 | | 3 | 82.3 |

NIS for a particular subject in one of the three treatments. Different subjects are represented by different symbols. All in all, there are 24 values for each of the first two treatments and 12 for the third.
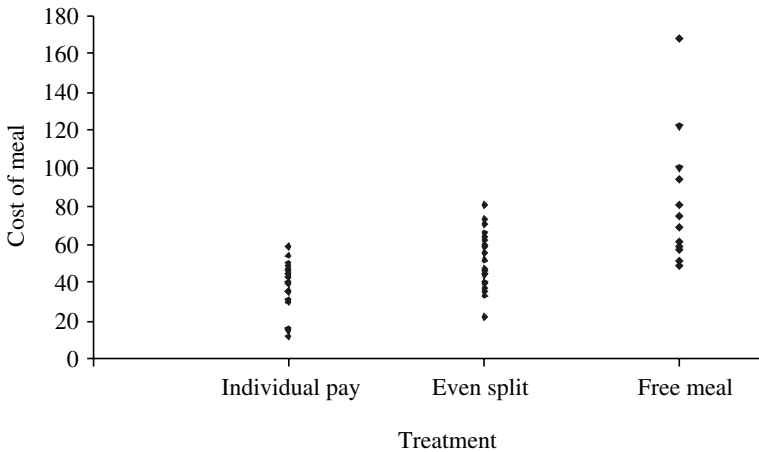


Fig. 1. *Summary of the Results*

Recall Hypotheses 1–3. These three hypotheses postulated differences between individual pay and even split, between even split and free meal, and between individual pay and free meal.

To test Hypotheses 1–3, we use two competing tests to determine if there are any reliable differences between each two independent groups – the parametric t-test and the nonparametric Mann-Whitney U-test. We find the 'number of items ordered' not informative in the Mann-Whitney U-test due to the large number of ties. We expect it to provide a rough statistic at best under the t-test as well since it can hardly be assumed to satisfy normality assumptions required by the t-test. We nonetheless report p-values for this test. The cost of the meals, however, provided clear-cut evidence to conclude that the samples are significantly different under all three hypotheses. Table 2 shows p-values for the three hypotheses, using the Mann-Whitney test on cost (column 1), the t-test on cost (column 2), and the t-test on number of items (column 3).

Finally, we use a graphical depiction of the population differences for the first three treatments to emphasise these results, using cumulative distribution graphs in Figure 2.

We see from the above graphs that in the individual-pay treatment costs tend to be substantially lower in all percentiles of the distribution relative to the even-split and free-meal treatments. Similarly, the even-split treatment costs tend to be substantially lower in all percentiles of the distribution relative to the free-meal treatment.

Table 2

*Hypothesis-tests for $H_o$: p-values for Hypotheses 1–3*

|  | Mann-Whitney U-test (one-sided) on cost of meal | Mann-Whitney U-test (one-sided) on number of items* | t-test (one-tail) on cost of meal | t-test (one-tail) on number of items ordered |
|---|---|---|---|---|
| Individual pay vs. even split | 0.0014 | 0.0948 | <0.0001 | 0.0818 |
| Individual pay vs. free | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| Even-split vs. free | 0.0008 | <0.0001 | 0.0003 | <0.0001 |

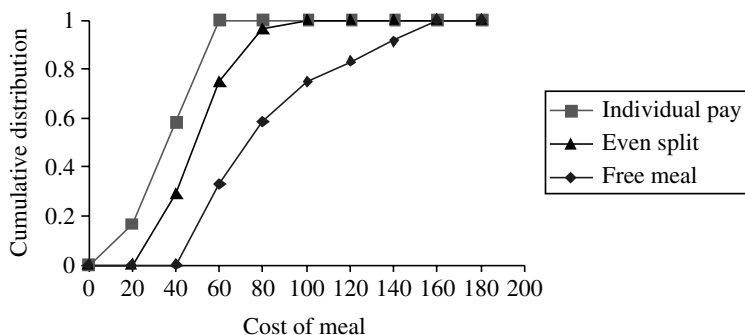*Mann-Whitney on the number of items ordered may be unreliable due to the large number of ties.



Fig. 2.  *Cumulative Distribution Graphs for Treatments 1–3*

3.2. *Regard for Others in the Restaurant Setting*

Though it is clear from the above results that an individual diner appears not to account fully for the cost her consumption imposes on her peers, the even-split treatment nonetheless leads to some other questions. In particular, does the individual ignore all of the cost she imposes on others, or does she account for some? The literature seems to present many approaches to answering questions of this kind. For example, we could suppose that the individual does not fully exploit her ability to consume at others' expense since her utility is increasing in the consumption of others; for a recent review of the altruism and fairness literature, see Camerer (2003).

The fourth restaurant treatment is introduced to examine the proposal that selfish considerations may not offer the best description of human agents. In that treatment, two groups of six diners, three males and three females each, were recruited by signs around the Technion campus. The groups were summoned to the same restaurant used in the other treatments. Also, as in the other treatments, the groups were balanced between men and women. Unlike the other treatments, the instructions specified to the subjects that at the end of meal each would be asked to pay individually only one-sixth of his individual meal cost (see Appendix A for the translation of the instructions from Hebrew).

By the 'selfish agent assumption', this treatment should not differ from the even-split treatment. The theories of altruism, equity, and reciprocity, however, would appear to suggest that agents are likely to consume more in this treatment relative to even split treatment. This is because no costs are imposed on others in the one-sixth-individual-pay treatment and hence regard for others does not play a role in this treatment,[3] whereas regard for others is expected to play some role in the even-split treatment. In other words, positive regard for others will raise the marginal cost of a meal unit and will therefore lower the optimal spending under the assumption of decreasing marginal benefit. It is important to note that whereas altruism or utilitarian motives are unequivocal in this prediction, distributive and reciprocity concerns depend crucially on expectations and could suggest predictions in either direction. We derive Hypothesis 4:

> *Hypothesis 4*: Diners will exhibit the same levels of consumption in the Even-Split treatment and the 'Pay 1/6' treatment.

The last Hypothesis addressed unselfish motives and postulated a difference between the even-split and 1/6 individual payment. Table 3 presents the results of the fourth treatment. Table 4 shows p-values for this Hypothesis, using the Mann-Whitney test on cost (column 1), the Mann-Whitney test on the number of items ordered (column 2), the t-test on cost (column 3), and the t-test on the number of items ordered (column 4).

---

[3] This is assuming of course that regard for the experimenter is weaker than that for the peers around the table. However, if we rely on the results reported in Section 3.1, we should not be concerned about this possibility: In the free-meal treatment of Section 3.1, subjects overconsumed relative to both other treatments by a phenomenal amount. Hence it seems that any concern for the experimenter's welfare is miniscule at best.

Table 3

*Summary of the Results Relevant to Hypothesis 4*

| | Even split | | | Pay 1/6 | |
|---|---|---|---|---|---|
| Sex | Items | Cost | Sex | Items | Cost |
| F | 2 | 81 | F | 3 | 101 |
| M | 2 | 73 | M | 3 | 85 |
| M | 2 | 71 | M | 3 | 74 |
| F | 1 | 66 | M | 2 | 59 |
| F | 2 | 64 | M | 2 | 58 |
| F | 3 | 62 | F | 2 | 57 |
| F | 2 | 60 | F | 2 | 57 |
| M | 2 | 59 | M | 2 | 51 |
| F | 2 | 59 | F | 2 | 50 |
| M | 3 | 56 | F | 2 | 46 |
| M | 2 | 52 | F | 1 | 26 |
| F | 2 | 47 | M | 1 | 25 |
| M | 2 | 46 | | | |
| F | 2 | 46 | | | |
| M | 2 | 45 | | | |
| M | 2 | 45 | | | |
| M | 2 | 44 | | | |
| M | 2 | 40 | | | |
| M | 2 | 40 | | | |
| F | 1 | 39 | | | |
| F | 1 | 37 | | | |
| F | 1 | 35 | | | |
| M | 2 | 33 | | | |
| F | 1 | 22 | | | |
| Avg. | 1.87 | 50.9 | Avg. | 2.08 | 57.4 |

Table 4

*p-values for Hypothesis 4*

| | Mann-Whitney U-test (one sided) on cost of meal | Mann-Whitney U-test (one sided) on number of items ordered | t-test (one-tail) on cost of meal | t-test (one-tail) on number of items ordered |
|---|---|---|---|---|
| Even-split vs. 1/6 Pay | 0.187 | 0.167 | 0.145 | 0.159 |

   Looking at both the cost of meals and at the number of items ordered, we find no significant differences between the even-split and 'Pay 1/6' treatments. This finding can be contrasted with the very significant differences between the even-split treatment and all other treatments we found in Section 3.1. We should caution, however, that the lack of significance could be due to the smaller sample size of the 'Pay 1/6' group.

### 3.3. *Gender Differences*

Studies have shown that males and females have different propensities for regard for others. In experiments, we find results in dictator games (Eckel and Grossman,

1998) and reward allocation games (Lane and Messe, 1971), which show more generosity in females than in males.[4] Such results would lead us to expect women to not exploit the ability to impose cost on others to the same extent as men.

Another plausible gender difference has to do with different physical limitations as well as discriminatory cultural norms related to eating. Both physical capacity and discriminatory social norms would lead one to expect women to face a lower ceiling on food consumption. Hence we would expect women to underutilise the ability to impose costs on others.

Surprisingly (or not, depending on one's prior), as Table 5 clearly shows, men and women do not differ in their consumption levels in three out of four treatments under any reasonably acceptable level of significance. In the free meal treatment, however, men tend to eat more than women in a manner marginally significant at the 5% significance level. However, given the lack of difference in the other three treatments, we tend to discount this finding.

## 4. A Laboratory Comparison

Cross-country studies (Roth *et al.*, 1991) have raised the possibility that subject pools in different countries may not share the same distribution of other-regarding preferences. Such cultural differences could affect the comparability of the present subject pool to other subject pools in the world. To exclude the possibility that our restaurant finding is driven by an odd subject pool, we briefly present the results of a simple negative externality experiment conducted in the lab with the same subject pool (Technion students) as in the restaurant study.

### 4.1. *The Laboratory Setting*

Subjects were recruited through signs around the Technion as in the previous setting. Instead of meeting at a restaurant, however, subjects were summoned to the laboratory. The show up fee was 80 NIS, the same as the restaurant setting. As before, there were six subjects per session. All subjects were in the same room and could see each other. Subjects were shown a 'production' table detailing the cost and revenue resulting from each production quantity, where

Table 5

*Reported t-test p-values (2-tail) for the Null Hypothesis of no Gender Effects*

|  | Gender effect |
| --- | --- |
| Individual pay | 0.9623 |
| Even split | 0.8470 |
| Free meal | 0.0568 |
| 1/6 payment | 0.8530 |

[4] There are also studies that reject claims of gender differences (Bolton and Katok, 1995).

Table 6

*Production Table for Laboratory Examination of the Diner's Dilemma*

| Units | Total cost | Total revenue |
|-------|-----------|---------------|
| 1 | 10 | 18 |
| 2 | 20 | 32 |
| 3 | 30 | 44 |
| 4 | 40 | 48 |
| 5 | 50 | 50 |
| 6 | 60 | 51 |
| 7 | 70 | 51.75 |
| 8 | 80 | 51.75 |
| 9 | 90 | 52 |

the quantity of production could vary from 1 to 9. The production table is presented in Table 6.

Three different treatments were run. In the first treatment, subjects bore the full cost and reaped the full revenue from each unit of production. In the second treatment, subjects reaped the full revenue but the total cost of production was added up over subjects and then divided equally among subjects. In the third treatment, subjects incurred only one-sixth of the full cost of their production but the remainder was not imposed on any of the participants; instead, it just vanished.

The parallel between this production problem and the unscrupulous diner's dilemma in the restaurant setting is evident. Namely, the first production treatment corresponds to the restaurant's individual-pay treatment, the second production treatment corresponds to the restaurant's even-split treatment and the third production treatment corresponds to the restaurant's one-sixth-pay treatment.

### 4.2. *Results*

The full list of 36 participants' choices is presented in Table 7.

Table 7

*Results of Laboratory Experiment*

| Observation | Individual-pay | Even-split | One-sixth-pay |
|-------------|---------------|------------|---------------|
| 1 | 3 | 2 | 3 |
| 2 | 3 | 2 | 4 |
| 3 | 3 | 3 | 4 |
| 4 | 3 | 3 | 5 |
| 5 | 3 | 3 | 5 |
| 6 | 3 | 3 | 5 |
| 7 | 3 | 3 | 5 |
| 8 | 3 | 4 | 5 |
| 9 | 3 | 4 | 5 |
| 10 | 3 | 4 | 5 |
| 11 | 3 | 5 | 5 |
| 12 | 5 | 5 | 5 |
| Average | 3.17 | 3.42 | 4.67 |

   In contrast to the restaurant results, the difference between the individual-pay and even-split treatments is not significant[5] (one-tail p-value = 0.23), whereas the difference between the even-split and one-sixth-pay treatments is highly significant (one-tail p-value = 0.0007). This evidence of unselfish motives might lead to the conclusion that the bill-splitting convention would not result in any significant social detriment. This conclusion does not appear consistent with what was observed in the restaurant.

   This finding shows that the results typically reported in the experimental literature are easily replicated with the subject pool in the restaurant study. Clearly, in terms of design, many things are different between the restaurant study and the lab experiment.

## 5. Conclusions

The literature on negative externalities is based on the prediction that an economic agent who is able to impose some of the cost of his consumption on others will over-consume relative to the socially efficient level. This is a direct result of the assumption that economic agents equate individual marginal costs and marginal benefits with complete disregard for the costs imposed on others or for social efficiency. However, there is an emergent volume of evidence placing in doubt some of the assumptions of classical economic theory. Such studies often demonstrate that small groups in the laboratory are likely to secure voluntary cooperation.

   The unscrupulous diner's dilemma gives us an opportunity to test this prediction in an environment close to real-life experiences. We find that the theoretical predictions work: people react to changes in incentives and they largely seem to ignore negative externalities. These results have great importance in the design of institutions. Institutions and rules that ignore the effect of negative externalities are inefficient – not only in theory, but also in practice. This inefficiency is the result of people playing the equilibrium of the game; even if they all prefer to be in a 'different game' (e.g., pay the bill individually). Interestingly, when asked which mechanism they would prefer, prior to informing them which mechanism they would face, 19 (80%) out of the 24 subjects we asked indicated they would prefer the individual pay over splitting the bill. However, when forced to play according to the less preferred set of rules (splitting the bill), subjects nevertheless minimise their losses by taking advantage of others.

   Given the clear preferences of the diners in our study, we are left wondering why we ever observe splitting of the bill in restaurants and more importantly, in economic institutions. We begin with the restaurant setting. Unlike our experiment, groups of diners eating together are generally not perfect strangers but rather friends or colleagues. Likewise, the custom of splitting the bill is generally

---

[5] It is interesting to note, however, that in the even-split treatment two subjects selected a quantity of two units, which is below the socially efficient level of production. Errors by subjects in the lab are not uncommon, nor are dominated choices unusual (as abundant evidence from second price sealed bid auctions shows). However, it is possible that the added level of complexity in understanding the even-split mechanism has resulted in a higher chance of making errors.

prevalent among friends or colleagues and not among strangers. This difference is critical since with friends and colleagues the game is repeated, so punishment strategies in response to excessive waste are feasible. Nevertheless, one would expect some waste to result even among friends, since monitoring and punishment are imperfect. In that case, is there any reason why one would prefer to split the bill?

Some cost is involved in paying individually. A part of it could be the mental cost of figuring out one's share of the bill, and calculating the portion of the tax and tip that apply to that share. Another part would be the social cost of appearing stingy or unfriendly. Given the cost of individually paying and the ability to reduce the inefficiency of splitting the bill through repeated game strategies, it may in fact be individually and socially optimal to split the bill among friends. However, the danger in customs which are based on rational decision making is that once they become conventions they are resistant to change, even when circumstances change. For example, when you find yourself dining with distant acquaintances you are not likely to encounter any time soon, it may nevertheless be rude in some settings (e.g., conferences) to suggest paying individually.

An argument of socially inefficient conventions could be made for larger and less personal economic institutions. For example, until the twentieth century, allocating fishing rights in coastal waters would have been a socially inefficient proposition. However, years of convention have produced the shibboleth of 'freedom of the seas' advocated by maritime nations, which is most certainly socially inefficient with large scale fishing methods and inexhaustible demand by a growing human population. Similarly, the practice of common grazing areas in fourteenth century English villages quickly became unsustainable once populations started growing. The practice of the commons is in fact not far different from the diner's problem. Though individual incentives for excess exist, in small communities the social mechanisms arising from repeated interaction and strong other-regarding preferences are in place to discourage any such excess. Once these social mechanisms are eliminated, the tragedy of the commons results.

Finally, small groups in the laboratory, including the laboratory experiment presented here, have been shown to arrive closer to the socially efficient level than models of selfish behaviour would. Given this common result, other-regarding preferences in many instances have been argued to be critical motives in decision making. Though the findings in the restaurant setting cannot preclude other-regarding considerations, they provide evidence in favour of other possible explanations for the results generally obtained in the laboratory. Such explanations include the concern of Kim and Walker (1984) that misunderstanding of the unfamiliar task could result in cooperation and the concern of Andreoni (1995) that some cooperation could be due to confusion and lack of experience with the task. In contrast to unfamiliar laboratory tasks, the restaurant is a familiar setting and ordering at a restaurant (as well as splitting the bill at a restaurant) is a familiar task. Another possibility is that the difference may be driven by the perception of the subjects regarding the task. In laboratory experiments the subjects may perceive that they were brought to the lab in order to test their attitudes towards

public goods, fairness etc. This perception is less likely in our restaurant setting, where subjects may behave in a somewhat more natural manner. Though these explanations and others remain to be studied, we hope this study has provided food for thought.

*University of Chicago and Technion*
*University of Texas at Dallas*

*Date of receipt of first submission: August 2002*
*Date of receipt of final typescript: July 2003*

## Appendix A

*Participant Instructions Sheet for the Restaurant Settings*

Welcome to 'Globes' Restaurant
   This experiment looks at emotions before and after eating. You therefore will be asked to eat.
   Within the next 10 minutes you must perform two tasks:

   1.  Fill out the questionnaire in front of you honestly and accurately.
   2.  Check the menu and write down your order on the empty sheet attached to the questionnaire. You will not have another opportunity to order. At the end of 10 minutes, the waitress will pick up your order.

   It is imperative that you **remain silent**. That is, do not communicate with the other participants at the table.
   Following the 10 minutes, before the meal, you will receive NIS 80 for your participation in the experiment.
   [Treatment 1:] At the end of the experiment you will receive a bill *for the food you order.* You will then have to pay the waitress. After that, you will be asked to fill out the same questionnaire.
   [Treatment 2:] At the end of the experiment you will receive a *bill for one-sixth of the entire bill of all participants at the table.* You will then have to pay the waitress. After that, you will be asked to fill out the same questionnaire.
   [Treatment 3:] At the end of the experiment you will be asked to fill out the same questionnaire. *You do not have to pay the bill.* The meal is on us!
   [Treatment 4:] At the end of the experiment you will receive a *bill for one-sixth of the cost of your individual order, which you will then have to pay the waitress.* After that, you will be asked to fill out the same questionnaire.
   **Bon appetit!**

## Appendix B

*Participant Instructions Sheet for the Laboratory Settings*

Welcome. This is an experiment in decision making. You will receive NIS 80 for showing up to the experiment, plus any amount that you earn in the course of the experiment. In the next 10 minutes we ask that you read the instructions and make your choice of number of units to purchase. This is the only decision you will have to make in the experiment. You have only one chance to make a choice, after which the experiment ends. Hence, it is crucial

that you make your choice carefully. If you have any questions, please raise you hand but do not exclaim out loud. We expect and appreciate your cooperation.

[Treatment 1:] Your choice is in terms of quantity, or *number of units*, you wish to purchase. At the end of the experiment, we will pay you according to how many units have purchased, but we will also charge you the *cost* for these units. So the earnings you take home at the end of the experiment, in addition to the show up fee, are your revenue from the units you bought, minus the cost of the units you bought. The Table below specifies the revenue and cost from each quantity you choose. The amount you earn in this experiment is independent of the choices and earnings of other participants.

[Treatment 2:] Your choice is in terms of quantity, or *number of units*, you wish to purchase. At the end of the experiment, we will pay you according to how many units have purchased, but we will also charge you the *cost* for these units as follows: the total cost of the quantity you choose will be added to the total costs of others' choices (there are six others in your group). You will then be asked to pay 1/6 of the total cost of everybody in your group. However, you revenue will be only the *revenue* corresponding to your individual choice. So the earnings you take home at the end of the experiment, in addition to the show up fee, are your individual revenue from the units you bought, minus the 1/6 the cost of the units *everybody in your group* bought. The Table below specifies the individual revenue and cost from each quantity you choose.

[Treatment 3:] Your choice is in terms of quantity, or *number of units*, you wish to purchase. At the end of the experiment, we will pay you according to how many units have purchased, but we will also charge you 1/6 of the *cost*. So the earnings you take home at the end of the experiment, in addition to the show up fee, are your individual revenue from the units you bought, minus the 1/6 the cost of the units. The Table below specifies the individual revenue and cost from each quantity you choose. The amount you earn in this experiment is independent of the choices and earnings of other participants (Table 8).

I choose to get a quantity of ———— units.

Table 8

*The Individual Revenue and Cost from Each Quantity*

| Units | Total cost | Total revenue |
| --- | --- | --- |
| 1 | 10 | 18 |
| 2 | 20 | 32 |
| 3 | 30 | 44 |
| 4 | 40 | 48 |
| 5 | 50 | 50 |
| 6 | 60 | 51 |
| 7 | 70 | 51.75 |
| 8 | 80 | 51.75 |
| 9 | 90 | 52 |

# References

Andreoni, J. (1988). 'Why free ride? Strategies and learning in public goods experiments', *Journal of Public Economics*, vol. 37(3), pp. 291–304.

Andreoni, J. (1993). 'An experimental test of the public-goods crowding-out hypothesis', *American Economic Review*, vol. 83(5), pp. 1317–27.

Andreoni, J. (1995). 'Cooperation in public goods experiments: kindness or confusion?', *American Economic Review*, vol. 85(4), pp. 891–904.

Asch, P., Gigliotti, G. A. and Polito, J. (1993). 'Free riding with discrete and continuous public goods: some experimental evidence', *Public Choice*, vol. 77(2), pp. 293–305.

Bohnet, I., Frey, B. and Huck, S. (2001). 'More order with less law: on contract enforcement, trust and crowding', *American Political Science Review*, vol. 95(1), pp. 131–44.

Bolton, G. and Katok, E. (1995). 'An experimental test for gender differences in beneficent behavior', *Economics Letters*, vol. 48, pp. 287–92.

Camerer, C. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.

Davis, D. and Holt, C. (1993). *Experimental Economics*. Princeton, NJ: Princeton University Press.

Dawes, R.M. and Thaler, R. (1988). 'Anomalies: cooperation', *Journal of Economic Perspectives*, vol. 2(3), pp. 187–97.

Eckel, C. and Grossman, P. (1998). 'Are women less selfish than men? Evidence from dictator experiments', ECONOMIC JOURNAL, vol. 108, pp. 726–35.

Erev, I., Bornstein, G. and Galili, R. (1993). 'Constructive intergroup competition as a solution to the free rider problem: a field experiment', *Journal of Experimental Social Psychology*, vol. 29, pp. 463–78.

Glance, N. S. and Huerman, B. A. (1994). 'The dynamics of social dilemmas', *Scientific American*, (March), pp. 76–81.

Gneezy, U. and Rustichini, A. (2000). 'A fine is a price', *Journal of Legal Studies*, vol. 29, pp. 1–17.

Isaac, M. and Walker, J. (1988). 'Group size and the voluntary provision of public goods: experimental evidence utilizing large groups', *Journal of Public Economics*, vol. 54, pp. 1–36.

Keser, C. (1996). 'Voluntary contributions to a public good when partial contribution is a dominant strategy', *Economics Letters*, vol. 50, pp. 359–66.

Kim, O. and Walker, M. (1984). 'The free rider problem: experimental evidence', *Public Choice*, vol. 43, pp. 3–24.

Lane, I. M. and Messe, L. A. (1971). 'Equity and distribution of rewards', *Journal of Personality and Social Psychology*, vol. 20, pp. 1–17.

Ledyard, J. O. (1995). 'Public goods: a survey of experimental research', In (J. Kagel and A. Roth, eds.), *The Handbook of Experimental Economics*, pp. 111–94, Princeton, NJ: Princeton University Press.

Roth, A. E., Prasnikar, V., Okuno-Fujiwara, M. and Zamir, S. (1991). 'Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: an experimental study', *American Economic Review*, vol. 81(5), pp. 1068–95.

Schulze, G. and Frank, B. (2003). 'Deterrence versus intrinsic motivation: experimental evidence on the determinants of corruptibility', *Economics of Governance*, vol. 4(2), pp. 143–60.

Thaler, R. (1980). 'Toward a positive theory of consumer choice', *Journal of Economic Behavior and Organization*, vol. 1, pp. 39–60.

Weimann, J. (1994). 'Individual behavior in a free riding experiment', *Journal of Public Economics*, vol. 54(2), pp. 185–200.