
Impulse Response Dynamics in Weakest Link Games

Sebastian J. Goerg

Florida State University,
Tallahassee

Abdolkarim Sadrieh

University of Magdeburg

Tibor Neugebauer

University of Luxembourg

Abstract. In a recent paper, Croson *et al.* (2015) experimentally study three weakest link games with multiple symmetric equilibria. They demonstrate that static concepts based on the Nash equilibrium (including multiple Nash equilibria, quantal response equilibria, and equilibrium selection by risk and payoff dominance) cannot successfully capture the observed treatment differences. Using Reinhard Selten's impulse response dynamics, we derive a proposition that provides a theoretical ranking of contribution levels in the weakest link games. We show that the predicted ranking of treatment outcomes is fully consistent with the observed data. In addition, we demonstrate that the impulse response dynamics perform well in tracking average contributions over time. We conclude that Reinhard Selten's impulse response dynamics provide an extremely valuable behavioral approach that is not only capable of resolving the indecisiveness of static approaches in games with many equilibria, but that can also be used to track the development of choices over time in games with repeated interaction.

JEL classification: C72, C90.

Keywords: Impulse balance; impulse learning; impulse matching; multiple equilibria; coordination game; public goods; weakest link; experiment.

1. INTRODUCTION

In a recent paper, Croson *et al.* (2015) experimentally study three weakest link games with multiple symmetric equilibria. They demonstrate that static concepts based on the Nash equilibrium – including multiple Nash equilibria, quantal response equilibria (McKelvey and Palfrey, 1995), as well as equilibrium selection based on risk and payoff dominance (Harsanyi and Selten, 1988) – cannot successfully capture the observed treatment differences. In this paper, we apply Reinhard Selten's *impulse response dynamics* to the weakest link games of Croson *et al.* (2015) and demonstrate that impulse response dynamics predict the relative ranking of average contribution levels in the three weakest link games.

We choose the weakest link game variants of Croson *et al.* (2015) for our study, because the small variations of the rules that distinguish these games with multiple equilibria lead to markedly different behavioral dynamics that are not easily explained by any rational equilibrium concept. Thus, the weakest link game

variants are a crucial test for any theoretical concept attempting to capture the dynamics of play. If we can show that Reinhard Selten's impulse response dynamics can account for the unexplained differences in behavioral dynamics observed in these variants of weakest link game, we can conjecture that the concept should be considered as an alternative theoretical approach for a wide range of games.

In a series of contributions, starting with Selten and Buchta (1999) and Selten and Stoecker (1986), Reinhard Selten developed a dynamic concept of behavioral adaptation toward the best response for the realized state of the world after receiving feedback. The impulse response concept has been applied to numerous decision and interaction situations in slightly different variations ('learning direction theory', 'impulse matching', 'impulse balance'), but always based on the basic idea of impulse response dynamics (Chmura *et al.*, 2012, 2014; Goerg and Selten, 2009; Neugebauer and Selten, 2006; Ockenfels and Selten, 2005, 2014; Selten, 2004; Selten and Chmura, 2008; Selten *et al.*, 2005). These publications have established that impulse response dynamics can not only predict final rest points of adaptation (i.e., equilibria), but can also be valuable for the analysis of the adaptation process.

In this paper, we demonstrate that impulse response dynamics can be successfully applied to the weakest link games of Croson *et al.* (2015). The described impulse response dynamic is a very simple process and yet we will demonstrate how helpful it is in explaining and understanding behavior in repeated experimental games.¹ We show that the ranking of treatment outcomes predicted by the impulse response dynamics is fully consistent with the observed data. We also show that the impulse response dynamics perform well in tracking the observed development of contributions over time. We conclude that Reinhard Selten's impulse response dynamics provide an extremely valuable behavioral approach that is capable of explaining outcome differences and behavioral decision patterns in repeated strategic interaction games, such as the ones described and discussed in this paper.

2. DESCRIPTION OF EXPERIMENTAL DESIGN

In this section, we briefly describe the three weakest link games that we consider. In each game, four individuals ($N = 4$) play 20 rounds ($T = 20$ with a restart at $t = 11$) of the game in a fixed matching. In each round t , each individual i submits a *contribution* x_{it} to the group account. Submissions in each round are simultaneous. The residual of each round's endowment ($50 - x_{it}$) automatically goes into the individ-

1. Repeated game play can also be investigated with more sophisticated learning models such as reinforcement learning (Erev and Roth, 1998; Roth and Erev, 1995), experience weighted attraction learning (Camerer and Ho, 1999; Ho *et al.*, 2007) or impulse matching learning (Chmura *et al.*, 2012). These models describe probabilistic decisions and adjustments of the underlying probabilities, require parameter estimations, and will generally exhibit much longer convergence paths than the number of rounds actually played in the experiment. These aspects make it difficult to compare the sophisticated models to the simple deterministic impulse response model that offers a parameter-free and within play time tracking of the choice dynamics. We therefore focus only on presenting the predictive power of the impulse response model without attempting to compare it to the virtually incomparable and much more complex probabilistic mechanisms.

Weakest Link - Impulse Response Dynamics

ual's private account. The individual's payoff π_{it} in round t is based on the amount in the private account, on the contributions into the group account and (in two of three treatments) on the comparison between the sizes of the contributions.

In the *weakest link baseline* treatment (WL-B), each individual's round t payoff consists of the sum of the amount in their private account plus two times the *minimum* amount allocated to the group account by any group member. Thus, the individual's payoff function in WL-B is given as follows:

$$\pi_{it}^B(x_{it}, x_{-it}) = 50 - x_{it} + 2 \min(x_{it}, x_{-it}), \quad x_{it}, x_{-it} \in [0, 50], \quad (1)$$

where $-i$ denotes the individuals in the group, who are not i .

In the second treatment, the *weakest link exclusion* treatment (WL-EX), each individual's round t payoff is the same as in equation (1) as long as his contribution exceeds the minimum of all contributions or is equal to the maximum. If that is not the case the individual is excluded from sharing the group benefits and is paid only the private account. Thus, the individual's payoff function in WL-EX is given as follows:

$$\pi_{it}^{EX}(x_{it}, x_{-it}) = \begin{cases} 50 - x_{it} & \text{if } \max(x_{-it}) > x_{it} = \min(x_{it}, x_{-it}), \\ \pi_{it}^B(x_{it}, x_{-it}) & \text{otherwise} \end{cases}, \quad x_{it}, x_{-it} \in [0, 50] \quad (2)$$

Finally, the third treatment is the *weakest link (exclusion and) redistribution* treatment (WL-RE), in which again the lowest contributors are excluded, but *all* of the group's benefits, including those from the excluded contributors, are distributed amongst the non-excluded individuals. The exclusion rule is exactly the same as in WL-EX, but the benefits from the contribution of the excluded individuals are redistributed instead of being lost. Thus, the individual's payoff function in WL-RE is given as follows:

$$\pi_{it}^{RE}(x_{it}, x_{-it}) = \begin{cases} 50 - x_{it} & \text{if } \max(x_{-it}) > x_{it} = \min(x_{it}, x_{-it}), \\ \pi_{it}^B(x_{it}, x_{-it}) + 2 \min(x_{it}, x_{-it}) \frac{N}{4-N} & \text{otherwise} \end{cases} \quad (3)$$

3. THE DATA

The data were originally collected by Croson *et al.* (2015) at LINEEX, the experimental laboratory at the University of Valencia. Subjects were students of economics or sociology. The experiments were computerized (z-Tree, Fischbacher, 2007) and lasted less than an hour. The subjects earned EUR 16 on average. At the outset of the experiment, the subjects were randomly assigned to groups of four and remained in this fixed matching for the entire experiment. In total, 96 subjects participated in the experiment, and each subject participated in one session only. In each of the treatments WL-B and WL-EX, 24 subjects were matched into six groups of four, i.e., there are six independent observations in WL-B and six in WL-EX. In the treatment WL-RE, 48 subjects were matched into 12 independent observations, i.e., there are 12 independent observations in WL-RE.²

- Further details on the experimental protocol including the instructions can be found in Croson *et al.* (2015) and are available from the authors upon request.

Figure 1 shows the average contribution data. We observe rather heterogeneous contribution paths in all three treatments.³ Contributions are greatest in WL-RE, followed by WL-EX, and WL-B. While we do see some stabilization of contribution behavior over time, there obviously is no clear point of attraction that absorbs choices in all groups and all treatments. The overall average contributions are 38, 22 and 11 cash units respectively. The differences between treatments are significant.⁴ In the following, we first discuss why Nash equilibria solution concepts cannot organize the data, before showing that the empirical differences are consistent with Reinhard Selten's impulse response dynamics.

4. NASH EQUILIBRIA SOLUTION CONCEPTS

The three investigated weakest link games have multiple symmetric Nash equilibria in pure strategies. In WL-B and WL-EX, every symmetric strategy profile constitutes a Nash-equilibrium.⁵ In WL-RE, both the strategy profiles with full contributions and with zero contributions by all players are Nash equilibria.⁶

Equilibrium selection with payoff dominance or risk dominance (Harsanyi and Selten, 1988) narrows down the space of predicted Nash equilibria. Payoff dominance selects the full contribution equilibrium in all three treatments. As Fatas *et al.* (2006) show, risk dominance allows no refinement in the WL-B, leaving us with the full set of symmetric equilibria. In WL-EX, risk dominance eliminates all intermediate equilibria, leaving us with the equilibria at zero and at full contribution. This is so, because upward deviations from a symmetric equilibrium by a player is always less risky than downward deviations due to the threat of exclusion to the deviating player. In WL-EX and WL-RE, both full contribution and zero contribution cannot be ordered according to the criterion of risk-dominance.⁷

An alternative approach using the Nash equilibrium concept is the quantal response equilibrium (QRE) that implements Nash equilibria allowing for errors on the part of all players (McKelvey and Palfrey, 1995). QRE models have been

3. Graphs showing the development of contributions for each independent observation are in Appendix S1.
4. The Jonckheere–Terpstra test rejects the null hypothesis of no differences in average contributions in favor of ordered alternatives across WL-B, WL-EX and WL-RE. The differences are highly significant; the p -value of the two-tailed test is 0.0002.
5. If all other individuals in the group choose \hat{x}_{-it} , i will not contribute more than \hat{x}_{-it} , because doing so would decrease i 's private account without increasing the payoff from the public good, which depends on the minimum contribution. Individual i would also refrain from contributing less than \hat{x}_{-it} , because doing so would decrease the public good payoff more than it would increase the private account.
6. If all other individuals in the group choose $0 < \hat{x}_{-it} < 50$, should contribute $\hat{x}_{-it} + 1$, because doing so will decrease i 's private account only by 1, but 'redistribute' to him the entire amount contributed by all other individuals. This incentive to increase contributions is present at any positive level of contributions by others. Hence, the only equilibrium with positive contributions is the equilibrium in which everyone contributes the maximum amount. The only other equilibrium is at $\hat{x}_{it} = 0$, where no one contributes and no one has an incentive to contribute, because increasing the contribution to $\hat{x}_{-it} = 1$ means giving up 1 from the private account, but receiving 0 from the redistribution of the contributions of others, since $\hat{x}_{-it} = 0$.
7. See Appendix S1 for details.

Weakest Link - Impulse Response Dynamics

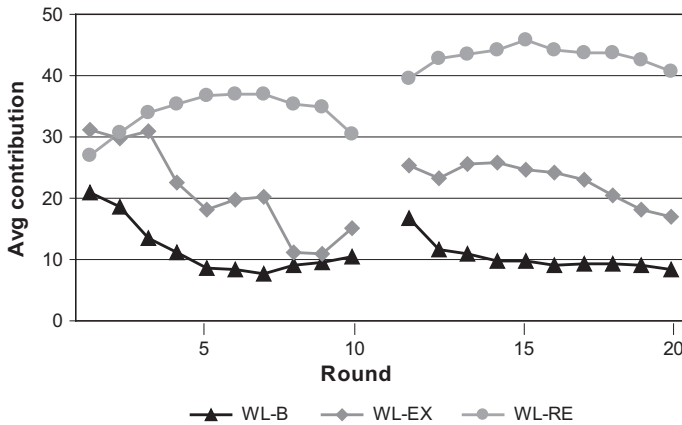


Figure 1 Average contributions in the three weakest link games (reproduced from Croson *et al.*, 2015)

shown to organize observed behavior successfully in a number of experimental studies. In the case of the weakest link game, however, Croson *et al.* (2015) show that when typical intermediate values of the noise parameter are considered, the QRE model does not organize the observed behavior well, because the predictions remain too low. Moreover, the predicted QRE dynamics lead to even lower predictions as the noise level is reduced, with the limiting logit equilibrium at zero contributions for all three treatments.

Obviously, none of the Nash equilibrium concepts are well in line with the observed behavior. While the multiplicity of Nash equilibria in WL-B and WL-EX provides room for heterogeneous outcome predictions, it neither provides clues on the adaptation paths nor does it explain the differences observed across treatments.

5. IMPULSE RESPONSE DYNAMICS

Observing that in some game situations the Nash equilibrium concept and its refinements are not very successful in capturing the diversity of behavioral outcomes, Reinhard Selten early on suggested using concepts that take the dynamics of behavior into account, where adaptation is based on *ex-post* rationality (Selten and Stoecker, 1986). The basic idea underlying this concept is that individuals making repeated decisions will adapt their behavior in the direction of the action that after receiving feedback (i.e., *ex post*) is recognized as what would have been the best response to the actually realized state of the world.⁸

8. Selten and Buchta (1999) illustrate this type of *ex-post* adaptation by their analogy of a marksman who learns how to hit a trunk with an arrow. 'If he misses the trunk to the right, he will shift the position of the bow to the left and if he misses the trunk to the left he will shift the position of the bow to the right. The marksman looks at his experience from the last trial and adjusts his behavior' (p. 86).

Note that the impulse response dynamics (i.e., the behavioral adaptation toward the *ex-post* best response that is called ‘learning direction theory’ in Selten and Buchtta (1999) and Selten *et al.*, 2005) describes boundedly rational behavior, because individuals are assumed to adapt their behavior to a state of the world (or to the strategy choices of others) that lies in the past, instead of giving a best response to their rational expectations of the state of the world that they will encounter next. In games with monotonically varying payoffs and an ordered action space, if the step size of adaptation is not too large and cycles do not ensue, this type of behavioral adaptation moves the system toward an ‘impulse balance point’ (or ‘impulse balance equilibrium’), in which the *ex-post* ‘impulses’ for upwards and downwards adaptations neutralize each other (Neugebauer and Selten, 2006; Ockenfels and Selten, 2005, 2014; Selten *et al.*, 2005). Usually and also in one of our two implementations, higher weights are put on impulses leading away from losses, where losses are defined with respect to the maximin.⁹ In the investigated weakest link games, the maximin strategy is to contribute zero and the maximin payoff is 50, because it is the maximum payoff a player can guarantee for himself no matter what strategy the others play. In the simple impulse response process that we describe in this paper, players adjust their action choices faster after experiencing a loss (i.e., a payoff smaller than 50) than after experiencing a gain (two-step adjustments for losses compared to one-step adjustments for gains).

The concept of impulse balance equilibria is not only applicable to games in pure strategies. In a number of studies, the impulse balance dynamics have also been applied successfully to games in mixed strategies, where impulses adjust the probabilities of playing the pure strategies (Chmura *et al.*, 2014; Goerg and Selten, 2009; Neugebauer *et al.*, 2015; Selten, 2004; Selten and Chmura, 2008). In games with certain outcomes and equilibria in pure strategies, Nash equilibria are generally also impulse balance equilibria. However, in games in which the Nash equilibria are in terms of expected payoffs, they do not necessarily coincide with impulse balance equilibria. Hence, the static predictions of impulse balance and Nash equilibria may fall together or fall apart, depending on the game. But, more importantly the impulse balance dynamics provide an adaptation path to equilibrium, while the Nash concept does not.

In the investigated weakest link games, the impulse balance equilibria generally coincide with the Nash equilibria. Specifically, this means that any symmetric strategy profile in WL-B and WL-EX is a Nash and an impulse balance equilibrium. In WL-RE this holds both for the symmetric zero and full contribution profiles. When instead of focusing on the static equilibria, we take the dynamics of adaptation into account, the impulse balance concept proves to be very useful. For any given first period strategy profile, we can use the impulse balance dynamics to derive the theoretical adaption paths (the ‘impulse matching’ paths) in the three treatments. Clearly, we observe that for any distribution of first period strategy profiles, the impulse matching path runs at the relatively lowest level of contributions in WL-B, followed by a higher contribution level in

9. In the spirit of the loss aversion literature (Kahneman and Tversky, 1979), the notion is that choices are more strongly driven away from losses than they are toward gains. Ding and Nicklisch (2013) provide additional experimental support for the stronger reaction to negative impulses.

Weakest Link - Impulse Response Dynamics

WL-EX, and the highest contribution level in WL-RE. Hence, we can generally show (in our Proposition) that the average level of contributions along impulse response paths have the same ordering as the contribution levels that we observe across treatments in the experiment.

Before we arrive at our Proposition and formally show that the ranking of impulse response paths is consistent with observed contribution levels, we need some formal definitions. Our analysis starts by considering the best response functions. We know the analysis technique of best response functions from Selten's game theoretic work (e.g., Selten, 1975). For the baseline treatment WL-B, the best response is the minimum contribution submitted by the others:

$$x_{it}^{BR-B} = \min(x_{-it}) \quad (4)$$

In the WL-EX treatment, the best response is zero if the minimum contribution of the others is zero. If all others play a certain contribution level greater than zero then that level is also the best response. For interior minimum contributions involving heterogeneous contributions the best response is the minimum contribution plus the smallest money unit, $\varepsilon > 0$.¹⁰ To avoid exclusion, obviously, the best response function responds with slightly higher contributions (on average) in the WL-EX treatment than in the WL-B:

$$x_{it}^{BR-EX} = \begin{cases} 0, & \min(x_{-it}) = 0 \\ \min(x_{-it}) & \text{if } \max(x_{-it}) = \min(x_{-it}) \\ \min(x_{-it}) + \varepsilon & \text{otherwise} \end{cases} \quad (5)$$

In the WL-RE treatment, the just described positive effect is even greater. The best response is almost the same as in the WL-EX treatment, but it involves a different response to joint contribution levels of the others. Unless all others play full contribution or zero, the best response of WL-RE is the minimum contribution plus a unit.

$$x_{it}^{BR-RE} = \begin{cases} 0, & \min(x_{-it}) = 0 \\ 50 & \text{if } \min(x_{-it}) = 50 \\ \min(x_{-it}) + \varepsilon & \text{otherwise} \end{cases} \quad (6)$$

The effects of exclusion and redistribution in the best response functions are subtle. The problem of incentives in the weakest link game is that output is vulnerable to deviation by one player. Any player has a kind of 'veto power' in competition, because a zero contribution voids all other contributions.

According to impulse response dynamics, subjects first observe the behavior from the previous round and then receive an impulse in the *direction* of the *ex-post* best response. Given this impulse, subjects then adjust their behavior in the direction indicated by the *ex-post* best response. However, it is important to note that a learning process based on impulses is not the same as playing the best response. Subjects in the investigated games, who are playing best response, will adjust their $t - 1$ behavior in such a way that the group most rapidly approaches a Nash equilibrium. In contrast, a subject responding to the impulses adjusts the behavior in the direction of the impulse, but only gradually. The predicted adjustment can be formalized for our environment as follows:

10. In the experiment subjects were enabled to contribute fractions of a Eurocent.

$$x_{it} = x_{it-1} + \varepsilon_{it}I(x_{it-1}^{BR-}), \quad \varepsilon_{it} \geq 0, \quad (7)$$

with $I(\cdot)$ being an indicator whether the adjustment toward the *ex post* best response is positive, negative, or zero and ε_{it} being the size of the adjustment. In Chmura *et al.* (2012) impulses change the probabilities of a mixed strategy depending on the size of forgone payoffs; the larger the forgone payoffs the larger are the adjustments of the probabilities. In the following, we keep things even simpler by assuming that subjects adjust their behavior in small constant equal steps upwards and downwards (as in Neugebauer *et al.*, 2015). Using equation (7) we can rank the contributions in the three games, explain the pattern over time and predict the selected impulse balance equilibrium.

Proposition Impulse response dynamics suggest the following ranking of contributions across the three weakest link games:¹¹

$$x_{it}^B \leq x_{it}^{EX} \leq x_{it}^{RE} \quad (8)$$

Proof. To see that the Proposition holds, first consider any initial contribution vector, including symmetric and asymmetric profiles. As a first example, let us start by examining an intermediate contribution vector where all subjects play $x_{i1} = 25$. For WL-B and WL-EX, this contribution profile is already an impulse balance point without positive or negative impulses. The impulse is positive, however, in case of the WL-RE where contributions symmetrically increase by one step in the next move. Repeating the increase step by step, we converge on the full contribution impulse balance equilibrium in the WL-RE treatment. The same dynamics follow from any other symmetric starting vector except at the extremes; zero and full contribution. At zero and at full contribution, impulses in all three games are zero and the dynamics come to rest. Hence, we have shown our proposition for all symmetric initial contributions.

Considering the asymmetric initial contributions, in WL-B, we find that there is an upward impulse to increase the minimum contribution and a downward impulse to decrease all other contributions. The upward adjustment process in WL-B stops when the lowest contribution and the second lowest contribution coincide. All contributions that are initially higher than the minimum contribution are adjusted downwards until they coincide with the minimum contribution, and remain at that level thereafter. The process rests when an impulse balance equilibrium is reached. In WL-B, the process selects the midpoint of the

11. We make the assumption that the initial contributions are equal in all treatments. However, if initial contributions increase from WL-B to WL-RE the proposition still holds. The data give support for this requirement confirming that initial contributions do not decrease from WL-RE to WL-B. The two-tailed pairwise Mann-Whitney test suggests that the initial individual contributions in WL-B are significantly different from (smaller than) the ones in WL-EX ($p = 0.008$) and WL-RE ($p = 0.009$), whereas the initial choices in the latter two treatments are statistically similar ($p = 0.123$).

initially lowest and second lowest contribution levels as the impulse balance equilibrium.

The initial adjustment dynamics in WL-EX are very similar to those in WL-B, the minimum contribution is adjusted upwards and the larger contributions are adjusted downwards. However, the adjustment of the smallest and the second smallest contributions does not stop once they coincide, because according to equation (7) the impulse is positive for all minimum contributors, as long as others choose higher contribution levels. Hence, the low contributions are adjusted upwards, and the high contributions are adjusted downwards, until all contributions coincide. Thus, in WL-EX, the predicted impulse balance point is the midpoint of the initially smallest and the initially largest contribution levels, which may be greater but never smaller than the predicted impulse balance equilibrium in WL-B.

In WL-RE, any minimum contribution is adjusted step-by-step upward in accordance with equation (7). As stated above, however, the impulse response dynamics in WL-RE rest in a symmetric contribution profile only when the full contribution equilibrium has been reached. Hence, with any homogeneous and heterogeneous initial contribution vector, except the singular case where all initial contributions are zero, the impulse response dynamics of the WL-RE move to higher levels than in WL-EX and WL-B. This completes the proof of our proposition.¹²

For example, assuming the initial contribution vector (5, 15, 25, 35), impulse response dynamics predict the impulse balance equilibrium strategy 10 (midpoint between 5 and 15) in WL-B, 20 (midpoint between 5 and 35) in WL-EX, and 50 (impulse balance equilibrium attracting all heterogeneous initial contributions) in WL-RE, respectively. But, impulse response dynamics do not only predict the order of impulse balance points across treatments, they also predict the adjustment paths (see the path descriptions in the proof of our proposition).

Our proposition is based on impulse response dynamics that have been verified to be descriptive of behavior in numerous experimental settings. We have demonstrated that this concept selects impulse balance equilibria that are ranked across the treatments in exactly the same order as observed in the experiment, with the contributions in WL-RE generally greater than those in WL-EX, which in turn are greater than those in WL-B. All in all, it seems that the impulse response dynamics are an extremely useful way of modeling behavior, especially when stationary concepts provide no theoretical guidance to explain observed behavior. In the next section, we will show that impulse response dynamics not only predict the observed order of contributions across the treatments, but the development of contributions over time.

12. In other words; with given positive initial contributions, initial impulse response dynamics are frequently the same in all treatments. The smallest contribution is adjusted upwards, and the larger contributions are adjusted downwards. In the WL-B treatment the upward impulses cease early, i.e., as soon as the two smallest contributions equalize. In the WL-EX treatment, the upward impulses cease later, i.e., as soon as the smallest and largest contributions equalize. In WL-RE, finally, the upward impulses cease at the boundary, i.e., where the efficient impulse balance equilibrium is reached.

6. IMPULSE RESPONSE DYNAMICS *VIS-À-VIS* THE DATA

As suggested above, the impulse response trajectories provide testable prediction on the adjustments across treatments. In this section, we examine how well the average trajectory based on the impulse response dynamics can track the observed behavior in the lab. We use the observed initial contribution profiles as starting point of the impulse response adjustment trajectories. For each independent session we develop the adjustment trajectories. In a first simulation, we employ a symmetric one-step adjustment approach; for each period the individual contribution is adjusted upwards or downwards in the direction of best response by one money unit, $\varepsilon_{it} = 1$. In a second simulation, we employ an asymmetric one-step gain/two-step loss adjustment process. According to this alternative process a double adjustment step is employed whenever the simulated contributor experiences a loss.¹³ Note again that the only input data required to develop this process are the vectors of the initial contributions of the original and the restart game. Everything else, i.e., the predictions for the rounds 2–10, follows from the impulse response dynamics.

In Figure 2, we present the resulting average impulse response trajectories for each of the treatments including original and restart games. The diamonds indicate observed average contributions while the lines indicate the impulse response trajectories. The solid line in Figure 2 represents the aggregate impulse response dynamics according to the symmetric one-step process, whereas the dashed line represents the one-step gain/two-step loss adjustment process. Where both processes coincide the dashed line is covered by the solid line. We see that the fit is not perfect, particularly not for the first 10 rounds, although impulse response trajectories move in the right direction. Comparing the fit of the impulse response dynamics in the original game and the restart game, we find that the average squared error per round is significantly smaller in the restart game.¹⁴ In the second 10 rounds, the simulated impulse response dynamics seem to track the data quite well.¹⁵

Figure 2 indicates that the one-step process (solid line) and the one-step gain/two-step loss process (dashed line) have a similar fit, but that the trajectory of

13. Reinhard Selten and collaborators propose that impulses to losses are double as strong as impulses to gains (Chmura *et al.*, 2012; Neugebauer *et al.*, 2015; Ockenfels and Selten, 2014). In this research, a payoff is a loss if it is smaller than the player's maximin payoff. In our setting, the maximin payoff is the player's endowment. Thus, for payoffs below the endowment, the corresponding adjustment in the following period involves two money units rather than one. Where no loss is experienced, the trajectory applies the one-step adjustment rule as before.
14. For each round of each session we compute the squared deviation of the observed average contributions from the predicted ones by impulse response. The period average squared error for the original game and the restart game is tested on the independent observations. According to the two-tailed Wilcoxon signed ranks test, the p -values are 0.116, 0.028, and 0.023 for WL-B, WL-EX, and WL-RE, respectively. The latter two results are significant, and the sign is correct for the WL-B.
15. The Spearman rank correlation coefficients between the average simulation outcome and the average observed outcome are larger in the second ten rounds of each treatment. For WL-EX and WL-RE the correlation is significantly positive in the second sequence ($p = 0.0362$ and 0.0081 one-tailed test; the p -value for the WL-B is 0.1987). For the first sequence only the correlation for the WL-RE is significant ($p = 0.0013$), whereas the other treatments are not ($p = 0.7589$ WL-B, and $p = 0.1040$ WL-EX).

Weakest Link - Impulse Response Dynamics

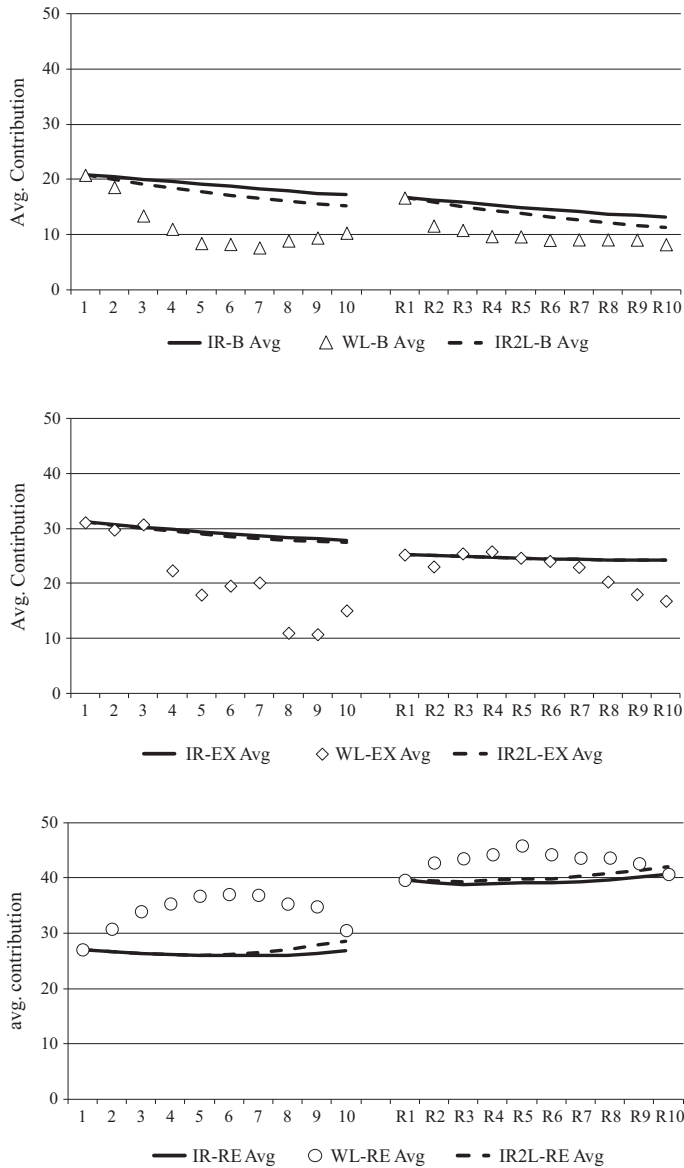


Figure 2 Average contributions and average impulse response trajectories (one-step adjustment IR; one-step adjustment for gains and double step adjustment for losses IR2L)

the latter seems a bit closer to the data. Indeed, the squared deviations of the models from the observation are smaller for the latter model but they are not significantly different. The considered trajectories are simplistic and owed to the one-step approach and generally do not allow for convergence on an equilibrium in ten rounds. For our data, nonetheless, it is not evident that converging trajectories would provide a much better fit. When we compare the fit of the one-step process's last round contributions to the fit of the impulse balance equilibrium

on which the trajectories converge, we find no significant differences in the corresponding squared errors.¹⁶

Finally, we report two more observations demonstrating that subjects adjust their round-by-round behavior consistent with the described impulse response dynamics. First, the *individual* contribution changes (between successive rounds) in the predicted direction significantly outnumber the changes in the opposite direction for each of the considered 24 independent sessions.¹⁷ Second, examining the change in group contributions between successive rounds, we find that the number of rounds in which the *group* contributions should increase according to the impulse response is significantly correlated with the number of rounds in which the observed group contributions increase.¹⁸

7. SUMMARY

Reinhard Selten has introduced many important theoretical concepts that can be successfully applied to analyze strategic interactions in the social sciences. In this paper, we touch on one of his most important and most recent contributions. We discuss the impulse response concept as a tool that offers a theoretical explanation both for the dynamics and for the impulse balance equilibria (i.e., the rest points of behavior) in situations with repeated interaction and multiple equilibria.

Using the impulse response dynamics, we derive a proposition that provides a theoretical ranking of contribution levels in weakest link games with exclusion and redistribution. The proposition establishes that the impulse response dynamics offer differentiated predictions even where static concepts based on Nash equilibria allow no such ordering. We show that the ranking of treatments, as predicted by our proposition based on the impulse response dynamics, is fully consistent with the data observed in a recent study by Croson *et al.* (2015). In that study, three game environments were considered under the following three conditions; baseline, exclusion and redistribution. Croson *et al.* (2015) demonstrate that static equilibrium concepts based on the Nash equilibrium (including multiple Nash equilibria, quantal response equilibria, as well as equilibrium selection by risk and payoff dominance) cannot successfully capture the observed treatment differences. In the present study, we show that Reinhard Selten's impulse response concept is quite successful in predicting behavioral outcomes.

Intuitively, our results suggest that the ranking of the impulse balance equilibria across treatments follow from the differences in the incentive structure of the treatment. The best response functions provide incentives for minimum contributors to increase their contributions in all treatments. Exclusion provides additional incentives for minimum contributors to increase contributions in

16. The average impulse balance equilibria on which our trajectories converge are 11, 25, 50 in the restart game of WL-B, WL-EX, and WL-RE. These compare to the observed last round contributions 8, 17, 41 and the simulated average contributions of our one-step process 13, 24, 41, respectively. According to the Wilcoxon signed ranks test, we have no significant differences between the squared deviations; p -values are 0.345, 0.833, 0.722.

17. The one-tailed probability that such an outcome occurs by chance is 0.5^{24} .

18. The Spearman rank correlation is 0.7498 and thus highly significant, $p < 0.001$.

Weakest Link - Impulse Response Dynamics

conditions of exclusion and redistribution, and redistribution provides incentives for high contributors to exceed the others' contributions in the game environment with redistribution. While the impulse response goes in the same direction as the best response, it is less extreme in its adjustments. Thus, the impulse response dynamics capture the inertia of human behavior as well as the convergence of the game play to an equilibrium.

We have demonstrated that the fit of the average choice is reasonably good, especially after a restart of the game (i.e., when subjects are more experienced in play). Concerning equilibrium selection in games with multiple Nash equilibria, our results suggest that using a simple dynamic approach such as the impulse response concept can help to identify the most probable outcomes. Hence, Reinhard Selten's impulse response dynamics present a powerful concept that can help to explain numerous behavioral patterns in strategic interaction settings, such as those discussed in this paper.

ACKNOWLEDGEMENTS

We thank Rachel Croson, Enrique Fatas, and Antonio Morales for providing data and valuable comments. We are also grateful for the superb editorial support and the constructive comments of the anonymous referees.

ADDRESS FOR CORRESPONDENCE

Address for correspondence: Abdolkarim Sadrieh, University of Magdeburg. Tel.: +49 391-67-18492; fax: ++49 391-67-11355; e-mail: sadrieh@ovgu.de

REFERENCES

- Camerer, C. F. and T. H. Ho (1999), 'Experience-Weighted Attraction Learning in Normal Form Games', *Econometrica* 67, 827–874.
- Chmura, T., S. J. Goerg and R. Selten (2012), 'Learning in Experimental 2×2 Games', *Games and Economic Behavior* 76, 44–73.
- Chmura, T., S. J. Goerg and R. Selten (2014), 'Generalized Impulse Balance: An Experimental Test for a Class of 3×3 Games', *Review of Behavioral Economics* 1, 27–53.
- Croson, R. A., E. Fatás, T. Neugebauer and A. J. Morales (2015), 'Excludability: A Laboratory Study on Forced Ranking in Team Production', *Journal of Economic Behavior and Organization* 114, 13–26.
- Ding, J. and A. Nicklisch (2013), 'On the Impulse in Impulse Learning', *Economics Letters* 121, 294–297.
- Erev, I. and A. E. Roth (1998), 'Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique Mixed Strategy Equilibria', *American Economic Review* 88, 848–881.
- Fatas, E., T. Neugebauer and J. Perote (2006), 'Within-Team Competition in the Minimum Effort Coordination Game', *Pacific Economic Review* 11, 247–266.

- Fischbacher, U. (2007), 'Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments', *Experimental Economics* 10, 171–178.
- Goerg, S. J. and R. Selten (2009), 'Experimental Investigation of Stationary Concepts in Cyclic Duopoly Games', *Experimental Economics* 12, 253–271. doi:10.1007/s10683-009-9218-8.
- Harsanyi, J. and R. Selten (1988), *A General Theory on Equilibrium Selection in Games*, MIT Press, Cambridge, UK.
- Ho, T. H., C. Camerer and J.-K. Chong (2007), 'Self-Tuning Experience-Weighted Attraction Learning in Games', *Journal of Economic Theory* 133, 177–198.
- Kahneman, D. and A. Tversky (1979), 'Prospect Theory: An Analysis of Decision Under Risk', *Econometrica* 47, 263–291.
- McKelvey, R. and T. Palfrey (1995), 'Quantal Response Equilibrium for Normal Form Games', *Games and Economic Behavior* 7, 6–39.
- Neugebauer, T. and R. Selten (2006), 'Individual Behavior of First-Price Auctions: The Importance of Information Feedback in Computerized Experimental Markets', *Games and Economic Behavior* 54, 183–204.
- Neugebauer, T., Sadrieh, A., and Selten, R. 2015, 'Taming Selten's Horse by Impulse Response', mimeo.
- Ockenfels, A. and R. Selten (2005), 'Impulse Balance Theory and Feedback in First-Price Auctions', *Games and Economic Behavior* 51, 155–170.
- Ockenfels, A. and R. Selten (2014), 'Impulse Balance in the Newsvendor Game', *Games and Economic Behavior* 86, 237–247.
- Roth, A. E. and I. Erev (1995), 'Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term', *Games and Economic Behavior* 8, 164–212.
- Selten, R. (1975), 'Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games', *International Journal of Game Theory* 4, 25–55.
- Selten, R. (2004), 'Learning Direction Theory and Impulse Balance Equilibrium', in: D. Friedman and A. Cassar (eds.), *Economics Lab—An Intensive Course in Experimental Economics*, Routledge, New York, pp. 133–140.
- Selten, R. and J. Buchta (1999), 'Experimental Sealed Bid First Price Auctions with Directly Observed Bid Functions', in: A. Rapoport, D. V. Budescu, I. Erev and R. Zwick (eds.), *Games and Human Behavior: Essays in Honor of Amnon Rapoport*, Lawrence Erlbaum Associates, Mahwah, NJ, pp. 101–116.
- Selten, R. and T. Chmura (2008), 'Stationary Concepts for Experimental 2×2 -Games', *American Economic Review* 93, 938–966.
- Selten, R. and R. Stoecker (1986), 'End Behavior in Sequences of Finite Prisoners' Dilemma Supergames: A Learning Theory Approach', *Journal of Economic Behavior & Organization* 7, 47–70.
- Selten, R., K. Abbink and R. Cox (2005), 'Learning Direction Theory and the Winner's Curse', *Experimental Economics* 8, 5–20.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Graphs for each of the independent observations separately.