Received: 28 February 2018 | Revised: 11 December 2018 | Accepted: 16 December 2018

DOI: 10.1111/jpet.12359

ORIGINAL ARTICLE

WILEY JOURNAL OF PUBLIC ECONOMIC THEORY

Impulse balance and framing effects in threshold public good games

Edward Cartwright[1] | **Anna Stepanova**[2] | **Lian Xue**[3,4]

[1]Department of Strategic Management and Marketing, De Montfort University, Leicester, UK

[2]School of Economics, Finance and Accounting, Coventry University, Coventry, UK

[3]Economic and Management School, Wuhan University, Wuhan, China

[4]Center for Behavioral and Experimental Research, Wuhan University, Wuhan, China

Correspondence
Edward Cartwright, Department of Strategic Management and Marketing, De Montfort University, Leicester, LE1 9BH, United Kingdom.
Email: edward.cartwright@dmu.ac.uk

Funding information
Engineering and Physical Sciences Research Council, Grant/Award Number: EP/P011772/1

In this paper, we revisit the evidence for framing effects in threshold public good games. Our particular focus is on why the probability of providing the public good appears to be higher in positive, give frames compared with negative, take frames. We show that the impulse balance theory can explain this effect. We also report a new experiment designed to test the predictions of the impulse balance theory. The results of the experiment fit well, both in quantitative and qualitative terms, with our predictions.

1 | INTRODUCTION

A large literature has looked at framing effects in social dilemmas (e.g., Andreoni, 1995; Cookson, 2000; Cox, 2015; Cubitt, Drouvelis, & Gächter, 2011; Dufwenberg, Gächter, & Hennig-Schmidt, 2011; Fosgaard, Hansen, & Wengström, 2014, 2017; Khadjavi & Lange, 2011; Park, 2000; Van Dijk & Wilke, 2000). For the most part, the literature has focused on linear public good games. Sonnemans, Schram, and Offerman (1998) provide a rare exception in looking at a threshold public good game. They show that the success rate in providing the public good is significantly higher if the game is framed in terms of provision of a public good as compared with the prevention of a public bad.[1] In this paper, we revisit that result and provide some new experimental evidence.

The main novelty in our approach is to apply the impulse balance theory. The impulse balance theory, building on learning direction theory, says that people have a tendency to move in the direction of ex-post rationality when learning from experience (Chmura, Goerg, & Selten, 2012, 2014; Selten, 2004; Selten, Abbink, & Cox, 2005; Selten & Buchta,

[1]Here, framing effect means a game is *presented* in two different ways—either provision of a public good or prevention of a public bad. An alternative line of work (see Buchholz, Cornes, & Rubbelke, 2018) looks at situations in which output is a public good for some and a public bad for others.

1999; Selten & Chmura, 2008). The theory has been applied in a number of contexts, including auctions (Ockenfels & Selten, 2005; Pezanis-Christou & Wu, 2018), duopoly games (Goerg & Selten, 2009), the newsvendor game (Ockenfels & Selten, 2014), weakest link games (Goerg, Neugebauer, & Sadrieh, 2016), and threshold public good games (Cartwright & Stepanova, 2017). We demonstrate here that the theory offers one explanation for the framing effect observed by Sonnemans et al. (1998).

To put our approach in context, it is important to recognize that the framing effect observed by Sonnemans et al. (1998) was a dynamic one. Initially, the success rate in both frames was similar, but in the public bad frame, it subsequently declined, whereas in the public good frame, it did not. The key to explaining the framing effect is, therefore, to explain why the success rate declined in the public bad frame. Sonnemans et al. (1998) proposed a theoretical model that can capture this decline and argued, through simulations, that the success rate in the public bad frame would typically converge to zero. The impulse balance theory is ideally suited to study such dynamic effects (Cartwright & Stepanova, 2017; Goerg et al., 2016). And interestingly, the model proposed by Sonnemans et al. (1998) is very similar in spirit to the impulse balance theory (which had not been developed at that time).

We demonstrate that the impulse balance theory allows us to make sharp predictions on when the framing effect will be large and small.[2] It can also predict when success rates will converge to zero. The results of Sonnemans et al. (1998) are consistent with our predictions but it is desirable to do further tests because the effect they observed, while statistically significant, was not particularly large. Moreover, Iturbe-Ormaetxe, Ponti, Tomás, and Ubeda (2011) and Kotani, Managi, and Tanaka (2008) obtain contrasting results. This can be explained by the fact they considered, respectively, stranger matching and a constantly changing environment, neither of which allows time for learning (which is critical to the impulse balance theory).[3] Even so, it is desirable to have more evidence. We, therefore, designed a new experiment in which we consider a setting where the framing effect is predicted to be small and one where it is predicted to be large.

Our experimental results (in both the United Kingdom and China) are consistent with the theoretical predictions. First, similar to Sonnemans et al. (1998), we found that the success rate in providing the public good was significantly higher with the public good than public bad framing. Second, we found, as predicted, that the size of the framing effect was attenuated by changes in the return from the public good. Specifically, when the return from the public good was "large," the framing effect was relatively small. But, when the return from the public good was "small," the framing effect was relatively large. As well as endorsing the predictions of the impulse balance theory, this result also shows that the size of framing effects can systematically depend on incentives in the game.

The paper proceeds as follows. In Section 2, we work through the public good frame, and in Section 3, we work through the public bad frame. Section 4 contains our main theoretical results. Section 5 describes our experiment design and Section 6 the results. Section 7 concludes.

2 | PUBLIC GOOD FRAME

We begin with the "standard" binary choice threshold public good frame. Following the approach of Sonnemans et al. (1998), we will call it the *public good provision frame*. It is characterized by players being asked how much they want to *give toward* a public good and the *positive externality* of contributing is emphasized (Cartwright, 2016). Specifically, there exists a set of players $i = \{1, \ldots, n\}$. Each player has to decide whether to contribute E units of a private good toward a public good. Players make choices simultaneously and independently. Let $a_i \in \{0, 1\}$ denote the choice of player i, where $a_i = 0$ indicates not contribute and $a_i = 1$ indicates contribute. Contribution profile $a = (a_1, \ldots, a_n)$ details the action of each player.

[2]The only occasion in which the framing effect does not exist is if success rates are predicted to be zero with both frames.

[3]Results more consistent with Sonnemans et al. (1998) are obtained by Bougherara, Denant-Boemont, and Masclet (2011). They found that contributions are significantly higher in a creating (public good) frame than a maintaining (public bad) frame when there is a provision point mechanism. Moreover, the effect is largely a dynamic one. The key difference is that Bougherara et al. (2011) consider a setting with continuous choice, whereas we focus on a setting with binary choice.

Given contribution profile a, let

$$c(a) = \sum_{i=1}^{n} a_i$$

denote the number of players who contribute. There exists an exogenously fixed threshold level $1 < t < n$. The payoff function of player i is given by

$$u_i(a) = \begin{cases} E(1 - a_i) + V & \text{if } c(a) \geq t \\ E(1 - a_i) & \text{otherwise} \end{cases},$$

where $V > E$ is the private return from the public good. Note that because $V > E$, it is efficient to provide the public good. Moreover, it is in the interests of a player to contribute if $t - 1$ others will contribute. The set of the Nash equilibrium for the game is characterized by Palfrey and Rosenthal (1984).

2.1 | Impulse balance equilibrium

To apply the learning direction theory and the impulse balance theory, we need to distinguish different experience conditions (Selten, 2004). In doing so, we take as given a contribution profile a and player $i \in N$. We distinguish three key experience conditions.[4]

Wasted contribution: Player i experiences the wasted contribution experience condition if $c(a) \leq t - 1$ and $a_i = 1$. The realized payoff of player i was 0 while it would have been E if he had not contributed. This implies that player i has a downward impulse of size E.

Lost opportunity: Player i experiences the lost opportunity experience condition if $c(a) = t - 1$ and $a_i = 0$. His realized payoff was E while it would have been V if he had contributed. This implies that player i has an upward impulse of size $V - E$.

Overcontribution: Player i experiences the overcontribution (OC) experience condition if $c(a) > t$ and $a_i = 1$. His realized payoff was V while it would have been $E + V$ if he did not contribute. Player i has a downward impulse of size E.

To calculate expected impulse suppose that each player independently chooses to contribute with probability p. The expected upward impulse of player $i = \overline{1, n}$ comes from the lost opportunity condition and is given by

$$\begin{aligned} IG^{+}(p) &= \Pr(i \text{ does not contribute}) \Pr(t - 1 \text{ others contribute})(V - E) \\ &= (1 - p)\binom{n - 1}{t - 1} p^{t-1}(1 - p)^{n-t}(V - E). \end{aligned}$$

The expected downward impulse of player i comes from the wasted contribution and OC experience conditions. From the wasted contribution, it is

$$\begin{aligned} IG^{-}_{WC}(p) &= \Pr(i \text{ contributes}) \Pr(t - 2 \text{ or less others contribute}) E \\ &= pE \sum_{y=0}^{t-2} \binom{n - 1}{y} p^{y}(1 - p)^{n-1-y}. \end{aligned}$$

And from the OC condition, it is

$$\begin{aligned} IG^{-}_{OC}(p) &= \Pr(i \text{ contributes}) \Pr(t \text{ or more others contribute}) E \\ &= pE \sum_{y=t}^{n-1} \binom{n - 1}{y} p^{y}(1 - p)^{n-1-y}. \end{aligned}$$

[4] In all other conditions, there is no impulse to change contribution.

So, the expected downward impulse of player i is

$$IG^{-}(p) = IG^{-}_{WC}(p) + IG^{-}_{OC}(p) = pE\left(1 - \binom{n-1}{t-1}p^{t-1}(1-p)^{n-t}\right).$$

An impulse balance equilibrium is a value for p such that $IG^{+}(p) = \lambda IG^{-}(p)$, where λ is an exogenously given weight on the downward impulse (Ockenfels & Selten, 2005). The interpretation of λ will be a recurring theme in the remainder of the paper. Here, we merely note that a value of $\lambda < 1$ implies a psychological *impulse toward contributing* (because downward impulse counts for less). The equilibrium condition equates to

$$\binom{n-1}{t-1}p^{t-1}(1-p)^{n-t}\left((1-p)\left(\frac{V}{E}-1\right)+\lambda p\right) = \lambda p. \tag{1}$$

We can distinguish stable and unstable impulse balance equilibria. We say that an equilibrium p^* is stable if $IG^{+}(p) > \lambda IG^{-}(p)$ for $p \in (p^* - \varepsilon, p^*)$ and $IG^{+}(p) < \lambda IG^{-}(p)$ for $p \in (p^*, p^* + \varepsilon)$ for some $\varepsilon > 0$.[5] To illustrate, Sonnemans et al. (1998) consider parameters $n = 5$, $t = 3$, $E = 60$, $V = 185$. Figure 1 plots the expected upward and downward impulses as a function of p. When $\lambda = 1$, we obtain three impulse balance equilibria: (a) a stable equilibrium $p^* = 0$, (b) unstable equilibrium $p^* = 0.11$, and (c) stable equilibrium $p^* = 0.55$. When $\lambda = 0.5$, we obtain equilibria: (a) $p^* = 0$, (b) $p^* = 0.05$, and (c) $p^* = 0.65$.

3 | PUBLIC BAD PREVENTION

We next consider, following the terminology of Sonnemans et al. (1998), a *public bad prevention* frame. This frame is characterized by players being asked how much they want to *take from* a public good and the *negative externality* of taking is emphasized (Cartwright, 2016). Although the public bad frame is equivalent to the public good frame, it is useful to start afresh and demonstrate equivalence, rather than assume it.

In the public bad prevention frame, each player has to decide whether to withrdaw E units of a private good from a public good. Let $w_i \in \{0, 1\}$ denote the choice of player i, where $w_i = 0$ indicates not withdraw and $w_i = 1$ indicates withdraw. Withdrawal profile $w = (w_1, \ldots, w_n)$ details the action of each player. With each withdrawal profile w, we associate a unique contribution profile a, and vice versa, using relation $w_i + a_i = 1$. Specifically, if $w_i + a_i = 1$ for all $i = \overline{1, n}$, then we write $w \leftrightarrow a$.

Given a withdrawal profile w, let

$$d(w) = \sum_{i=1}^{n} w_i$$

denote the number of players who withdraw. Note that $d(w) + c(a) = n$ if $w \leftrightarrow a$. The threshold of $c(a) \geq t$ can be reinterpreted as $d(w) \leq n - t$. The payoff function of player i is given by

$$u_i(w) = \begin{cases} Ew_i + V & \text{if } d(w) \leq n - t \\ Ew_i & \text{otherwise} \end{cases}.$$

It is trivial that the set of the Nash equilibrium in the public bad prevention frame is equivalent to that in the public good provision frame. The same is not true, as we shall now show, for impulse balance equilibrium.

[5] If $p^* = 0$ or $p^* = 1$, the definition is amended as appropriate.

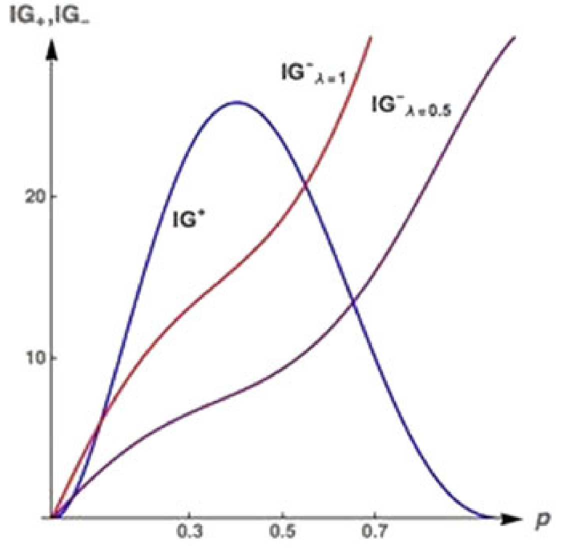

FIGURE 1 The expected upward and downward impulses as a function of p in a public good frame

3.1 | Impulse balance equilibrium

We again distinguish three experience conditions. In doing so, and this is key to our whole approach, we now think of upward as to withdraw (because it is to go from $w_i = 0$ to $w_i = 1$). Take as given a withdrawal profile w and player $i \in N$.

Wasted contribution: Player i experiences the wasted contribution experience condition if $d(w) > n - t$ and $w_i = 0$. The realized payoff of player i was 0 while it would have been E if he had withdrawn. There is an upward impulse of strength E.

Lost opportunity: Player i experiences the lost opportunity experience condition if $d(w) = n - t + 1$ and $w_i = 1$. The realized payoff was E while it would have been V if he had not withdrawn. There is a downward impulse of strength $V - E$.

Overcontribution: Player i experiences the OC experience condition if $d(w) < n - t$ and $w_i = 0$. The realized payoff was V and would have been $E + V$ if he had withdrawn. There is an upward impulse of strength E.

Retaining a homogeneity assumption, let q denote the probability that each player withdraws. The expected downward impulse of a player i comes from the lost opportunity experience condition,

$$\text{IB}^{-}(q) = q\binom{n-1}{n-t}(1-q)^{t-1}q^{n-t}(V-E).$$

The expected upward impulse of player i comes from the wasted contribution and OC experience conditions,

$$\text{IB}_{B}^{+}(q) = \text{IB}_{WC}^{+}(q) + \text{IB}_{OC}^{+}(q) = (1-q)E\ \left(1 - \binom{n-1}{n-t} q^{n-t}(1-q)^{t-1}\right).$$

An impulse balance equilibrium is a value for q such that $\text{IB}^{+}(q) = \lambda \text{IB}^{-}(q)$, where λ is the exogenously given weight on the downward impulse. This equates to condition

$$1 - q = \binom{n-1}{n-t}(1-q)^{t-1}q^{n-t}\left(\lambda q\left(\frac{V}{E} - 1\right) + 1 - q\right). \tag{2}$$

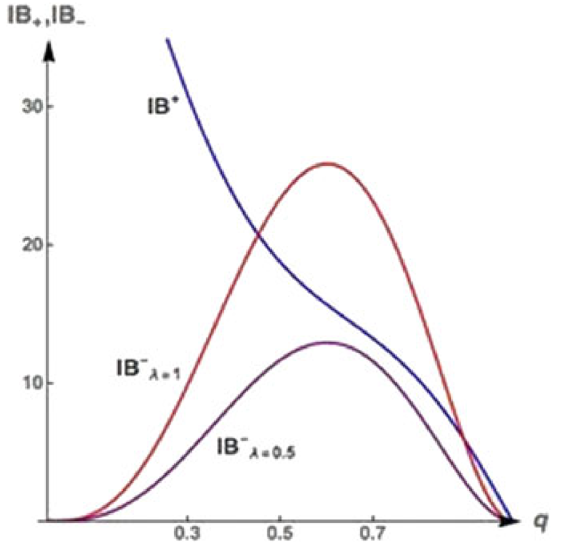

FIGURE 2 The expected upward and downward impulses as a function of q in a public bad frame

Figure 2 plots the upward and downward impulse when $n = 5$, $t = 3$, $E = 60$, $V = 185$. When $\lambda = 1$, we obtain three impulse balance equilibria: (a) a stable equilibrium $q^* = 0.45$, (b) unstable equilibrium $q^* = 0.89$, and (c) stable equilibrium $q^* = 1$. When $\lambda = 0.5$, we obtain a unique equilibrium of $q^* = 1$.

4 | FRAMING EFFECT

In this section, we explore the connection between the public good and public bad frames. Let PG*(λ) denote the set of impulse balance equilibria as a function of λ, found by solving condition (1). In terms of the public bad frame, if we set $1 - q = p$ and use $\binom{n-1}{t-1} = \binom{n-1}{n-t}$, then condition (2) can be rewritten as

$$\binom{n-1}{t-1} p^{t-1}(1-p)^{n-t}\left((1-p)\left(\frac{V}{E}-1\right)+\frac{p}{\lambda}\right)=\frac{p}{\lambda}, \tag{3}$$

where p is the probability of contribution. Let PB*(λ) denote the set of impulse balance equilibria in the public bad frame.

Comparison of conditions (1) and (3) yields the following:

$$\mathrm{PB}^*\left(\frac{1}{\lambda}\right) = \mathrm{PG}^*(\lambda). \tag{4}$$

This allows us to directly compare the set of equilibria in the public good and public bad frames and give insight into a predicted framing effect. If $\lambda = 1$, we would expect no framing effect on contributions. Previous estimates of λ are, however, significantly less than one (Cartwright & Stepanova, 2017; Ockenfels & Selten, 2005). This suggests that we could expect a framing effect. To explore this more, we need to look at PG*(λ) and PB*(λ) for different values of λ (focusing on the case of $\lambda < 1$). Our first proposition concerns the multiplicity of equilibria.

Proposition 1. *If $\lambda > \bar{\lambda}$, where*

$$\bar{\lambda} = \left(\frac{V}{E} - 1\right) \max_{p \in (0,1)} \left\{ \frac{\binom{n-1}{t-1} p^{t-2} (1-p)^{n-t+1}}{1 - \binom{n-1}{t-1} p^{t-1} (1-p)^{n-t}} \right\},$$

then $PG^*(\lambda) = \{0\}$. *Otherwise,* $|PG^*(\lambda)| > 1$. *Equivalently, if* $\lambda < \frac{1}{\bar{\lambda}}$, *then* $PB^*(\lambda) = \{0\}$. *Otherwise,* $|PB^*(\lambda)| > 1$.

Proof. It is easy to verify that $p = 0$ is an impulse balance equilibrium for any value of λ. If $p > 0$, then we can rewrite condition (1) as

$$\lambda = \left(\frac{1-p}{p}\right)\left(\frac{\binom{n-1}{t-1} p^{t-1} (1-p)^{n-t}}{1 - \binom{n-1}{t-1} p^{t-1} (1-p)^{n-t}}\right)\left(\frac{V}{E} - 1\right).$$

This is feasible if and only if $\lambda \leq \bar{\lambda}$ giving the desired properties of PG*. If $\lambda < \frac{1}{\bar{\lambda}}$, then $\frac{1}{\lambda} > \bar{\lambda}$. This coupled with $PB^*(\lambda) = PG^*(\frac{1}{\lambda})$ gives the desired properties of PB*. ■

To illustrate Proposition 1, consider the case $n = 5$, $t = 3$, $E = 60$, $V = 185$. We obtain

$$\bar{\lambda} = \frac{125}{60} \max_{p \in (0,1)} \left\{ \frac{6p(1-p)^3}{1 - 6p^2(1-p)^2} \right\} = 1.76.$$

To explain what this means, suppose $\lambda = 0.25$. In the public good frame, we clearly have $\lambda < \bar{\lambda}$ implying multiple equilibria. Indeed $PG^*(0.25) = \{0, 0.02, 0.72\}$. For the corresponding public bad frame, we note that $\lambda < \frac{1}{\bar{\lambda}} = 0.56$. So, Proposition 1 implies that there is a unique impulse balance equilibrium $PB^*(0.25) = \{0\}$.

More generally, Proposition 1 shows that if $\lambda < \bar{\lambda} < \frac{1}{\lambda}$, there are multiple equilibria in the public good frame but a unique equilibrium of zero contributions in the public bad frame. In the case of multiple equilibria, we cannot say to which equilibrium a particular group may converge. It does not seem unreasonable, however, to assume that when there are multiple equilibria, at least some groups will converge on an equilibrium with positive contributions. This leads to the following hypothesis.

Hypothesis 1. *If* $1 < \bar{\lambda} < \frac{1}{\lambda}$ *or* $\lambda < \bar{\lambda} < 1$, *then the probability of the public good being provided is higher in a public good than public bad frame.*[6]

Our second hypothesis details another consequence of Proposition 1. If $\bar{\lambda} < \lambda$, then zero contributions is the unique equilibrium in the public good frame. If, in addition, $\lambda < 1$, then $\lambda < \frac{1}{\bar{\lambda}}$ and so zero contributions is the unique equilibrium in the public bad frame.

Hypothesis 2. *If* $\bar{\lambda} < \lambda < 1$ (*or* $1 < \lambda < \frac{1}{\bar{\lambda}}$), *then the probability of the public good being provided is zero in the public good and public bad frames.*

We now consider the cases not covered by Hypotheses 1 and 2. If $\lambda < 1$, then hypotheses 1 and 2 cover all cases except $\lambda < 1 < \frac{1}{\lambda} < \bar{\lambda}$. In this case, there are multiple impulse balance equilibria in both the public good and public bad frames. Let

$$p^G(\lambda) = \max\{p \in PG^*(\lambda)\}$$

[6] We can also say that If $1 < \bar{\lambda} < \lambda$ (or $\frac{1}{\lambda} < \bar{\lambda} < 1$) the probability of the public good being provided is higher in a public bad than a public good frame. Each of these conditions can only hold, however, if $\lambda > 1$.

TABLE 1 Values of $p^G(\lambda)$ and $p^B(\lambda)$ and the probability of the public good are provided for different values of λ

	Public good		Public bad	
λ	$p^G(\lambda)$	$\Pr(c \geq t)$	$p^B(\lambda)$	$\Pr(w \leq n - t)$
0.25	0.72	87	0	0
0.5	0.65	77	0	0
0.75	0.60	68	0.48	47
1	0.55	59	0.55	59

be the equilibrium in the public good frame with the largest probability of contribution. We now show that $p^G(\lambda)$ is a decreasing function of λ.

Proposition 2. *If $\lambda_1 < \lambda_2 < \bar{\lambda}$, then $p^G(\lambda_1) > p^G(\lambda_2)$.*

Proof. Both $I^+(p)$ and $I^-(p)$ are continuous functions. So, the function $\Delta(\lambda_1, p) = \lambda_1 I^-(p) - I^+(p)$ is also a continuous function. Consider $\lambda_1 < \bar{\lambda}$. By construction $\Delta(\lambda_1, p^G(\lambda_1)) = 0$. Setting $p = 1$, it is simple to see that $I^+(1) = 0 < I^-(1)$ implying that $\Delta(\lambda_1, 1) > 0$. Moreover, by construction, there is no $p \in (p^G(\lambda_1), 1)$ such that $\Delta(\lambda_1, p) = 0$. So, (observation 1) $\Delta(\lambda_1, p) > 0$ for any $p > p^G(\lambda_1)$. Consider now $\lambda_2 > \lambda_1$. It is simple to show that for $p > 0$, it must be that $I^-(p) > 0$. Hence, (observation 2) $\Delta(\lambda_2, p) = \lambda_2 I^-(p) - I^+(p) > \lambda_1 I^-(p) - I^+(p) = \Delta(\lambda_1, p)$ for all $p > 0$. Combining observations 1 and 2, we get $\Delta(\lambda_2, p) > 0$ for any $p \geq p^G(\lambda_1)$. Hence, $p^G(\lambda_2) < p^G(\lambda_1)$. ■

To put Proposition 2 in context, consider again the example $n = 5$, $t = 3$, $E = 60$, $V = 185$. The left-hand side of Table 1 details $p^G(\lambda)$ for four different values of λ. We see that the equilibrium probability of contributing falls within range 0.72–0.55 for λ between 0.25 and 1. The corresponding probability of the public good being provided ($\Pr(c \geq t)$) ranges between 87 and 59%.

Proposition 2 focuses on the public good frame. It can, however, easily be applied to the public bad frame. Consider the case $\lambda < \frac{1}{\bar{\lambda}}$. Then there are multiple impulse balance equilibria in the public bad frame. Let

$$p^B(\lambda) = \max\{p \in PB^*(\lambda)\}.$$

Relation (4) implies that $p^B(\lambda) = p^G(\frac{1}{\lambda})$. An immediate corollary of Proposition 2 is, therefore, that $p^B(\lambda)$ is an increasing function of λ. This is apparent in Table 1, where $p^B(\lambda)$ varies between 0 and 0.55 for λ between 0.25 and 1.

More generally, consider what happens when $\frac{1}{\bar{\lambda}} < \lambda < \bar{\lambda}$. Then there are multiple equilibria in both the public good and public bad frames. Again, with multiple equilibria we cannot say to which equilibrium a particular group will converge. Let us assume, however, that a group converges on either the $p = 0$ equilibrium or the respective p^B or p^G equilibrium.[7] Moreover, assume that a group is as likely to converge on the $p = 0$ equilibrium in the public good as public bad frame.[8] Then we obtain a third hypothesis (which nests Hypothesis 1).

Hypothesis 3. *If $\lambda < 1 < \bar{\lambda}$, then the probability of the public good being provided is higher in a public good than public bad frame.*

[7]There are almost always three impulse balance equilibria. The middle of these is unstable in the sense that $I^-(p) > I^+(p)$ for $p < p^*$ and $I^-(p) > I^+(p)$ for $p > p^*$, where p^* is the equilibrium. So, this assumption is very mild.

[8]If $p^G > p^B$, then one could argue that the basin of attraction of the p^G equilibrium is larger than that of the p^B equilibrium. To be more precise, define the basin of attraction of an equilibrium p^* as the values of p such that the net impulse $I^+(p) - I^-(p)$ is in the direction of the equilibrium. For instance, in the example when $\lambda = 0.75$, the basic of attraction of the p^G equilibrium is $p \in (0.07, 1]$ and that of the p^B equilibrium is (0.16, 1].

It is a simple matter to extend Hypotheses 1–3 to other cases such as $\bar{\lambda}$, $\lambda > 1$. We, though, have focused on the cases that seem most relevant.

4.1 | Interpretation and discussion

In the interpretation of our results, consider first the public good frame. If $\lambda < 1$, then a player has an impulse to contribute, suggesting a stronger "psychological" desire to contribute than to not contribute. Technically speaking, this means that to obtain an equilibrium we need the strength of downward impulse to be compensatingly larger than that of the upward impulse. For example, setting $\lambda = 0.25$, we obtain $p^G(0.25) = 0.72$ which corresponds to an expected upward impulse of $I^+(0.72) = 8.26$ and expected downward impulse of $I^-(0.72) = 33.04$. A stronger desire to contribute than not is consistent with the approach of Rapoport (1987) and the idea of a warm glow from giving (Andreoni, 1995). Sonnemans et al. (1998) also find that in the public good frame, satisfaction is lowest in the lost opportunity (or what they call critical noncontribution) experience condition. This is also consistent with an impulse to contribute.

Consider now the public bad frame. Our analysis is based on the notion that the change in frame leaves λ unchanged but flips the meaning of upward and downward. So now the impulse is to withdraw, suggesting a stronger desire to withdraw (or not contribute) than to not withdraw (or contribute). To put this in context, again set $\lambda = 0.25$ and $p = 0.72$. Then the expected upward impulse is $I^+(0.72) = 33.04$ and expected downward impulse is $I^-(0.72) = 8.26$. This is not consistent with equilibrium and suggests a tendency toward everyone withdrawing. One possible explanation for an impulse to withdraw is that a person feels an entitlement to withdraw. For instance, Sonnemans et al. (1998) find that in the public bad frame, satisfaction is lowest in the wasted contribution (or futile nonwithdrawal) experience condition. This suggests that subjects were "frustrated" if they missed an opportunity to withdraw.

Let us clarify that we view λ as a psychological parameter, which is *not* influenced by the frame. Framing has the effect of changing the notion of upward impulse so as to influence behavior, taking λ as given. In particular, the impulse changes from contribute to withdraw. More generally, however, one could envisage that the change in frame may also change λ. For instance, the notion that contributing can give warm glow would seem relatively straightforward, but the idea that withdrawal can give a similarly sized psychological boost may seem less intuitive. This might suggest that, say, $\lambda < 1$ in the public good frame (because of a warm glow from contributing) while $\lambda = 1$ in the public bad frame (because of an absence of warm glow). If a change in frame influences λ, then the story is "less neat." The framing effect picked up in Hypotheses 1–3 would, however, still exist as long as we do not see an extreme jump in λ.

4.2 | Loss aversion and impulse

It is not our objective in this paper to consider alternatives to impulse balance. Our focus is to see whether impulse balance can reliably predict framing effects. It is, however, worth reflecting on a slightly different approach to formulating impulse balance equilibrium introduced by Selten and Chmura (2008), which we will call SC impulse balance equilibrium. This approach does away with λ but does add loss aversion into the mix. To see how this works, consider first the public good provision frame. A natural reference point in this frame is to not contribute and keep the endowment E (Sonnemans et al., 1998). Player i's payoff in this case is E, and so a payoff below E is interpreted as a loss.

The only scenario in which player i earns less than E is the wasted contribution experience condition. This condition, because of loss avoidance, is therefore given a *higher* weight. Specifically, SC impulse balance equilibrium is obtained by setting $\gamma IG^-_{WC}(p) + IG^-_{OC} = IG^+(p)$, where γ is a parameter that measures loss avoidance (and was set to 2 by Selten & Chmura, 2008). Recall that in our setting, the equilibrium is obtained by setting $\lambda IG^-_{WC}(p) + \lambda IG^-_{OC} = IG^+(p)$. So, the SC equilibrium will differ from the equilibrium we are using (unless

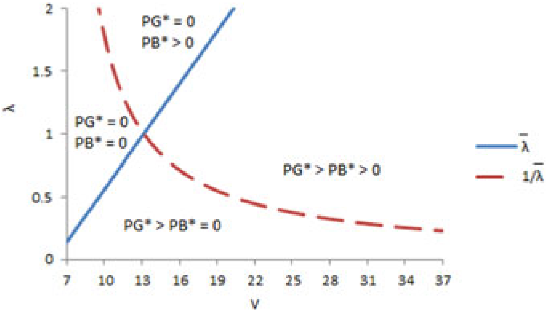

FIGURE 3 Predicted framing effects as a function of V and λ

$\lambda = \gamma = 1$). If $\lambda = \gamma$, then the differences will typically be small.[9] If, however, γ measures loss aversion, then we should have $\gamma > 1$, and yet we have been suggesting $\lambda < 1$. So, the two equilibria will diverge. Indeed, when $\gamma = 2$, the only SC equilibrium is $p = 0$. Essentially, the "fear" of losing E deters people from contributing. This is hard to reconcile with observed contributions.

The key question for us is whether SC impulse balance equilibrium can capture a framing effect. Here we encounter the further problem of determining an appropriate reference point. If we adopt the same reference point in the public bad frame as in the public good frame then we should observe no framing effect (because the equilibrium condition is identical in the public good and public bad frame). If, as discussed by Sonnemans et al. (1998), we put E as the reference point in the public good frame and 0 as the reference point in the public bad frame, we should obtain higher, and not lower, contributions in the public bad frame. A further alternative is to use V as a reference point in the public bad frame to capture the idea of taking from an existing public good. This gives a result in the desired direction but only at the cost of seeming too ad hoc.

In short, as Van Dijk and Wilke (1995) highlighted, the appropriate reference point in threshold public good games is open to interpretation. This is why we prefer to use impulse balance equilibrium rather than the SC equilibrium. But the arguably more important thing to take from this discussion is that λ is distinct from loss aversion. In particular, λ is picking up a "bias" toward being influenced by upward impulse more than downward impulse. Flipping the interpretation of upward and downward, thus, causes the framing effect. Adding loss aversion into the mix may further enrich the analysis. But our results are not driven by a change in the reference point or loss aversion.

5 | EXPERIMENTAL DESIGN AND RESULTS

Our objective with the experiment is to test the predictions of Hypothesis 3 and Proposition 1. Setting $n = 5$ and $t = 3$, we get that $\bar{\lambda} \approx 0.84(v/E - 1)$. If, therefore, we fix $E = 6$ and vary V, then we obtain the values for $\bar{\lambda}$ depicted in Figure 3. Prior studies suggest that we would expect a value of λ between around 0.25 and 1 (Cartwright & Stepanova, 2017; Ockenfels & Selten, 2005). This leaves a region for smaller values of V where we expect the success rate to be positive with a public good frame and zero with a public bad frame.[10] There is then a region for larger values of V where we expect the success rate to be positive in both the public good and public bad frame. Note that Sonnemans et al. (1998), after rescaling, had $V = 18.5$ and so were seemingly on the boundary where we might expect a positive success rate in the public bad frame depending on λ.

[9]For instance, in the example $n = 5$, $t = 3$, $E = 60$, $V = 185$, the SC equilibria, when $\gamma = 0.5$, are $p = 0, 0.45, 0.58$ compared with impulse balance equilibria of $p = 0, 0.05, 0.65$. Recall that the 0 and 0.58 versus 0.65 equilibria are our primary focus.

[10]There is a region for V around 6 where we expect the success rate to be zero in the public good frame. Results from a separate study of ours (yet to be published) confirm this result.

TABLE 2 Value of $P^G(0.5)$ and $P^B(0.5)$ for different values of V, and in brackets, the probability of the public good are provided

V	Public good	Public bad
31	0.72 (0.87)	0.55 (0.59)
13	0.57 (0.63)	0 (0)

We used a 2 × 2 design in which we vary the frame—public good or public bad—and the return from the public good —$V = 13$ or 31.[11] Our basic hypotheses are that there should be a large framing effect when $V = 13$ and a smaller framing effect when $V = 31$. To quantify this, consider Table 2, in which we detail P^G, P^B and the corresponding probability of the public good being provided when $\lambda = 0.5$. When $V = 13$, we are predicting a large framing effect. When $V = 31$, we predict a smaller but still significant framing effect. This provides straightforward testable hypotheses.

Each session was organized as follows. Subjects were given instructions for either the public good or public bad frame. The instructions were written for a specific value of V, either 13 or 31. The instructions in the public good frame talked of the *positive* benefit to the group if group members *gave* tokens. By contrast, the instructions in the public bad frame talked of the *negative* consequences for the group if group members *withdrew* tokens. Subjects were then assigned to groups of five and played the relevant game in fixed groups for 20 rounds. After 20 rounds, subjects were given a second set of instructions in which the value of V was changed from 13 to 31 or vice versa. They were then reassigned to new groups of five and played the relevant game in fixed groups for a further 20 rounds. Note that the frame remained the same over the 40 rounds. This design allows a between-subject comparison of frame and a between- and within-subject comparison on the effect of V.

After participants had finished the 40 rounds of choices, they completed a questionnaire. Other than standard demographic questions, such as gender, it included a number of questions directly related to the task. Specifically, subjects were first asked to indicate on an 11-point Likert scale how much they agreed with the statements, "I wanted to do the best for the group" and "It was important the outcome was fair to all members of the group." They were then given six outcomes and asked to rate on an 11-point scale how satisfied they would feel with each of the outcomes. The six outcomes included an example of wasted contribution, lost opportunity, and OC as well as the case where exactly three group members contribute.

The experiment was run using z-Tree (Fischbacher, 2007). We initially ran four sessions corresponding to the four conditions. So, there were two sessions with a public good frame, one with $V = 31$ in the first 20 rounds and $V = 13$ in the last 20 rounds and one with $V = 13$ in the first 20 rounds and $V = 31$ in the last 20 rounds. Similarly, there were two sessions with a public bad frame. There were 15 subjects (or three groups) in each session meaning a total of 60 subjects. Subsequently, following the comments of a referee, we ran further sessions to boost the number of observations. A total of 120 subjects took part in these further sessions equally split across the four conditions. This means a further 30 subjects (or six groups) per condition. We highlight that the initial sessions took place in the United Kingdom and the subsequent sessions in China, and so we shall only pool the data where analysis (not always reported here) shows no significant differences. Subjects were paid in cash at the end of the session based on the outcome of two randomly selected rounds from the experiment.

6 | EXPERIMENTAL RESULTS

We summarize our empirical findings with seven results. The first of these reaffirms the main finding of Sonnemans et al. (1998), namely that contributions are higher in the public good provision than public bad prevention frame. Note that throughout the following, we will define contributors as participants who choose to contribute in the

[11]Let us highlight that we see λ as a psychological parameter over which we do not have experimenter control.

TABLE 3 Contribution rate and success rate by framing and *V*

	Pooled			United Kingdom			China		
Value	PG	PB	Diff	PG	PB	Diff	PG	PB	Diff
$V = 31$	73.1	62.4	10.7**	64.0	52.0	12.0	77.6	67.7	9.9*
	(90.6)	(79.4)	(11.2)	(83.3)	(63.3)	(20.0)	(94.2)	(87.5)	(6.7)
$V = 13$	56.3	40.7	15.6*	55.2	36.5	18.7	51.5	39.9	11.6
	(64.7)	(42.2)	(22.5*)	(63.3)	(42.5)	(20.8)	(65.4)	(42.1)	(23.3)
Diff	16.8***	21.7***		8.8	15.5		16.1**	27.8**	
	(25.9***)	(37.2**)		(20.0)	(20.8)		(28.8***)	(45.4**)	

Note. The contribution rate is the average percentage of cooperative choices (contribute or not withdraw) across 20 rounds. The success rates are in parenthesis, which report the percentage of groups that have reached the threshold of public good provision or public bad prevention. "Diff" reports the differences in outcome variable between the two framings and between high and low values. Significance levels of Diff are derived from the Mann–Whitney test for equality of distribution across the two framings (values) using group averages. In total, we have 36 observations per condition (12 groups in the United Kingdom and 24 groups in China). In this and the following tables, the symbols represent ***$p < 0.01$, **$p < 0.05$, and *$p < 0.1$.

public good frame and opted not to withdraw in the public bad frame. Also, if three or more members of the group contribute, then we say the group was successful in providing the public good.

Result 1. *The contribution rate is significantly higher in public good provision framing than in public bad prevention framing.*

Evidence for Result 1 can be seen from Tables 3 and 4. Table 3 reports the average contribution rate over 20 rounds in each condition. We report results for the United Kingdom and China and the pooled results. It also reports, in brackets, the proportion of times the threshold was met. Note that these numbers can be directly compared with those predicted in Table 2. As expected the contribution rate and success rate are higher in the public good frame compared with the public bad frame. This holds in both the United Kingdom and China and is very similar to the effect observed by Sonnemans et al. (1998), where the contribution rate was 51.5 in the public good frame and 39.9 in the public bad frame.

The statistical significance levels reported in Table 3 are obtained from the Mann–Whitney–Wilcoxon test using group average across 20 rounds as independent observations. As you can see the framing effect only shows up as marginally significant at best. This is, though, a relatively conservative test given that we only have 36 observations in total. Table 4 provides more compelling evidence in support of Result 1. In Column (1), we report the results of a probit regression with individual contribution as the dependent variable. In Column (2) we report the results of an ordered probit with number of contributors in the group as the dependent variable. And in Column (3), we report the results of a probit regression with group success as the dependent variable. The baseline treatment is the public bad frame with a low V.

Post-regression tests from Column (1) confirm that the framing effects are significant in both $V = 31$ ($p < 0.01$) and $V = 13$ ($p = 0.02$). Table 5 details the post-regression results. In addition, we report power analysis calculations using the mean contribution rates between frames and individual averages as units of observation.[12] Note that we obtain similar results with panel regressions with random effects at the session level, and the additional robustness check is reported in the online Supporting Information.[13]

[12]The power size of the framing effect (PG > PB) is 0.758 in $V = 13$ and 0.919 in $V = 31$, which means that there is a 75.8%/91.9% probability that the results we found will not commit a Type II error.

[13]For a further robustness check, we also run panel regression with random effects with standard errors clustering at subject level and group level, and the treatment effects (framing and V) remain robust. In addition, as discussed in the later results, the framing effect remains robust with a bootstrap regression using individual average contribution rate as the unit of observation (specification 1 in Table 8).

TABLE 4 Probit regression on cooperative outcome variables (the United Kingdom and China)

Variables	(1) Contribution	(2) No. contribute	(3) Success rate
PG, *V*=31	0.856***	1.378***	1.542***
(*bl: PB, V*=13)	(0.110)	(0.121)	(0.206)
PB, *V*=31	0.556***	0.855***	1.037***
	(0.0817)	(0.0936)	(0.224)
PG, *V*=13	0.395**	0.612**	0.580***
	(0.163)	(0.262)	(0.205)
Place	0.250**	0.456***	0.315**
(*bl: UK*)	(0.108)	(0.161)	(0.154)
Order	0.0574	0.129	0.176
(*bl: V*=13 first)	(0.0899)	(0.132)	(0.143)
Observations	7,200	1,440	1,440
Pseudo *R*-squared	0.0485	0.0736	0.141

Note. Specification 1 and 3: probit regression with clustered standard errors at session level. Specification 2: ordered probit regression with clustered standard errors at session level.[14] Independent variable for Specification 1: whether participants choose to contribution in each period; Specification 2: numbers of participants within a group (out of five) that choose to contribute (or not withdraw); Specification 3: whether the group contribution pass the threshold. Robust standard errors in parentheses.

TABLE 5 Robustness check of treatment difference by framing and V (the United Kingdom and China)

	Framing			
Value	PG		PB	r
V = 31	73.1	>***	62.4	0.758
	(0.31)		(0.30)	
	∨***		∨***	
V = 13	56.3	>**	40.7	0.919
	(0.34)		(0.38)	

Note. The outcome variable is the mean contribution rate (in percentage). Standard errors are in the parenthesis, where unit observation is individual average across the 20 rounds. Significance levels are derived from the regression results from Column (1) in Table 4. *r* reports the one-sided power analysis of the means of contribution rate between the two frames with high or low values, where unit observation is individual average across 20 rounds (in total 60 observations in the United Kingdom and 120 in China).

Result 2. *Contribution rates are significantly higher with a high value of $V = 31$ than with a low value of $V = 13$.*

Support for Result 2 comes again from Tables 3 and 4. In both the public good and public bad framing, we found significantly higher contribution rates when V was high than when it was low (although the effect is more pronounced in China). Similar results are found in terms of success rate or numbers of contributors in the group. This is as predicted and consistent with the general evidence from threshold public good games (Croson & Marks, 2000). One thing we did not predict was the relatively high contribution and success rate in the public bad frame with a low V. We will come back to this later.

[14]The regression results remain robust if the standard errors are clustered at group level. In total there were 36 clusters (12 groups in the United Kingdom and 24 groups in China).

Result 3. *The public good versus bad framing effect is larger with a high value of V than with a low value.*

The evidence for this has already been discussed to some extent. When $V = 31$, the contribution rate in the public good treatment is 17.1% higher than that in the public bad treatment. When $V = 13$, the relevant difference is 38.3%. The effects of framing in each round can be seen in Figure 4. The vertical axis of the gray line (triangle symbol) represents the difference in average contribution rate between public good and public bad frame. In most rounds, you can see that the differences are positive, supporting Result 1. Also, for most rounds, the difference in $V = 13$ is above that in $V = 31$, supporting Result 3.

6.1 | Results on experience conditions

In this section, we focus more specifically on experience conditions and the tendency of individuals to change contribution. In doing so, we follow Sonnemans et al. (1998) in distinguishing between whether a contribution is futile, critical, or redundant for a particular subject. A contribution is *futile* if less than $t - 1$ others contribute; if a subject contributes when their contribution is futile then they experience the wasted contribution experience condition. A contribution is *critical* if exactly $t - 1$ others contribute; if a subject does not contribute when their contribution is critical, they experience the lost opportunity experience condition. Finally, a contribution is *redundant* if more than t others contribute; if a subject contributes when their contribution is redundant, they experience the OC condition.

In all, this gives us six conditions to consider: contribute when futile (wasted contribution), contribute when critical, contribute when redundant (OC), not contribute when futile, not contribute when critical (lost opportunity), and not contribute when redundant. Note that impulse balance only picks up three of these possibilities because in the other three there is no impulse to change.

Result 4. *In the public good frame, we see relatively more instances of the OC condition (contribute when redundant), and in the public bad frame, we see more instances of the not contribute when futile condition.*

Figure 5 plots the proportion of each condition distinguishing by treatment. We can readily see that the proportions are almost identical except for the contribute when redundant and not contribute when futile conditions. To be more specific, there is an around 16% switch from FutileNot to RedunCon as the frame changes from public bad to public good. The difference is statistically significant (Mann–Whitney test with individual average as independent observations, $p < 0.001$ for RedunCon and $p = 0.002$ for FutileNot).[15] This is consistent with an impulse to contribute in the public good frame (subjects are less eager to move away from OC and wasted contribution) and withdraw in the public bad frame (subjects avoid wasted contribution). To investigate this further, we can look at the dynamics of individual choice.

Result 5. *Participants who did not contribute in the previous round have a significantly higher probability of contributing in the current round in the public good frame.*

In Table 6, we report the results of probit regressions with the dependent variable of whether a subject contributed in round t, specification 1, whether a subject switched to contribute, specification 2, or deviated to not contribute, specification 3. Note that switching behavior happened 20.3% of the time, among which 47.3% switched to contribute. The main independent variables are the framing, V, whether the public good was successfully provided in round $t - 1$ and control variables such as China and the ordering of the sessions.

[15]The results remain robust in separate tests for the United Kingdom and China.

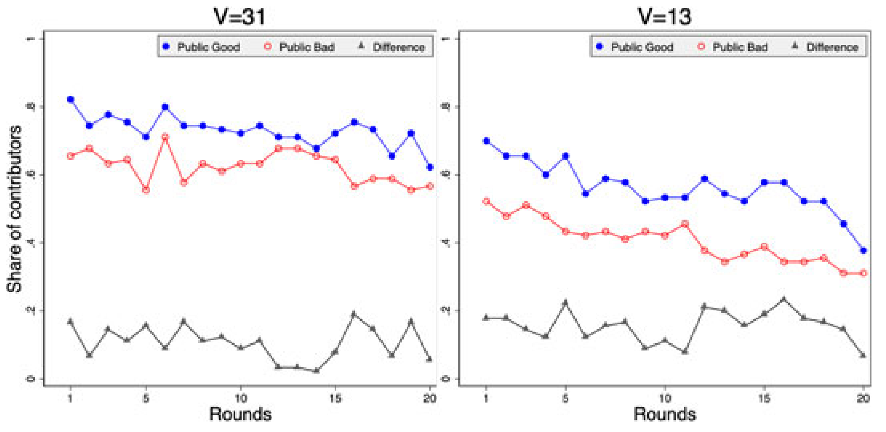

FIGURE 4 Proportion of contributors in public good and public bad frames by round for high and low value (pooled data)

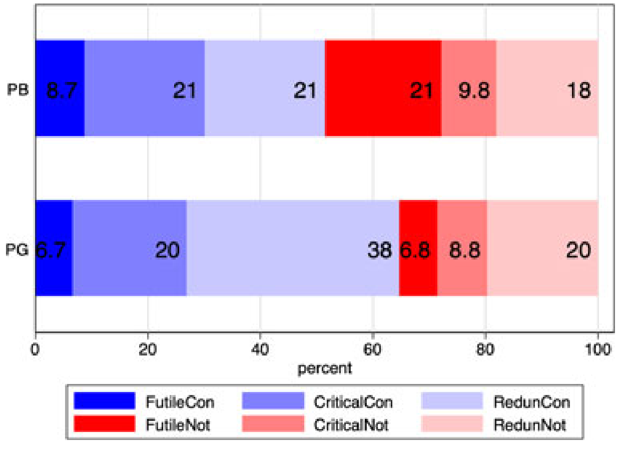

FIGURE 5 Proportion (%) of six possible outcome conditions by treatment (pooled)

Column (1) of Table 6 reaffirms Results 1 and 2 with contributions higher in the public good frame and when $V = 31$. Column (2) shows that subjects in the public good frame have a higher probability to switch to contribute if they did not contribute in the last round, ceteris paribus.[16] Indeed, if the condition in the last round condition was Critical-Not, that is, two others contribute and the participant does not contribute, then the predicted tendency to switch is significantly higher in the public good frame (33.9%) than in the public bad frame (17.9%; Fisher's exact test, $p < 0.001$). Similarly, we find a significantly higher tendency to switch to contribute in the Futile-Not condition in the public good frame compared with the public bad frame ($p < 0.001$), but not in the Redun-Not condition ($p = 0.48$). This starkly illustrates the impulse to contribute in the public good frame as compared with the public bad frame. Note that in Column (3), we also examine the tendency to deviate to not contribute, but do not find significant framing effects.

6.2 | Results on stated preferences

Recall that at the end of the experiment, subjects filled in a questionnaire in which they were asked how much they valued the group and fairness and how satisfied they were with various outcomes.

[16]Note that we obtain similar results with panel regressions with random effects at the session level (see the online Supporting Information).

TABLE 6 Probit regression on tendency to contribute

Variables	(1) Contribution	(2) Switch_con	(3) Switch_dev
PG	0.216***	0.242**	0.115
	(0.0626)	(0.117)	(0.0836)
V = 31	0.288***	0.183*	−0.0933
	(0.0574)	(0.106)	(0.0937)
lag-success	0.784***	−0.604***	0.153**
	(0.145)	(0.0851)	(0.0679)
Place (bl: *UK*)	0.193**	−0.000847	−0.0881
	(0.0808)	(0.134)	(0.101)
Ordering *(bl: V=13 first)*	0.00615	−0.0193	−0.0482
	(0.0701)	(0.111)	(0.0862)
Constant	−0.731***	−1.130***	−1.286***
	(0.173)	(0.146)	(0.110)
Observations	6,840	6,840	6,840

Note. Probit regression with clustered standard errors at session level. Dependent variable for Specification (1): Contribution; (2): Switch to contribute; and (3): Switch to not contribute.

Result 6. *Participants' attitudes toward fairness and OC and critical contribution conditions are sensitive to framing but concerns for group payoff are not.*

Table 7 summarizes the average responses of the post-experimental questionnaire. We see no significant framing effect in terms of "I wanted to do the best for the group." In China (but not the United Kingdom), we see significantly more support in the public bad frame for "It was important the outcome was fair to all members of the group." Subjects were then asked their satisfaction with possible outcomes. Overall, we see a much lower level of satisfaction in the wasted contribution condition. More relevant for us, though, are the framing effect differences. Here, the effects are mixed and differ in the United Kingdom and China. We see relatively higher levels of satisfaction in the OC condition in the public good frame and lower levels of satisfaction in the Critical-Con condition. Our interpretation of the mixed results is that although framing has an effect on attitudes these are not the main drivers of the framing effect we observe in contributions.

Result 7. *Participants who state higher preference for group payoff/fairness contribute more frequently. The correlation between contributions and preferences for group payoff and fairness is stronger in the public bad than public good frame.*

Table 8 reports a linear regression in which the dependent variable is the number of times a subject contributed over the 40 rounds. The independent variables include a dummy for the public good frame, the extent to which the subject agreed with either the statement about group payoff or fairness, and the interaction between treatment and stated preference. Here, we separate report the results for the United Kingdom and China due to different stated preference, as shown in Table 8. Consistent with Result 1, the frame is highly significant. As one would expect, subjects with a higher preference for group payoff or fairness contributed significantly more often.

Most interesting for our purposes is that a preference for group payoff or fairness was more strongly correlated with contributions in the public bad than public good frame. Consistent with our interpretation of Result 6, this

TABLE 7 Agreement and satisfaction levels from post-experiment questionnaire in the United Kingdom and China

	United Kingdom			China		
	PG	PB	p	PG	PB	p
Group	6.73	6.00	0.53	7.90	8.00	0.76
Fairness	6.70	5.43	0.26	6.38	7.78	0.01***
WC	1.83	3.30	0.19	2.10	1.95	0.67
LO	4.53	4.07	0.58	3.77	3.37	0.43
OC	7.10	6.63	0.71	7.95	6.60	0.03**
CC	6.33	6.23	0.82	6.33	7.15	0.03**
RN	7.17	8.07	0.13	7.02	6.78	0.65

Note. Participants were asked how much they value group payoffs and fairness among group members, as well as their satisfaction levels under the three key experience conditions, namely wasted opportunity (WC), lost opportunity (LO) and overcontribute (OC). In addition, participants also report their satisfaction levels in Critical-Con (CC) and Redundant-Not (RN) conditions. All questions are in the 11-point Likert scale from 0 to 10. *p*-values are reported from the Mann–Whitney U test, 60 individuals in the United Kingdom and 120 in China.

suggests that the framing effect we observe in contributions goes beyond the framing effect observed in stated attitudes. Specifically, in the public good frame, it seems as though subjects had a "bias" toward contributing irrespective of social preferences or attitudes. In the public bad frame, by contrast, we see that contributions are lower and more aligned with attitudes. We explore possible interpretations of this in the concluding discussion.

TABLE 8 Linear regression on contribution rate over 40 rounds

Variables	(1) United Kingdom	(2) United Kingdom	(3) China	(4) China
PG frame	0.223***	0.221***	0.148	0.346***
	(0.0511)	(0.0412)	(0.0916)	(0.105)
Best for group	0.0541***		0.0726***	
	(0.00194)		(0.0101)	
Best group* PG	−0.0163***		−0.00264	
	(0.00336)		(0.0118)	
Fairness		0.0408***		0.0254***
		(0.00630)		(0.00945)
Fairness* PG		−0.0178**		−0.0299***
		(0.00879)		(0.0101)
Constant	0.118**	0.221***	−0.0286	0.355***
	(0.0494)	(0.0395)	(0.0880)	(0.0809)
Observations	60	60	120	120
R-squared	0.302	0.197	0.257	0.075

Note. Linear regression with bootstrapped standard errors. To control for potential interaction effect within a session or any other session effect, standard errors are clustered at session level. The unit observation is the average contribution rate of each individual over the 40 periods in our experiment. Specifications (1) and (2) use data from the UK experiments; Specifications (3) and (4) use data from China. The results remain robust after controlling for demographic variables. Standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

7 | CONCLUSION

In this paper, we have revisited the framing effect observed by Sonnemans et al. (1998). They found that subjects were more willing to contribute to a threshold public good if it was framed in terms of public good provision rather than public bad prevention. One key contribution of our work is to demonstrate that impulse balance theory provides a framework with which to predict the size of the framing effect. We then test those predictions with a new experiment and find broad support for them. This adds to the growing evidence that impulse balance theory can help predict behavior in games (e.g., Cartwright & Stepanova, 2017; Goerg et al., 2016; Ockenfels & Selten, 2005; Selten & Chmura, 2008).

Our headline findings are relatively straightforward and provide a consistent picture. In particular, we observe a framing effect similar to that of Sonnemans et al. (1998). The size of the framing effect is also attenuated, as predicted by the impulse balance theory, by the size of return from the public good. If the return from the public good is large then the framing effect is relatively small because the success rate at providing the threshold public good is relatively high in both the public good and public bad frame. By contrast, if the return from the public good is small, then the framing effect is large because the success rate drops in the public bad frame. The experimental results we observed were also quantitatively similar to those predicted.

Although the headline results were straightforward, it is interesting to reflect in this concluding discussion on some of the more subtle findings that are open to interpretation. Particularly relevant is to look at the role played by the weight on the downward impulse. Our approach is to assume that a change in frame flips the notion of upward impulse from that of contributing in the public good frame to withdrawing in the public bad frame. This makes a difference because the weight on the downward impulse, λ, is assumed to be less than one meaning that upward impulses "count for more." So, in the public good frame, there is a tendency toward contributing, whereas in the public bad frame, there is a tendency toward withdrawing.

A tendency toward contributing in the public good frame is consistent with the notion of warm glow (Andreoni, 1995; see also Rapoport, 1987), which may be related to group identity (Chakravarty & Fonseca, 2017). And our experimental results in the public good frame are close to those predicted, when $\lambda = 0.5$. So, the approach we are taking as regards the public good frame would seem relatively uncontroversial. The notion that group members have a tendency to withdraw in the public bad frame would seem more open to question. This posits that not only is there an absence of warm glow but an additional force pushing toward withdrawal. This may be the case if people feel an entitlement to withdraw. But our results cast some doubt on the strength of this force. In particular, we predicted a zero contribution rate for a low return from the public good and yet the observed contribution was rate was a sustained 41%.

The evidence for sustained contributions in the public bad frame suggests that the impulse to withdraw was not as strong as the impulse to contribute. This, in turn, would mean that the frame changes the weight on the downward impulse. Indeed, our results are consistent with a value of λ around 0.5 in the public good frame and around 1 in the public bad frame. This picks up the impulse to contribute in the public good frame, from warm glow or similar, but no strong impulse to withdraw in the public bad frame. Such application of the impulse balance theory could be criticized for the somewhat arbitrary changes in λ. Let us highlight, therefore, the main counter-argument to this criticism—namely, that we were able to make clear testable predictions without arbitrary changes in λ and find evidence in support of those hypotheses. So, the impulse balance theory still has bite.

What our approach is not able to do is unpick why a warm glow exists (and possibly only exists) in the public good frame. This could partly be due to confusion (Fosgaard et al., 2017) but that seems unlikely to be the whole story. For instance, our Results 4–7 suggest that in the public good frame, subjects are more drawn toward contribution than in the public bad frame. Also, Park (2000) and Kotani et al. (2008) find that those with an individualistic value orientation are particularly influenced by frame. This points to the possibility that the public good frame better “incentivizes” less cooperative types. Similarly, beliefs and emotions may play a part (Cubitt et al., 2011; Dufwenberg et al., 2011). The value in our approach is that, taking warm glow as given, we can make reliable predictions on the size of framing effect observed.

We finish with a potential application of our work to cybersecurity. The main source of cyber-breach within organizations are phishing attacks in which cybercriminals use sophisticated social engineering to get victims to reveal sensitive information or transfer money (Department for Digital, Culture, Media and Sport, 2018). Cybersecurity within an organization can, therefore, be viewed as a form of threshold public good game in which employees can exert effort to avoid an attack, and a significant breach is averted if and only if sufficiently many employees exert effort. Effort may, for instance, include avoiding distractions when checking emails and seeking advice on nonroutine financial requests. Framing effects can occur in the training and advice given to employees. A positive frame would focus on the positive externality of security that results from effort (e.g., “checking all nonroutine financial requests with your line manager before acting makes your company safe”). A negative frame would focus on the damage that can result from a lack of effort (e.g., “clicking on unsolicited links risks damaging the company’s computer systems”). Our research suggests that the positive frame will be most effective. Future work aims to explore this in more detail.

ACKNOWLEDGMENT

The authors want to thank EPSRC for project EP/P011772/1, on the EconoMical, PsycHologicAl and Societal Impact of RanSomware (EMPHASIS).

ORCID

Edward Cartwright http://orcid.org/0000-0003-0194-9368

REFERENCES

Andreoni, J. (1995). Warm-glow versus cold-prickle: The effects of positive and negative framing on cooperation in experiments. *The Quarterly Journal of Economics*, *110*, 1–21.

Bougherara, D., Denant-Boemont, L., & Masclet, D. (2011). Cooperation and framing effects in provision point mechanisms: Experimental evidence. *Ecological Economics*, *70*, 1200–1210. https://doi.org/10.1016/j.ecolecon.2011.01.023

Buchholz, W., Cornes, R., & Rubbelke, D. (2018). Public goods and public bads. *Journal of Public Economic Theory*, *20*, 525–540.

Cartwright, E. (2016). A comment on framing effects in linear public good games. *Journal of the Economic Science Association*, *2*(1), 73–84.

Cartwright, E., & Stepanova, A. (2017). Efficiency in a forced contribution threshold public good game. *International Journal of Game Theory*, *46*, 1–29.

Chakravarty, S., & Fonseca, M. A. (2017). Discrimination via exclusion: An experiment on group identity and club goods. *Journal of Public Economic Theory*, *19*, 244–263.

Chmura, T., Goerg, S. J., & Selten, R. (2012). Learning in experimental 2 × 2 games. *Games and Economic Behavior*, *76*, 44–73.

Chmura, T., Goerg, S. J., & Selten, R. (2014). Generalized impulse balance: An experimental test for a class of 33 games. *Review of Behavioral Economics*, *1*, 27–53.

Cookson, R. (2000). Framing effects in public goods experiments. *Experimental Economics*, *3*(1), 55–79.

Cox, C. A. (2015). Decomposing the effects of negative framing in linear public goods games. *Economics Letters*, *126*, 63–65.

Croson, R. T., & Marks, M. B. (2000). Step returns in threshold public goods: A meta-and experimental analysis. *Experimental Economics*, *2*(3), 239–259.

Cubitt, R. P., Drouvelis, M., & Gächter, S. (2011). Framing and free riding: Emotional responses and punishment in social dilemma games. *Experimental Economics*, *14*(2), 254–272.

Department for Digital, Culture, Media and Sport (2018). Cyber security breaches survey. *UK Government, Ipsos MORI and University of Portsmouth*. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/702074/Cyber_Security_Breaches_Survey_2018_-_Main_Report.pdf

Dufwenberg, M., Gächter, S., & Hennig-Schmidt, H. (2011). The framing of games and the psychology of play. *Games and Economic Behavior*, *73*, 459–478.
Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171–178.
Fosgaard, T. R., Hansen, L. G., & Wengström, E. (2014). Understanding the nature of cooperation variability. *Journal of Public Economics*, *120*, 134–143.
Fosgaard, T. R., Hansen, L. G., & Wengström, E. (2017). Framing and misperception in public good experiments. *Scandinavian Journal of Economics*, *119*, 435–456.
Goerg, S. J., Neugebauer, T., & Sadrieh, A. (2016). Impulse response dynamics in weakest link games. *German Economic Review*, *17*(3), 284–297.
Goerg, S. J., & Selten, R. (2009). Experimental investigation of stationary concepts in cyclic duopoly games. *Experimental Economics*, *12*(3), 253–271.
Iturbe-Ormaetxe, I., Ponti, G., Tomás, J., & Ubeda, L. (2011). Framing effects in public goods: Prospect Theory and experimental evidence. *Games and Economic Behavior*, *72*, 439–447.
Khadjavi, M., & Lange, A. (2011). Doing good or doing harm: Experimental evidence on giving and taking in public good games. *Experimental Economics*, *18*, 1–10.
Kotani, K., Managi, S., & Tanaka, K. (2008). Further investigations of framing effects on cooperative choices in a provision point mechanism. *Economics Bulletin*, *3*(51), 1–9.
Ockenfels, A., & Selten, R. (2005). Impulse balance equilibrium and feedback in first price auctions. *Games and Economic Behavior*, *51*, 155–170.
Ockenfels, A., & Selten, R. (2014). Impulse balance in the newsvendor game. *Games and Economic Behavior*, *86*, 237–247.
Palfrey, T. R., & Rosenthal, H. (1984). Participation and the provision of discrete public goods: A strategic analysis. *Journal of Public Economics*, *24*, 171–193.
Park, E. S. (2000). Warm-glow versus cold-prickle: A further experimental study of framing effects on free-riding. *Journal of Economic Behavior and Organization*, *43*, 405–421.
Pezanis-Christou, P., & Wu, H. (2018). A non-game-theoretic approach to bidding in first-price and all-pay auctions. Available at https://ssrn.com/abstract=3233164
Rapoport, A. (1987). Research paradigms and expected utility models for the provision of step-level public goods. *Psychological Review*, *94*(1), 74.
Selten, R. (2004). Learning direction theory and impulse balance equilibrium, *Economics lab—An intensive course in experimental economics* 133. New York, NY: Routledge.
Selten, R., Abbink, K., & Cox, R. (2005). Learning direction theory and the winner's curse. *Experimental Economics*, *8*(1), 5–20.
Selten, R., & Buchta, J. (1999). Experimental sealed bid first price auctions with directly observed bid functions. In D. V. Budescu, I. Erev, & R. Zwick (Eds.), *Human behavior and games: Essays in honor of Amnon Rapoport*. Mahwah, NJ: Lawrence Erlbaum.
Selten, R., & Chmura, T. (2008). Stationary concepts for experimental 2 × 2 games. *American Economic Review*, *98*, 938–66.
Sonnemans, J., Schram, A., & Offerman, T. (1998). Public good provision and public bad prevention: The effect of framing. *Journal of Economic Behavior and Organization*, *34*(1), 143–161.
Van Dijk, E., & Wilke, H. (1995). Coordination rules in asymmetric social dilemmas: A comparison between public good dilemmas and resource dilemmas. *Journal of Experimental Social Psychology*, *31*(1), 1–27. https://doi.org/10.1006/jesp.1995.1001
Van Dijk, E., & Wilke, H. (2000). Decision-induced focusing in social dilemmas: Give-some, keep-some, take-some, and leave-some dilemmas. *Journal of Personality and Social Psychology*, *78*, 92–104.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article. Also available at https://figshare.com/authors/Edward/_Cartwright/4438759

How to cite this article: Cartwright E, Stepanova A, Xue L. Impulse balance and framing effects in threshold public good games. *Journal of Public Economic Theory*. 2019;1–20. https://doi.org/10.1111/jpet.12359