Experiments on the Fly*

By Aleksandr Alekseev[a], James Alm[b], Vjollca Sadiraj[c], David L. Sjoquist[c]

October 2020

Abstract

How do exogenous increases in resources to a government affect its expenditure decisions? Economic theory typically predicts that a lump-sum grant will have the same impact on government expenditures as an increase in income. However, empirical studies consistently find that government spending is stimulated far more by grants than by income; that is, grants have a "flypaper effect" because the money "sticks where it hits". We conduct a laboratory experiment that controls for the most important factors that have been suggested in explaining the existence of the flypaper effect. Our experimental design crosses four transfer delivery methods with three voting frameworks. We examine three payoff-equivalent transfer delivery methods, all relative to a fourth baseline treatment with no transfer: an increase in income, a subsidy (repayment) for expenditures on the public good, and a lump-sum grant. Our two alternative voting frameworks are voting over levels of expenditures and voting over changes with information on public good externalities, each relative to a third baseline treatment where voting is over changes from a default (reference) level of expenditures. We find robust evidence of a flypaper effect: both the subsidy and the lump-sum grant increase expenditures more than does an equivalent increase in income. Our results are largely consistent with, and explained by, theoretical models that rely upon behavioral economics.

Keywords: Laboratory experiment; intergovernmental grants; flypaper effect; reference dependence; public goods; incremental budgeting.

JEL Codes: C9; H4; H8.

---

*We thank Kevin Ackaramongkolrotn for his assistance in conducting the experiment, and James Cox for his helpful comments. The experiment was completed while Alekseev was a graduate student at Georgia State University.

[a]University of Regensburg, [b]Tulane University, [c] Georgia State University

1. Introduction

What are the effects of financial transfers (like intergovernmental grants) on the level of public expenditures? Conventional economic theory predicts that a lump-sum grant will have the same impact on government expenditures as an equivalent increase in income. However, a common finding in the empirical literature is that grants have a much larger effect on the level of public expenditures than an increase in income, a result labeled the "flypaper effect" by Arthur Okun because "money sticks where it hits" (Inman 2009). There have been numerous attempts to explain the flypaper effect, by developing alternative theories of government behavior and by critiquing the empirical methodologies. However, none of these attempts is wholly convincing, and the flypaper effect remains an "anomaly" (Hines and Thaler 1995).

In this paper we conduct a laboratory experiment that controls for the most important factors that have been used theoretically and empirically to explain the existence of the flypaper effect. We induce a political environment in which payoff functions are single-peaked, the feasible policy options lie on a straight line, and the median proposal is chosen for implementation. We then cross four transfer delivery methods with three voting frameworks. We examine three payoff-equivalent transfer delivery methods, all relative to a fourth baseline treatment with no transfer: an increase in income, a subsidy (repayment) for expenditures on the public good, and a lump-sum grant. We also examine two voting frameworks – voting over levels of expenditures and voting over changes with information on public good externalities – each relative to a third baseline treatment where voting is over changes from a default (reference) level of expenditures. We find robust evidence of a flypaper effect: both the subsidy and the lump-sum grant increase expenditures more than does an equivalent increase in income. These results are largely consistent with, and explained by, theoretical models that rely upon behavioral economics mechanisms.

Our experimental design allows us to make several novel contributions to the literature on voting over expenditures, in ways that disentangle possible explanations for the flypaper effect.

First, and most importantly, our experiment allows us to detect if there is a flypaper effect in an environment in which we control for factors (e.g., an improperly specified econometric model, political agenda control, endogenous grants) that have been suggested as the primary causes for the flypaper effect. Our experiment also allows us to determine in the presence of a flypaper effect whether the impact is robust to the framing of the decisions (e.g., voting over changes or levels, individual valuations of public good).

Second, we are able to explore the efficiency aspects of the flypaper effect; that is, does the existence of the flypaper effect increase or decrease social efficiency of public expenditures? Because we induce public good benefits for each and every participant, we are able to identify the economically efficient level of public expenditures, a challenging – even an impossible – task for observational data. Thus, we can explore the efficacy of alternative voting frameworks and transfer types in producing a socially efficient level of expenditures.

Third, in our baseline voting frame subjects propose changes to an existing level of the public expenditure, which represents a strategy of incremental budgeting. There is a substantial behavioral economics literature that finds that decisions are affected by the status quo (Kahneman, Knetsch, and Thaler 1990) and by reference points in general (Crawford and Meng 2011). In our experiment, the default expenditure is the level of expenditures last implemented, and subjects propose changes from the default level. Thus, we can explore the effects of reference dependence on public good levels.

Fourth, one proposed explanation for the flypaper effect is the presence of fiscal illusion, or the inability of voters to perceive the true price of the public good (Courant, Gramlich, and

Rubinfeld 1979; Turnbull 1998), possibly due to political agenda control (Filimon, Romer, and Rosenthal 1982) or to inefficient political institutions (Chernick 1979). However, these papers do not provide a direct test of fiscal illusion. Our experimental design is able to examine the role of fiscal illusion in the flypaper effect, especially the ways in which behavioral economics considerations (e.g., reference points, loss aversion, and mental accounting) may contribute to fiscal illusion.

Fifth, under-provision of public goods that is driven by positive externalities not integrated in an individual's decision problem is a standard public good problem. Providing information on social benefits may ameliorate this problem, as suggested by an extensive research body that studies individual propensity to forego one's own monetary benefit for increased efficiency (Andreoni and Miller 2002; Charness and Rabin 2002; Cox and Sadiraj 2012). Therefore, in one of our treatments we provide information about total group benefits that result from public expenditures, and we then explore the effect of this treatment on voting outcomes.

We find robust evidence of a flypaper effect: expenditures are significantly greater with a lump-sum grant or a subsidy than with an equivalent increase in income. We also find that expenditures are closer to the efficient level when proposals are framed as changes to a given expenditure level from a default level and when subjects are prompted to consider social benefits by the provision of information about the benefits from public expenditures. Finally, we find that the institutional nature of the transfer (e.g., a lump-sum transfer versus a subsidy) affects the voting outcome, but we find no evidence that fiscal illusion contributes to the flypaper effect. We are able to explain these results using theoretical models that incorporate behavioral economics mechanisms.

The remainder of the paper proceeds as follows. In the next section we briefly discuss the relevant literature on the effect of grants on public expenditures, as well as applications of reference dependent and social preferences models in the public finance literature. In section 3 we describe the experimental treatments, while in section 4 we present theoretical analyses and hypotheses for the game played in our experiment. Section 5 contains the results of the experiments. A summary section concludes the paper.

2. Related Literature

We consider three literatures relevant to our experiment: the effect on public expenditures of grants, the effect of default levels (i.e., reference dependence) on spending, and the effect of information (through other regarding preferences) on expenditures. In each case, while there has been a surge in the use of behavioral economics in recent years, the trend has not caught up in the public budgeting literature (Alm and Bourdeaux 2013), and there are also few experimental studies in this literature (Alm and Jacobson 2007).

2.1. Increasing Community Resources: Is There A Flypaper Effect?

There are substantial theoretical and empirical literatures that explore the effect of increases in income and grant funding on public goods expenditures. A flypaper effect (i.e., the marginal propensity to spend on a public good out of grant funding is greater than out of income) has been documented in many studies of the effects of intergovernmental grants on public expenditures (Fisher 1982; Hines and Thaler 1995; Inman 2009; Leduc and Wilson 2017).[1]

[1] Inman (2009) reports that a Google search revealed over 3500 papers documenting and seeking to explain the flypaper effect A flypaper effect has also been found in other settings. For example, Singhal (2008) uses the data on the windfall gains states received from the tobacco settlement, and estimates a marginal propensity to spend on tobacco control from these disbursements of 0.2 against zero from private income. Moreover, a flypaper effect appears to be present outside of public finance as well. Abeler and Marklein (2010) find a flypaper effect in individual decision-making tasks, and Blanchard, de Silanes, and Shleifer (1994) document a similar phenomenon for private firms that received cash windfalls as a result of lawsuits.

Several explanations have been advanced to explain why studies find a flypaper effect (Inman 2009). Some of the explanations focus on econometric issues. It has been argued that grants have not been properly modeled (Moffit 1984), that the empirical models have been incorrectly specified because they have omitted important variables (Hamilton 1983), that the models ignore the distortionary effects of taxes (Vegh and Vuletin 2015), and that grants are not exogenous because they are determined through a political process (Knight 2002). Others have suggested that the flypaper effect is due to fiscal illusion, i.e., the inability of voters to perceive the true price of the public good when there is a lump-sum grant (Courant, Gramlich, and Rubeinfeld 1979; Turnbull 1998). Some authors have tied this fiscal illusion to political agenda control (Filimon, Romer, and Rosenthal 1982) or to inefficient political institutions (Chernick 1979). Hines and Thaler (1995) and Ashworth and Heyndels (1999) suggest that the flypaper effect can be explained by behavioral economics. They base their explanation on three concepts from behavioral economics: reference points, loss aversion, and mental accounting. Neither paper provides a formal treatment or a direct test of their explanations.

Our study offers contributions on all of these issues. Our laboratory experiment allows us to explore the effects of grants in an environment that preserves private and public goods valuations, budget constraints (relative price and income), and vote aggregation procedures. We thus measure the effect of grants controlling in our experiment for both the non-behavioral and the behavioral factors that have been suggested as explanations for the existence of a flypaper effect. Becker, Hopp, and Kriebel (2018) is the only prior laboratory experiment we are aware of that explores the effects of grants on public expenditures, although in a very different experimental design than the one that we employ.[2]

[2] In their experiment, three players, one poor and two rich, vote on the amount to transfer to the poor. There are two accounts (a Common Account and an Individual Account) and two treatments (transfers made from the Individual

2.2. Specifying The Default Level

In some of our treatments a default expenditure level is specified, and individuals make proposals over increasing or decreasing it. While there is a large experimental and behavioral economics literature that explores reference dependence (O'Donoghue and Sprenger 2018), we know little about such effects in the context of public budgeting. One exception here is work by Crain and Crain (1998), who explore empirically the effects on budget levels under alternative baseline budget rules and who find that expenditures increase with the value of the baseline. Their study implies that in our experiment the default level will positively affect the level of public good expenditures selected.

2.3. Incorporating Preferences for Efficiency

An externality is a defining feature of a public good, leading to inefficient under-provision of a public good due to free-riding. Indeed, there is an enormous experimental literature that explores this inefficiency and the ways in which individual preferences affect the efficiency. For example, Ackert, Martinez-Vazquez and Rider (2007) conduct an experiment in which subjects trade off increased deadweight loss for a more progressive tax system in the context of a fixed budget, using the framework of Charness and Rabin (2002). Fréchette, Kagel, and Morelli (2012) conduct experiments in which social decisions are made regarding the allocation of a fixed social budget to public and private goods. Bosman et al. (2013) conduct an experiment that investigates the effect of a rotation scheme (that restricts the right to vote to a rotating basis) on collective and individual welfare, and the committee ability to reach an agreement.

Account and transfers made for the Common Account). After each round, the money left in the Common Account is allocated to the rich. In this set up, although it should make no difference to the rich whether the transfer to the poor is from the Common Account or the Individual Account, Becker, Hopp, and Kriebel (2018) find that the size of the transfer is larger when transfers are from the Common Account. Importantly, however, they do not examine other aspects of the flypaper effect, the ones that we emphasize here.

In all of these cases, the experiment is based on the idea that voters have to make a tradeoff between individual and common interests. Similarly, in our experimental design subjects are sometimes prompted to consider the total group benefit of the public expenditures. However, in our design the level of public good expenditure is determined within a median voter framework.

3. Experimental Design

Our experiment uses a 4X3 full factorial design that crosses four transfer delivery conditions with three framing (or voting) conditions. Our four transfer delivery conditions consist of three alternative ways in which additional resources are provided, each relative to a fourth baseline (or a No-Transfer condition): a direct payment to individuals (a Lump-Sum treatment); a re-payment for expenditure on a public good (a Subsidy treatment); and a direct payment that is added to expenditures on a public good (a Grant treatment). We use the term "transfer" to refer to all treatments in which additional resources are provided. We cross these four transfer delivery conditions with three framing (or voting) conditions, or voting over *changes* in the level of expenditures (Default condition) and voting over *changes* in the level of expenditures with information on public good externalities (Default with Information condition), each relative to a third baseline condition where voting is over only a reference *level* of expenditures (No-Default condition). We use the term "framing" to refer to these voting conditions.

The first block of five rounds is always the No-Transfer condition, while the order of the remaining three transfer delivery conditions is randomized. We do not randomize the order of the No-Transfer condition because we wanted the transfer conditions to be in terms of an additional monetary transfer (of 200 points) from the same baseline.

In the framing conditions subjects make proposals over changes from the expenditure level

in the previous round, mimicking how actual government budget decisions are most commonly made, i.e., using incremental budgeting in which changes are made to the prior year's budget.[3] The transfer delivery conditions are assigned within subjects, while the framing conditions are assigned between subjects; that is, each subject experiences all the transfer conditions but only one framing condition.

Each group consists of five individuals, and the group remains fixed during the entire experiment. The income (or endowment) level is the same across subjects, but public good valuations vary, as discussed more fully below. All five individuals in the group simultaneously make proposals on public good expenditures, and the median proposal is selected for implementation.[4] Public good cost is shared equally. Proposals are either in terms of expenditure *levels* (No-Default) or *changes* (increase or decrease) from a default level (Default and Default with Information frames). The default level is set equal to the group's expenditures in the preceding round, except for the first round of each block when the default is 200. In the Grant treatment, the group's expenditure decision does not include the transfer since the grant is later added to the group's expenditures.

In the No-Transfer baseline condition, each individual's endowment is 100 points. In the other three transfer delivery conditions, an additional monetary transfer of 200 points is introduced. We consider three different implementations of the additional monetary transfer of 200 points: increasing individual endowments from 100 to 140 (Lump-Sum condition); reducing (by 40 points) the individual payment for the public good (Subsidy condition); and adding 200 points to the group's proposed expenditure level (Grant condition). Note that in each of these three transfer

[3] An alternative budgeting method is zero-based budgeting, in which the budget is built up from zero each year.
[4] Participant instructions are in Appendix I. We use a meaningful context in the instructions to enhance subjects' understanding of the decision-making environment (Alekseev, Charness, and Gneezy 2017).

conditions the group endowment increases by 200 points.

The experiment provides us with a direct test of the flypaper effect because we are able to control for the environment (e.g., benefits from the public good and private consumption) and keep the social choice mechanism (e.g., the median voter) constant. Note especially that, because we specify subjects' valuations of public and private goods, we are able to measure directly social welfare, which is a challenging problem with observational data. Our knowledge of the socially efficient expenditures level allows us to measure the efficiency of each transfer delivery and framing treatment in addressing the under-provision of public goods. Furthermore, because we know each individual's public good valuations, we are able to explore possible behavioral explanations for the variation in expenditures across treatments, such as social preferences or reference dependence.

In summary, we study choices over pure public goods expenditures in payoff-equivalent implementations of transfer delivery methods, we consider alternative framings of proposals, we specify the status quo, and we assign (monetary) public good benefits to the participants. These design features allow us to examine how the amount of a public good and how economic welfare vary with the four transfer delivery conditions and the three framing conditions.

The next sections provide more detailed information both on the transfer delivery conditions and on the framing conditions. Our experimental design is summarized in Table 1.

[Insert Table 1]

3.1. Transfer Delivery Conditions

At the beginning of each block, subjects perform an easy real-effort task by attempting to correct five misspelled words in a paragraph. Subjects receive 100 points for the attempt.[5] After

[5] We decided to have a real-effort task and reward subjects for their effort rather than their performance for three reasons. First, we wanted subjects to feel entitled to the endowment. Second, we wanted to keep the initial endowment

the working task, the experiment proceeds to the voting phase.

The voting phase consists of the four blocks with different transfer delivery conditions. Each block starts with a practice round followed by five paying rounds. In the voting phase, each subject privately makes a proposal on group expenditures, and the median of the proposals is selected for implementation. In the Lump-Sum condition, the experimenter adds an additional 40 points to each of the five subject's earned income of 100 points. In the No-Transfer and Lump-Sum conditions, the public good expenditures and the group payment are the same as the median proposal. In the Subsidy condition, the public good expenditure level is the median proposal, but the group payment is 200 points less (or 40 points less per individual). In the Grant condition, the group payment is the median proposal, but the public good expenditure level is 200 points more.

To be clear, in the No-Transfer and Lump-Sum conditions, the level of expenditures, the median proposal, and the total taxes paid are all equal. In the Subsidy condition, expenditures and the median proposal are equal, while total taxes equal expenditures less the subsidy. In the Grant condition, expenditures equal the median proposal plus the grant, and total taxes equal the median proposal. We define an *adjusted* proposal as the observed proposal in all but the Grant condition, for which the adjusted proposal equals the observed proposal plus grant.

Individual i's payoff in each round is the income (w) remaining after paying the (equal) share of the group payment (E/n), where n is the size of the group, plus the minimum of either group expenditures (E) or the individual's public good valuation (v_i), which we assign. Thus, the payoff to individual i is given by $[w - E/n + \min(v_i, E)]$.[6] Note that an individual's benefit from each one-point increase in the public good is a one-point increase in the public good benefit up to

constant across subjects to control for the income effect. Third, we did not want subjects to be frustrated during the working task.

[6] This function is a common one; for example, see Hewitt et al. (2005). We discuss this function in more detail later.

the (assigned) valuation of v_i.[7] In the No-Transfer condition, the valuations v_i are randomly assigned in each round without replacement from the set {50, 150, 250, 350, 450}. To induce an income effect on the demand for the public good, and more importantly to reduce binding constraints on proposals, the valuations in the three other transfer delivery conditions are sampled without replacement from the set {70, 210, 350, 490, 630}.[8] Thus, the larger valuations in these three conditions will by design elicit larger group expenditures than in the No-Transfer condition. The question upon which we focus is whether the increase in expenditures matches the increase in the median valuation (about 100 points more) or whether it is affected by the transfer delivery method of the additional resource (i.e., 200), as suggested by the flypaper effect.[9]

3.2. Framing Conditions

There are three framing conditions. In the No-Default frame, each subject in every round makes a proposal for the group expenditure level, and the group expenditure is the median of all five adjusted proposals. In the Default frame, the default expenditures in the first round of each block are set equal to 200, whereas at the beginning of the subsequent four rounds the default is set at the group's proposed expenditures in the preceding round. Each subject makes a proposal on how much to *increase* or *decrease* the default expenditures, and the median of the proposed changes is selected for implementation and added to or subtracted from the default level. In the Grant treatment, subjects are making proposals over changes from the group's proposal in the preceding round, which does not include the grant of 200 points added to the proposed expenditure.

[7] For example, if the group expenditure is 400 points, then the return to an individual with valuation 50 is 50 points whereas the return to an individual with valuation 450 is 400 points. The cost for either individual is 80, which is subtracted from the income earned during the working task.

[8] In the Subsidy and Grant block, this set of valuations may impose a binding constraint on proposals for subjects with the lowest or the highest valuation (see Section 4.2).

[9] Note that monetary incentives for individuals with valuation 350 are the same in the No-Transfer treatment and the other three transfer treatments. Data from such individuals will be informative for inferring other-regarding behavior because 350 is the (group) efficient expenditure level in the No-Transfer treatment but not in treatments with transfers. See Table 2.

Finally, in the Default with Information frame, the default is set as in the Default frame, but subjects are provided with information on the total group benefit of expenditures, in addition to information on the personal benefit and cost (tax payment).[10] The information conveys the public good nature of group expenditures, and makes salient the externalities of proposals. It therefore may elicit other-regarding preferences and nudge group expenditures toward group efficient levels.

4. Theoretical Predictions and Hypotheses

To derive hypotheses and guide empirical analysis, we present simple theoretical models that mimic the structure of our experimental design. The level of public good expenditures is determined by a voting procedure in which each voter proposes a level of public good expenditures, the median of those proposals is selected, and each individual pays an equal share of public good cost; we refer to this as the *median-voter* decision procedure. Under these assumptions, it is weakly dominant for each individual to propose her most preferred public good level. Therefore it suffices to look at the (non-strategic) individual's decision-problem.

Let $z=(x,E)$ denote an ordered pair of private good expenditures (x) and public good expenditures (E), and let w denote the total resources available to an individual for expenditures on the two goods. The individual's budget constraint, in the case of no grant or subsidy, is given by $w=x+E/n$, where n is the number of subjects in the group. We first look at the effects (if any) of different implementations of transfer delivery on the preferred level of expenditures, and then examine alternative models of decision-making.

4.1. Public Goods With Transfers

[10] For example, if an individual is considering proposing 100, the screen for that individual would show the value of social benefits, or $\sum_{i=1}^{5} \min(v_i, 100)$, in addition to own valuation and tax payment.

We first consider the implications of transfers on public good provision. Suppose that additional transfer funds (G) are available. We consider three methods of delivering the funds: distributing them equally among individuals (Lump-Sum method); subsidizing individuals for group expenditures (Subsidy method); or adding funds to the total group expenditures (Grant method). Without loss of generality, we index decision makers by their public good satiation levels; that is, $v_k \leq v_j$ for all $k < j$. We are interested in whether there is a flypaper effect and which of the three transfer delivery methods induce expenditures that are closer to the socially efficient expenditure level, which is v_{n-1} in our experimental game (see below for an explanation).

If we apply a behavioral economics framework, then Hines and Thaler (1995) suggest that the flypaper effect can be explained by mental accounting. As described by Thaler (1999), mental accounting implies that voters view a transfer as being different from an increase in income. Furthermore, an individual could also view a subsidy as different from a grant. If so, voters may respond differently to each type of transfer delivery.

To implement this concept, assume that individual preferences over the public good and payments for financing it are represented by the following utility function

$$u(w, E) = w - f(E/n) + g(E), \qquad [1]$$

where $f(.)$ is an increasing convex function that captures individual's disutility from paying taxes and $g(.)$ is an increasing concave function that captures individual's utility from the public good provision. For the standard model, $f(.)$ is the identity function, and $g(E) = \min(v_i, E)$. For linear social preferences, for individual i it is the case that

$$g(E) = (1-\theta)\min(v_i, E) + \theta\left(\frac{1}{n-1}\sum_{j\neq i}\min(v_j, E)\right)$$

for some $\theta \in [0,1)$. In section 4.2, we look closely at each of these special cases using a reference dependent extension in the spirit of Koszegi and Rabin (2006).

Lump-Sum Method: Suppose that the transfer *G* is equally distributed among *n* individuals, so that the after-transfer income is *w + G/n*. The individual's decision problem becomes

$$\max_E u(w,E) = w + G/n - f(E/n) + g(E),$$

and therefore the optimal expenditure level E^o is determined by

$$ng'(E^o) = f'(E^o/n). \qquad [2]$$

Subsidy Method: Suppose that the government subsidizes expenditures by providing additional resources *G*. The individual's decision problem is thus

$$\max_E u(w,E) = w - f((E-G)/n) + g(E).$$

By the convexity of *f*(.) and concavity of *g*(.), the first-order condition is also sufficient for optimality. Hence, the optimal expenditure level, E^s, is determined by

$$ng'(E^s) = f'((E^s - G)/n). \qquad [3]$$

Grant Method: Suppose that the government adds *G* to the public good expenditures. The individual's decision problem in this case is

$$\max_E u(w,E) = w - f(E/n) + g(E+G),$$

and therefore the optimal expenditure level E^g is determined by

$$ng'(E^g + G) = f'(E^g/n). \qquad [4]$$

It is easy to show (see Appendix II) that first order conditions [2], [3], and [4] imply that

$$E^o \le E^s \le E^g + G;$$

that is, optimal expenditures are largest in the Grant implementation and smallest in the Lump-Sum implementation. The following statement is therefore true:

Proposition 1. Suppose that preferences over expenditures are represented by utility function [1] and that the aggregation procedure is the median proposal. If additional transfers are available, then

$$E^o \leq E^s \leq E^g + G.$$

We call E^o, E^s *and* $E^g + G$ the adjusted proposals. Our first null hypothesis for individual behavior is:

> **H1:** ***The Subsidy method elicits larger adjusted proposals than the Lump-Sum method but smaller than the Grant delivery method.***

It follows that our first null hypothesis for implemented expenditures is an implication of H1 for the median aggregation rule of proposals:

> **HA1:** ***The expenditure level in the Subsidy method is between the levels in the Grant and Lump-Sum methods.***

4.2. Public Goods With Defaults And Information

We next consider alternative treatments under different theoretical frameworks. Let *n* be the group size and λ_i be individual *i*'s loss aversion parameter. We say that a group is sufficiently large if $n \geq \max_i(\lambda_i)$ and that public good benefits are moderate if $\max_i(v_i) \leq nw$. Our main result is:[11]

[11] Proposition 2 is based on Observations 1 and 2 derived below. Predictions of social preferences are the same with and without information. In information treatment externalities are made salient, so social preferences become more prevalent.

Proposition 2. Let v_i denote individual *i*'s public good satiation value. For moderate public good benefits, individual *i*'s optimal expenditure level depends on information as follows:

a. **Absent information, individual *i*'s optimal expenditure level is v_i for the standard conventional model and for the reference dependent model if groups are sufficiently large.**
b. **With information, individual *i*'s optimal expenditure level for social preferences is weakly larger than the individual's public good satiation value for *i* = 1,...,*n*-2 but weakly smaller than the individual's public good satiation value v_n for individual *i* = *n*.**

Standard Model. This is the special case of specification [1] with $f(.)$ equal to the identity function and $g(E) = \min(v_i, E)$. By positive monotonicity in private good *x*, the budget constraint is binding, and therefore individual *i*'s decision problem is

$$\max_{E \in [0, nw]} \pi(E) = w - E/n + \min(v_i, E) \qquad [5]$$

where $v_i \in \{v_1, \ldots, v_n\}$. Note that the individual *i*'s marginal cost of public good expenditure is always 1/*n*, whereas the marginal benefit is 1 up to v_i but zero above that. Therefore, for individual *i* the optimal level of public good expenditure E^s is given by

$$E_i^s = \begin{Bmatrix} v_i, & if\ v_i \le nw \\ nw, & otherwise \end{Bmatrix}$$

which concludes the proof of Proposition 2 for the standard model. For the parameters chosen in our experiment, the first inequality is always satisfied, hence for the standard model our second null hypotheses for individual behavior and implemented expenditures are:

H2: *Individual i's proposal is the valuation v_i.*

HA2: *The expenditure level is the median valuation*.

Reference Dependent Model. With the exception of the No-Default frame treatments, subjects in our experiment make proposals over changes from a default expenditure level (E^r) that is determined by the implemented public good expenditures in the preceding round within each block. Following Kőszegi and Rabin (2006), we specify the reference dependent utility $U^r(.)$ as a weighted average of consumption utility (specification [5]) and gain-loss utility $R(.)$ or

$$U^r(E \mid E^r) = \pi(E) + \eta R(E \mid E^r), \qquad [6]$$

where $\eta > 0$, $R(.)$ is the gain-loss utility with respect to $z^r = (w - E^r/n, \min(v, E^r))$ being determined by the default expenditure E^r. It is straightforward to verify that the gain-loss utility is

$$\begin{aligned} R(E \mid E^r) &= \min(v,E) - \min(v,E^r) + \lambda_1[E^r - E]/n, \quad E \geq E^r \\ &= \lambda_2[\min(v,E) - \min(v,E^r)] + [E^r - E]/n, \quad E < E^r \end{aligned} \qquad [7]$$

where $\lambda_1, \lambda_2 (>1)$ denote loss aversion parameters. Note that, at any expenditures E exceeding the default expenditure level E^r, there is always a loss in terms of private good payoff because the tax payment is larger. However, whether a gain occurs in terms of the payoff from the public good depends on the position of the individual's public good satiation value v. In contrast, at any E that falls short of E^r, there is always a gain in private good payoff, but whether a loss in the payoff from the public good happens depends again on the position of v. It can be shown that:

Observation 1. With reference dependent utility [6] (and [7]), the optimal level of expenditures E^{rd}

a. is the individual's satiation value v if $E^r \geq v$

b. and, if $E^r < v$, is given by

$$\begin{aligned} E^b &= v, \qquad if \quad (n-1) > (\lambda_1 - n)\eta \\ &= E^r, \qquad if \quad (n-1) < (\lambda_1 - n)\eta \\ &\in [E^r, v], \quad if \quad (n-1) = (\lambda_1 - n)\eta \end{aligned}$$

PROOF. See Appendix II.

Note that $\lambda_1 \leq n$ is a sufficient condition for $(n-1) > (\lambda_1 - n)\eta$. Therefore, an implication of the reference dependent model for individual proposals (as well as for implemented expenditures) is the same as for the standard model for sufficiently large *n* (i.e., $\lambda_1 \leq n$), which concludes the proof of Proposition 2.a for the reference dependent model. Most studies (Charles-Cadogan 2018) find that estimated loss aversion parameter is less than 5, which is the group size in our experiment. Therefore, hypotheses for the reference dependent model for our experiment are:

H3: ***Individual i's proposal is the valuation v_i.***

HA3: ***The expenditure level is the median valuation*.**

Social Preferences Model. There is a sizable literature that favors social preferences against the conventional (own payoff) specification [5]. In the Default with Information frame, subjects are given information about total group benefit from the public good, which could make other-regarding preferences more salient. Following the previous literature, we introduce such other-regarding preferences in the model specification [5] by letting individual *i*'s objective function be defined over own and others' payoffs as follows:

$$u_i(E) = (1-\theta)\pi_i(E) + \theta \frac{1}{n-1}\sum_{j \neq i} \pi_j(E) \qquad [8]$$

for some $\theta \in (0,1)$. The following observation concludes the proof for Proposition 2.b.

Observation 2. With social preferences [8],

a. individual *i*'s optimal expenditure level satisfies[12]

$$E_i^c \in [v_i, v_{n-1}]$$

b. the individual with the largest benefit v_n proposes less than own benefit if $\theta \in (1-1/n, 1)$

PROOF. See Appendix II.

It follows from v_{n-1} being the group efficient expenditure level (see above) that:

H1.a: ***Information on group public good benefits nudges proposals towards the efficient expenditure level.***

Furthermore, even when the individual with median valuation has standard preferences, the median voter proposal can exceed the median value of v (and therefore be more efficient) if individuals with lower valuations are sufficiently altruistic to propose more than the median valuation. Thus, an implication of H1.a for the median aggregation rule is:

HA1.a: ***Providing information on group benefits increases expenditures*****.**

The next section presents our experimental results.

5. Results

A total of 155 subjects participated in the experiment with the following distribution between the three framing conditions: 65 (No-Default), 45 (Default), and 45 (Default with Information).[13] The average payoff to the subjects increased from 241 points in the No-Transfer condition to 344 points in the Grant condition. The largest average payoff (360 points) is observed

[12] Individual *i* gets larger utility at any expenditure level $v_k > v_i$ such that $\theta > (n-1)/n(n+1-k)$. For example, in a group of 5, an individual with $\theta = 0.25$ and valuation v_1 prefers v_2 over the own valuation but prefers v_1 over v_4.

[13] Subjects earned on average \$31.50. One round was selected at random to determine payoff and 10 experimental points were worth \$1. In addition, each subject received a \$5.00 show up fee.

in the Grant condition with the Default with Information framing, and the smallest average payoff (312 points) is found in the Lump-Sum condition with the No-Default framing. Recall that an additional 200 points (or 40 points per individual) is made available in the Lump-Sum, Subsidy, and Grant transfer delivery methods relative to the No-Transfer condition.[14] Note that the increase in mean subjects' payoffs in the three transfer treatments over the No-Transfer treatment is larger than 40 points.

In the first subsection, we report the experimental results that relate most directly to the existence of the flypaper effect. In the second subsection, we explore additional features of the flypaper effect related to efficiency, budgeting and framing, fiscal illusion, and social preferences. In both subsections we examine both aggregate and individual data, with individual allowing us to study variations in the effect of the interaction of grant delivery and decision frame across individuals.

5.1. Is There A Flypaper Effect?

Focusing first on aggregate results, Table 2 shows the main summary statistics and least squares estimates of group expenditures, and Figure 1 shows the full distributions of group expenditures in each of the 12 (=4X3) treatment cells.[15] The least square regressions in Part 1.A Transfer Effect of Table 2 measure the effect of the transfer delivery method relative to the Lump-Sum condition for each frame, while the regressions in Part 1.B Frame Effect of Table 2 measure the effect of the frame relative to the Default frame for each transfer delivery method.

[Insert Table 2]

[14] For these calculations, we do not include data from the first round of each block in order to reduce noise coming from switching between grant implementation treatments. If we include first round data, the average payoff increases from 238 points in the Baseline to 341 points in the Grant treatments.

[15] Results reported in Table 2 and Figure 1 do not include data from the first round of each block in order to reduce noise coming from switching between transfer treatments. The overall patterns remain the same if we add data from the first round in each block.

[Insert Figure 1]

Table 2 and Figure 1 clearly demonstrate the existence of the flypaper effect. Subsidizing payments to the public good (the Subsidy row) induces expenditure levels that are between the ones observed in the Lump-Sum and Grant implementations. Note also that providing information (Default + Info column) on total group public good benefit boosts expenditures, while replacing voting over changes with voting over levels (No-Default column) has a negative effect on group expenditures, as discussed in more detail in section 5.2.

Importantly, these results are consistent with the flypaper effect, and this flypaper effect is robust across the three frames: whether or not voting is over levels from a default,[16] whether or not voting is over changes from a default, and whether or not information on group benefits is made salient. Finding a flypaper effect in our experiment raises important questions about any of the econometric explanations for the empirically observed flypaper effect; it also challenges political agenda control and inefficient political institutions explanations, given that these considerations are not present in our experimental design.

The pattern that we observe in the data is predicted by the behavioral model ([1]) in the spirit of "mental accounting", an implication of which is that the Subsidy condition elicits intermediate public good expenditures (HA1). Regarding mental accounting, the presumption in the flypaper literature (Hines and Thaler 1995; Becker, Hopp, and Kriebel 2018) is that grants will have a larger effect on expenditures than an increase in income. However, previous discussions of mental accounting provide no reason to expect that the Grant condition will lead to larger public good expenditures than the Subsidy condition. In contrast, our behavioral model explicitly predicts that the Grant condition will have a larger effect on the public good expenditures than the Subsidy

[16] Recall that, with the exception of round 1 in each block, the default expenditure at round *n* is the group's expenditure level implemented at round *n*-1.

condition. The overall estimated treatment effects on group expenditures, reported in Part 1.A Transfer Effect in Table 2, are 68.55 points and 124.0 points for the Subsidy and the Grant treatments, respectively, as compared to a payoff equivalent increase in income (the Lump-Sum treatment).[17, 18]

Note that in Table 2 the median of adjusted proposals in the Default frame is larger than the median individual satiation value (250 points and 350 points in treatments with and without transfers, respectively).[19] We find only a small percentage of proposals (E) that match the own value of v: 22.6 percent in the No-Transfer treatment, 15.6 percent in the Lump-Sum treatment, 11.9 percent in the Subsidy treatment, and 10.6 percent in the Grant treatment.[20] These patterns of group expenditures in Table 2 suggest that individual proposals are lowest in the Lump-Sum condition and largest in the Grant condition.

We are also able to examine the flypaper effect at the individual level, including the factors that might explain individual differences in the impact of transfers on proposals. Not unexpectedly, we find that the flypaper effect differs across subjects. Let P_{LS}, P_S, and P_G denote a subject's adjusted proposal in the Lump-Sum, Subsidy, and Grant conditions, respectively. Table 3 shows,

[17] For detailed transfer delivery effects on expenditures, see Part 1.A Transfer Effect in Table 2. In each frame, the differences between the expenditures in the Lump-Sum and Subsidy conditions, as well as between the Subsidy and Grant conditions, are statistically significant at conventional levels. The only exception is the Default with Information frame, where expenditures in the Subsidy condition are (weakly) larger than the expenditures in the Lump-Sum treatment; that is, the estimated coefficient for the Subsidy condition (34 points increase in expenditure) is weakly statistically different from 0 (one-sided p-value=0.059).

[18] Note that it is of interest to determine the effects (if any) of subjects' demographic characteristics. Because there are 31 groups in the full dataset, there is too little variation to include demographic variables in each of the regressions reported in Table 2. Therefore, in results presented in Appendix III we include just two regressions, one that considers alternative transfer delivery methods and one that considers alternative frames. The inclusion of the demographic variables does not materially change the estimates of the treatment effects; if anything, the treatment effects for the Frame condition become sharper. Only three of the coefficients on the demographic variables are statistically significant.

[19] Recall that the adjusted proposal equals the individual's proposal, with the exception of the Grant treatment, where it is the Proposal plus the Grant.

[20] There are 620 proposals in each treatment. There are little differences in these percentages if we consider the number of proposals that are within a small deviation of v (e.g., 20 points).

for each value of v, the relationship between the adjusted proposals and the transfer delivery method.

[Insert Table 3]

The majority of subjects (ranging from 57.4 percent to 68.4 percent, depending on the value of v) made larger adjusted proposals in the Grant condition than in the Subsidy condition (column 1 in Table 3). Similarly, the majority of subjects (ranging from 56.1 percent to 81.3 percent) proposed larger expenditures in the Subsidy condition than in the Lump-Sum condition (column 2 in Table 3). An even larger percentage of subjects (ranging from 74.2 percent to 89.7 percent) preferred larger expenditures in the Grant condition than in the Lump-Sum condition (column 3 in Table 3).[21]

Figure 2 shows the average adjusted proposals for each satiation value v across all transfer delivery methods and frames. Several patterns emerge from Figure 2. First, mean proposals are increasing in v. Second, for values of v that are below the socially efficient values, mean proposals exceed v (i.e., they lie above the 45-degree line in Figure 2). Third, mean adjusted proposals at the maximum values of v fall short of v (i.e., they lie below the 45-degree line).

[Insert Figure 2]

Proposals are expected to increase in v, regardless of the frame, because the larger the individual's maximum benefit from the public good, the larger is the individual's demand for the public good, given that the marginal benefit (up to v) is 1 but the marginal cost is only 1/5. The positive monotonicity is visible in Figure 2, and it is robust across all transfer delivery methods

[21] For those subjects for which $P_S < P_{LS}$ or $P_G < P_{LS}$, i.e., cases that are counter to what we observe in the aggregate, the percentage differences in the values of the proposals for each comparison were small, on average less than 8.0 percent.

and frames. Also, for each transfer delivery method the differences in individual adjusted proposals between frames get smaller as v increases.

To disentangle the effects of alternative transfer delivery methods and frames on individual proposals, we look at the percentage difference between adjusted proposals (denoted P) and valuations (v), measured as $D = 100(P\text{-}v)/v$, and we use linear regression analysis to explain variations in this variable. D is interpreted as the excess proposal P relative to v. We use D rather the P as the dependent variable to capture the curvature suggested by the descriptive results in Figure 2. Our list of regressors includes the satiation value v, the level of default expenditures for the Default and Default with Information frames, the selected expenditures in the prior round for the No-Default frame, and indicators for the No-Transfer, Subsidy, and Grant conditions, where the excluded (control) condition is the No-Transfer condition. The regressions control for potential order effects of transfer delivery methods, for learning (via the dummies for rounds), as well as for subjects' demographic characteristics. We estimate a linear regression specification with robust standard errors clustered at an individual level. Since the effect of default may depend on whether it exceeds or falls short of valuation v, least squares estimates are separately derived for the two situations. The estimation results are reported in Table 4.

[Insert Table 4]

In general, the regression results for individuals are consistent with the patterns observed for average proposals in Figure 2. The coefficients on v are negative and statistically significant in all 6 regressions in Table 4. The negative coefficient implies that, as v increases, the excess proposal rate gets smaller. It also implies that the percentage increase in P as v increases is smaller than the percentage increase in v. This behavior was observed earlier, and it is expected because the larger is the valuation the closer it is to the efficient level and the smaller is the increase that is

needed to improve efficiency. In addition, the coefficients on Default expenditures in Table 4 are positive and larger when the default is larger than own valuation; that is, informing subjects of the expenditures in the prior round (i.e., the default) increases *P* relative to *v*. This effect is quite large. For example, a 100-point increase in the value of the default increases *D* by 52.2 percentage points when the default exceeds own valuations. For the No-Default frame, we include the prior round expenditure level, and the coefficient on this variable is only marginally statistically significant and much smaller (27.5 percentage points) than the coefficient on Default. This result is consistent with intuition because in the No-Default frame the expenditure in the prior round is not salient. These results suggest that the default level is not particularly relevant unless it is made salient.

If there is an income effect, then with larger resources subjects should make larger proposals relative to *v*, meaning that *D* should increase with additional resources. Consistent with an income effect, the coefficients in Table 4 on the No-Transfer dummy, which capture the effect of the smaller income in the No-Transfer condition than in the Lump-Sum condition, are all negative. However, the coefficients on the No-Transfer dummy are statistically significant only when own valuation exceeds the default. In contrast, the coefficients on the Grant condition are positive and statistically significant in all six regressions, which is also consistent with an income effect. The Subsidy condition elicits larger excess proposals than the Lump-Sum condition in cases when a default is absent or when a default is present but exceeds own public good valuation. The coefficients on Grant are larger than the coefficients on Subsidy, and both sets of coefficients are larger in absolute value than the coefficients on No-Transfer.

Once again, all of these aggregate and individual results are consistent with a flypaper effect, and with the effect being larger for the Grant condition than for the Subsidy condition.

Switching to a Subsidy in the Default with Information frame has no statistically significant effect on *D*.

5.2. Other Results[22]

Efficiency. While empirical studies that use observational data also find that grants are associated with larger public good expenditures than comparable increases in income, it is ambiguous whether these larger expenditures are socially desirable. The primary reason for this ambiguity is asymmetric information about the valuations of public goods, which are known only to individual households. An advantage of our study over observational studies is that by design we know individuals' public good valuations. When information about group benefit is not available, our data favor the Grant transfer because it induces expenditures that are closer to the socially efficient level than the expenditures in the Lump-sum and Subsidy conditions.[23] For the Grant condition, the Default frame results in almost perfectly efficient expenditures (488 points versus the optimal level of 490 points), while the mean expenditures in the No-Default frame fall short of the efficient level only by 9 percent (445 points versus 490 points).[24] However, when information on the group benefit is available, the Grant condition elicits too much of the public good (544 points), while the Subsidy (474 points) and the Lump-Sum (440 points) transfer delivery methods come closer to the efficient level of public good provision.

Under-provision of public goods is a chronic economic problem, and costly interventions,

[22] As in section 5.1, we exclude data from the first round in each treatment block in all data analysis and reported statistics in this section.

[23] For the Default treatments (middle column in Table 2), based on one sample *t*-test on mean group expenditures (one observation for each group), the null hypothesis for the mean (expenditure) level of 490 points is rejected at the 5 percent significance level for data in the Subsidy (two-sided p-value=0.027) and the Lump-Sum treatments (two-sided p-value=0.012) but not for the Grant treatment (two-sided p-value=0.925).

[24] For the Grant treatments (last row in Table 2), based on one-sample *t*-tests conducted on mean group expenditures (one observation for each group), the null hypothesis for the mean expenditure level of 490 points is rejected by data in the Default with Information (two-sided p-value = 0.057) and the No-Default treatments (two-sided p-value = 0.015) but not for the Default treatment (two-sided p-value = 0.925).

such as subsidies, are often used to ameliorate the problem. Our results favor the Grant method as a preferred transfer delivery method for solving the problem of under-provision of public goods, unless providing additional information on group benefits is viable. In this case, the Subsidy or the Lump-Sum methods are more promising in addressing the under-provision problem.

Budgeting and Decision Frames. Comparing the No-Default and Default frames allows us to measure the effect of eliminating a default on public good expenditures, and comparing the Default to the Default with Information frames allows us to measure the effect of making the public good nature of expenditures salient when decisions are made over changes. While such information is rarely available to decision makers because of information asymmetry, the nature of the externality associated with the public good (positive or negative) is available in our experimental design. Recall that our theoretical analysis predicts larger expenditures in the Default with Information frame (Hypothesis HA1.a), driven by social preferences (model specification [8]). This behavior is clearly supported by the data.

The Default frame is comparable with observational data since in that frame subjects are voting over changes in expenditures from a default, which is the common practice in setting government budgets. The estimated treatment effect on group expenditures, reported in Part 1.B Frame Effect of Table 2, is -66.11 points (p-value=0.008) when voting is over levels, driven by data in No-Transfer and Lump-Sum conditions. In the No-Transfer condition, we find that the mean group expenditure in the No-Default frame is comparable to the level predicted by the standard model (i.e., the median value of v of 250 points). However, in the Default treatments, group expenditures increase by 45.2 percent, reaching levels that are not statistically different from the efficient level (i.e., the total group payoff maximizer = 350 points). We see a similar, though less dramatic, increase of 28.1 percent (296 points up to 380 points) in the Lump-Sum treatment.

Recall that predictions for the reference dependent model are the same as for the standard selfish model, which are inconsistent with our data. In Appendix 4 we also look at whether our subjects' decisions exhibit preference for status quo (i.e., the default, as in Tversky and Kahneman (1991)), and we find supporting evidence when its predictions are the same as predictions of social preferences ([8]) but limited evidence when the predictions are different.

Fiscal Illusion. For $v = 350$,[25] we find that 60.0 percent of the subjects made larger adjusted proposals in the Lump-Sum treatment than in the No-Transfer treatment.[26] The percentages are even larger for the other two treatments as both subsidy and grant implementations are eliciting larger proposals than the lump-sum mode for the majority of subjects: 64 percent (Subsidy) and 80 percent (Grant) as reported in the 350 row of Table 3.

These findings are not consistent with fiscal illusion. Voters' proposals result in larger expenditures in the Lump-Sum transfer delivery method than the No-Transfer method. However, there is no reason why the Lump-Sum transfer should affect the voter's view of the marginal tax price. Likewise, the Subsidy and Grant transfer delivery methods should not have a differential effect on the voter's calculation of the marginal tax price. Even so, proposals differ between the two treatments.

Information and Social Preferences. As noted earlier, providing information on public good benefits at the group level elicits slightly larger expenditures (column [3] in Table 2), with increases in expenditures of 15.7 percent (Lump-Sum), 9.9 percent (Subsidy), and 11.6 percent (Grant). While the coefficients for the least square estimates reported in Part 1.B Frame Effect in

[25] We can only compare the proposals in the three transfer conditions to the No-Transfer condition for v equal to 350 points because the values of v in the No-Transfer condition for each of the respective rank of v are smaller than the values in transfers conditions other than for $v = 350$.

[26] This is consistent with preferences for efficiency. Recall that in the No-Transfer treatment proposing more than 350 points does not increase economic efficiency. In contrast, however, in all of the transfer conditions 350 points falls short of the efficient expenditure level of 490 points, so that proposing larger expenditures increases efficiency in the transfer conditions.

Table 2 have the expected positive sign, they are not statistically different from zero, with the exception of the Grant condition. Our data suggest that providing information on the positive externalities of individual payments may nudge socially-oriented individuals toward internalizing externalities.

An implication of the model of social preferences in section 4 is that an individual with social preferences will propose more than the own valuation v if $v < v_{n-1}$ but less if $v > v_{n-1}$. We construct an individual's excess proposal as adjusted proposal less v (i.e., $P - v$). Figure 3 shows the histograms of $P - v$ separately for $v < v_{n-1}$ (top row) and $v > v_{n-1}$ (bottom row). The columns in Figure 3 correspond to three different framing of proposals: No-Default (left column), Default (middle column), and Default with Information (right column). We observe a total of 1488 proposals for which $v < v_{n-1}$ and 496 proposals for which $v > v_{n-1}$. When $v < v_{n-1}$, 74 percent (1105 out of 1488 proposals) of the adjusted proposals exceed v; when $v > v_{n-1}$, 71 percent (350 out of 496 proposals) of the adjusted proposals fall short of v_n. This mirror pattern of the excess proposals between low and high beneficiaries of the public good is robust across the different frames. Prompting consideration of group benefits through the Default with Information frame nudges subjects away from the selfish behavior, which is visible in the drop of the fraction of proposals that are equal to valuations (observations at 0).

[Insert Figure 3]

6. Conclusions

Prior theoretical and empirical work has been unable to resolve the existence of the flypaper effect in government expenditures. In this paper, we conduct a laboratory experiment to investigate

possible explanations for the flypaper effect. Our experimental design allows for choice of public good expenditures that are consistent with the standard model of grants, i.e., we control for the non-behavioral factors that have been suggested as explanations for the flypaper effect observed in empirical studies of grants. Our design also allows for behavioral economics mechanisms that have been suggested as explanations for the flypaper effect. Our goal is to examine whether a flypaper effect exists in the laboratory and whether alternative institutional arrangements affect its existence.

Our main result is that we find that increases in public good expenditures are largest when the transfer is a grant added to the public good and smallest for an (payoff equivalent) increase in income, with a subsidy increasing expenditures by more than an increase in income and less than the grant. These results are consistent with a flypaper effect.

We are also able to examine the effects of transfers and framing on the efficiency of public good provision. In the treatment in which there is no transfer and voting is over the level of public good expenditures, proposed expenditures are close to the expenditures predicted by a model in which subjects maximize their own payoff (utility). Under the median voter procedure, the expenditures that a utility maximizing (selfish) subject would propose will generally not be at the socially efficient level. We find that when proposals are changes to a given expenditure level (e.g., the expenditure level chosen in the prior round) and when subjects are prompted to consider social efficiency, adopted expenditures are close to the socially efficient level. In addition, we find that the institutional nature of the transfer (or a lump-sum transfer versus an expenditure reimbursement) affects the voting outcome. Further, we find that proposed public good expenditure level is larger if the proposals are changes from a reference (default) level, regardless of the nature of the transfer, and that prompting subjects to consider the effect of the expenditures

on other participants (e.g., other regarding preferences) results in larger proposed expenditures. We find no evidence consistent with a role for fiscal illusion in the flypaper effect.

What can explain our results? Recall that explanations for the presence of a flypaper effect in empirical studies fall into roughly three categories: econometric issues, political control, and voter behavior (fiscal illusion). It is essential to recognize that we find a flypaper effect despite there being nothing in our experimental design that reflects these three possible explanations; that is, there are no econometric issues, there are no political control issues (e.g., agenda control or inefficient political institutions), and there are no voter behavior issues. For example, one of the suggested causes of the flypaper effect is fiscal illusion, where voters misperceive the marginal tax price of government services in the presence of government transfers. Our results are not consistent with fiscal illusion. While our experiment cannot of course rule out that these issues explain the findings in published empirical research, our results strongly suggest that there are other causes of the flypaper effect.

In contrast, our results suggest that mechanisms consistent with behavioral economics may drive the flypaper effect. Indeed, we are able to explain our results with an additive separable nonlinear behavioral model. In addition, we looked at a linear social preference model in which subjects have preferences over efficiency, beyond the subject's direct payoff from the public good provision. We also looked at a linear model of reference dependence in which a subject's utility depends on gain-loss in utility, whose value depends on the difference between the reference level of expenditure and the proposed expenditure. Individual proposals are consistent with the former (linear social preferences) but not the latter (linear reference dependence).

Understanding the effects of transfers is important for both theoretical and practical considerations. The former tries to answer how communities make the optimal choice between

private and public consumption, an issue on which our knowledge is far from perfect. The latter tries to measure the impact of different implementations of grants on local expenditures, which is important for designing policies especially in the current environment. Our results suggest that, if one is concerned about the efficiency of public good provision, the Default with Information frame should be preferred over other frames. More broadly, our results suggest that the optimal design for intergovernmental transfers must consider that money typically "sticks where it hits", regardless of the specific mechanism by which this mechanism exists.

References

Abeler, Johannes and Felix Marklein. 2017. "Fungibility, labels, and consumption." *Journal of the European Economic Association* 15 (1): 99-127.

Ackert, Lucy, Jorge Martinez-Vazquez, and Mark Rider. 2007. "Social preferences and tax policy design: Some experimental evidence." *Economic Inquiry* 45 (3): 487-501.

Alekseev, Aleksandr, Gary Charness, and Uri Gneezy. 2017. "Experimental Methods: When and Why Contextual Instructions are Important." *Journal of Economic Behavior & Organization* 134: 48-59.

Alm, James and Carolyn J. Bourdeaux. 2013. "Applying behavioral Economics to the public sector." 2013. *Hacienda Pública Española/Review of Public Economics* 206 (3/2013): 91-134.

Alm, James and Sarah Jacobson. 2007. "Using laboratory experiments in public economics." *National Tax Journal* 60 (1): 129-152.

Andreoni, James and John Miller. 2002. "Giving according to GARP: An experimental test of the consistency of preferences for altruism." *Econometrica* 70 (2): 737-753.

Ashworth, John and Bruno Heyndels. 1999. "Reference point effects in local taxation: It all depends on how you look at it." *Proceedings of the Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association,* Atlanta, GA: 335-341.

Becker, Johannes, Daniel Hopp, and Michael Kriebel. 2018. "Mental accounting of public funds-The flypaper effect in the lab." Beiträge zur Jahrestagung desVereins für Socialpolitik: Digitale Wirtschaft – Session: Experiments – Decision Theory, No. A19-V3, ZBW - Leibniz-Informationszentrum Wirtschaft, Kiel, Hamburg.

Blanchard, Olivier Jean, Florencio Lopez de Silanes, and Andrei Shleifer. 1994. "What do firms do with cash windfalls?" *Journal of Financial Economics* 36 (3): 337-360.

Bolton, Gary E. and Alex Ockenfels. 2000. "ERC – A theory of equity, reciprocity, and competition." *The American Economic Review* 90 (1): 166-193.

Charles-Cadogan, Godfrey. 2018. "Estimation and inference for loss aversion in cross sectional regressions of happiness and consumption on economic growth." Paper presented at Royal Economics Society, Sussex 2018 and at the European Econometrics Society, Cologne 2018.

Charness, Gary and Matthew Rabin. 2002. "Understanding social preferences with simple tests." *The Quarterly Journal of Economics* 117 (3): 817-869.

Chernick, Howard. 1979. "An economic model of the distribution of project grants." In Peter Mieszkowski and William Oakland (eds.), *Fiscal Federalism and Grants-in-Aid*. Washington, D.C.: Urban Institute Press, 81-103.

Crain, W. Mark and Nicole Verrier Crain. 1998. "Fiscal consequences of budget baselines." *Journal of Public Economics* 67 (3): 421-436.

Crawford, Vincent P. and Juanjuan Meng. 2011. "New York City drivers' labor supply revisited: Reference-dependent preferences and rational expectations targets for hours and income." *The American Economic Review* 101 (5): 1912-1932.

Courant, Paul N, Edward M Gramlich, and Daniel L Rubinfeld. 1979. "The stimulative effects of intergovernmental grants: Or why money sticks where it hits." In Peter Mieszkowski and William Oakland (eds.), *Fiscal Federalism and Grants-in-Aid*. Washington, D.C.: Urban Institute Press, 5-21.

Cox, James C. and Vjollca Sadiraj. 2012. "Direct tests of individual preferences for efficiency and equity." *Economic Inquiry* 50 (4): 920-931.

Fehr, Ernst and Klaus M. Schmidt. 1999. "A theory of fairness, competition, and cooperation." *The Quarterly Journal of Economics* 114 (3): 817-68.

Filimon, Radu, Thomas Romer, and Howard Rosenthal. 1982. "Asymmetric information and agenda control: The bases of monopoly power in public spending." *Journal of Public Economics* 17 (1): 51-70.

Fisher, Ronald C. 1982. "Income and grant effects on local expenditure: The flypaper effect and other difficulties." *Journal of Urban Economics* 12 (3): 324-345.

Fréchette, Guillaume R., John H. Kagel, and Massimo Morelli. 2012. "Pork versus public goods: An experimental study of public good provision with a legislative bargaining framework." *Economic Theory* 49: 779-800.

Hamilton, Bruce W. 1983. "The flypaper effect and other anomalies." *Journal of Public Economics* 22 (3): 347–361.

Hines, James R. and Richard H. Thaler. 1995. "Anomalies: The flypaper effect." *The Journal of Economic Perspectives* 9 (4): 217-226.

Hewett, Roger, Charles A. Holt, Georgia Kosmopoulou, Christine Kymn, Cheryl X. Long, Shabnam Mousavi, and Sudipta Sarangi. 2005. "A classroom exercise: Voting by ballots and feet." *Southern Economic Journal* 72 (1): 253-263.

Inman, Robert P. 2009. "The flypaper effect." *The New Palgrave Dictionary of Economics.* London, UK: Palgrave Macmillan.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. 1990. "Experimental tests of the endowment effect and the Coase Theorem." *The Journal of Political Economy* 98 (6): 1325-1348.

Knight, Brian G. 2002. "Endogenous federal grants and crowd-out of state government spending: Theory and evidence from the federal highway aid program." *The American Economic Review* 92 (1): 71-92.

Kőszegi, Botond and Matthew Rabin. 2006. "A model of reference-dependent preferences." *The Quarterly Journal of Economics* 121 (4): 1133-1165.

Leduc, Sylvain and Daniel Wilson. 2017. "Are state governments roadblocks to Federal stimulus? Evidence on the flypaper effect of highway grants in the 2009 recovery act." *American Economic Journal: Economic Policy* 9 (2): 253-292.

Moffitt, Robert A. 1984. "The effects of grants-in-aid on state and local expenditures: The case of AFDC." *Journal of Public Economics* 23 (3): 279-305.

O'Donoghue, Ted and Charles Sprenger. 2018. "Reference-dependent preferences." In B. Douglas Bernheim, Stefano Dellauigna, and David Laibson (eds.), *Handbook of Behavioral Economics: Foundations and Applications*. Amsterdam, The Netherlands: North Holland, 1-78.

Palfrey, Thomas R. 2016. "Experiments in political economy." In John H. Kagel and Alvin E. Roth (eds.), *The Handbook of Experimental Economics, Vol 2*. Princeton, NJ: Princeton University Press, 347-434.

Rabin, Matthew. 2000. "Risk aversion and expected utility theory: A calibration theorem." *Econometrica* 68 (5): 1281-1292.

Singhal, Monica. 2008. "Special interest groups and the allocation of public funds." *Journal of Public Economics* 92 (3): 548-564.

Turnbull, Geoffrey K. 1998. "The overspending and flypaper effects of fiscal illusion: Theory and empirical evidence." *Journal of Urban Economics* 44 (1): 1-26.

Thaler, R. H. 1999. "Mental accounting matters." *Journal of Behavioral Decision Making* 12: 183-206.

Tversky, Amos and Daniel Kahneman. 1991. "Loss aversion in riskless choice: A reference-dependent model." *The Quarterly Journal of Economics* 106 (4): 1039-1106.

Vegh, Carlos A. and Guillermo Vuletin. 2015. "Unsticking the flypaper effect in an uncertain world." *Journal of Public Economics* 131: 142-115.

Figure 1. Group Expenditures across Treatments

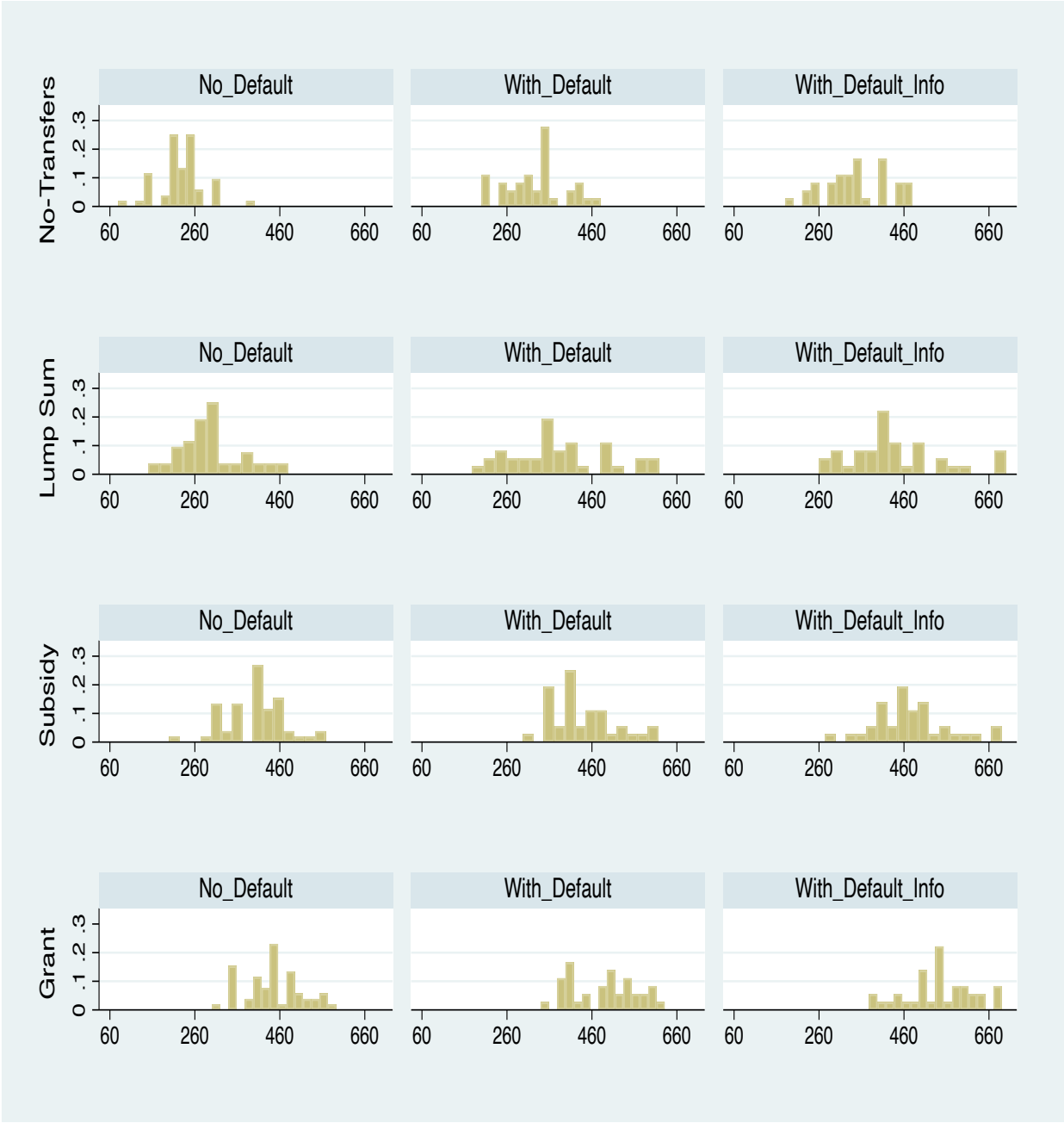
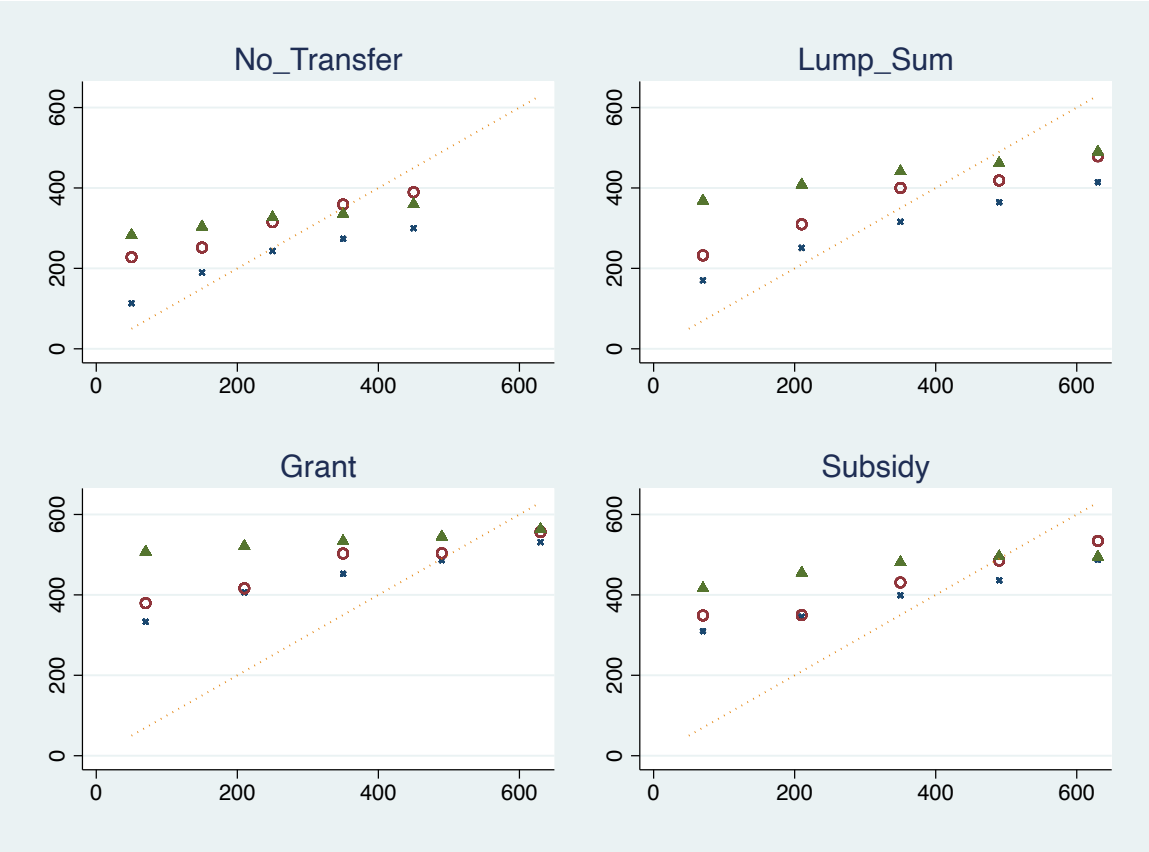
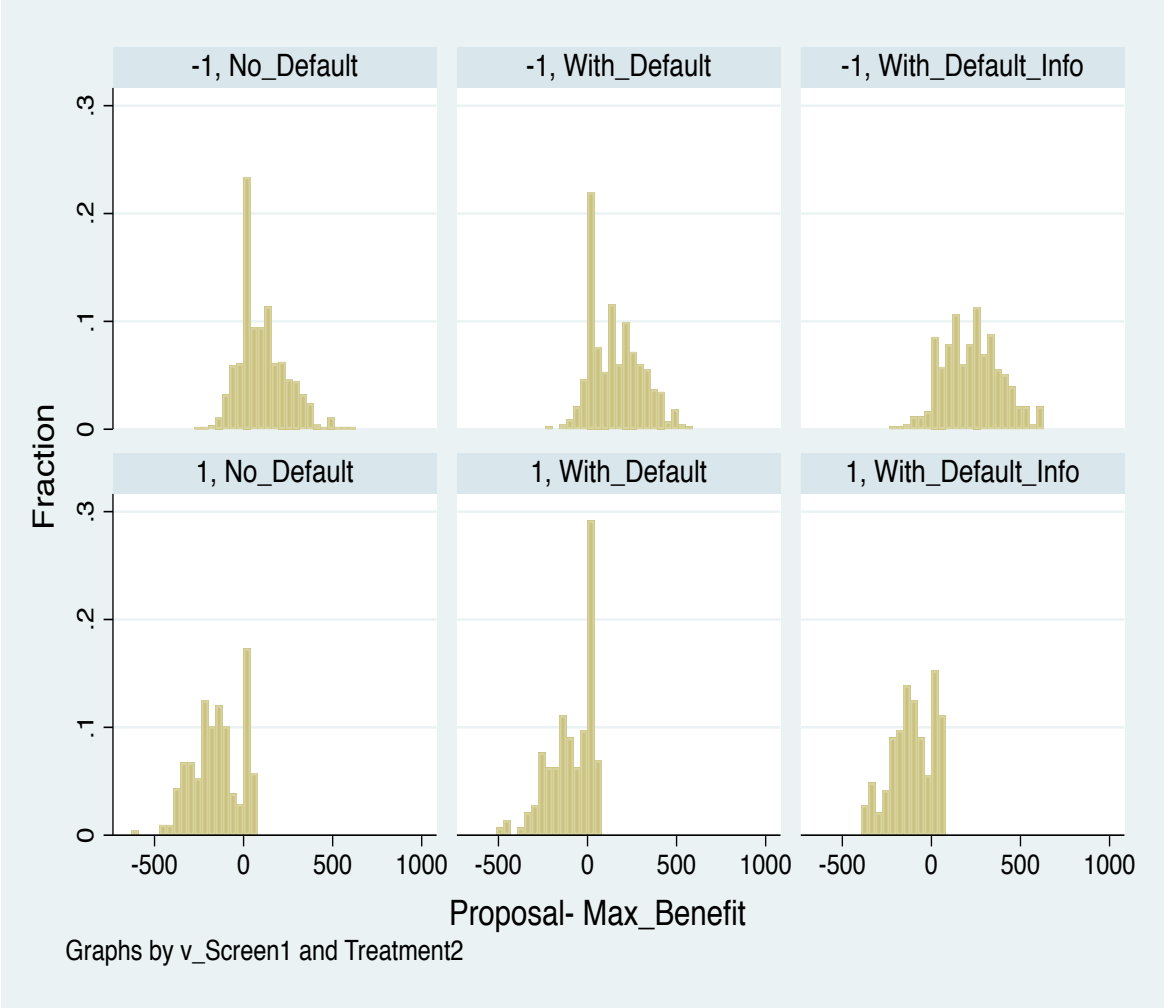


Figure 2. Mean Adjusted Proposals for Each Satiation Value (v) across Treatments



Notes: With the exception of Grant treatment, adjusted proposal equals the observed proposal; In the Grant treatment, Adjusted Proposal equals (observed proposal) + 200. The vertical axis is the mean adjusted proposal and horizontal axis is satiation value, *v* (maximum benefit from the public good). Round 1 data in each block are not included. Legend: x (No-Default), Circles (Default), Triangles (Default with Information). 45-degree line: Adjusted Proposal equals Satiation Value (v).

Figure 3. Histograms of Observed Adjusted Proposal less than Satiation Value (*v*)



Notes: In the top row, the Maximum Benefit (v) is smaller than v_{n-1} (which is group efficient expenditures level). In the bottom row, the Maximum Benefit (v) is larger than v_{n-1}.

Table 1. Experimental Design

Framing Conditions	
No-Default	Proposal is in levels of expenditures.
Default	Proposal is in changes in expenditures from a benchmark. The benchmark is set equal to the group's expenditures in the preceding round, except for the first round of each block when the default is 200.
Default with Information	This is same as the Default frame except that the subjects are told the total group benefit of the expenditures for any proposal.
Transfer Conditions	
No-Transfer	Subjects earn 100 units. The public good level = the adopted proposal = expenditure level = tax payment.
Lump-Sum	Subjects are given an addition 40 units, for a total of 140 units. Public good level = adopted proposal = expenditure level = equals tax payment.
Subsidy	A 200-unit, i.e., 40 units per subject, reimbursement of the adopted proposal is provided. Public good level = adopted proposal = equals expenditure level. Tax payment = adopted proposal - the 200 - unit subsidy.
Grant	The adopted proposal is increased by 200 units, i.e., 40 units per subject. The adopted proposal = tax payment = expenditure level. Public good level = adopted proposal + 200 units.
Proposals	Subjects propose total expenditure level. The median proposal is selected. Taxes for each subject are an equal share of the cost of the net expenditures. Adjusted proposal = proposal + grant.

Notes: The framing conditions are assigned between subjects. The transfer delivery conditions are assigned within subjects; the order of three transfer delivery conditions (Lump-Sum, Subsidy, Grant) is randomized.

Table 2. Summary Statistics of Group Expenditures

	[1] No-Default	[2] Default	[3] Default + Info	[4] Predictions	
Observations (groups)	**208 (13)**	**144 (9)**	**144 (9)**	**Egoistic**	**Efficient**
No-Transfer	223.27	**324.17**	**345.00**	250	350
	(53.90)	(71.57)	(75.35)		
Lump-Sum	296.54	380.00	**440.00**	350	490
	(72.27)	(114.99)	(113.14)		
Subsidy	397.88	435.56	**474.17**	350	490
	(70.27)	(77.59)	(88.85)		
Grant	445.38	**487.78**	544.44	350	490
	(67.98)	(79.47)	(82.41)		

Notes: Standard deviations are in parentheses. Observations from the first round in each block are not included. Entries in the last two columns show group expenditures for the egoistic and the reference dependent models, along with the socially efficient expenditure level (with the least payoff inequality). Figures in bold indicate that expenditures are not statistically different (at 5% significance level) from the efficient expenditures level, based on one sample *t*-test.

Least Squares Estimates

(Dependent Variable: Total Group Expenditures)

1.A Transfer Effect	**All Data**	**No-Default**	**Default**	**Default + Info**
No Transfer	-74.52***	-73.27***	-55.83*	-95.00***
	(12.09)	(13.244)	(26.915)	(26.670)
Subsidy	68.55***	101.35***	55.56*	34.17
	(11.77)	(14.446)	(24.604)	(19.598)
Grant	124.0***	148.85***	107.78***	104.44***
	(13.39)	(18.939)	(23.082)	(28.976)
Constant	362.4***	296.54***	380.00***	440.00***
	(18.75)	(16.854)	(34.603)	(34.412)
Observations	496	208	144	144
R-squared	0.389	0.633	0.334	0.390
Number of clusters	31	13	9	9

1.B Frame Effect	**All Data**	**No Transfer**	**Lump-Sum**	**Subsidy**	**Grant**
No-Default	-66.11***	-100.9***	-83.46**	-37.67	-42.39
	(23.26)	(21.56)	(36.96)	(24.57)	(26.94)
Default + Info	44.03	20.83	60	38.61	56.67*
	(29.85)	(29.00)	(46.66)	(30.98)	(32.36)
Constant	406.9***	324.2***	380***	435.6***	487.8***
	(20.44)	(19.39)	(33.08)	(19.53)	(22.06)
Observations	496	124	124	124	124
R-squared	0.153	0.421	0.278	0.145	0.231
Number of clusters	31	31	31	31	31

Notes: The control treatments are Lump-Sum in Part 1.A and Default in Part 1.B. Robust standard errors (clustered by group) are in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 3. Relationship between Adjusted Proposals under Alternative Treatments

	Percent of Subjects for Whom:		
Value of *v*	**[1]** $P_G > P_S$	**[2]** $P_S > P_{LS}$	**[3]** $P_G > P_{LS}$
70	57.4%	81.3%	89.7%
210	64.5%	60.6%	80.0%
350	68.4%	63.9%	80.0%
490	59.3%	66.4%	74.2%
630	63.9%	56.1%	76.8%

Notes: The total number of subjects in all cells is 155.

Table 4. Determinants of Excess Adjusted Proposal over Maximum Benefit (v)

Dependent Variable: 100*(Adjusted Proposal - Maximum Benefit) / Maximum Benefit						
	No-Default		**Default**		**Default with Information**	
	***v* ≤ default**	***v* > default**	***v* ≤ default**	***v* > default**	***v* ≤ default**	***v* > default**
Benefit (v)	-1.139***	-0.153***	-1.227***	-0.170***	-1.533***	-0.189***
	(0.108)	(0.009)	(0.114)	(0.020)	(0.172)	(0.015)
Default			0.522***	0.141***	0.722***	0.163***
			(0.172)	(0.017)	(0.176)	(0.026)
Previous Round	0.275	0.146***				
	(0.174)	(0.021)				
No Transfer (D)	-24.071	-7.456**	15.229	-8.460***	-14.079	-14.755***
	(20.407)	(3.412)	(17.678)	(2.445)	(17.355)	(4.203)
Subsidy (D)	98.889***	-0.289	47.140***	3.696	20.317	-1.115
	(19.546)	(3.188)	(16.962)	(2.737)	(12.895)	(2.006)
Grant (D)	163.163***	36.339***	145.793***	31.517***	150.496***	30.937***
	(15.489)	(2.316)	(21.390)	(3.070)	(20.610)	(3.119)
Constant	38.595	-2.943	-165.887	30.297*	324.374*	4.839
	(221.931)	(27.212)	(108.535)	(16.573)	(188.128)	(23.631)
Demographics?	Yes	Yes	Yes	Yes	Yes	Yes
Order Control?	Yes	Yes	Yes	Yes	Yes	Yes
Round Control?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	456	568	395	325	416	304
Adjusted R^2	0.484	0.511	0.536	0.560	0.588	0.646

Note: Robust standard errors (clustered by subject) are in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Round 1 data in each block (treatment) are not included. (D) indicates a dummy variable. The Lump-Sum treatment is the control treatment. Demographics include Age, GPS, CRT Score, and dummies for Female, Black, Freshman, Christian, Liberal, and Did Not Vote. All model specifications control for block order and round (within each block) effects.

Appendix I: Instructions

1. Treatments with No Group Information

WELCOME

No Talking Allowed
Now that the experiment has begun, we ask that you do not talk. If you have a question while you are reading the instructions, please raise your hand and the experimenter will approach you and answer your question in private.

Turn Off Personal Electronics
Turn off and put away all mobile phones and other electronic equipment. Anyone seen using such items during the experiment will be excluded from participation in future experiments.

Payment Amount
Your total payment will consist of a participation fee of $5, and the amount you earn in the course of the experiment. The earnings during the experiment will be in points and converted to U.S. dollars at a rate of 10 points = $1. You will be paid in cash privately at the end of the experiment.

Payment Procedure
There will be 20 rounds in today's session, split into four blocks. One round will be picked at random for payment. It is, therefore, worth your effort to think carefully about each decision, since you don't know which decision round will be selected.

Time
Today's session will consist of two types of decision tasks (working and voting) and a demographic survey. The whole session should take up no more than 2 hours.

Fixed Matching and Anonymity
At the beginning of the experiment you will be randomly matched with four other people in the room to form a group of five. Groups remain fixed during the entire experiment but no one will learn the identity of other people in his or her group. Your decisions will be completely anonymous. No participant will be able to link your choices to your identity.

Decision Tasks
The experiment today consists of 20 rounds organized in 4 blocks of 5 rounds each. A block consists of a working task and five rounds of a voting task.

<u>Working Task</u>
In the working task, you will be given a paragraph of text that contains 5 misspelled words. You will be asked to try to correct all of them. You will earn a certain number of points upon completion of this task. You will perform the working task only once in each block, but your earned income will be used in all 5 rounds of the voting task. The maximum earnings in the working task will differ across the blocks.

Voting Task

In the voting task you will be asked to make a proposal for the level of group expenditures. These expenditures will be used to finance a public good (to be explained below).

Block 1. In this block you will have a chance to earn 100 points in the working task.
In the voting task your group makes a decision on the group expenditures. One point of group expenditures is worth 1 unit of the public good. The public good benefits all five group members, although by different numbers of points.

How does the public good benefit you?
Your benefit from the public good will equal the group expenditures, up to a maximum benefit. The maximum benefit will differ across members of your group and across rounds. For each member, a tax equal to 1/5 of the group expenditures is deducted from the earnings from the working task.

> Example Suppose that your assigned maximum benefit is 150 points.
> If the group expenditures are 100 points then: (1) your benefit from the public good is 100 points and (2) the tax is 20 points (=100/5)

If the group expenditures are 250 points, then your benefit from the public good is simply 150 points (your assigned maximum benefit) and the tax is 50 points (=250/5).

How are the group expenditures on the public good determined?
Each group member privately proposes a level of expenditures on the public good. The proposal can be up to five times the earnings in the working task. The median value of the proposals is selected as the group expenditures.

> Example Suppose that the proposed expenditures are 50, 100, 200, 300, 300. Then the group expenditure level is the median value, 200 points.

How do you submit your proposal?
You will make your proposal on group expenditures by using a slider on a choice screen, which moves in 10 point increments. You confirm your selection by clicking the "Submit" button.

The screen will show your earnings and your maximum benefit. As you move the slider the screen will also show your benefit and your tax (share of group expenditures) IF your proposal is selected as the group expenditures.

What is your round payoff?
Your total payoff for each round is the sum of your benefit from the public good and your earnings in the working task minus a tax equal to 1/5 of the group expenditures.

$$\text{Payoff} = (\text{public good benefit}) + (\text{income} - \text{tax})$$

Example Suppose that:
1. your earnings in the working task are 100 points,
2. your assigned maximum benefit from the public good is 150 points and

3. the group expenditures are 200 points.
   In this scenario, the tax is 40 (=200/5) points, so your remaining earnings are 60 (=100-40) points. Your benefit from the public good is 150 points, and therefore your round payoff is 210 (=150+60) points.

End of round. After all decisions in a round are made you will see the results screen with the information about all the proposals, your benefit, your tax (share of group expenditures), and your total payoff. After that a new round within a block will begin.

Blocks 2 to 4. In two of these three blocks you will earn 100 points, whereas in one of them you will earn 140 points for attempting to correct all misspelled words.

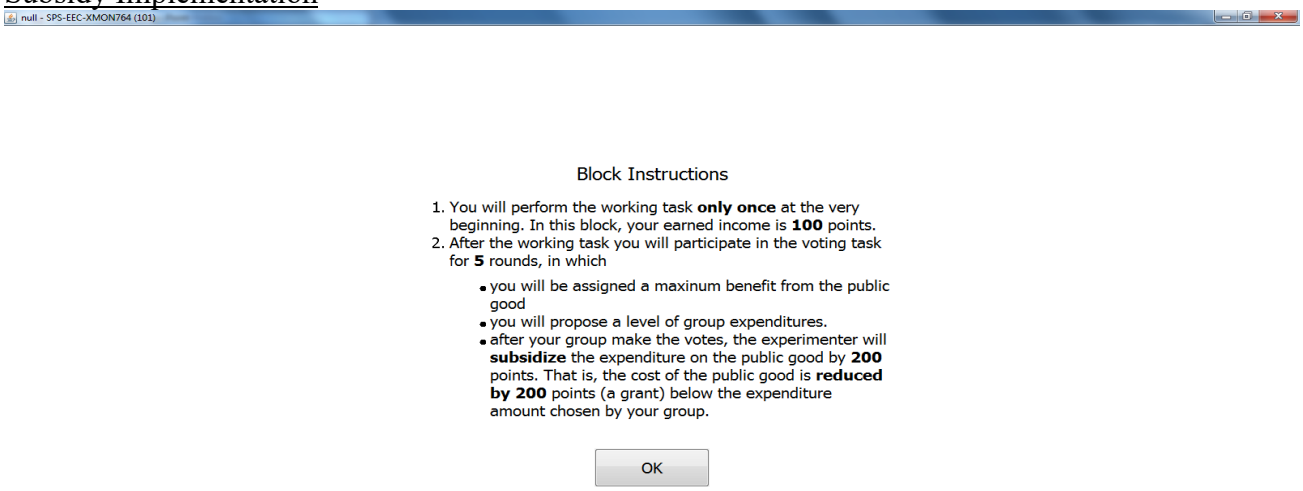
2. Treatments With Group Information

Instructions are the same as above but with an additional sentence added at the end of the paragraph above *What is your round payoff ?*

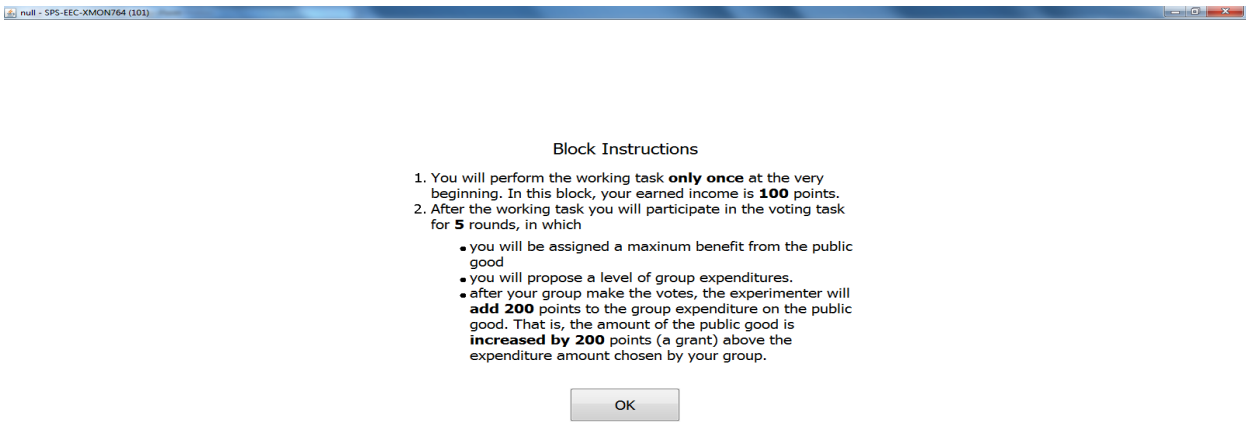
You will also see the total benefit your group will get IF your proposal is selected as the group expenditure level.

3. No-Default Treatments: On screen Instructions at the beginning of Blocks 2 to 4

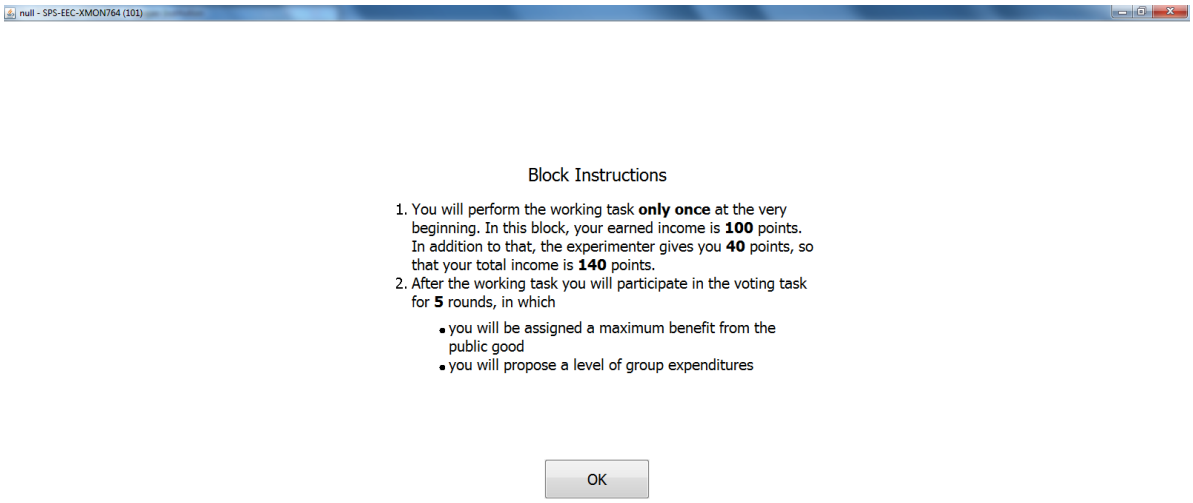
Subsidy Implementation



Grant Implementation



Lump-Sum Implementation



4. **Default treatments: On screen Instructions at the beginning of each grant delivery method with Default**

Same as above (in No-Default treatments) but replace

- You will propose a level of group expenditures

with

- You will propose a change in group expenditures from a given default level

Appendix II: Proofs

PROOF OF PROPOSITION 1

The Lump-Sum scenario elicits smaller expenditures than the Subsidy scenario, $E^o \le E^s$. We show that the opposite supposition, optimal expenditures in the Lump-Sum scenario being larger, or $E^o > E^s$, leads to a contradiction. Indeed,

$$f'((E^s - G)/n) = ng'(E^s) \ge ng'(E^o) = f'(E^o/n) \ge f'(E^s/n)$$

where the equalities follow from first-order conditions ((2) and (4)), the first inequality follows from concavity of $g(.)$, and the second one from convexity of $f(.)$. Thus, $f'((E^s - G)/n) \ge f'(E^s/n)$. Together with convexity of f(.), this implies that

$$E^s - G \ge E^s$$

which contradicts with G>0.

The Grant scenario elicits larger expenditures than the Subsidy scenario, or $E^s \le E^g + G$. Suppose that this is not true. If $E^s > E^g + G$ (*), then by convexity of f(.) and by first-order conditions ((3) and (4)) we have

$$ng'(E^g + G) = f'(E^g/n) \le f'((E^s - G)/n) = ng'(E^s)$$

Dividing by n, we get that $g'(E^g + G) \le g'(E^s)$, and by concavity of $g(.)$, we have that $E^g + G \ge E^s$, which contradicts the supposition (*).

PROOF OF OBSERVATION 1 (Reference Dependent Choices)

The following two cases are possible depending on whether the default expenditure level E^r exceeds the individual's valuation v.

Case 1: $E^r > v$. Note that expenditures E larger than the default level E^r cannot be optimal since both (consumption and gain-loss utility) terms of U^r at such E are smaller than at E^r. Expenditure levels smaller than v also cannot be optimal because

$$U^r(v) - U^r(E) = (v - E)(1 - \frac{1}{n} + \eta(\lambda_2 - \frac{1}{n})) > 0, \quad \forall E \in [0, \min(v, E^r))$$

At any expenditures level E between v and E^r there are no gains and no losses with respect to the public good since the benefit is v for all $E > v$, but there is a gain in the private good payoff because

at $E < E^r$ more resource (out of w) is left to spend on the private good. Verify that own valuation, v is optimal as

$$U^r(v) - U^r(E) = \frac{1}{n}(E - v)(1+\eta) > 0, \quad \forall E \in (v, E^r]$$

Case 2: $E^r \le v$. Compared to expenditure level v, larger expenditures cannot be optimal since they are associated with losses from the private good and no gains from the public good. Expenditures below the default level E^r also cannot be optimal because

$$U^r(E) - U^r(E^r) = (E - E^r)(1 - \frac{1}{n} + \eta(\lambda_2 - \frac{1}{n})) < 0, \quad \forall E \in (0, \min(v, E^r))$$

Expenditures between the reference level E^r and the public good satiation level v come with gains ($E^r < E < v$) in the payoff from the public good but losses ($(E^r - E)/n < 0$) in private good payoff. Using statements [3] and [4], we get

$$U^r(E \mid E \in (E^r, v)) = w + E(1 - \frac{1}{n}) + \eta(E - E^r)(1 - \frac{1}{n}\lambda_1)$$

which is increasing in E iff $(n-1) > (\lambda_1 - n)\eta$. Summarizing, when the default expenditure is below the individual's valuation v, the optimal expenditures, E^b is given by

$$\begin{aligned} E^b &= v, && \text{if} \quad (n-1) > (\lambda_1 - n)\eta \\ &= E^r, && \text{if} \quad (n-1) < (\lambda_1 - n)\eta \\ &\in [E^r, v], && \text{if} \quad (n-1) = (\lambda_1 - n)\eta \end{aligned}$$

PROOF OF OBSERVATION 2 (Social Preferences)

To facilitate the discussion on the trade-off between own payoff and group payoff maximization, write the right hand side of [8] as a weighted sum of own payoff plus the average group payoff

$$u_i(E) = (1 - \theta\frac{n}{n-1})\pi_i(E) + \theta\frac{1}{n-1}\sum_j \pi_j \qquad \text{[A.1]}$$

We begin by noting that any public good expenditure level between v_{n-1} and v_n maximizes total group payoff (i.e., such values of E maximize the last term in specification [A.1]), and thus are socially efficient. Indeed, for any integer $k < n-1$ at any $E \in (v_k, v_{k+1})$ the marginal cost of E for

the group (is 1) and less than the marginal benefit of E for the group (= *n-k)*. Hence, efficiency requires $E \in [v_{n-1}, v_n]$.

Next, it is straightforward to verify that, for any expenditure $E \in (v_{n-1}, v_n]$, one can move from the payoff distribution at any such E to the payoff distribution when expenditure is v_{n-1} by transferring $(E - v_{n-1})/n$ from individual with valuation v_n to every other individual with a smaller valuation. By the Dalton principle, the payoffs are more equally distributed when the expenditure level is v_{n-1} than when $E \in (v_{n-1}, v_n]$. This, together with i's payoff decreasing for $E > v_i$, implies that the solution for E to problem [A.1] for an individual for whom $v_i < v_n$ cannot exceed v_n However, the solution also cannot be smaller than v_i because both terms in [A.1] (i's own payoff and group payoff) increase in E for $E < v_i (\leq v_{n-1})$. Verify that individual i gets larger utility at any expenditure level $v_k > v_i$ such that $k < n + \left(1 - (n-1)/n\theta_i\right)$.[27] Therefore, i's optimal expenditure level $E^o_{i \in \{1,\ldots,n-1\}}$ for social preferences [8] satisfies

$$E_i^o \in [v_i, v_{n-1}] \qquad \text{[A.2]}$$

On the other hand, for the individual with the largest benefit v_n, the marginal utility at expenditure level E between v_{n-1} and v_n is $(1 - \theta\frac{n}{n-1})(1 - \frac{1}{n})$, which is negative if $\theta \in (1 - 1/n, 1)$. If so, then he will propose less than his own benefit v_n.

[27] For example, in a group of 5 an individual with $\theta = 0.25$ and valuation v_1 prefers v_2 over own valuation but prefers v_1 over v_4.

Appendix III: Additional Regressions

Dependent Variable: **Total Group Expenditure**	**Transfer Method**	**Framing Method**
No Transfer	-74.52***	
	(-12.2)	
Subsidy	68.55***	
	(-11.88)	
Grant	124.03***	
	(-13.51)	
No Default		-56.60**
		(-24.72)
Default + Infor		43.90**
		(-21.10)
Mean Group Age	34.17***	14.01
	(-11.42)	(-14.22)
Mean Group GPA	37.07	22.44
	(-28.07)	(-28.22)
Proportion Female	-95.17	-71.74
	(-66.71)	(-64.14)
Proportion Black	96.64**	40.77
	(-42.45)	(-46.36)
Proportion Freshman	123.11	27.33
	(-75.82)	(-68.4)
Proportion Christian	-52.13	-26.98
	(-57.32)	(-47.48)
Proportion Liberal	90.06	84.68
	(-70.27)	(-55.9)
Proportion Did Not Vote	-39.53	54.29
	(-46.81)	(-50.98)
CRT Score	41.22*	5.95
	(-22.60)	(-25.35)
Constant	-524.89	-9.53
	(-327.50)	(-350.44)
Observations	496	496
Adjusted R^2	0.51	0.17
Residual Standard Error	83.53 (df = 483)	108.58 (df = 484)
F Statistic	43.98*** (df = 12; 483)	10.39*** (df = 11; 484)

Note: *** p<0.01, ** p<0.05, * p<0.1. Since the regressions consider groups level outcomes, we aggregate the demographic characteristics to a group level. In column 2, the excluded treatment is Lump-Sum Transfer, while in column 3 the excluded frame is Default.

Appendix IV. Preference for Status Quo

A selfish subject who places a special value on the status quo (i.e., the default, as in Tversky and Kahneman (1991)) may propose some P_i that rests between the default E^r and the own value of v; that is, $P_i > v$ if $E^r > v$; $P_i = v$ if $E^r = v$; and $P_i < v$ if $E^r < v$. We explore the extent to which each subject's behavior is consistent with this hypothesis in the table below.[28]

Consistency of Adjusted Proposals with Status Quo Effect

	[1] $E^r < v$ **Number [% of column total] {% of row total}**	**[2]** $E^r = v$ **Number [% of column total] {% of row total}**	**[3]** $E^r > v$ **Number [% of column total] {% of row total}**	**[4] Total**
$E_i < v$	24	4	29	57
	[21.1%]	[9.8%]	[4.1%]	
	{42.1%}	{7.0%}	{50.9%}	
$E_i = v$	16	13	73	102
	[14.0%]	[31.7%]	[10.3%]	
	{15.7%}	{12.7%}	{71.6%}	
$E_i > v$	74	24	607	705
	[64.9%]	[58.5%]	[85.6%]	
	{10.5%}	{3.4%}	{86.1%}	
Total	114	41	709	864

Notes: E^r is the default expenditure, and E_i is the observed subject's (adjusted) expenditure proposal. The table does not include observations from round 1 in each block, observations from treatment with No-Default, and observations from cases in which $v \geq v_{n-1}$. The percent of the column total is in square brackets, and the percent of the row total is in curly brackets.

Entries on the diagonal running down from left to right represent cases in which each subject's behavior is consistent with the status quo hypothesis. These entries (644 observations) account for 74.5 percent of the 864 total observations. However, a closer look reveals that 607 (out of 644 diagonal) observations come from the cases where the default expenditure E^r is larger than v. In these cases, both social preferences and status quo effects predict that adjusted proposals P_i exceed v. However, if the default is smaller than v, then the status quo effect predicts adjusted proposals to fall below v, whereas social preferences yield the opposite prediction (unless v is the largest value v_n).

Note that there appears to be an asymmetry in that a larger percentage of cases in which the default value is greater than v (the third column) is consistent with the status quo hypothesis (85.6 percent) than for the first column in which the default value is less than v (21 percent). Indeed, 65 percent (74 out of 114) of observations in column 1 are consistent with preferences for efficiency but inconsistent with the status quo hypothesis. Note also that all 57 observations in row 1 violate both selfish and social preferences.[29] While these observations constitute only 6.6 percent

[28] We exclude observations for which v equals either v_{n-1} or v_n because in those cases proposing less than v can be driven in part by altruism. We also exclude observations from the No-Default treatments because there is no salient status quo there.

[29] Selfish preferences imply that the adjusted proposal will equal v, which is not the case in column 1. Preferences for efficiency imply that the adjusted proposal should be larger than v.

(57 out of 864) of all observations, 42.1 percent (24 out of 57) of these violations are consistent with a preference for the status quo. The default is larger than the maximum benefit value v in 82 percent (709 out of 864) of the cases (see Table 4, last row), and an overwhelming percentage (86 percent) of the observations are consistent with both the status quo effect and preferences for efficiency.