Transmuting Samaritan's Dilemmas in Irrigation Aid:

An Application of the Topology of 2x2 Ordinal Games

Bryan Bruns[1]

TRANSMUTING 2X2 GAMES WORKING PAPER NUMBER 2

SEPTEMBER 10, 2010

Abstract: Aid risks discouraging or "crowding out" local effort in commons such as irrigation systems, posing problems for international development programs, including attempts to promote participatory irrigation management (PIM) and irrigation management transfer (IMT). James Buchanan used game theory models to analyze structures of payoffs and preferences that create what he named Samaritan's Dilemmas. The topology of 2x2 ordinal games developed by Robinson and Goforth offers a useful tool for examining the relationship between Samaritan's Dilemmas and other problems of collective action, and the potential for institutional solutions through changing payoffs. In the case of irrigation aid, switches in payoffs that realign incentives to favor joint investments, and thereby transmute Samaritan's Dilemma into a Win-win Commons game, show the potential for counter-intuitive solutions through increased attention to co-management and joint investment in commons.

Keywords: irrigation, water resources governance, institutional analysis, infrastructure investment, game theory, ordinal 2x2 games,

A. SAMARITAN'S DILEMMAS IN INTERNATIONAL DEVELOPMENT

A tale of perverse incentives and disappointing outcomes in irrigation aid starts when farmers build canals and dams, (often temporary structures of rocks, logs, and earth) which divert water to their fields. This makes them better off, and creates a situation that a government or other potential donor prefers. Aid for further improvements, such as a permanent weir and concrete lined canals, may make things even better. However, if farmers then leave further maintenance and repair to the donor, and the donor would rather invest alone than do nothing, this creates a Samaritan's Dilemma. Aid can reduce local investment. The outcome can arise from farmers acting according to their preferences in the situation they face. However, it frustrates those who think farmers

---

[1] Bryan Bruns is an independent consulting sociologist, specializing in community participation in water resources management. This paper develops ideas from an earlier paper (Bruns 2008), and from work done while a Visiting Scholar at the Workshop in Political Theory and Policy Analysis at Indiana University Bloomington. Support provided by the Workshop is gratefully acknowledged. bryanbruns@bryanbruns.com

should share responsibility, and that lack of local investment contributes to problems in irrigation operation and maintenance, including unfair water allocation and neglect of preventive maintenance. If intervention occurs first, the story may be even simpler: the potential for local resource mobilization may never be fully realized.

Worse yet, inaccurate expectations by farmers, government budget constraints, and other problems may lead to a situation where neither water users nor government invests in maintenance and repairs, leading to deterioration or collapse. Part of this situation can be analyzed as a moral hazard problem, where the potential for future aid discourages local preventive maintenance (Bruns 1992). Government agencies may also have weak incentives to ensure proper preventive maintenance, especially if neglect leads to benefits from rehabilitation projects (Ostrom, Schroeder, and Wynne 1993; Araral 2006).

Farmers may feel that government is much wealthier and more capable, and will have to step in and make repairs. If investment by one or the other alone would be sufficient, then this would resemble a game of "Chicken," where the conflict is about who will invest. Most analysis in the literature using Chicken and other game theory models of irrigation investment has focused on interactions between farmers, rather than between farmers and government. (Taylor and Ward 1982; Bardhan 1993)

If both have incentives to neglect maintenance, even if the other is also neglecting maintenance, then the preference structure could resemble the game of Prisoner's Dilemma. Both would be better off with proper preventive maintenance, but each follows short-run incentives to neglect it, resulting in a result neither likes, a Pareto-inferior outcome. Multiplayer versions of Prisoner's Dilemma have been analyzed as public goods and commons dilemmas (see Kollock 1998 for a review of research on social dilemmas).

Perverse incentives may perpetuate a vicious cycle of poor irrigation performance, with low returns for farmers and infrastructure investments, where the presence and prospect of government aid reduces local investment (Araral 2005). More generally, government investment may displace or "crowd out" community collective action. These may be worsened if investments are done in ways that make it difficult or impossible for irrigators to participate, and their roles as coproducers of irrigation services are ignored or misunderstood (Ostrom 1997).

Irrigation management transfer projects have typically sought to escape such problems through "one last rehabilitation" after which farmers are supposed to take over operation and maintenance, at least for part of the system. This is usually done as part of policies for participatory irrigation management (PIM) or irrigation management transfer (IMT). These projects have typically been justified by anticipated savings in government budgets for operation and maintenance, and less frequent need for irrigation, and reinforced by ideas about reducing the scope of government action and shifting responsibilities to beneficiaries.[2] Performance of PIM and IMT projects has often been mixed at best (Vermillion 2006; Rap 2006).

Specific incentive problems in how aid can discourage local effort are usually embedded within larger sets of institutional problems (Gibson et al. 2005). These include conflicting interests between those providing funding and the organizations and individuals acting for them (principal-agent conflicts) and difficulties in organizing to provide collective goods. For irrigation, construction of permanent headworks may dissolve shared interests of head and tail-end users who had been united by the need to repair and rebuild works. Bureaucratic agencies may lack strong incentives to ensure good performance and maintenance for public infrastructure investments. Gibson, Andersson, Ostrom, and Shivakumar's research (2005) provided important insights into the array of problems facing international development efforts, but had less to say about specific solutions to Samaritan's Dilemma.

A particularly crucial problem for Samaritan's Dilemmas, and one that does not seem to have received much attention, is that it is hard for governments to make credible commitments to withhold further aid, especially in the context of:

- fungibility between preventive maintenance, repairs, and rehabilitation;
- likely occurrence of force majeure damage from floods and landslides;
- limited capacity of irrigators to make expensive investments infrastructure;
- ambiguous or contested ideas concerning user and government responsibilities;
- incomplete rehabilitation due to budget constraints and other limitations,
- inadequate performance of remaining government tasks, such as headworks management and river basin water allocation (giving farmers a basis to claim that aid is still needed); and
- continuing incentives for officials and politicians to provide aid.

This paper concentrates on Samaritan's Dilemmas in irrigation aid, and how they might be transformed to achieve better outcomes. The next section discusses models of Samaritan's Dilemmas, and the difficulty of finding better solutions within the framework of earlier game theory analysis. The following sections introduce the topology of 2x2 games, and potential solutions through swapping payoff rankings to transmute one game into another. Changes to make joint investment in irrigation more attractive are then discussed as part of policies for co-management and incremental improvement.

[2] Economic justifications for IMT typically neglect aggregate effects that lower prices and allow consumers to reap most of the benefits from investment in irrigation, and usually fail to analyze whether shifting O&M costs to farmers actually reduces total societal (economic) costs for irrigation, rather than just hypothetical financial reductions in government budgets.

BUCHANAN'S MODELS

James Buchanan (Buchanan 1977) offered an influential critique of how aid could discourage self-help, using game theory models to analyze what he named Active Samaritan's Dilemma. Figure 1 shows the payoff structure, which matches the story told above about government aid to irrigation. The row player has a dominant strategy, Within the narrow rules of standard game theory, there is no escape (Schmidtchen 2002). As long as the donor prefers the result from investing alone to stopping aid, the donor is trapped by their own preferences.

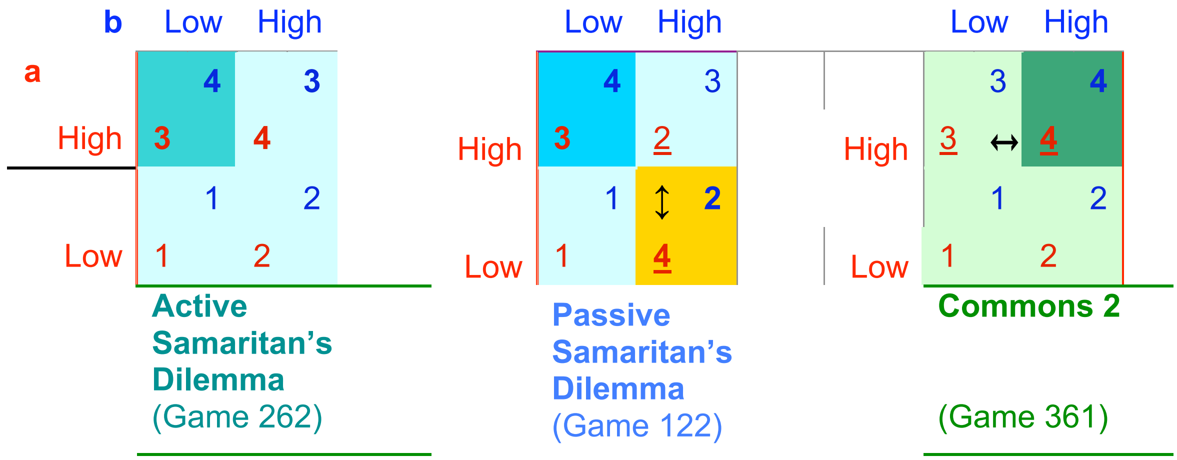
Unlike Prisoner's Dilemma, there is no Pareto- superior alternative that both would prefer. Instead, the situation is more like a game of Chicken or Battle of the Sexes, where the two players have conflicting preferences. However, in Samaritan's Dilemma, the donor's preferences create dominant strategy for the donor, which then allows the recipient to follow their preference to do less.

Buchanan suggested the donor might try to deceive the recipient into believing the donor will withhold aid unless the recipient also invests. Or, the donor could try threats, attempting to convince the recipient that the donor will not invest alone (even though this is contrary to the donor's real preference structure). Or, the donor might delegate action to a third party, instructed to stop aid unless the recipient contributes. Buchanan suggested that the feasibility of such solutions with repeated play might depend on the donor's discount rate for trading off short-term pain for future gains.

Later analysis finds that for conditions of full information and standard rationality assumptions, repeated play does not yield a better outcome (Schmidtchen 2002). The Nash Equilibrium is already Pareto-efficient (so there is no equivalent to the Folk theorem solution in repeated play for Prisoner's Dilemma). Thus, threats would not be credible, so it is not simply a matter of having the "strategic courage" to declare and, if necessary, apply a harsh policy of threats to withhold aid. A contract delegating a third party to implement a strategy would not be self-enforcing (due again to the prospective loss for the donor) and would only work if third party enforcement were available for the contract. Thus, for conventional non-cooperative game theory (where external enforcement is not an option), there would be no alternative solution. As Buchanan suggests, the situation may require acceptance that compassion comes with costs, so donors are unable to achieve their most preferred outcome of joint investment.

Buchanan also explored an alternative, Passive Samaritan's Dilemma, which swaps two payoffs. Here, the donor prefers that the recipient invest alone, and joint investment is the donor's third choice. In formal terms, this game has two Nash Equilibria, with unequal outcomes, an asymmetric cross between Chicken and Battle of the Sexes. Either irrigators invest alone, and get their third choice outcome, or the donor invests alone, getting their second choice, and the irrigators their top preference, as in the original Active Samaritan's Dilemma. So, switching the preferences does not yield a mutually satisfactory solution, instead either leaving the donor investing alone again, or the recipient stuck in their third choice outcome.

Figure 1. Samaritan's Dilemmas for Irrigation Investment



Furthermore, Passive Samaritan's Dilemma actually models a somewhat strange sort of Samaritan, one who prefers not to help (Schmidtchen 1999:5), and prefers dependency to joint effort. This might be seen as a reluctant patron, who would prefer to do nothing but insists on doing everything if they do act. It does not change the situation sufficiently to ensure self-help, and still risks creating a dependency trap.

As noted by Schmidtchen, Samaritan's Dilemma is a personal dilemma, but is not a social dilemma in the narrow definition of a situation where individually rational action makes both players are worse off. Instead, in the two-player case, it is a personal dilemma in that the donor is unable to achieve their preferred outcome even though both are acting rationally. It thus resembles a Battle of the Sexes Game, where on or the other can achieve their best outcome, but not both. However, if the resulting behavior has wider impacts, such as on the demand for and supply of government aid and local investment, then it can also be seen as a form of social dilemma.

TOPOLOGY OF 2X2 GAMES

The transformation between Active and Passive Samaritan's Dilemma by swapping preferences does raise the interesting question of whether there are other swaps in payoffs that might lead to better outcomes, which is the core concern of this paper. The brief answer is that if joint investment can be made more attractive to the irrigators, switching their top two payoffs, then Samaritan's Dilemma could be transformed into a win-win game where both sides' incentives would encourage them to cooperate.

The topology of 2x2 ordinal games (Robinson and Goforth 2005) can be used to examine the potential outcomes of changes that switch the ranks in adjoining payoffs, and particularly whether this would align interests so as to achieve mutually satisfactory solutions. This section briefly introduces the topology, and the next section discusses moves between games, that transmute one game into another, after which the paper returns to the substantive implications for Samaritan's Dilemmas in irrigation aid.

In Samaritan's Dilemma, for each player there are three possible swaps. Three of the six switch adjoining payoffs, (1↔2; 2↔3; and 3↔4) and so seem empirically more likely. As discussed by Robinson and Goforth (2005) changes in ranking of different outcomes might occur as a result of small changes in information, preferences, technology, or minor errors in identifying games. Swaps could come from changes in attitudes, such a greater concern for impacts on the other player. They could also result from deliberate design, changes in rules that affect the distribution of benefits and costs, punishments for violations, monitoring, and enforcement, and thereby change expected payoffs. Payoffs are not necessarily fixed, but instead are subject to design. The topology maps the network of possible changes in incentive structures.

While three swaps involve adjoining payoff ranks, the other three involve non-adjoining ranks (1↔3; 2↔4; and 1↔4), as in the (2↔4) switch between Buchanan's Active and Passive Samaritan's Dilemmas. They thus actually involve changing the relative ranks of three different outcomes, something that seems less likely, or would require more complications and higher transaction costs. The discussion in this paper concentrates on the more feasible cases of swaps between adjoining payoffs, and on strict ordinal games, i.e. those with four distinctly ranked payoffs for each player, i.e. no ties.[3]

For Samaritan’s Dilemma, the two swaps that create win-win games represent either raising the recipient's payoff from joint investment enough for them to prefer that, or the donor shifting to prefer investing alone. The second swap could be seen as a result of changing expectations about what level of effort is realistic for the recipient. This might be easier to understand if they would be making a low (but non-zero) level of investment feasible for them. Such a preference structure might also involve recognizing how acting as a Samaritan yields benefits for the donor and the kind of society the donor wants to live in, rather than ideological insistence on the individual self-reliance (Stone 2008). Policy recommendations thus depend on the empirical situation, including the feasibility of making joint investment more attractive.

For an individual game, it is possible to examine the alternatives created by payoff swaps one by one. However, since there are many games, it would be useful to have a more systematic framework for analysis. The potential for multiple models, including Active and Passive Samaritan’s Dilemma, Battle of the Sexes, Chicken, Prisoner’s Dilemma and Stag Hunt (also called Assurance or Coordination games), makes it useful to understand the relationships between games; how they differ or might be transformed into each other. The topology of 2x2 ordinal games developed by David Robinson and David Goforth (2005) offers an elegant structure for understanding how games are linked by swaps in adjoining payoffs, and the relationships among games. This paper uses a modified version of their Periodic Table of 2x2 Ordinal Games, showing numerical payoffs and payoff families (Bruns 2010).

[3] The topology can be extended to non-strict games (Robinson et al. 2007), and provides a framework for analyzing equivalent games with interval or cardinal (real) payoff scales (Robinson and Goforth 2005).

Swaps in the two lowest-ranked payoffs (1 and 2) connect games that are closest to each other in the topology. Substantively, it can be argued that such swaps are likely to be the least significant, and easiest to achieve in terms of transaction costs. A process to generate the topology starts with swapping the two lowest payoffs for Row or Column or both, to generate a *tile* of four similar games. Figure 2 shows the four similar games that compose the Samaritan's Dilemma Tile, (and those composing the Prisoner's Dilemma Tile). The two Samaritan's Dilemma games on the right have a dominant strategy for the Row player, while those on the left have a dominant strategy for both.

Figure 2. Tiles for Samaritan's Dilemma and Prisoner's Dilemma-Chicken

Swapping Column's middle two payoffs then generates adjoining games, with further one two swaps completing two more tiles (Figure 3). All three tiles and twelve games share the basic structure of Samaritan's Dilemma, where Row has a dominant strategy that leads to getting Row's second choice payoff.

Swapping Row's middle payoffs, and completing the resulting tiles generates additional games, forming a *layer* of nine tiles and thirty-six games. Since further row or column swaps return to the same games, the layer itself is a torus. The side edges connect to each other, as do the top and bottom. For display, this can be rotated, step-by-step, to position a particular game at the corner. For displaying game properties, it turns out to be convenient to put the region with two dominant strategies in the upper right of the layer, as shown in Figure 3. This means that some tiles, including the Samaritan's Dilemma Tile, are split, (as if the layer had been scrolled on step vertically and one step horizontally), but facilitates showing other relationships.

In addition to Samaritan's Dilemmas, this layer, labeled Layer Two in Figure 3, includes:

- Nine Cyclic games with no Nash Equilibria (for ordinally-ranked payoffs),
- Six Unfair games along the top where Row does even worse, with a dominant strategy leading to getting their next to worst outcome, while Column gets their top choice,
- Four Second Best games where both get their second-ranked outcome, and
- Five asymmetric cousins to Prisoner's Dilemma, members of the Prisoner's Dilemma Family with Pareto-inferior Nash Equilibria, in which neither player gets their top choice, and one or both have to settle for their next to worst choice.

Swapping the top two payoffs from Samaritan's Dilemma creates a game on another layer, and the same process of swapping lowest and middle payoffs generates the layer. Layer Three is composed of win-win games all of which have a Nash Equilibrium where both players get their top choice (4,4), sometimes referred to as "boring" or "no conflict games."[4] This layer includes nine Stag Hunt games with two Nash Equilibria, and twenty-seven Harmonious games with a single Nash Equilibrium.

Swapping Column's payoffs for the other three games on the same Tile as Samaritan's Dilemma links to games on a win-win tile (labeled the Commons tile on Figure 4d). It turns out that swapping Row's top two payoffs also leads to the same tile. This set of double links forms what Robinson and Goforth call a hotspot, which doubly joins two layers.

Swapping Row's middle payoffs in Samaritan's Dilemma creates an adjacent Unfair game (Game 212, in the upper left corner of Layer Two). From there, a swap in the top two payoffs creates the game of Chicken (Game 222, in the lower left corner of Layer One). From there, the rest of Layer One can be generated. Swapping the lowest row or column payoffs in Chicken makes two more unfair games. Swapping both Row's and Column's lowest payoffs transforms Chicken into Prisoner's Dilemma. Swapping Row's middle payoffs in Chicken transmutes it into the game Buchanan called Passive Samaritan's Dilemma. Swapping Column's middle payoffs then produces a Battle of the Sexes. From that Battle of the Sexes game, swaps in the two lowest-ranked payoffs complete a tile of four Battle of the Sexes Games (the other symmetric Battle of the Sexes game has also been called Hero (Rapoport 1967). Swapping middle payoffs from Hero creates a biased game where one player's dominant strategy produces a Nash Equilibrium that gives them their top-ranked preference. These Self-serving Games complement the Samaritan's Dilemma games, and together with the Battles of the Sexes make up a family of Biased Games, where one player gets their top choice but the other only gets their second-ranked outcome.

---

4 As discussed by Robinson and Goforth (2005), most of the games on Layer Three are actually games of mixed interests. Thus the term "no conflict" is a misnomer. Only nine games are fully aligned games of pure cooperation (Figure 4f). In the rest of the games on Layer Three, if the players are not at the Nash Equilibrium, then a change in strategy by one player may reduce the payoff for the other, i.e., the inducement correspondences are not always positive. Even within the nine pure cooperation games, where moves by one player to increase her own payoffs always improve payoffs for the other, the four Coordination games are Stag Hunts, where the Pareto-inferior Nash Equilibrium may be risk-dominant. Stag Hunt games are definitely not "no conflict" or "boring" and pose interesting challenges for collective action (Skyrms 2004). While in symmetric Stag Hunt games the Pareto-inferior Nash Equilibrium is the Maximin outcome, and, arguably, risk-dominant, it is interesting to note that for two pairs of asymmetric Stag Hunt games (Assurance 4 and 4', 342, 324; and Coordination 2 and 4, 334 and 343), the Pareto-inferior Nash Equilibrium is *not* the Maximin outcome.

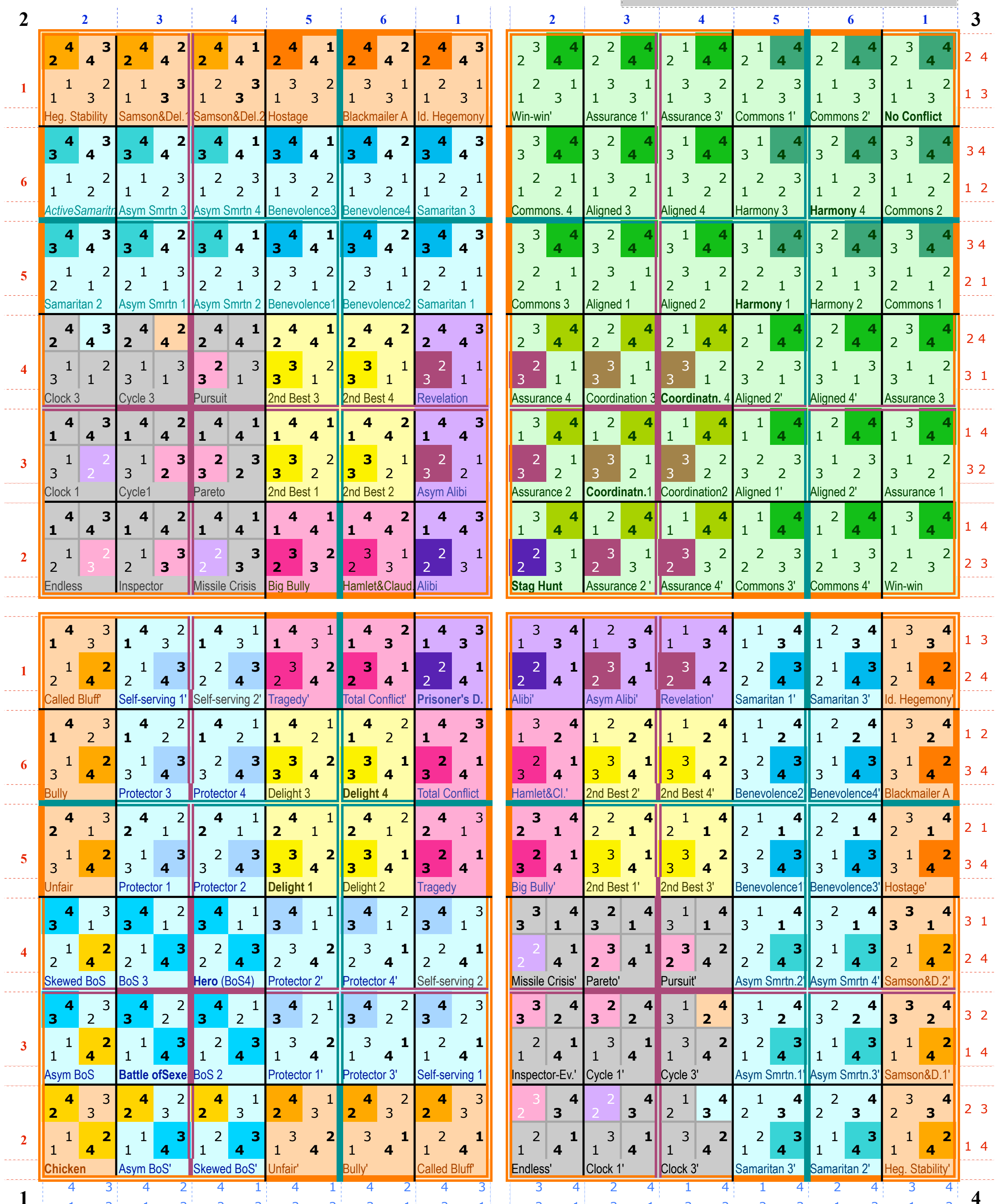
3. Topology of 2x2 Ordinal Games, with Payoff Families

Adjacent games are neighbors
by swapping adjoining payoff ranks
1-2 swaps link tiles of 4 games
2-3 swaps join tiles in 4 layers
3-4 swaps weave across layers
through 6 Pipes and 6 Hotspots
Robinson-Goforth Numbering: Layer-Row-Column Pd=111

Column Payoffs
Row payoffs
Pareto optimal
Pareto-inferior
Prisoner's Dilemma
Nash equilibrium
(Maximin for cyclic)

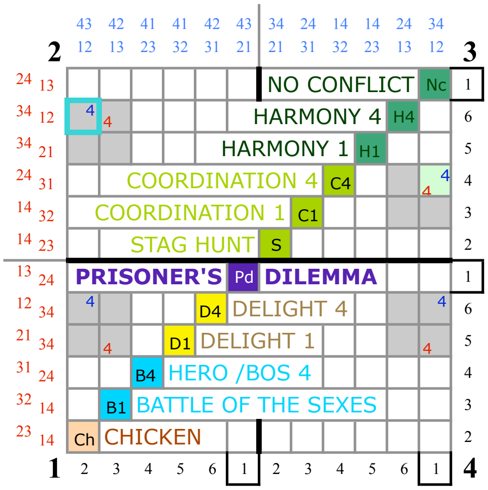
2 3
1 4
Each layer is a torus
table is a torus

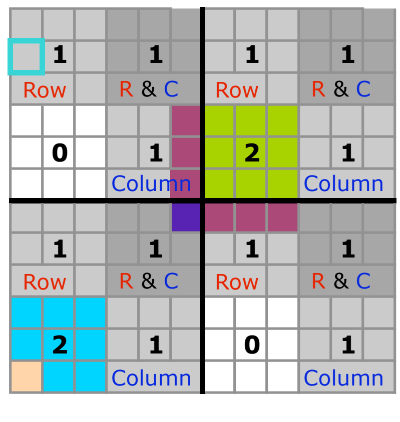
1.WIN-WIN 4-4
Stag Hunt | Harmonious
2.BIASED 4-3
Samaritan | Self-serving
Batttles of the sexes
3.SECOND BEST 3-3
4.UNFAIR 4-2
5.Pd Family | Pd 2-2
Alibi 3-2 | Tragic 3-2
6. CYCLIC

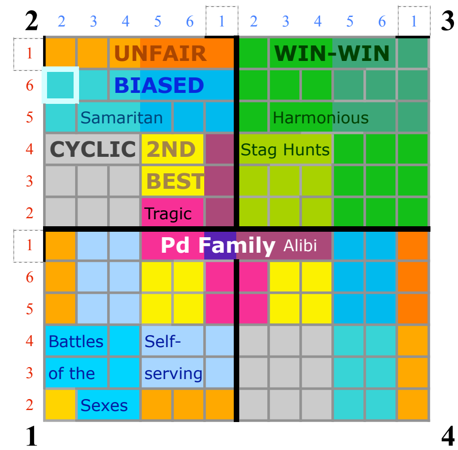


To find a game: Make ordinal 1<2<3<4. Put Row's 4 right; Column's 4 up. Row's 4 up=Layers 2&3; down=1&4. Column's 4 left=1&2. Right=3&4.Find Row & Column payoffs. Game is at intersection.

4. Structures in the Topology of 2x2 Ordinal Games







a. Twelve Symmetric Games on Diagonal

Row payoffs same across row, column same down

See David Robinson and David Goforth. 2005

The Topology of the 2x2 Games: A New Periodic Table

b. Major Regions in the Topology

Number of Nash Equilbria; and **Row** and **Column** dominant strategies

c. Payoff Familes and Subfamilies

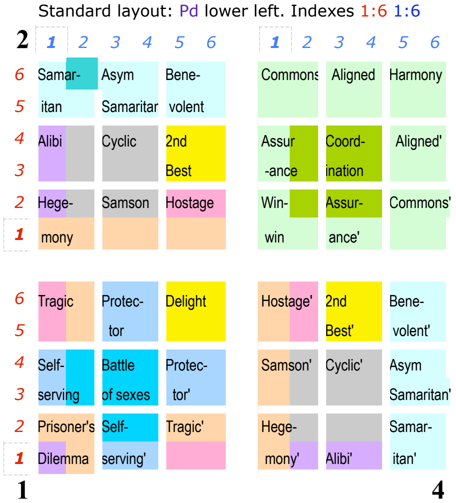
Tragic subfamily added to Robinson-Goforth Pd family

Added families: Second Best, Biased, and Unfair

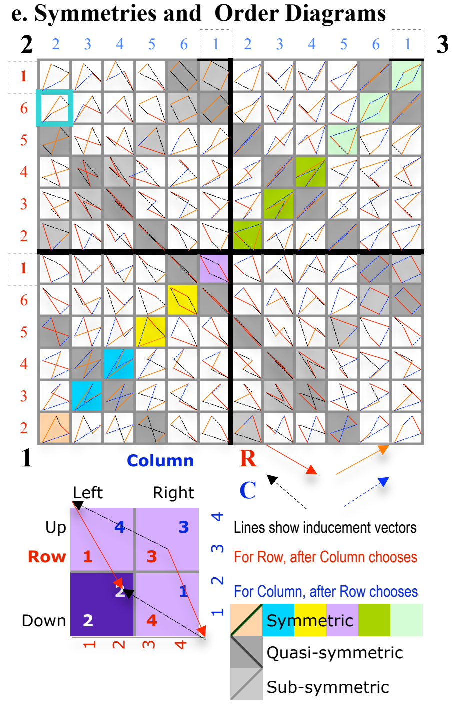
Subfamilies:Samaritan, Self-serving, and Harmonious

d. 4 Layers and 36 Tiles

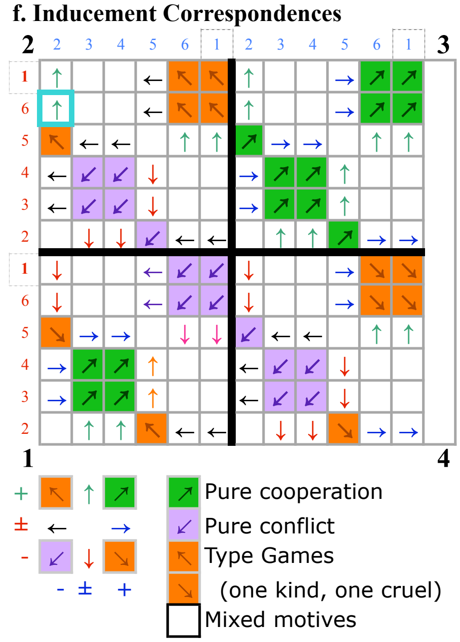
Standard layout: Pd lower left. Indexes 1:6 1:6



e. Symmetries and Order Diagrams



f. Inducement Correspondences



For inducement correspondences, see:

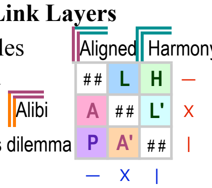
Greenberg, J. 1990. *The Theory of Social Situations*

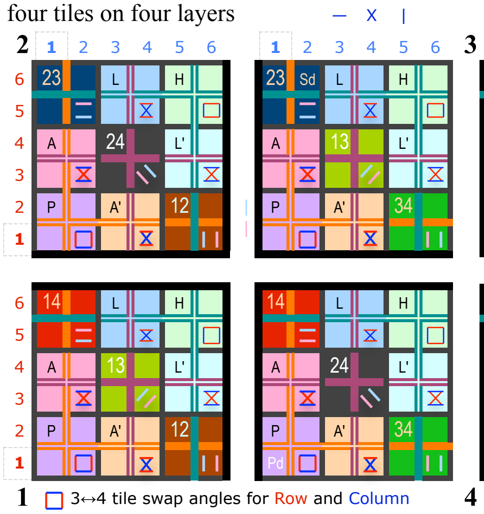
For Type games, see Robinson & Goforth 2005

g. Hotspots and Pipes Link Layers

6 hotspots double-link tiles on two layers, numbered by layers linked ##

Six pipes link four tiles on four layers

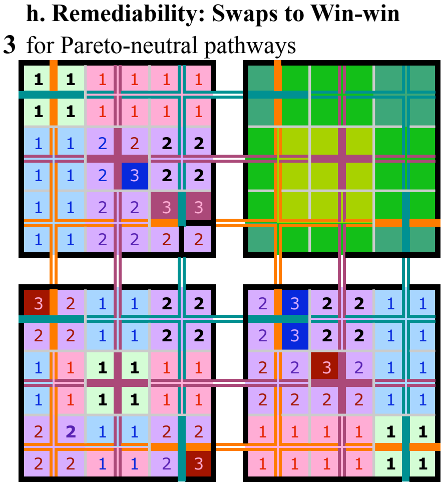




3↔4 tile swap angles for Row and Column

Inside tiles: ↕ row swaps switch row. ↔ column swaps switch column

h. Remediability: Swaps to Win-win for Pareto-neutral pathways



Bold - both players have pathways

Active Samaritan's Dilemma is Game 262 on the left edge of the second row from top. Row has a dominant strategy. There is one Nash Equilibrium, which is biased against Row (3,4). Twelve games, on three tiles on Layer 2, share the same outcome, as do twelve reflections on Layer 4, composing the Samaritan subfamily. In this game, Row's inducements are all positive,Column's mixed Swapping two top payoffs (3↔4) transmutes it into a Commons Game (362) on Layer 3. Equivalent Row swaps for other games on the same tile form a hotspot linking Layers Two and Three. Games on the Samaritan's Dilemma tile can be transmuted to Win-win games through 3-4 swaps for either player.

Layer One has the games which have been most interesting for game theorists, those where players' top preferences are most misaligned, located in opposing combinations of strategies, in diagonally opposite cells of the normal form matrix display. Layer One includes most of the famous symmetric games that game theorists have concentrated on, and a variety of others, mostly asymmetric, which have received less attention:

- Prisoner's Dilemma
- Chicken
- Eight Battle of the Sexes games, only two of which are symmetric
- Four Second Best games, including the game Binmore (Binmore 2007) calls Prisoner's Delight
- Four tragic, little-known cousins to Prisoner's Dilemma, with poor payoffs and which lack any Pareto-superior alternative to the Nash Equilibrium
- Seven Unfair games[5]
- Twelve Self-serving games

Layer Four is a mirror image of Layer Two, switching the payoffs for Row and Column. The symmetric games are located on Layers One and Three along a diagonal from lower left to upper right, which forms an axis of symmetry for the table. The twelve symmetric games on the diagonal (Figure 4a), together with the sixty-six above or below, make up the seventy-eight "unique" 2x2 games identified by Rapoport and Guyer (1966; Rapoport, Guyer, and Gordon 1976). Showing the full topological structure requires all 144 games formed by the combinations of payoffs for Row and Column. These in turn represent the full set of 576 strict ordinal games that could be created by swapping rows or columns, or both.

The topology groups games with similar properties. As in Figure 3, this can best be shown by putting Prisoner's Dilemma near the center, in the upper right corner of the layer. Each layer has three upper rows of games with a dominant strategy for Row, and three columns on the right where Column has a dominant strategy, and the overlapping area in the upper right where both have dominant strategies. This creates regions with zero, one or two dominant strategies. The four blocks with no dominant strategies include Battle of the Sexes and Stag Hunt games with two Nash Equilibria and cyclic games with no Nash Equilibria.

The colored strands running between tiles reveal the links between layers created by swaps in the top two preferences. On Layer One, for the left two Battle of the Sexes games, swapping Column's top two payoffs transforms them into Stag Hunt games on the right side of the Coordination Tile. Similarly, the right two games on the Battle of the Sexes Tile (to the right of the vertical purple line) transform into the left two Coordination

---

5 Chicken and the Battles of the Sexes variants with unequal payoffs at Nash Equilbria form a border zone between the pure Battles of the Sexes and the Unfair games. In categorizing payoff families, Chicken has been grouped with the other games that have 4-2 Nash Equilibria, while the impure Battles of the Sexes, which have a 4-3 Nash Equilibrium are included in the Battle of the Sexes Family.

games (to the left of the purple line running through the tile). Swaps in the top two games also transmute them into the bottom two games on the Coordination Tile, and the bottom two games in the Battle of the Sexes Tile transmute into the top two games on the Coordination tile. These double-links in swaps between the Battle of the Sexes and Coordination Tiles form a *hotspot*. The links weave around the colored strands, as if woven around warp and weft strands in cloth. The two cyclic tiles are similarly double-linked in another hotspot, joining Layers Two and Four.

The Samaritan's Dilemma Tile is part of a hotspot that links to the Common Interest Games on Layer Three (hotspots are shown Figure 4g). (Again, note that in the display with Prisoner's Dilemma near the center, the Samaritan's Dilemma and Common Interest tiles appear split across the edges of the (torus-shaped) layer.) A reflected equivalent to the hotspot linking the Samaritan’s Dilemma tile and the Commons tile hotspot links Layers Three and Four. The other two hotspots are similarly positioned on the lower two layers. One hotspot links unfair and tragic games on Layers One and Four, and its reflected equivalent links Layers One and Two.

While hotspots double-link two layers, *pipes* link a pile of four tiles across four layers. From the Harmony Tile on Layer Three (located above and to the right of the Coordination Tile) (Figure 4d), column swaps in the top payoffs transform into games on the Benevolence Tile on Layer Two, while row swaps transform into games on the Benevolence' Tile on Layer Four. From the Benevolence Tile on Layer Two, row swaps lead to games on the Delight Tile. Similarly, from the Benevolence' Tile on Layer Four, column swaps in the top two payoffs also lead to the Delight tile, completing the process of stitching together four tiles on four layers, in what is here labeled the Harmony Pipe. (Robinson and Goforth refer to this as the Anti-PD Pipe.) Two similar pipes lie to the left of and below the Harmony Pipe, joining the Aligned Tiles, the Asymmetric Samaritan's Dilemma Tiles, and the Protector Tiles.

Another pipe links the Prisoner's Dilemma Tile with pairs of Unfair and Alibi games on Layers Two and Four and with a Stag Hunt and three Harmonious games on Layer Three. In the display with Prisoner's Dilemma near the center, the games in this Prisoner's Dilemma Pipe lie on the corners of each layer.

The relationships among hotspots and pipes are easier to see in the display with Prisoner's Dilemma in the lower left corner, as in Figure 4g. This display format is also the basis for the numbering system, in which Prisoner's Dilemma is game 111, with indices proceeding according to layer, row, and column. In this "standard" display, the hotspots make a diagonal from upper left to lower right in each layer. The Harmony Pipe is in the upper right corner of each layer, with the two Aligned Pipes to the left and below. Symmetrically, The Prisoner's Dilemma is in the lower left corner, with the two Alibi pipes lie above and to the right.

The structure of hotspots and pipes maps the transmutations between games created by swaps in the highest two payoffs, and is further discussed in the next section. Overall, the topology display not only shows relationships among game properties, but also provides a tool for understanding the changes that would result from swaps in adjoining payoffs, whether accidental or due to deliberate institutional design. For any game, it is possible to see the full set of possible transformation through row or column swaps in adjoining payoffs, with swaps in the lowest two payoffs on the same tile, swaps in middle payoffs on adjacent tiles, and colored strands marking the swaps in top payoffs in pipes or hotspots that weave together layers.

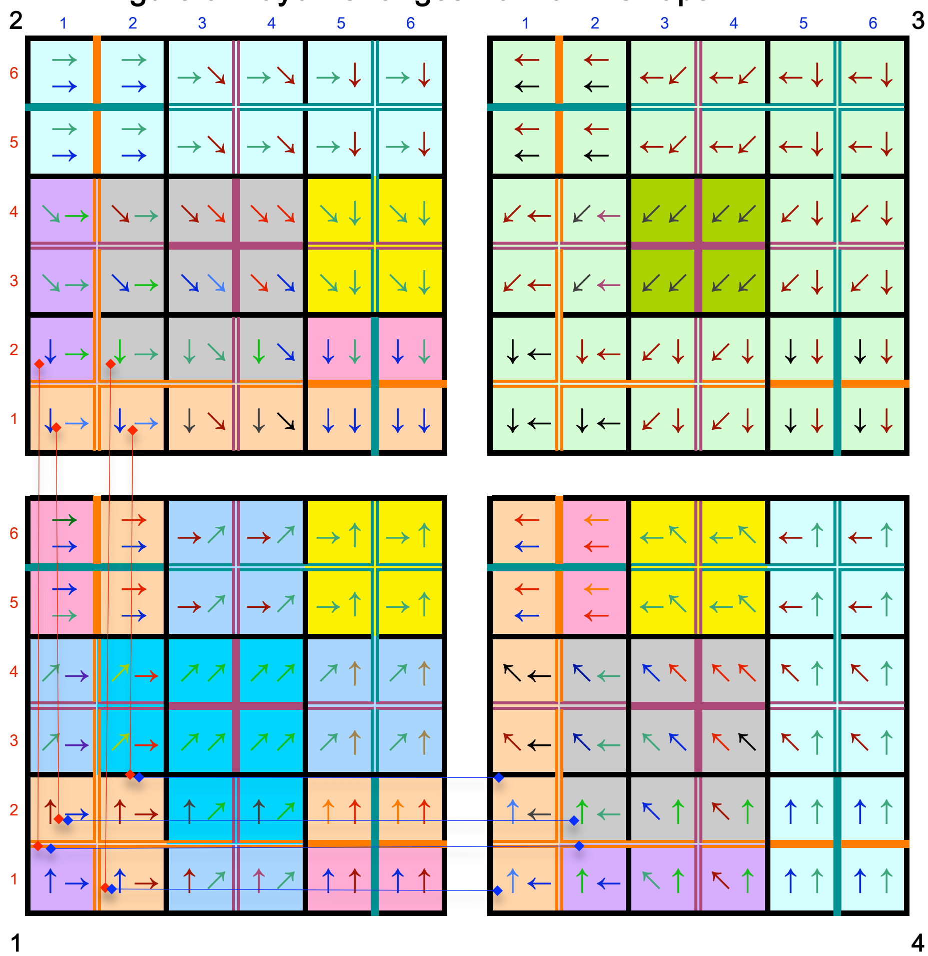
TRANSMUTATIONS

The topology provides a systematic way of looking at transmutations between games resulting from switching adjoining payoff ranks. As discussed earlier, swapping the recipient's (Column's) highest two payoffs converts the Active Samaritan's Dilemma (Game 262) into a game of Common Interest (Game 361). The other games on the same tile also transmute into games on the Commons Tile for swaps in the donor's (Row's) payoffs, forming a hotspot that links Layers Two and Three. (The destination for the row swap is located on the opposite side of the tile, as indicated by the colored strand running through the middle of the tile).

For the games on the adjoining Asymmetric Samaritan Tile, swaps in Column's top two payoffs transmute into win-win games on the Aligned Tile. Games on the Benevolent tile have dominant strategies for both players, and swaps in the top two column payoffs transmute into Win-win games on the Harmony Tile, which also has two dominant strategies. Thus, all twelve games in the Samaritan's Dilemma Family on Layer Two can be converted into Win-win games through a single swap in the donor's (Row's) Payoffs. Equivalently, for the reflected versions on Layer Four, where Row and Column exchange positions, a single column swap in the highest payoffs transmutes them into Win-win games. The topology shows that this potential for transformation is not a unique or isolated potential, but instead characterizes a whole family of Samaritan's Dilemma games. However, only the games on the Samaritan's Dilemmas tile are part of the hotspot with double-links. It is only for these games that a swap for Column also creates a win-win game.

The same potential for conversion exists for the two Unfair games located above the Samaritan games, and the other games that share the same pattern of column payoffs. These make up two vertical *slices,* with twelve games. So a slice of three tiles, Unfair, Alibi and Cyclic games, all can be transformed by a single row swap in the two highest payoffs. Three of these turn into Stag Hunt games, and the rest become Harmonious Games. While the potential for transformation may be obvious simply by looking at the shared structure of column payoffs, the pattern of which ones become Stag Hunts, with two Nash Equilibria, and which ones become Harmonious games, with a single Nash Equilibrium, is made visible by the topology display.

Figure 5. Payoff Changes from 3↔4 Swaps



↑→ Arrows show direction of swaps between tiles in pipe or hotspot, for Row and Column

	—	X	|
—	##	L	H
X	A	##	L'
|	P	A'	##

Within tiles:

row swaps switch rows ↕

column swaps switch columns ↔

Arrow colors show payoff changes for players		other player				
		+	+/	=	-/=	-
player with swap	+	+:+		+:=		+: -
	+/=		↗↗		←	
	=	=:+		=:=		=: -
	-/=		↑		↙↙	
	-	- :+		- :=		- : -

Green and blue swaps help one or both players without hurting either, (are Pareto-efficient) and so are likely to be acceptable to both players and provide pathways to Win-win games

The payoff changes from transmutations in the two highest payoffs are shown in Figure 5. Arrows indicate the direction of transmutations between tiles within hotspots and pipes (as in the structure shown in Figure 4g). Within tiles, row swaps switch rows, and column swaps switch column, i.e. end up on the other side of the colored strands running through the middle of the tile. Arrows are colored according to the change in payoff for the player with the swap. Green and blue arrows show transformations that help one or both player and hurt neither, and so are likely to provide acceptable pathways to Win-win games.

For Layer Three, all swaps in the highest two payoffs would result in a lower Nash Equilibrium payoff for one or the other player. This is of course inherent in the structure of the topology, in which Layer Three has all the games where both top payoffs are in the same cell (4,4), and emphasizes how this layer offers an attractive zone of stability. For Layer Two, as described earlier, the upper row of tiles (upper two rows of games), and left column of tiles all transmute to Win-win games on Layer Three (as do the lower row and right column of tiles on Layer Four). On Layer One, the middle row and column of tiles, making up the Battles of the Sexes and Protector games, transmute to Layer Three, while the four tiles on the corners do not (Prisoner's Dilemma, Delight, and Tragic tiles). While swaps for an individual game can be found directly by looking at the game (most obviously for those with a biased, 4-3 Nash Equilibrium, but less easily for other payoff patterns) the display makes it possible to see the pattern of hotspots and pipes, and how transmutations to Win-win games relate to other game characteristics.

On Layer Two, over half the games (20/36; 55%) can be transmuted into Win-win games through a single swap. The other games transmute to Layers One and Four, mostly obviously for the 12 hotspot and 24 (cyclic) hotspot tiles. For the Delight games, column swaps go to Protector games on Layer One, while row swaps go to Samaritan games on Layer Four. For the remaining tile of Cyclic and Unfair games, the patterns of payoffs mean that row swaps turn the cyclic games turn into Protector games, and the Unfair games into impure Battles of the Sexes, while column swaps link to cyclic and Prisoner's Dilemma family games on Layer Four. Transmutation linkages for Layer Four are the mirror images of those on Layer Two.

For Layer One, the Protector and Battle of the Sexes (pure and impure) games transmute directly to Layer Three, (also 20/36; 55%) while none of the other games does. (This is consistent with the classification of Chicken as an Unfair game, based on its payoff outcome, even though it has two Nash Equilibria, like the Battles of the Sexes.) The Delight games transmute into Samaritan games, through a row or column swap. From there another swap, for the other player, converts into Win-win games.

As discussed by Robinson and Goforth, the Prisoner's Dilemma pipe has the most diverse mix of games and transformations. As shown in Figure 5 (standard display), and in Figure 3 (periodic table display, with Prisoner's Dilemma near the center), single row or column swaps from Prisoner's Dilemma transform it into an Alibi game (221 or 212), still in the same Prisoner's Dilemma family. A second swap then transmutes into a precarious Stag Hunt (322). If both defect, they get their third-ranked payoff (2,2), an outcome as bad as in Prisoner's Dilemma or the Alibi games (221, 412). If their strategies are not coordinated, or if one player defects, the defecting player gets their second-ranked outcome and the other their lowest.

Chicken requires swaps in the highest payoffs for both row and column before it can be converted into the Win-win game of No Conflict (311). This pathway first converts Chicken into an asymmetric Unfair game with only one Nash Equilibrium (212 or 421), after which swaps for the other player create a well-aligned Win-win game of No Conflict (411). If the situation for irrigation aid were equivalent to Chicken, where investment by one party is sufficient, then a solution through transforming the game to align incentives would require changes for both players. The relevant comparison would be with the potential for the second-best (3,3) solution, where investment might be less efficient but more balanced, and with a lower risk of falling into the worst outcome.

For the other two unfair games on the Prisoner's Dilemma tile (112 or 121), one swap converts it into a cyclic game, and the other into an unfair game on Layer Two or Four (212 or 411). A further swap then transmutes into a Win-win game. For all four games on the Prisoner's Dilemma tile, two swaps in the highest payoffs can turn it into a Win-win game. Prisoner's Dilemma

Passive Samaritan's Dilemma (Game 132) can be converted into a Win-win game (Assurance 3: 343) through a single swap. However, this is equivalent to the solution where the donor reconciles themselves to investing alone (or, more realistically, making a much higher level of investment). An alternative pathway would require two column swaps, first transforming into a Battle of the Sexes, and then into a Win-win Stag Hunt game (Coordination 2: 334). These two swaps would be equivalent to a single swap in non-adjoining preferences, 2↔4). This pathway would represent a situation where the farmers end up preferring to invest alone, perhaps due to an increase in their capability and demand, or much lower expectations about the results if the government invests alone. This would be in accordance with the original preference structure for Passive Samaritan's Dilemma, where the government prefers that farmers invest alone (rather than preferring joint investments, as in Active Samaritan's Dilemma).

Figure 4h summarizes the minimum number of swaps needed to transmute different games into Win-win games. On Layers Two and Four, thirteen games can be transformed into Win-win games through two Pareto-neutral swaps, pathways that improve or preserve payoff ranks for both players. For the Cyclic, Samson and Hegemonic Tiles, there is no feasible switch directly to Win-win through swaps in the top payoffs, instead the pathway passes first through a swap in middle payoffs onto an adjoining tile. For two Tragic games (Big Bully, 225 and Hamlet and Claudius, 226), and one Cyclic Game (pure Pareto, 234) three swaps are required to reach Win-win. A pair of Tragic games on Layer One (Total Conflict, 116 and 161) also require three swaps to reach Win-win. These are thus, in terms of Pareto-efficient swaps, the hardest games to improve, the most difficult to remedy. In elementary terms, comparison of the number of swaps needed to create a Win-win game illustrates *remediability*, in Williamson's (1986) sense of the actual feasibility of improvement.

The most difficult games to transmute to Win-win are not Prisoner's Dilemma and Chicken, as might have been expected given the preoccupation of game theory with the challenges posed by these games. Prisoner's Dilemma and Chicken require two swaps, though each player must have a swap. Instead, the most difficult games to transmute to Win-win are Tragic games on hotspots linking Layers One to Layers Two and Four, plus a cyclic game composed purely of Pareto-optimal payoffs. On Layer one, the pair of Total Conflict games require three swaps. The remaining Tragic game (Tragedy, 115 and 151) lies next to a Protector game, to which it can be converted through a middle swap, and then another swap, for the other player, converts it into a Win-win game (Assurance 3', 314).

Three fourths of the most difficult games to remedy, (Pareto, Big Bully and Total Conflict, and their mirror images: 234 and 443, 225 and 452, and 416 and 461) are games of pure conflict. Interests are completely opposed; each gain of one payoff rank by one player reduces the payoff to the other by one rank. This is most visible in the order diagram display, where these games form a line, as a move up by one player are matched by moves down by the other.

By contrast, the Samaritan's Dilemma games, located on the hotspots linking Layer Three to Layers Two and Four, offer the easiest opportunities for improvement, with swaps in the highest payoff for either player converting into a Win-win game. The topology structure helps to put this potential into context, and show that it is relatively unique, with only the Battle of the Sexes Games having a similar pair of transmutations available for either Row or Column.

SOLVING SAMARITAN'S DILEMMAS IN IRRIGATION AID

The thesis of this paper is that solving Samaritan's Dilemmas in irrigation aid might be accomplished by improving the payoffs to irrigators from joint investment, as examined both in terms of game theory models and policy recommendations. As stated earlier, in most cases, once government has provided aid, it is unrealistic for farmers or government to expect that government will never again provide aid. Making a credible commitment to restrict future eligibility for aid is difficult, since maintenance, repair, and rehabilitation are fungible, floods and landslides create force majeure events that farmers may be unable to repair on their own, farmers and agency officials will have incentives to find loopholes in restrictions, and farmers may use incomplete or inadequate government performance by government in rehabilitation and river water allocation to claim they need aid. For Samaritan's Dilemma, more fertile ground for solutions lies in shaping joint investments, so that these will be attractive for irrigators.

Transforming Samaritan's Dilemma into a game of common interest provides a formal model for how the solution would work, by raising the payoff to joint investment until it is higher than the payoff where farmers make little or no investment. This could occur through a range of factors, which are summarized in Figure 6 and discussed below. Depending on conditions, it may be possible to adjust one or several of these factors sufficiently to swap payoffs. Large internationally-funded irrigation rehabilitation projects seem to exhibit the most pathological forms of perverse incentive structures, while other programs, funded by national budgets or smaller donors, often already contain more elements that make joint investment productive.

Make irrigation more profitable. Irrigation management transfer has been more successful in areas where agriculture is more commercialized. Farmers have more financial capacity, and more at stake to lose if irrigation performance is poor. This is likely to increase the productivity of joint investment (as well as encouraging more investment by farmers on their own).

Improve policies and procedures for future joint investment. Many IMT projects have concentrated on building local capacity, and "one last rehabilitation" without clarifying the specific criteria, procedures, and resources that would be used for any joint investment in the future. Recognizing that there will be continued interaction, both for financial aid and for other matters such as technical advice and river basin water allocation, makes it clearer why institutions for co-management need to be developed. Developing co-management institutions, including those for joint investment, may increase the expected benefits to farmers from joint investment, and reduce the risk that, deliberately or not, they shirk on preventive maintenance and minor repairs.

Invest incrementally. A one-shot interaction maximizes the temptation to behave opportunistically (Bruns 2008). Repeated interaction alone is not sufficient to solve Samaritan's Dilemmas. However, it can help build the social capital of working relationships, increase the importance of a reputation for keeping commitments, and make threats of sanctions more credible. Furthermore, large rehabilitation projects overwhelm the more limited resources that farmers can mobilize in a short period of time. Smaller incremental investments make farmer cost-sharing more feasible, and make it easier for irrigators to have a greater role in carrying out and supervising works.

*Strengthen cost-sharing mechanism*s. The simplest and most effective kind of conditionality occurs where government provides some crucial materials, such as cement, steel reinforcing rods, and gabion wire, and leaves it up to communities to provide sand, gravel, stones, and labor. This works most easily in mountainous areas and more subsistence oriented economies, where local materials and labor are more available. Matching formulas based on cash contributions are more susceptible to manipulation, but can be strengthened by requiring communities to mobilize funds in advance, providing government funds in tranches based on construction progress and fulfillment of local cost-sharing commitments, and by the broader shift to a strategy of incremental investment, to that local cost sharing is more feasible.

Make budget constraints transparent. A two-player model of subsidies, whether Samaritan's Dilemma or Chicken, may be framed as if government's resources and capacity to provide aid were unlimited. However, in the face of multiple demands, government budgets are usually far less than needs. Increased transparency can reduce the risks of opportunism based on mistaken assumptions that government will always do what farmers cannot. Competitive budget allocation, in an open, participatory process further encourages comparison of alternatives and awareness of budget constraints.

Improving the payoffs from joint investment offers a pathway for solutions, but still leaves plenty of scope for bargaining about how joint investment may be arranged. The relevance of joint investment as a solution depends on the underlying production functions, whether the situation can be arranged so that joint investment will pay off better (as contrasted with a situation where investment by one party is enough, for which Chicken is a better model, and a transformation into a Win-win solution would require changing payoffs for both sides, not just one).

Potential solutions through realigning incentives for joint investment would also still be vulnerable to other problems. Principal-agent conflicts may still lead agencies and specific office-holders to try to extract rents, keep budgets secret, avoid accountability, and prefer large projects implemented by big contractors. Collective action problems may make it hard for communities to participate, mobilize resources, and manage infrastructure. Nevertheless, given the importance of irrigation for farmers' livelihoods, the political benefits of aid spread widely, and the potential benefits of joint investment, it seems like it should be feasible under some circumstances to craft solutions that encourage joint investment.

Figure 6. Comparison of Irrigation Rehabilitation and Co-Adaptive Co-management

Irrigation Rehabilitation	**Adaptive Co-management**
Single-play	Repeated transactions
Lumpy rehabilitation	Incremental aid
Single solution	Problem-solving
Pre-commitment/Ultimatum	Negotiate joint decision
No budget constraint (or not revealed)	Transparent competition
Independent action	Shared investments
Two-player	Multi-player

Adapted from Bruns 2008

CONCLUSIONS: TRANSMUTING SAMARITAN'S DILEMMAS

Aid may discourage self-help, posing a challenge for institutional design. In the case of irrigation aid, perverse incentives contribute to vicious cycles of inefficient rehabilitation.

Attempts to escape this through transferring irrigation management to irrigators have been difficult, since it is hard for governments to make credible commitments to not provide aid in the future.

Buchanan's models of Samaritan's Dilemma illustrate incentive structures that can lead aid recipients to reduce investment. However, his proposed solutions depend on deception, implausible threats, or agents bound by externally-enforced contracts. Buchanan's comparison of two games that differ by swapping two payoffs raises the question of how incentives might be transformed by other swaps in payoffs.

The topology of 2x2 ordinal games found by Robinson and Goforth provides an elegant framework for understanding the relationships among games by swapping adjoining payoffs, including the pathways for transforming Samaritan's Dilemmas and other collective action problems into Win-win games. Buchan's Active Samaritan's Dilemma is part of a large family of similar games, where one player has a dominant strategy that leads to getting their second-choice outcome. For Samaritan's Dilemma and the most similar games on the same tile, a swap in the two highest payoffs *for the other player*, transmutes the game into a Commons game where both get their most preferred outcome, a Harmonious Win-win game.

Analysis of swap transmutations for Samaritan's Dilemma and similar games provides a simple model not only for comparing incentive structures within games, but also for analyzing the potential for transmutations between games. Samaritan's Dilemma turns out to be one of the easiest games to change to Win-win, since this can be done through a single swap for either player. Most games can be converted into Win-win games by one or two swaps. However, eight games require three swaps to reach Win-win: three Tragic pairs and a pair of cyclic games. Three fourths of those are games of pure conflict, where the interests of players are diametrically opposed, with a gain in payoff rank for one player being precisely offset by a reduction in payoff rank for the other. These make up the most difficult, or most irremediable, games, in terms of transmutation into Win-win games.

In the case of irrigation aid, rather than emphasizing policies for local maintenance, and hard-to-believe threats of no future aid, as has occurred in many IMT projects, a better alternative might be to make joint investment more attractive, i.e. changing the game into one where incentives are better aligned. This takes advantage of the relatively close alignment of interests which already exists, and the somewhat unique potential to make a transition not only through changing payoff ranks for the donor, but alternatively through changes that would swap the two top preferences for the farmers. In practical terms, this could be done by making irrigated agriculture more profitable; clarifying policies and procedures for joint investment, particularly by strengthening cost-sharing mechanisms; developing capacity for joint investment during project and program implementation; increasing transparency about budget constraints; and shifting from Infrequent large-scale rehabilitation projects to programs that aid incremental joint investment.

LITERATURE CITED

Araral, Eduardo. 2005. Bureaucratic Incentives, Path Dependence, and Foreign Aid: An Empirical Institutional Analysis of Irrigation in the Philippines. *Policy Sciences* 38: 131-157.

Araral, Eduardo K. 2006. Decentralization Puzzles: A Political Economy Analysis of Irrigation Reform in the Philippines. Ph.D. Dissertation, Indiana University Bloomington.

Bardhan, P. 1993. Analytics of the Institutions of Informal Cooperation in Rural Development. *World Development* 21, no. 4: 633–639.

Binmore, K. G. 2007. *Game theory: a very short introduction*. Oxford University Press, USA.

Bruns, Bryan Randolph. 2008. *Aiding Adaptive Co-management in Irrigation*. Presented at Governing Shared Resources: Connecting Local experience to Global Challenges. 12th Biennial Global Conference of the International Association for the Study of the Commons. Cheltenham, England, July 14.

Bruns, Bryan Randolph. 2010. Transmuting Social Dilemmas into Win-win Games: Payoff Families in the Topology of 2x2 Ordinal Games. Transmuting 2x2 Games Working Paper Number 1. http://dlc.dlib.indiana.edu/dlc/handle/10535/5969

Buchanan, James. 1977. The Samaritan's Dilemma. In *Freedom in Constitutional Contract: Perspectives of a Political Economist*, ed. James Buchanan, 169-185. College Station: Texas A&M University Press.

Gibson, Clark C, Krister Andersson, Elinor Ostrom, and Sujai Shivakumar. 2005. *The Samaritan's Dilemma: The Political Economy of Development Aid*. Oxford: Oxford University Press.

Kollock, P. 1998. Social dilemmas: The Anatomy of Cooperation. Annual review of sociology 24, no. 1: 183–214

Ostrom, Elinor. 1997. Crossing the Great Divide: Coproduction, Synergy, and Development. In *State-Society Synergy: Government and Social Capital in Development*, ed. Peter Evans, 85-118. Berkeley: University of California Press.

Ostrom, Elinor, Larry Schroeder, and Susan Wynne. 1993. *Institutional Incentives and Sustainable Development: Infrastructure Policies in Perspective*. Boulder: Westview Press.

Rap, Edwin. 2006. The Success of a Policy Model: Irrigation Management Transfer in Mexico. *Journal of Development Studies* 42, no. 8: 1301 - 1324.

Rapoport, A. 1967. Exploiter, Leader, Hero, and Martyr: the Four Archetypes of the 2 Times 2 Game. *Behavioral science* 12, no. 2: 81.

Rapoport, A., and M. Guyer. 1966. A Taxonomy of 2 x 2 games. *General Systems* 11, no. 1: 203–214.

Rapoport, A., M. Guyer, and D. G Gordon. 1976. *The 2 x 2 Game*. Univ of Michigan Pr.

Robinson, D. R., and D. J. Goforth. 2005. *The Topology of the 2x2 Games: a New Periodic Table*. Routledge.

Robinson, D., D. Goforth, and M. Cargill. 2007. Toward a Topological Treatment of the Non-strictly Ordered 2x2 Games

Schmidtchen, D. 2002. To Help or Not to Help: The Samaritan's Dilemma Revisited. In

Method and Morals in Constitutional Economics: Essays in Honor of James M. Buchanan, 470.
Skyrms, B. 2004. *The Stag Hunt and the Evolution of Social Structure*. Cambridge Univ Pr.
Stone, D. 2008. *The Samaritan's Dilemma: Should Government Help Your Neighbor?* Nation Books.
Taylor, M., and H. Ward. 1982. Chickens, Whales, and Lumpy Goods: Alternative Models of Public-Goods Provision. *Political Studies* 30, no. 3: 350–370.
Vermillion, Douglas. 2006. *Lessons Learned and to be Learned about Irrigation Management Transfer*. Ubud, Bali.