 Experimental Evidence on Tax Compliance and Voluntary Public Good Provision

Abstract - Existing experimental literature on tax compliance and crowding–out examines either the incentive to evade tax or the incentive to give, but not both. This paper provides an experimental examination of the behavior of tax evasion and voluntary contributions when both publicly and privately provided public goods are present. The experimental evidence suggests that the privately provided public good is a substitute for the publicly provided public good, but the converse does not hold, and that the level of compliance may be underestimated, ceteris paribus, if private contributions are not taken into account.

INTRODUCTION

The issue of tax compliance has attracted extensive attention since Allingham and Sandmo’s (1972) pioneering theoretical paper. Due to the advantage of quarantining unwanted disturbances outside the laboratories, in recent years more and more researchers have relied on experimental methods to explore the behavior of tax evasion. Among them, Alm and his coauthors have conducted a series of inspiring tax compliance experiments (Alm, McKee, and Beck, 1990; Alm, Jackson, and McKee, 1992a, 1992b, 1993; Alm, McClelland, and Schulze, 1992; Alm, Sanchez, and de Juan, 1995; and Alm and McKee, 1998). Their experiments are of particular interest since they consider the tax–funded public good that is incorporated in the theoretical model posed by Cowell and Gordon (1988), while earlier experimental studies, for instance, Spicer and Becker (1980), Spicer and Thomas (1982), Spicer and Hero (1985), Robben, Webley, Elffers, and Hessing (1990), and Collins and Plumlee (1991), were generally developed from the theoretical work by Allingham and Sandmo (1972) and a later extension by Yitzhaki (1974), in which the expenditure in relation to tax revenue is neglected.

Alm and his coauthors apply the voluntary contribution mechanism, which is frequently used in experiments related to the voluntary provision of public goods, to tax compliance experiments. In this mechanism, subjects voluntarily contribute some or all of their endowments to a privately provided public good. The extent to which they benefit from the public good depends on the total contributions made by the group. Instead of contributing to the public good
directly, in Alm and his coauthors’ tax compliance experiments, subjects decide on the amount of income reported to the tax authority. Then the declared income is taxed and the taxes are used to fund the public good. The level of the public good is determined based on the total amount of taxes paid by the group. The rest of their story is similar to the theoretical settings and other experimental designs in the literature: subjects’ true incomes are audited by a certain or an uncertain probability. The subject who is audited and caught underreporting will pay the tax evaded and a fine, which is some multiple of the evaded tax. A major finding of their studies is that the level of compliance is higher when the taxes collected are used to fund the public good than when the taxes are discarded.

Although neglecting the use side of tax revenue is unrealistic in exploring the behavior of tax evasion, the tax–funded public good is not the only category of public good in reality. As observed in the real world, many public goods are provided by the private sector alone and receive no support from the government. Churches, the Red Cross, private shelters, private dog pounds, and private lighthouses are examples. Since public goods are not restricted to being provided by the public sector, the employment of the voluntary contribution mechanism in Alm et al.’s experiments actually gives rise to some interesting questions. Besides paying taxes to fund the publicly provided public good, if individuals also contribute voluntarily to the public good provided by the private sector, how will they allocate their income between the two public goods? Furthermore, will compliance be lower if the privately provided public good is valued relatively more highly than the publicly provided public good? Will taxes crowd out private contributions and will the crowding–out become more severe if the publicly provided public good is valued relatively more highly than the privately provided public good?

The effect of taxes on private contributions has been extensively discussed in the crowding–out literature. The theoretical papers by Warr (1982, 1983), Roberts (1984, 1987), and Bergstrom, Blume, and Varian (1986) show that government contributions to charity completely crowd out private contributions. The reason for this result is that they assume that individuals are purely altruistic, caring only about the total contributions to the third party, but not the relative magnitudes of the private and government contributions that the total contributions are composed of. This setting implies that private and government contributions are perfect substitutes and, therefore, individuals will simply substitute government contributions for private contributions dollar for dollar as their taxes increase. As a result, taxes are neutral, having no effect on total contributions.

This neutrality or complete crowding–out result has been challenged by some theoretical examinations based on impure altruism. For instance, besides caring about the total contributions Andreoni (1989, 1990) assumes that individuals also obtain some private benefit from the act of giving, which he refers to as “warm–glow.” The “warm–glow” component makes government contributions imperfect substitutes for private contributions and, therefore, individuals will reduce their own contributions less than dollar for dollar as government contributions to charity increase. This partial crowding–out result has been supported by many empirical studies (Abrams and Schmitz, 1978; Reece, 1979; Schiff, 1985; Kingma, 1989; Khanna, Posnett, and Sandler, 1995; Payne, 1998, 2001; Khanna and Sandler, 2000; Okten and Weisbrod, 2000; and Hungerman, 2005) and some experimental examinations (Andreoni,
Experimental Evidence on Tax Compliance and Voluntary Public Good Provision

1993; Chan, Godby, Mestelman, and Muller, 2002), in which the evidence often exhibits slight or modest levels of crowding–out, and sometimes even a crowding–in.

Despite these findings, a common feature of the above–cited studies is that they examine the crowding–out hypothesis in a climate without tax evasion. In particular, in the experimental studies cited above, the public goods are funded by lump–sum taxes. Furthermore, since in the theoretical settings of Warr (1982, 1983), Roberts (1984, 1987), and Bergstrom et al. (1986) the public good consumption comes from the sum of government and private contributions, subsequent experimental studies generally follow this setup, implicitly assuming that the two sources are used to fund the “same” public good. This implies that private and government contributions yield the same marginal benefits to taxpayers, excluding any possibility that taxpayers may benefit more from either one than the other, and the subsequent influences that taxes may have on private contributions.

In sum, although Alm and his coauthors apply the voluntary contribution mechanism to tax compliance experiments, in their experiments the public good is funded by taxes alone, not private contributions. Therefore, in their experiments the interaction between taxes and private contributions cannot be investigated. In the experimental literature on crowding–out, the public good is funded by both taxes and private contributions, but the taxes are lump–sum taxes. Therefore, only the effect of taxes on private contributions can be examined, while how the incentive to pay taxes is affected by private contributions remains unanswered. Since in the real world some people do have the incentive to evade tax, and besides paying taxes to fund the publicly provided public goods, people often make contributions to the privately provided public goods, the literature on tax compliance, crowding–out, and the voluntary provision of public goods is complementary.

This paper attempts to connect the three avenues of the experimental literature by incorporating voluntary contributions into a tax compliance framework. Unlike the crowding–out literature in which taxes and private contributions are used to fund the same public good, this paper assumes that taxes are used to fund a publicly provided public good, and private contributions are used to provide a privately provided public good. The reason for this setting is that people may give more and evade more tax if the privately provided public good is relatively more valuable as compared with the publicly provided public good. On the contrary, people may evade less tax and give less if the publicly provided public good is operated more efficiently than the privately provided public good. Assuming that there is only one public good fails to explore the interactions between tax compliance and private giving when the publicly or the privately provided public good is valued relatively more highly than the other.

Three baseline treatments are employed here. The first baseline treatment includes only the public good funded by taxes and the second one includes only the public good established through voluntary contributions. The third treatment involves both public goods. The three baseline treatments have the same marginal per capita returns (MPCRs) of the publicly and privately provided public goods, and the same tax rate, audit probability, and fine rate whenever these fiscal variables appear. The MPCRs of the publicly and privately provided public goods, the tax rate, and the audit probability are each varied in the other four treatments. Through this experimental setting, this paper aims to explore the effects of private contributions on compliance, the
crowding–out effects of taxes on private contributions, and the effects of changes in fiscal variables on compliance and private contributions.

The main findings of this paper are summarized as follows. First, compliance is significantly higher with private contributions than without private contributions, suggesting that the levels of compliance may be underestimated, ceteris paribus, if private contributions are not taken into account. Second, a higher MPCR of the publicly provided public good will significantly improve compliance, but has no significant effect on private contributions. However, an increase in the MPCR of the privately provided public good improves private contributions significantly, but compliance first increases and then falls as the MPCR of the privately provided public good increases to a level exceeding the MPCR of the publicly provided public good. These findings suggest that the privately provided public good is a substitute for the publicly provided public good, but the converse does not hold. Third, crowding out is slight and statistically insignificant when the tax rate is moderate and when moderate taxes are used to fund the publicly provided public good that is more valuable as compared with the privately provided public good. However, a high tax rate will significantly crowd out private contributions, and sometimes the extent of the crowding–out is more than dollar for dollar. It is also found that a higher audit probability leads to a significantly higher level of compliance, but has no significant effect on private contributions, and that a higher tax rate significantly lowers both compliance and private contributions.

The following section of this paper explains the experimental design for exploring tax compliance, private contributions, and crowding–out. The third section discusses the results of the experiments. The fourth section concludes.

EXPERIMENTAL DESIGN

The experimental design conforms to the basic structures of the income tax systems in the United States and Taiwan and is similar to those in the existing public goods experiments and Alm et al.’s tax compliance experiments. In the experiment, subjects receive income, they pay taxes on the income declared, and they make contributions to a privately provided public good. After declarations and contributions, their true incomes are audited according to a certain probability. They pay the evaded taxes and a fine if they are caught cheating. The taxes are used to fund a publicly provided public good whose size depends on the tax payments of all group members. The size of the privately provided public good depends on the total contribution made by the group. The group members’ payoffs come from the benefits derived from the two public goods and a private good, which is their income net of taxes, contributions, and the evaded taxes and penalties if there are any. To prevent subjects’ emotional reactions from being aroused or subjects’ incentives from being affected, neutral terms are used in the instructions. For instance, in the Subjects’ Instructions they are informed that their true income will be “checked” instead of “audited.” The terms “taxes” and “contributions” are also avoided and are replaced by “investments.”

Table 1 summarizes the magnitudes of the experimental parameters used in the seven treatments conducted in this research. The first three treatments, Only–Y, Only–Z, and YZ–base, serve as the baseline treatments. Y and Z denote, respectively, the publicly and privately provided private goods. Public good Y is funded by taxes and public good Z is funded via voluntary contributions. The Only–Y treatment corresponds to those experiments conducted by Alm et al., in which only the publicly provided public
good Y exists. The Only–Z treatment is a typical public goods experiment, in which only the privately provided public good Z is involved. The YZ–base treatment includes both public goods.

In the three baseline treatments, with a publicly provided public good, the tax rate is 0.2, the audit probability is 0.02, and the fine rate is three times the evaded taxes. The MPCRs of Y and Z are both set at 0.5, i.e., each dollar of taxes paid and each dollar of voluntary contributions made by anyone yields, respectively, a return of 0.5 dollars to each group member. Notice that the MPCR must meet the condition \((1/\text{group size}) < \text{MPCR} < 1\). The reason for this is that if the MPCR is greater than one, then everyone will have the incentive to pay as much tax as he or she can or will donate all of his or her income. By contrast, if the MPCR is less than the reciprocal of the group size, then no one will have the incentive to pay taxes or give any income.

Just as in the case of the YZ–base treatment, both public goods Y and Z co–exist in the remaining four treatments. In one of the four treatments, public good Y is valued more highly than public good Z and this is represented by a higher MPCR of Y (0.7). We call this treatment the high–value–Y treatment, or HY. By contrast, in another treatment the MPCR of public good Z is 0.7 and this treatment is called the high–value–Z treatment, or HZ. In the remaining two treatments, either a higher tax rate (0.4) or a higher audit probability (0.05) is adopted. The two treatments are referred to as the high–tax–rate and the high–audit–probability treatments and are abbreviated as HT and HP, respectively. The magnitudes of these fiscal parameters were chosen to be comparable with those used in existing experimental studies or in institutional settings.¹

The issues of tax compliance, crowding–out, and the voluntary provision of public goods can be examined in the following three directions. First, the Only–Y treatment can be used as the control group to explore the effects of the existence of the privately provided public good on compliance. The YZ–base and HZ treatments are the same as the Only–Y treatment in every respect except that they also involve the privately provided public good Z. Comparing the YZ–base treatment with the Only–Y treatment tells

¹ For instance, Alm et al. have in various experiments set the MPCRs of the publicly provided public good to be 0.25, 0.4, and 0.75, the tax rates to be 0.1, 0.3, 0.4, and 0.5, the fine rates to be 1, 2, 3, and the audit probabilities to be 0.02, 0.04, 0.05, 0.06, and 0.10. Furthermore, as reported by Andreoni, Erard, and Feinstein (1998), the audit probability in the United States was 4.75 percent in 1965, but dropped to only 0.8 percent in 1990. Although it increased in 1995, a low level of 1.7 percent still remained.
us how the existence of the privately provided public good affects compliance and thereby government contributions when both public goods are equally valuable. Contrasting the HZ treatment with the Only–Y treatment helps us understand how the results change if the privately provided public good is relatively more valuable as compared with the publicly provided public good.

Second, to examine how private contributions are crowded-out by government contributions (taxes), the treatments with different levels of tax rates can be compared. Notice first that in the Only–Z treatment the tax rate is zero and the MPCR of the privately provided public good (Z) is 0.5. The YZ–base, HY, HT and HP treatments all have taxes and the same MPCR of Z as that in the Only–Z treatment. Hence, comparing the Only–Z treatment with the YZ–base treatment helps us find out how government contributions crowd out private contributions when both public goods are equally valuable. Examining the differences between the Only–Z treatment and each of the HY, HT and HP treatments helps us understand how the degree of crowding-out changes when the tax-funded public good Y is valued more highly than the privately provided public good Z, when the tax rate is higher, and when the audit probability is higher. Furthermore, since the only difference between the YZ–base and HT treatments is the levels of the tax rates, the degree of crowding-out can also be measured by contrasting these two treatments.

Third, and finally, the YZ–base treatment can serve as the control treatment for the four treatments with both publicly and privately provided public goods, that is, the HY, HZ, HT and HP treatments. Comparing the differences between the YZ–base treatment and each of these four treatments tells us how the changes in the MPCR of Y, MPCR of Z, tax rate, and audit probability affect both tax compliance and private contributions.

The experimental procedures were run as follows. Two sessions, with 20 subjects in each session, were conducted in the HP treatment. All other treatments consisted of only one session, in which 32 to 44 subjects were used. The subjects were recruited from economics courses at National Chengchi University in Taiwan. None of them had ever participated in any public goods or tax compliance experiments.

Subjects in each treatment were randomly and anonymously assigned to groups of four and they remained in the same groups for all 12 rounds. At the beginning of each round, the four group members were randomly assigned the income levels NT$15, NT$20, NT$25, and NT$30, respectively. When a new round started, these four levels of income were reassigned to the four group members. Subjects knew their own income and the distribution of income, but not the income levels of the other three group members. In order to be fair to all subjects, each of them was assigned each level of income three times, but to prevent any expectations that might bias the subjects’ decisions, they were unaware of this setting during the experiments.

In all treatments except Only–Y and Only–Z, subjects were required to make two decisions in each round. First, they decided to declare some or all of their

---

2 In Alm et al.’s experiments, subjects play the game for 20–25 rounds. There are two justifications for the shorter play employed in this study. First, in many public goods experiments, the lengths of the games are ten to 15 rounds. The length of 12 rounds used here is comparable to the public goods experiments. Second, as will be described later, subjects were encouraged to perform some calculations, and except in the Only–Y and Only–Z treatments in which subjects made only one decision, subjects in all other treatments were required to make two decisions. Since the decision time for each round is longer than that in other tax compliance experiments, fewer rounds can reduce problems arising from impatience.
incomes. The income declared could not be below zero, nor could it exceed the income assigned. A certain proportion of the declared income was invested in the Y account (the publicly provided public good). This proportion served as the tax rate and was 0.4 in the HT treatment and 0.2 in other treatments. The payoff from Y for each subject was the product of the MPCR of Y and the total investment in Y by the group. Second, they decided to invest some or all of their remaining income in the Z account (the privately provided public good). The payoff from Z to each subject was the product of the MPCR of Z and the total investment in Z by the group. The rest of the income was invested in each subject’s own X account (the private good). To make sure that the subject’s investment in the Z account would not exceed the subject’s income net of the investment in Y, every subject was encouraged to calculate his (her) investment in Y given his (her) declared income and the income left after investing in Y. In the Only–Y treatment, subjects only declared income. They were encouraged to calculate the investment in Y (taxes) given their declarations and the income left in their X accounts after paying taxes. In the Only–Z treatment, subjects only made investments in the Z account (private contributions) and calculated the income left in their X accounts. During the experiments, the computers could automatically pick up those subjects whose declared incomes or investments in the Z account exceeded the upper limits indicated above. Fortunately, no one made such mistakes.

In the Only–Z treatment, the subject’s payoff was the money left in the X account plus the earnings from the Z account. In the Only–Y treatment, the subject’s payoff before his (her) true income was audited was the money left in the X account plus the earnings from the Y account. In all other treatments, the subject’s payoff before auditing was the money left in the X account plus the earnings from the Y and Z accounts.

The auditing procedure was as follows. At the beginning of each session of the HP treatment, subjects were informed that they were each randomly assigned a subject number, which ranged from one to 20. Twenty balls each with a different number from one through 20 written on it were shown publicly and were put into a transparent bag in order. In each round, one ball was drawn from the bag after all subjects finished their decisions. This ball was then placed back into the bag. The subject whose subject number coincided with the number on the ball was chosen and his (her) true income was audited. Doing so made the audit probability 0.05. The same procedure was run for the Only–Y, YZ–base, HY, HZ and HT treatments, except that subjects were informed that they were each randomly assigned a subject number, which ranged from one to 50, and that 50 balls were used in these treatments to make the audit probability 0.02. If the subject was caught under-reporting, his (her) payoff this round would be either reduced by the sum of the under-investment to the Y account (i.e., the taxes evaded) and three times this amount or zero if his (her) payoff this round became negative after such a reduction.

Subjects were given written instructions in Chinese. The experimenter read the instructions aloud, performed the calculations of the examples in the instructions on the whiteboard, and answered any questions raised by the subjects. At the end of each round, each subject received an earnings report. The earnings report for all treatments except Only–Y and Only–Z indicated the information regarding his (her) reported income and investment in the Y account given his (her) income reported, his (her) investment in the Z account, the money left in his (her) X account after investing in Y and Z, the total investments in the Y and Z accounts.
made by the other three members, the total investments in the Y and Z accounts including his (her) own, whether his (her) true income was audited and the amount of the reductions, his (her) payoff this round, and his (her) cumulated payoff until this round. Any information regarding the Y account and income auditing was absent from the earnings report for the Only–Z treatment. Likewise, any information regarding the Z account was absent from the earnings report for the Only–Y treatment.3

The Only–Y and Only–Z treatments lasted about 60 minutes. Since subjects in the remaining five treatments were required to make two decisions and perform more calculations, the five treatments each took about 80 minutes. The average payoff of all participants was NT$463.75 (with a standard deviation of NT$74.60, a maximum of NT$708, and a minimum of NT$350.9).4

RESULTS

Table 2 summarizes the data resulting from the entire 12 rounds in each treatment. Since the subjects who are endowed with higher income inherently have a greater ability to pay taxes and to make contributions, the reported income and private contributions are divided by the true income when performing statistical analyses. The two variables after such manipulations are referred to as the compliance rate and the private contribution rate. Since the government contributions (taxes) are calculated as the reported income times the tax rate, the government contribution rate is equivalent to the compliance rate times the tax rate. The total contribution rate for an individual is defined as the sum of his (her) private contribution and taxes paid divided by his (her) true income.

Before exploring the issues of tax compliance, voluntary contributions, and crowding–out in detail following the three directions described in the second section, I first take a brief look at the compliance rates and private contribution rates in all treatments. Table 2 and Figures 1 and 2 illustrate this information. Figure 1 shows that the average compliance rate in each experiment starts high in the beginning. The highest average compliance rate is 0.805 in the HY treatment and the two lowest are 0.522 in the Only–Y treatment and 0.525 in the HT treatment. A downward trend over the 12 rounds is observed in every treatment except in the HY treatment. At the end of the experiments, the average compliance rates reach the minimum in almost all treatments, ranging from the highest, 0.464, in the HY treatment and the two lowest are 0.522 in the Only–Y treatment and 0.525 in the HT treatment. A downward trend over the 12 rounds is observed in every treatment except in the HY treatment. At the end of the experiments, the average compliance rates reach the minimum in almost all treatments, ranging from the highest, 0.464, in the HY treatment to the lowest, 0.194, in the HT treatment. A glimpse of Table 2 reveals some evident differences in the average compliance rates between the HY and HT treatments. Over the entire 12 rounds, the highest average compliance rate was found in the HY treatment, with a maximum of 0.805, and the lowest was in the HT treatment, with a minimum of 0.194. For the only–Y treatment, the highest average compliance rate was 0.522, and the lowest was 0.525.

3 The subject’s investment in the Y account, the money left in his (her) X account, and his (her) payoffs were calculated based on the subject’s decisions using computers. Subjects’ calculations on the decision forms were not used to determine their payoffs, nor were they used in data analyses. Subjects’ calculations were checked after the experiments. This helps us to understand to what extent the subjects comprehended the experiments. It was found that only ten subjects did not calculate the investments in the Y accounts or the money left in the X accounts in six or more rounds. Furthermore, no subjects made errors of calculations in more than six rounds. Although it was often the case that three to six subjects made mistakes in each round of the first four rounds in each treatment, at most three subjects made mistakes in each round of the last half of each treatment, and this number was often only one or zero. These observations suggest that subjects learned from the correct information on the earnings reports provided to them and their behavior in the earlier rounds may have been influenced by the learning effects.

4 When these experiments were conducted, the exchange rates between the NT (New Taiwan) dollar and the US dollar were in the range of 32:1 to 33:1. The part–time hourly wage rate for an undergraduate student in Taiwan is about NT$120.
Experimental Evidence on Tax Compliance and Voluntary Public Good Provision

rate is 0.648 in the HY treatment, which is more than twice the lowest average compliance rate of 0.302 in the HT treatment. These magnitudes are comparable with those found by Alm, Jackson, and McKee (1992a, 1992b, 1993), Alm, McClelland, and Schulze (1992), and Alm, Sanchez, and de Juan (1995), who report that the average compliance rates under the condition of public good provision generally fall into the range of 0.266 to 0.675.

Figure 2 shows that the highest average private contribution rate in each treatment occurs in the first round, ranging from the highest, 0.445, in the Only–Z treatment to the lowest, 0.252, in the HT treatment. In all treatments other than the HY treatment, the average private contribution rates generally decrease across rounds and reach the minimum in the final two rounds. At the end of the experiments, the highest average private contribution rate is 0.251 in the HZ treatment and the lowest is 0.082 in the HT treatment. As indicated in Table 2, the average private contribution rates over all 12 rounds range from the highest, 0.319, in the HZ treatment to the lowest, 0.143, in the HT treatment. While the results from existing public goods experiments are diversified under different values of the experimental parameters (the magnitudes of the MPCRs, the levels of endowments, the sizes of groups, etc.), the magnitudes and the downward trend of the private contribution rates found here are comparable with some experimental studies, such as Andreoni (1988), Isaac and Walker (1988a, 1988b), and Isaac, Walker, and Williams (1994).

The Effect of Private Contributions on Government Contributions

Contrasting the Only–Y treatment with the YZ–base and HZ treatments in Figure 1 gives us a preliminary sketch of how the existence of the privately provided

| TABLE 2 | DESCRIPTIVE STATISTICS |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                | Only–Y           | Only–Z           | YZ–base         | HY              | HZ              | HT              | HP              |
| Average compliance rate | 0.307 (0.215) | —                | 0.490 (0.238) | 0.648 (0.201) | 0.435 (0.292) | 0.302 (0.250) | 0.524 (0.315) |
| Average government contribution rate | 0.062 (0.043) | —                | 0.098 (0.048) | 0.130 (0.040) | 0.087 (0.058) | 0.121 (0.060) | 0.105 (0.063) |
| Average private contribution rate | —                | 0.292 (0.214) | 0.272 (0.198) | 0.249 (0.122) | 0.319 (0.213) | 0.143 (0.211) | 0.239 (0.219) |
| Average total contribution rate | 0.062 (0.043) | 0.292 (0.214) | 0.370 (0.204) | 0.379 (0.140) | 0.406 (0.232) | 0.263 (0.252) | 0.343 (0.259) |
| Percent of subjects declaring zero income | 52.60% (0.319) | —                | 29.73% (0.299) | 11.57% (0.214) | 35.21% (0.349) | 36.34% (0.367) | 26.04% (0.337) |
| Percent of subjects making no private contribution | —                | 33.13% (0.293) | 35.98% (0.272) | 22.22% (0.266) | 24.38% (0.322) | 49.31% (0.386) | 30.63% (0.353) |
| Percent of subjects completely free riding | 52.60% (0.319) | 33.13% (0.293) | 18.18% (0.209) | 9.26% (0.198) | 16.46% (0.270) | 30.09% (0.358) | 30% (0.315) |
| Percent of subjects fully complying | 13.54% (0.181) | —                | 22.92% (0.244) | 25.23% (0.260) | 11.88% (0.180) | 6.94% (0.133) | 26.67% (0.369) |
| Average earnings (in NT$) | 384.30 (12.24) | 446.58 (55.29) | 465.31 (55.15) | 497.95 (39.31) | 544.68 (86.28) | 436.59 (67.63) | 455.53 (67.83) |

Notes: The observations are the average values of all individual subjects’ average choices across all 12 rounds. The numbers in parentheses are the standard deviations of all subjects’ average choices across all 12 rounds.
Figure 1. The Average Compliance Rates by Round

Figure 2. The Average Private Contribution Rates by Round
public good affects government contributions. Since the three treatments all have the same tax rate of 0.2, looking at the government contribution rate is identical to looking at the compliance rate. One can see immediately from Figure 1 that the compliance rate starts at 0.522 in the Only–Y treatment, which is much lower than the 0.779 in the YZ–base treatment. The profound differences between the two treatments remain in the early and middle rounds of the experiments, but are reduced in the final three rounds. This suggests that subjects’ behavior may be very different after they have gained some experience and learned the reactions from their partners. Therefore, it is better to examine the differences between treatments by using the data from different stages of the experiments. A two–sided Mann–Whitney U test shows that, except in rounds 3, 4, 11, and 12, the compliance rate is significantly lower in the Only–Y treatment than in the YZ–base treatment in every other round (p ≤ 0.047). If we divide the experiments into three stages, each containing four rounds, then the differences are still highly significant for rounds 1–4 (p = 0.009) and rounds 5–8 (p = 0.000), and become marginally significant for the last four rounds (p = 0.052).5

One may expect that compliance will be lower if the privately provided public good is valued more highly than the publicly provided public good since under this situation individuals may substitute private contributions for taxes. Even so it is still found that the average compliance rates are lower in the Only–Y treatment than in the HZ treatment in all rounds, but only that the differences are less significant. A two–sided Mann–Whitney U test shows that the compliance rate is still lower in the Only–Y treatment than in the HZ treatment in every round, but the differences are significant only in round 3 (p = 0.072), round 5 (p = 0.019), and round 10 (p = 0.038), or if we look at the stages of rounds 1–4 (p = 0.033) and rounds 5–8 (p = 0.060), or the entire 12 rounds (p = 0.073).

The evidence above suggests that introducing private contributions into the tax compliance experiments will improve, or at least will not reduce, compliance. Therefore, the level of compliance may be underestimated, ceteris paribus, if private contributions are not taken into account. The intuition for this result is as follows. In the public goods game, the group payoff is maximized if everyone cooperates, although the one who unilaterally deviates can earn a higher payoff than if he (she) does not deviate. When there is no privately provided public good, subjects can cooperate only through paying taxes. Since the taxes account for at most 20 percent of their income in the Only–Y treatment, subjects may not have the incentive to cooperate or may not clearly see the advantage of cooperation. The existence of a privately provided public good increases the stakes and the incentive for cooperation. When subjects make voluntary contributions, they also increase their involuntary contributions by behaving more compliantly.

The Crowding–Out Effects of Taxes on Private Contributions

To examine how private contributions are crowded–out by government contributions to charity, in the experimental studies of the crowding–out literature the treatments with different levels of tax rates are compared. Here, the Only–Z treatment has no taxes involved and can serve as the control treatment for all other treatments with taxes and the same MPCR of the privately provided public good Z, namely, the YZ–base, HY, HT, and

5 This test is performed by using the individual subjects’ choices as the observations. When looking at a subject’s choices over several rounds, the average of these choices is used as the observation.
HP treatments. Moreover, since the only difference between the YZ-base and HT treatments is the levels of the tax rates, the degree of crowding-out between the two treatments can be also examined.

The degree of crowding-out is measured by one minus the change in total contributions/the change in taxes. Since total contributions are composed of private contributions and taxes, the degree of crowding-out is also equivalent to minus the change in private contributions/the change in taxes. Complete crowding-out would predict that private contributions are completely decreased by taxes so that total contributions are not affected. In this case the degree of crowding-out is one. On the contrary, if private contributions are not affected by taxes at all so that total contributions increase completely by taxes, then the degree of crowding-out is zero.

A preliminary sketch of the crowding-out effect can be observed from Figure 2 since it illustrates the private contribution rates across rounds in various treatments. As Figure 2 shows, the private contribution rates in the HT treatment are not only the lowest in all 12 rounds, but are also substantially below those in the Only-Z and YZ-base treatments. Hence, it is expected that a substantial crowding-out will exist in the HT treatment. For the YZ-base, HY, and HP treatments, the levels of crowding-out seem smaller and different between the early and later rounds of the experiments. When contrasted with the Only-Z treatment and by using the data from all rounds, Table 3 shows that the degree of average crowding-out is 19.95 percent for the YZ-base treatment, 27.40 percent for the HY treatment, 50.89 percent for the HP treatment, and 119.79 percent for the HT treatment. Except in the HT treatment, the levels of crowding-out in other treatments are modest. A two-sided Mann-Whitney U test that compares the private contribution rates in the Only-Z treatment with each of the YZ-base, HY, HT and HP treatments confirms these results: the differences are significant only for the HT treatment ($p = 0.000$) but not for the other three treatments ($p \geq 0.172$). The degree of average crowding-out for the contrast between the HT and YZ-base treatments soars to 558.20 percent, indicating that an NT$1 increase in taxes will crowd out private contributions by NT$5.58. The difference in the private contribution rates between the two treatments is highly significant ($p = 0.000$).

Since subjects’ behavior in the earlier rounds may be influenced by the learning effects, it may be more appropriate to look at only later rounds of the game. Measuring the degree of crowding-out by using the data from only the last six rounds leads to slight differences. When contrasted with the Only-Z treatment, the degree of average crowding-out becomes –29.53 percent for the YZ-base treatment, –14.96 percent for the HY treatment, 9.59 percent for the HP treatment, and 87.68 percent for the HT treatment. Notice that the negative values of the crowding-out actually indicate crowding-in. For instance, the magnitude –29.53 percent for the YZ-base treatment indicates that an NT$1 increase in taxes will not crowd out private contributions, but will increase private contributions by NT$0.29. In performing again a two-sided Mann-Whitney U test, it is found that the differences in the private contribution rates between the Only-Z treatment and each of the YZ-base, HY, and HP treatments remain insignificant ($p \geq 0.390$), and a high tax rate (0.4 in the HT treatment) still significantly reduces the private contribution rate ($p = 0.005$). The degree of average crowding-out is 556.71 percent for the contrast between the HT and YZ-base treatments, and the difference between the private contribution rates in the two treatments remains highly significant ($p = 0.000$).

These findings suggest that the level of crowding-out depends on the magnitudes
of the fiscal variables. If the tax rate is moderate or if taxes resulting from a moderate tax rate are used to fund a more highly valued publicly provided public good, then the crowding–out or crowding–in is slight and statistically insignificant, suggesting that taxes have no significant effect on private contributions. However, the degree of crowding–out may exceed one if a heavy tax rate (0.4) is imposed. In this case, the subjects’ tax burden becomes so heavy that they have to cut private contributions drastically to pay taxes. As a result, private contributions are not only completely crowded out by taxes, but the drastic reduction in private contributions also exceeds the increase in taxes so that total contributions also fall.

The modest and insignificant effects of crowding–out or crowding–in in the YZ–base, HY, and HP treatments are consistent with some of the experimental and empirical findings in the studies cited above. Although the result where the degree of average crowding–out in the HT treatment exceeds one has not appeared in the existing literature, this result is complementary to that of Chan et al. (2002), who used three different tax rates (0, 0.15, and 0.25) in their experiments and found

### TABLE 3

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Mean total contributions per subject (in NT$)</th>
<th>Mean private contributions per subject (in NT$)</th>
<th>Mean taxes per subject (in NT$)</th>
<th>The degree of average crowding–out using Only–Z as the control treatment</th>
<th>The degree of average crowding–out using YZ–base as the control treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Only–Z</td>
<td>6.381 (4.607)</td>
<td>6.381 (4.607)</td>
<td>0</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>YZ–base</td>
<td>8.127 (4.550)</td>
<td>5.946 (4.391)</td>
<td>2.181 (1.092)</td>
<td>19.95%</td>
<td>—</td>
</tr>
<tr>
<td>HY</td>
<td>8.505 (3.113)</td>
<td>5.580 (2.710)</td>
<td>2.925 (0.911)</td>
<td>27.40%</td>
<td>—</td>
</tr>
<tr>
<td>HT</td>
<td>5.852 (5.714)</td>
<td>3.174 (4.724)</td>
<td>2.678 (2.254)</td>
<td>119.79%</td>
<td>558.20%</td>
</tr>
<tr>
<td>HP</td>
<td>7.528 (5.701)</td>
<td>5.194 (4.836)</td>
<td>2.334 (1.427)</td>
<td>50.89%</td>
<td>—</td>
</tr>
</tbody>
</table>

Rounds 1–12

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Mean total contributions per subject (in NT$)</th>
<th>Mean private contributions per subject (in NT$)</th>
<th>Mean taxes per subject (in NT$)</th>
<th>The degree of average crowding–out using Only–Z as the control treatment</th>
<th>The degree of average crowding–out using YZ–base as the control treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Only–Z</td>
<td>4.442 (4.447)</td>
<td>4.442 (4.447)</td>
<td>0</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>YZ–base</td>
<td>6.693 (5.133)</td>
<td>4.955 (4.815)</td>
<td>1.738 (1.247)</td>
<td>–29.53%</td>
<td>—</td>
</tr>
<tr>
<td>HY</td>
<td>7.676 (3.907)</td>
<td>4.863 (3.421)</td>
<td>2.813 (1.125)</td>
<td>–14.96%</td>
<td>—</td>
</tr>
<tr>
<td>HT</td>
<td>4.709 (5.883)</td>
<td>2.537 (4.678)</td>
<td>2.172 (2.477)</td>
<td>87.68%</td>
<td>556.71%</td>
</tr>
<tr>
<td>HP</td>
<td>6.328 (5.975)</td>
<td>4.242 (5.040)</td>
<td>2.086 (1.564)</td>
<td>9.59%</td>
<td>—</td>
</tr>
</tbody>
</table>

Notes: The observations are the average values of all individual subjects’ average choices across all 12 rounds or the final six rounds. The numbers in parentheses are the standard deviations of all subjects’ average choices across all 12 rounds or the final six rounds. The degree of average crowding–out is positive if taxes reduce private contributions, and is negative if taxes increase private contributions.
evidence that the average crowding–out increases from 70.4 percent with a tax rate change from zero to 0.15 to 75.5 percent with a tax rate change from 0.15 to 0.25. While the authors do not find a crowding–out effect greater than 100 percent, their findings indicate that the degree of crowding–out increases with tax rates.

**Tax Compliance and Private Contributions**

The above two subsections have provided a preliminary sketch of compliance and private contributions. This subsection explores in depth the effects of changes in the fiscal variables on compliance, private contributions and, hence, total contributions. To discern these effects, each of the HY, HZ, HT, and HP treatments can be compared with the YZ–base treatment. The Mann–Whitney U tests presented above can play a useful role in the analysis. Another approach is to compile the data from these five treatments, and include data from the Only–Y and Only–Z treatments. Then regressions can be run by setting the compliance rate, private contribution rate, and total contribution rate as the dependent variables and the relevant fiscal variables as the independent variables. Since the three dependent variables are censored at zero, Tobit maximum likelihood estimations are suitable.6

The test results from the Tobit estimation differ slightly from those using the Mann–Whitney U test. However, given the following two advantages for the Tobit estimation over the Mann–Whitney U test, only the Tobit estimation results are reported. First, when a variable changes across three or more treatments, the Tobit estimation can give us a definite effect of the change in the variable. In contrast, the Mann–Whitney U test can only compare the difference between every two treatments and, therefore, may provide inconsistent results.7 Second, the Tobit estimation can tell us if there is a clear impact of income on the three dependent variables. A Mann–Whitney U test, which can only be used to compare the differences between different income groups within the same treatments, may result in very different results from different treatments, which it in fact does.

Table 4 provides the results by using the data from the entire 12 rounds and from the final six rounds. Table 4 shows that a higher tax rate leads to a significant reduction in compliance if we look at the entire 12 rounds, and this result remains if we look at only the final six rounds. A higher tax rate also significantly reduces the private contribution rate, and as a consequence, the total contribution rate also decreases significantly. This finding is consistent with the result above where the degree of average crowding–out is close to or exceeds one for the HT treatment.

The audit probability has no significant effect on compliance if we look at the entire 12 rounds, but this effect becomes marginally significant ($p = 0.057$) if we focus our attention on the final six rounds. A higher audit probability has no significant effect on the private contribution rate. Since the deterrence effect

---

6 Table 2 presents descriptive statistics on all seven treatments. On average, over the entire 12 rounds, more than 30 percent of the observations of each dependent variable are zeros. As indicated by Greene (1997, pp. 959–62), the Tobit model is suitable for this type of data since it can account for the qualitative difference between the limit (zero) observations and the non–limit (continuous) observations.

7 For instance, by using the data from rounds 7 to 12, the Mann–Whitney U test shows that the compliance rate increases significantly as the MPCR of the public good Z increases from zero in the Only–Y treatment to 0.5 in the YZ–base treatment, decreases (insignificantly) as the MPCR of Z increases from 0.5 in the YZ–base treatment to 0.7 in the HZ treatment, and increases (insignificantly) as the MPCR of Z increases from zero in the Only–Y treatment to 0.7 in the HZ treatment.
is only marginally significant in the final six rounds, a higher audit probability in turn has no significant effect on the total contribution rate.

Next, the existence of the publicly provided public good (Y) has no significant effect on the private contribution rate regardless of whether we look at the entire 12 rounds or only the final six rounds. This suggests that the overall crowding–out effect of taxes on private contributions is insignificant. Notice that this does not conflict with the results above to the extent that a higher tax rate will significantly crowd out private contributions and a higher MPCR of the publicly provided public good will have the opposite effect. The crowding–out effects from different treatments simply offset each other.

Table 4 reports that the effect of a higher MPCR of Y on the compliance rate is positive and highly significant. The intuition behind this result is straightforward: players are more willing to pay taxes if the taxes are used to fund a highly valued public good. Table 4 also shows that a higher MPCR of Y eventually (in the final six rounds) has a positive and marginally significant effect on the total contribution rate.

On the other hand, a higher MPCR of the privately provided public good (Z) will not only significantly raise the private contribution rate, but will also reduce the compliance rate. Furthermore,
the existence of the public good Z has a positive and highly significant effect on the compliance rate and thereby the total contribution rate. These two results associated with the compliance rate indicate that the anti-compliance effect is larger and more significant when the MPCR of the privately provided public good is large relative to that of the publicly provided public good. However, having a privately provided public good still increases compliance as compared with having no privately provided public good at all. Figure 1 provides some complementary observations. The compliance rates in the HZ treatment are generally lower than those in the YZ-base treatment, but the compliance rate in every round of the HZ treatment is still higher than that in the Only-Y treatment.

A higher MPCR of the publicly or privately provided public good implies that the individual can pay the same price for more public good consumption or, alternatively, a lower price for the same amount of public good consumption. Hence, the positive own effects of the changes in the MPCRs of both the publicly and privately provided public goods exhibit the fact that the demand curves for both public goods are negatively sloped. The cross effects are asymmetric: an increase in the MPCR of the publicly provided public good has no significant effect on voluntary contributions, but an increase in the MPCR of the privately provided public good will significantly reduce the involuntary contributions. This finding indicates that the privately provided public good is a substitute for the publicly provided public good, but the reverse situation does not hold.

Finally, the relationship between income and the compliance rate is negative. Although it is insignificant for the entire 12 rounds, a marginally significant effect is found for the final six rounds (p = 0.055). Furthermore, income has a significantly negative effect on the private contribution rate and, therefore, a significantly negative effect on the total contribution rate regardless of whether we look at the entire 12 rounds or the final six rounds.

CONCLUSION

This paper presents an experimental design to explore the behavior of tax evasion and voluntary contributions. Subjects not only pay taxes to fund the publicly provided public good, but they also voluntarily contribute to a privately provided public good. The experimental evidence suggests that when the government imposes a heavy tax, it not only deters compliance, but also crowds out private contributions significantly. Raising audit probabilities improves compliance but not private contributions. Furthermore, income has negative and significant relationships with compliance and private contributions.

The most interesting evidence reported in this paper is that an increase in the MPCR of the publicly provided public good significantly improves compliance, but has no significant effect on private contributions. By contrast, an increase in the MPCR of the privately provided public good not only increases private contributions, but also reduces compliance. The asymmetric cross effects suggest that, besides the monetary rewards, players seem to derive some private benefit or psychological satisfaction from voluntary contributions. The private benefit or psychological satisfaction helps maintain the same level of private contributions even as the MPCR of the publicly provided public good increases even if it increases to a level exceeding the MPCR of the privately provided public good. On the contrary, compliance first increases but then falls as the MPCR of the privately provided public good increases and exceeds the MPCR of the publicly provided public good. However, having a privately provided public good still improves compli-
ance significantly as compared with not having one at all.

Acknowledgments

I thank editor Therese McGuire and two anonymous referees for providing me with many valuable comments and suggestions. The financial support from the National Science Council in Taiwan (grant number: NSC 93–2415–H–004–008) is gratefully acknowledged.

REFERENCES


Warr, Peter G.  

Warr, Peter G.  

Yitzhaki, Shlomo.  