Social tipping our way – or maybe not – to some kind of future

Charles Efferson[1]

w/ Sönke Ehret[1], Sara Constantino[2,3], Elke Weber[3], Ernst Fehr[4], Sonja Vogt[1]

[1]**University of Lausanne**

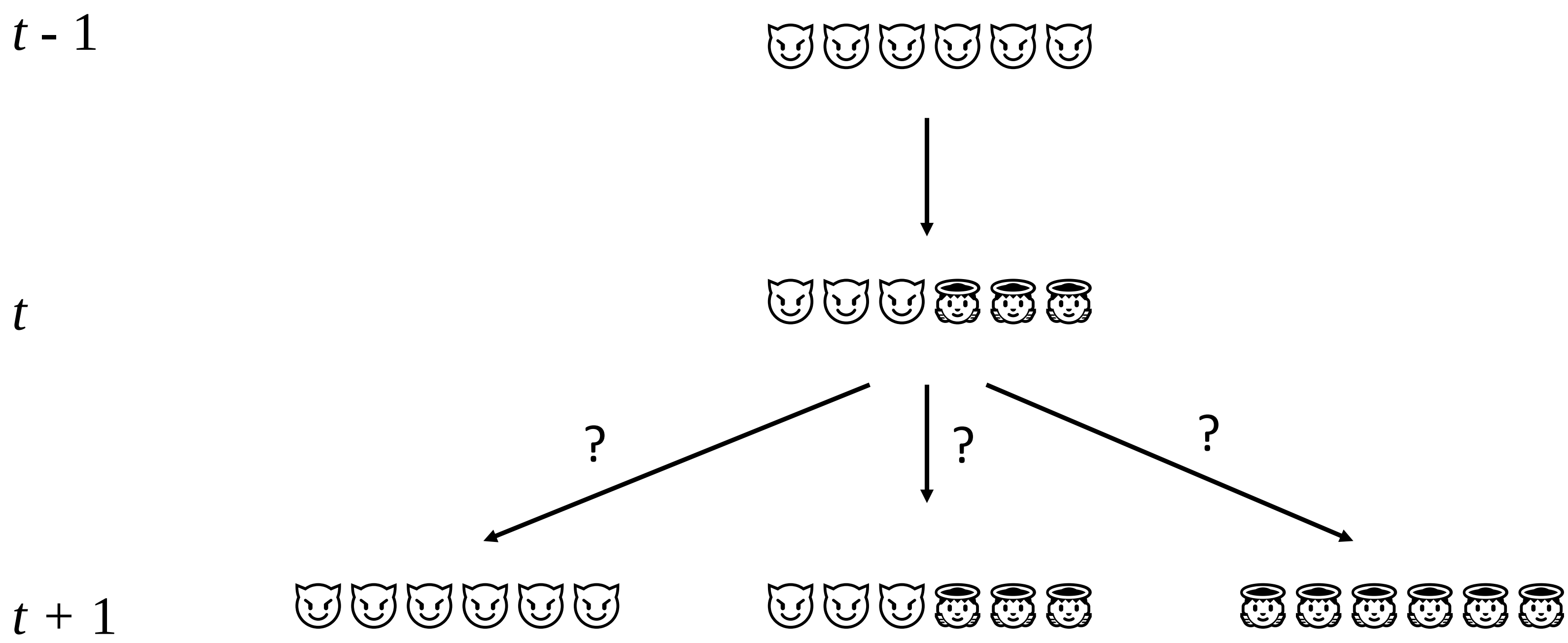
[2]**Northeastern University**

[3]**Princeton University**

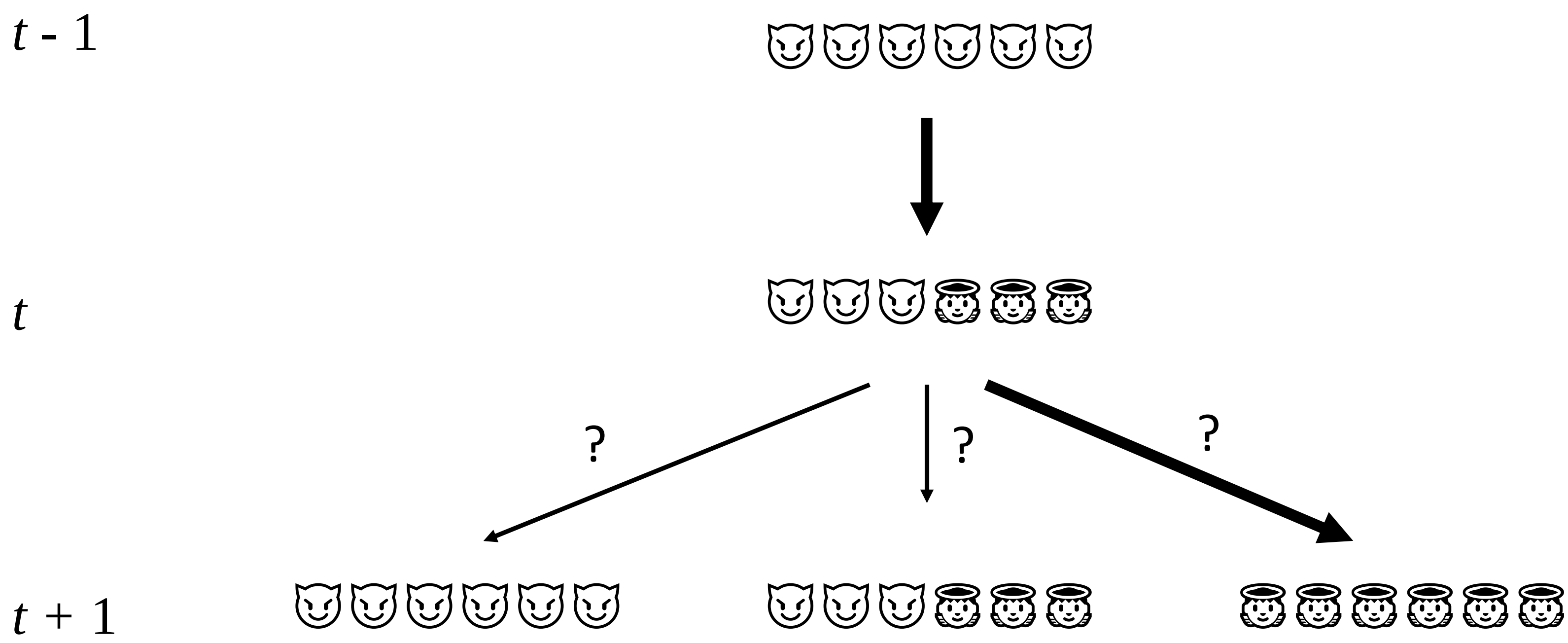
[4]**University of Zurich**

ASSA, San Antonio, January 2024

Applied Cultural Evolution in General

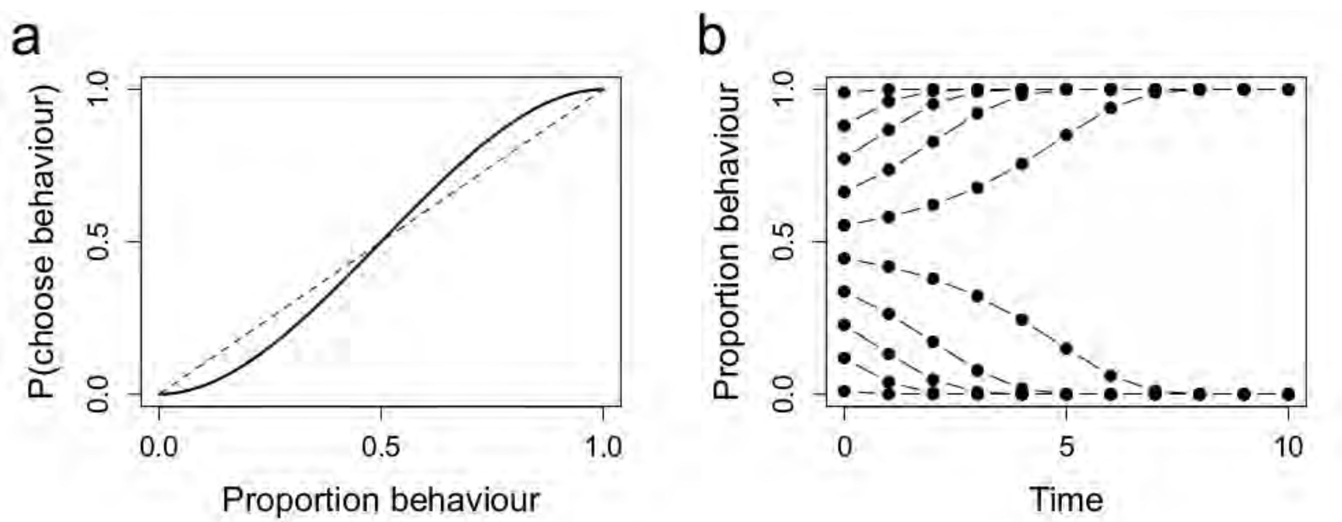


Social Tipping as One Possible Outcome



Intervention as Exogenous Trigger for Endogenous Norm Change

- Social forces working against you can switch to working for you (Nyborg *et al.*, 2016).
- Recruit cultural evolutionary processes to steer cultural evolution in a direction consistent with policy objectives.



$\Rightarrow$ Potential for stagnation and potential for rapid change go hand-in-hand!

Norms as an equilibrium selection mechanism.

One Idea, Many Domains

An Expansive Set of Domains

- Potential informs policy related to **female genital cutting** (Shell-Duncan and Hernlund, 2000; UNFPA-UNICEF, 2015; Cloward, 2016; Camilotti, 2016; Platteau *et al.*, 2018), **child marriage** (Malhotra, Anju and Warner, Ann and McGonagle, Allison and Lee-Rife, Susan, 2017; Delneuville, Amy, 2017; Lee-Rife *et al.*, 2012; Bicchieri *et al.*, 2017; Cloward, 2016), **open defecation** (Shakya *et al.*, 2015), **domestic violence** (World Health Organization, 2017; Platteau *et al.*, 2018), and a **preference for sons** (Schief *et al.*, 2021).

- Research has also highlighted the role of social influence, and in some cases its policy relevance, with respect to **smoking** (Christakis and Fowler, 2008), **foot binding** (Mackie, 1996), **alcohol consumption** (Prentice and Miller, 1993), **obesity** (Christakis and Fowler, 2007), **bullying** (Paluck *et al.*, 2016), **energy conservation** (Allcott, 2011), **tax compliance** (Hallsworth *et al.*, 2017), **resource conservation** (Castilla-Rho *et al.*, 2017; Koch and Nax, 2017; Travers *et al.*, 2021), and **climate change** (Nyborg *et al.*, 2016; Farmer *et al.*, 2019; Dávila-Fernández and Sordi, 2020; Otto *et al.*, 2020).

Tipping Experiments

Tipping at 25% (Centola *et al.*, 2018, *Science*)

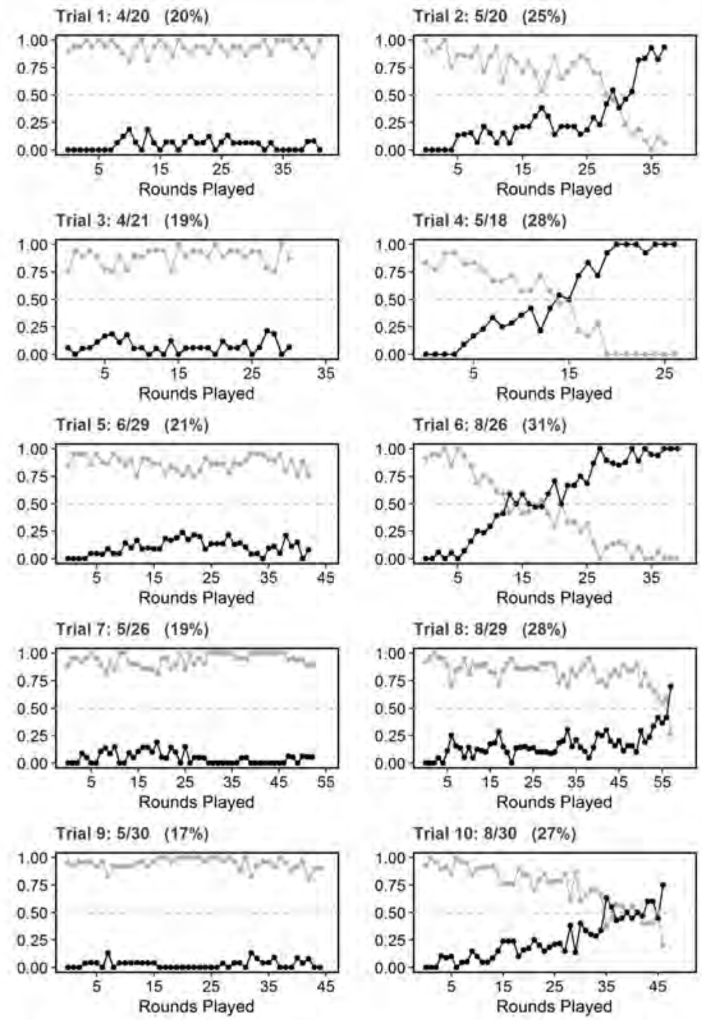


Fig. 2. Time series showing adoption of the alternative convention by noncommitted subjects (i.e., experimental subjects). Gray lines indicate the popularity of the established convention; black lines show the adoption of the alternative convention. Success was achieved when more than 50% of the noncommitted population adopted the new social convention. Trials in the left column show failed mobilization, whereas trials in the right column show successful mobilization. A transition in the collective dynamics happens when C reaches ~25% of the population. Each round is measured as $N/2$ pairwise interactions, such that each player has one interaction per round on average.

What if people differ from each other?

Heterogeneous Preferences (Andreoni *et al.*, 2021, *PNAS*)

A

Initial situation: all players prefer Blue in period 1.

Row player: PREFERS BLUE / Column player: PREFERS BLUE	*Blue*	*Green*
Blue	$v_H \, , v_H$	$v_H - p * g_t \, , v_L - p * (1 - g_t)$
Green	$v_L - p * (1 - g_t) \, , v_H - p * g_t$	$v_L \, , v_L$

B

Transition: both preferences are common in intermediate periods.

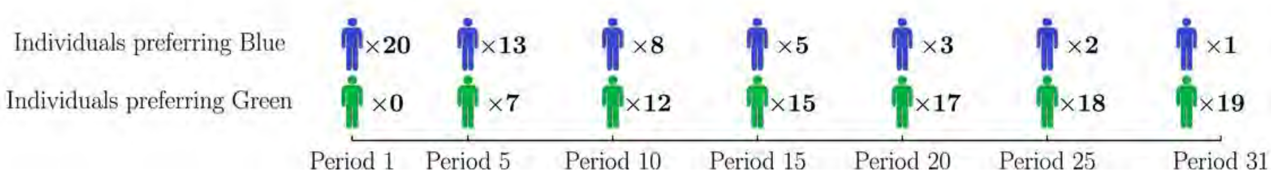
Row player: PREFERS BLUE / Column player: PREFERS GREEN	*Blue*	*Green*
Blue	$v_H \, , v_L$	$v_H - p * g_t \, , v_H - p * (1 - g_t)$
Green	$v_L - p * (1 - g_t) \, , v_L - p * g_t$	$v_L \, , v_H$

C

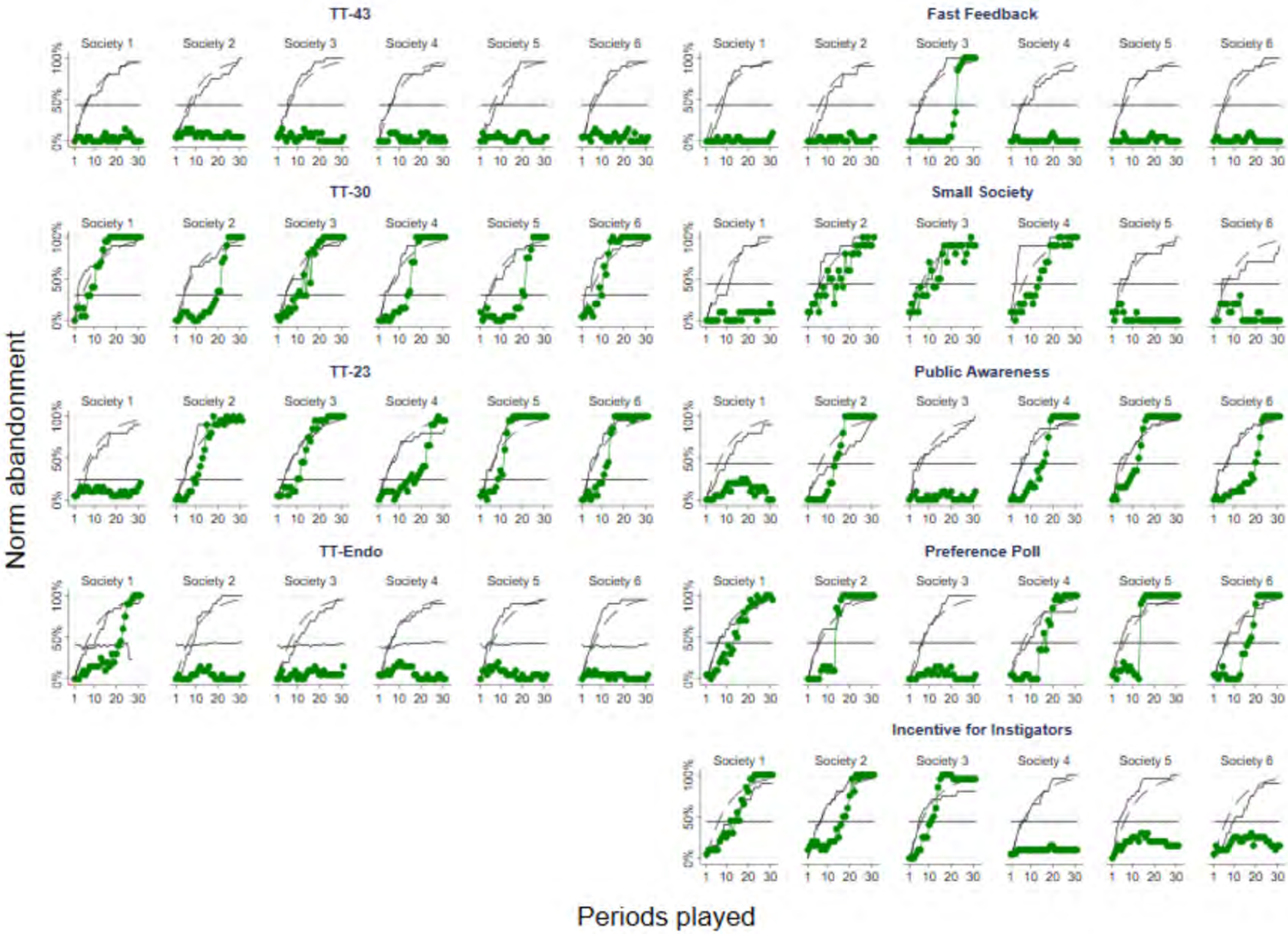
Final situation: nearly all players prefer Green in late periods.

Row player: PREFERS GREEN / Column player: PREFERS GREEN	*Blue*	*Green*
Blue	$v_L \, , v_L$	$v_L - p * g_t \, , v_H - p * (1 - g_t)$
Green	$v_H - p * (1 - g_t) \, , v_L - p * g_t$	$v_H \, , v_H$

D



A Nuanced View (Andreoni *et al.*, 2021, *PNAS*)



Mundane heterogeneity and the individual-population disconnect

(Efferson *et al.*, 2023, *Oxford Handbook of Cult Evo*)

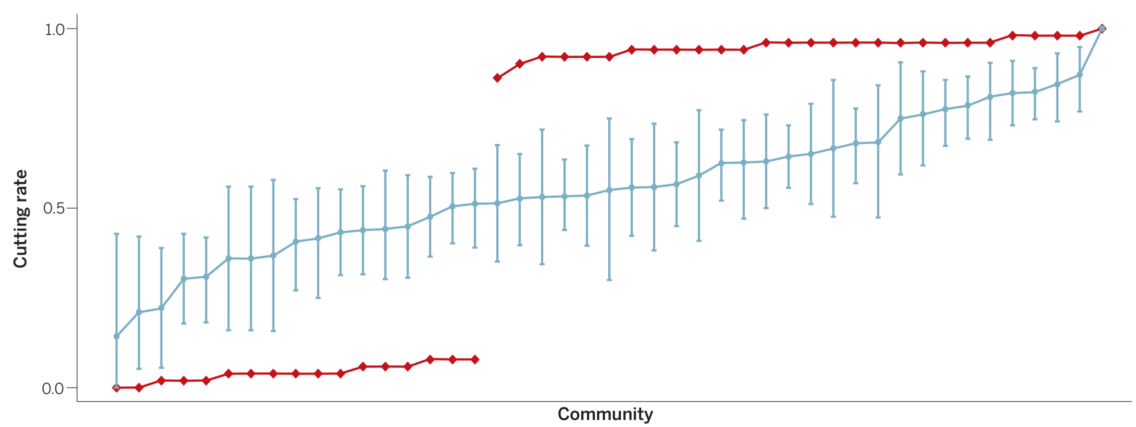
Social Tipping and Female Genital Cutting (Efferson *et al.*, 2015, *Science*)

BEHAVIOR

Female genital cutting is not a social coordination norm

New data from Sudan question an influential approach to reducing female genital cutting

By **Charles Efferson,[1*†] Sonja Vogt,[1*†] Amy Elhadi,[2] Hilal El Fadil Ahmed,[2] Ernst Fehr[1†]**



Cutting rates in Gezira communities. Red diamonds show ordered cutting rates as predicted by the coordination game model (*12*). Blue dots show actual cutting rates across the 45 communities with 95% boot-strapped confidence intervals.

Social Tipping and Female Genital Cutting (Vogt *et al.*, 2016, *Nature*)

LETTER

doi:10.1038/nature20100

Changing cultural attitudes towards female genital cutting

Sonja Vogt[1*], Nadia Ahmed Mohmmed Zaid[2], Hilal El Fadil Ahmed[3], Ernst Fehr[1§] & Charles Efferson[1*§]

Social Tipping and Sex Selection in Armenia

(Schief *et al.*, 2021, *Demography*)

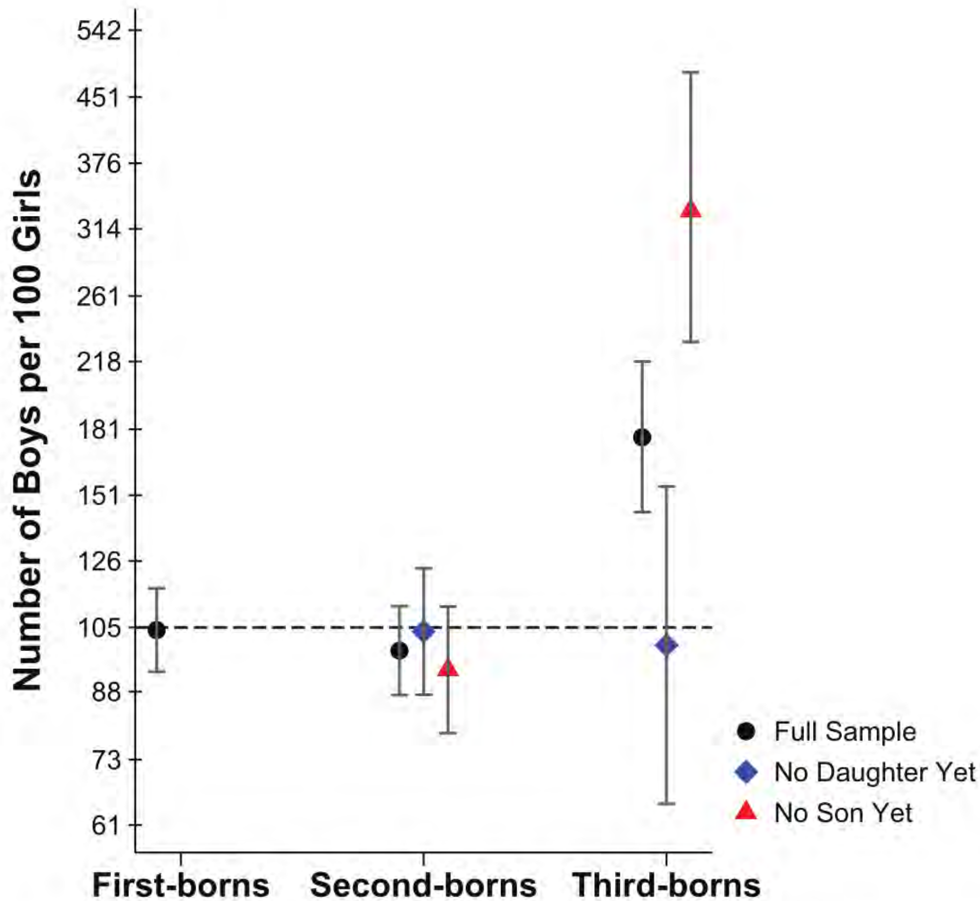


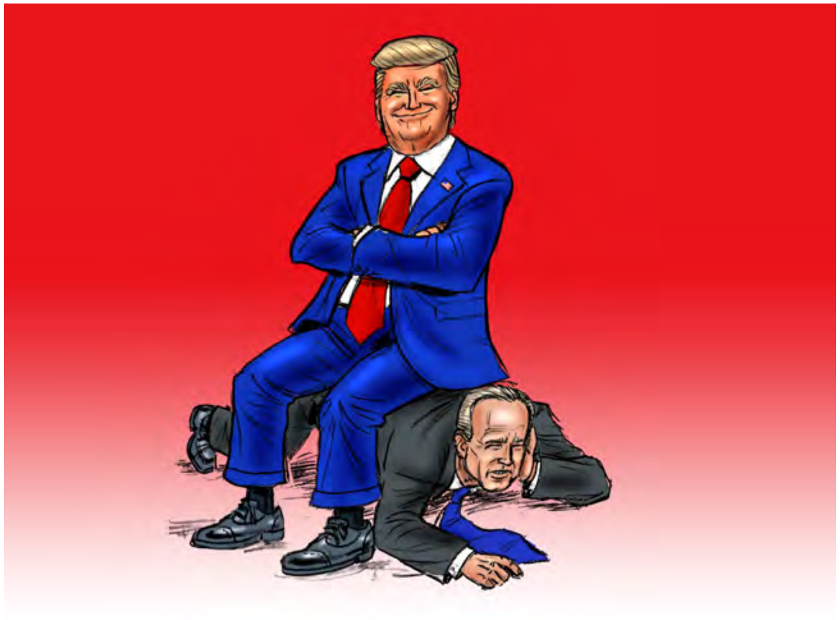
Fig. 1 The ratio of boys to girls in our sample, computed separately by birth order. We compute 95% confidence intervals using the Clopper-Pearson method for calculating binomial proportion confidence intervals.

What if people differ from each other? (Efferson *et al.*, 2020, *NHB*)

- Group identities can completely undermine tipping that would otherwise occur (Ehret *et al.*, 2022).
- If the social planner has an extremely effective intervention, targeting the segment of society most resistant to change is the best strategy *for behavior change*.
- If the intervention is likely to have heterogeneous effects (Vivalt, 2015; Vogt *et al.*, 2016), the social planner can expect a fundamental trade-off.
  - Targeting the amenable will maximize the direct effect and minimize the secondary indirect cultural evolutionary effect.
  - Targeting the resistant will minimize the direct effect and maximize the secondary cultural evolutionary effect conditional on a direct effect of a given size.

The Tiny Step from Social Tipping to Chronic Disagreement

(Ehret *et al.*, 2022, *NHB*)



Today's Experiments:

Different Intervention Strategies in Heterogeneous Populations

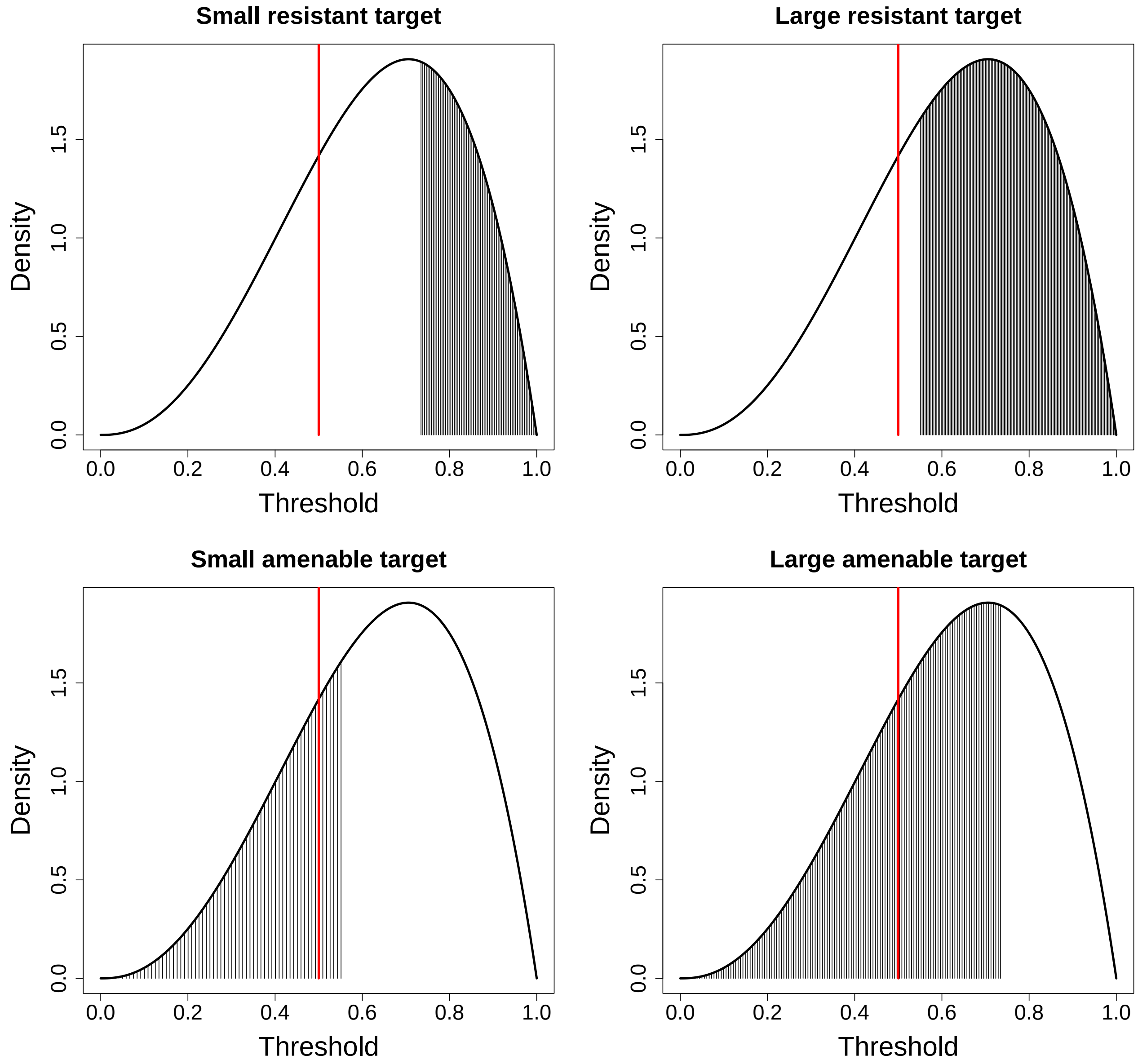
Experimental Framework in a Nutshell

- Experiment is based on a repeated play of coordination games with heterogeneous preferences and stranger matching (groups of 12).
- Before intervention, majority of players viewed coordinating on one option (expected status quo) as risk dominant, cf. coordinating on the other option (social planner's alternative).
- After 15 periods, we targeted a subgroup of players and incentivized them to choose the alternative behavior. Players continued for 25 rounds.
- In practice, as we will see, almost all groups did in fact converge on the expected status quo, and so the intervention was in fact almost always initiating a process of behavior change from the expected status quo to the alternative.
- We manipulated the size of the intervention (1/3 versus 2/3) and whom we targeted (amenable versus resistant players).

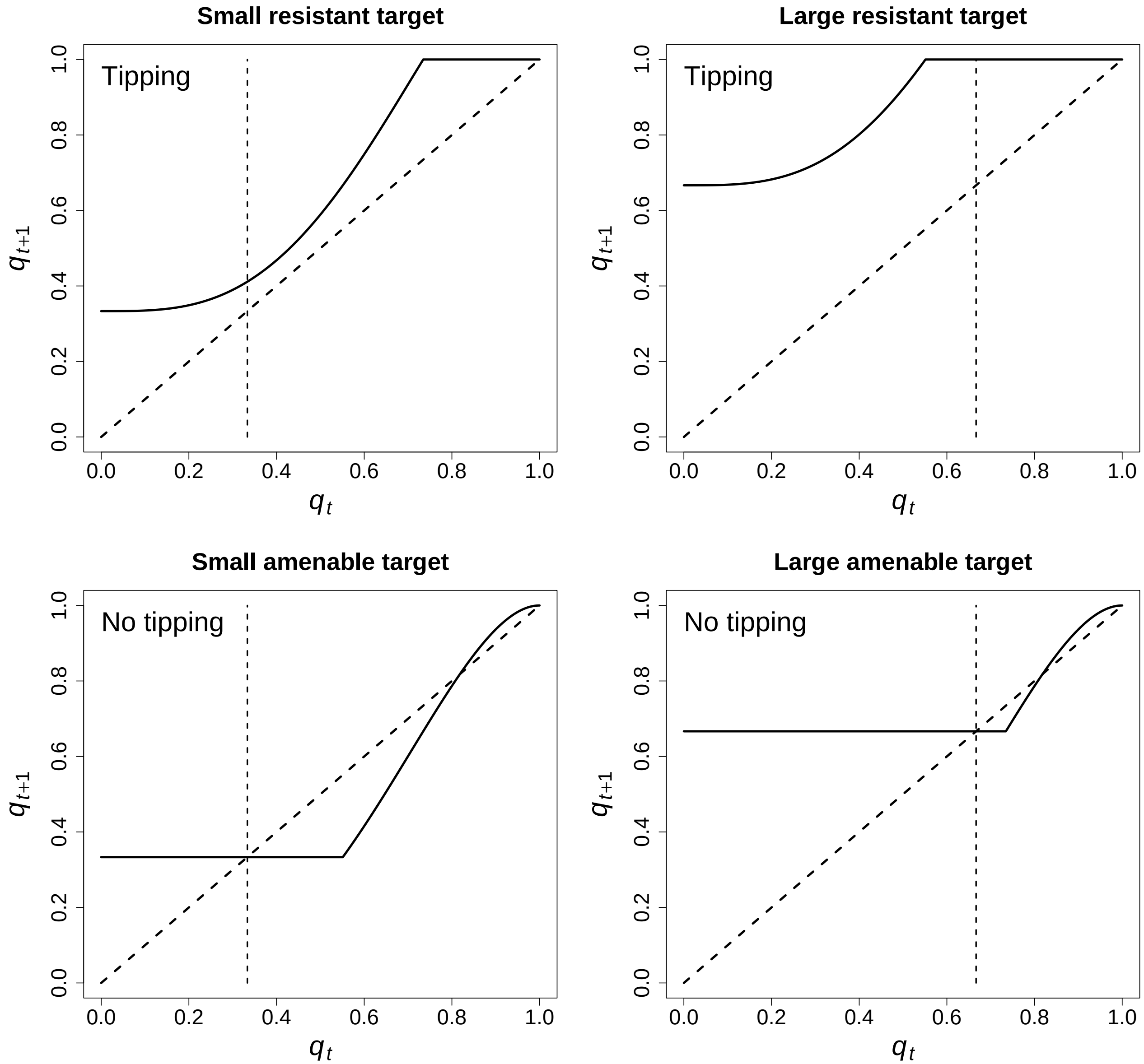
A Highly Effective Intervention:

Experiment #1

Experimental Design in Principle, Pre-intervention



Experimental Design in Principle, Post-intervention



Experimental Design in Practice

- Before intervention, repeated play with stranger matching and x_i randomly and uniquely assigned to individual players from $\{52, 78, 92, 104, 116, 124, 134, 142, 152, 160, 170, 184\}$:

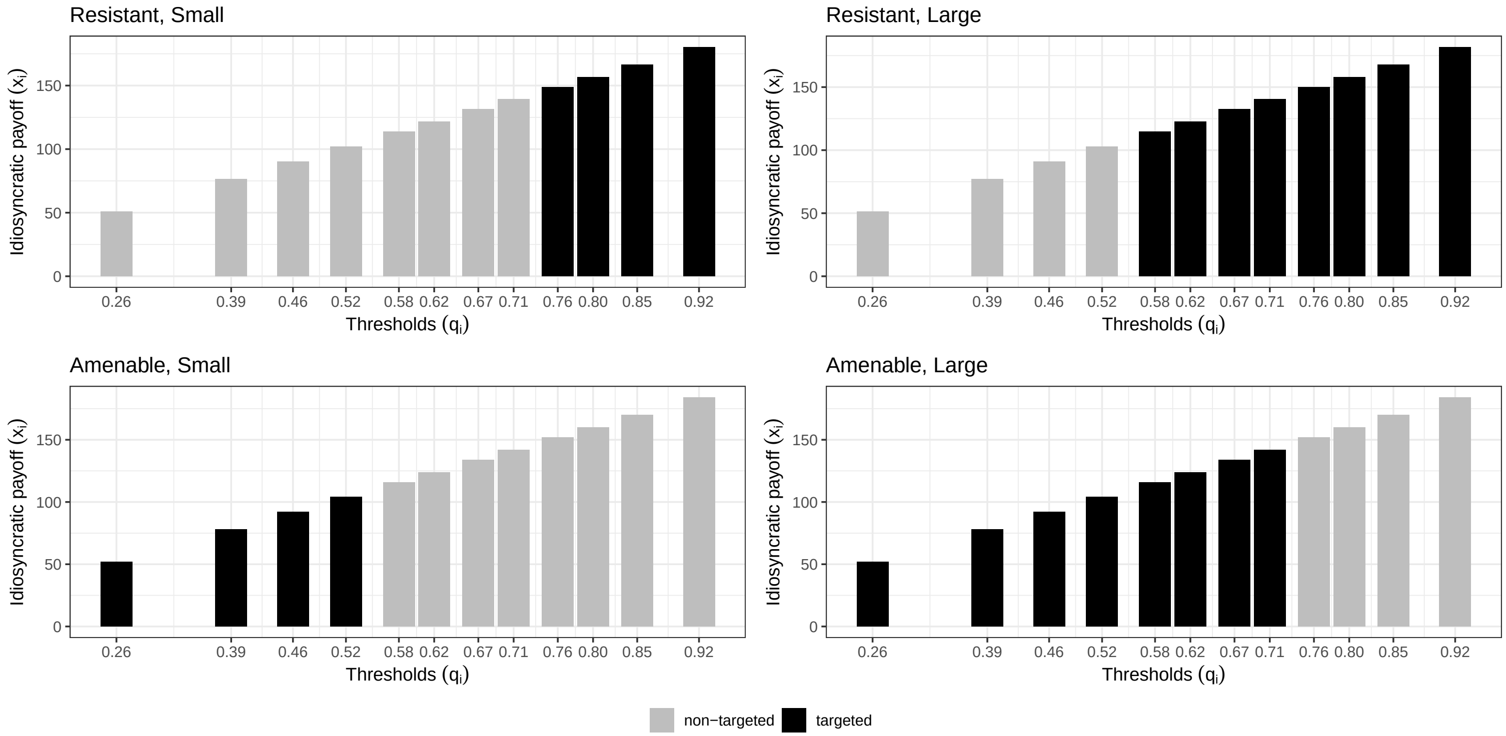
	E[SQ] (#)	Alt (@)
E[SQ] (#)	$100 + x_i$	x_i
Alt (@)	100	200

- After 15 periods, which we estimated would be enough for convergence, intervene by targeting some (T) but not all (NT) participants with new incentives:

	(a) Pre-int (all)		**(b)** Post-int (T)		**(c)** Post-int (NT)	
	E[SQ] (#)	Alt (@)	E[SQ] (#)	Alt (@)	E[SQ] (#)	Alt (@)
E[SQ] (#)	$100 + x_i$	x_i	0	0	$100 + x_i$	x_i
Alt (@)	100	200	300	300	100	200

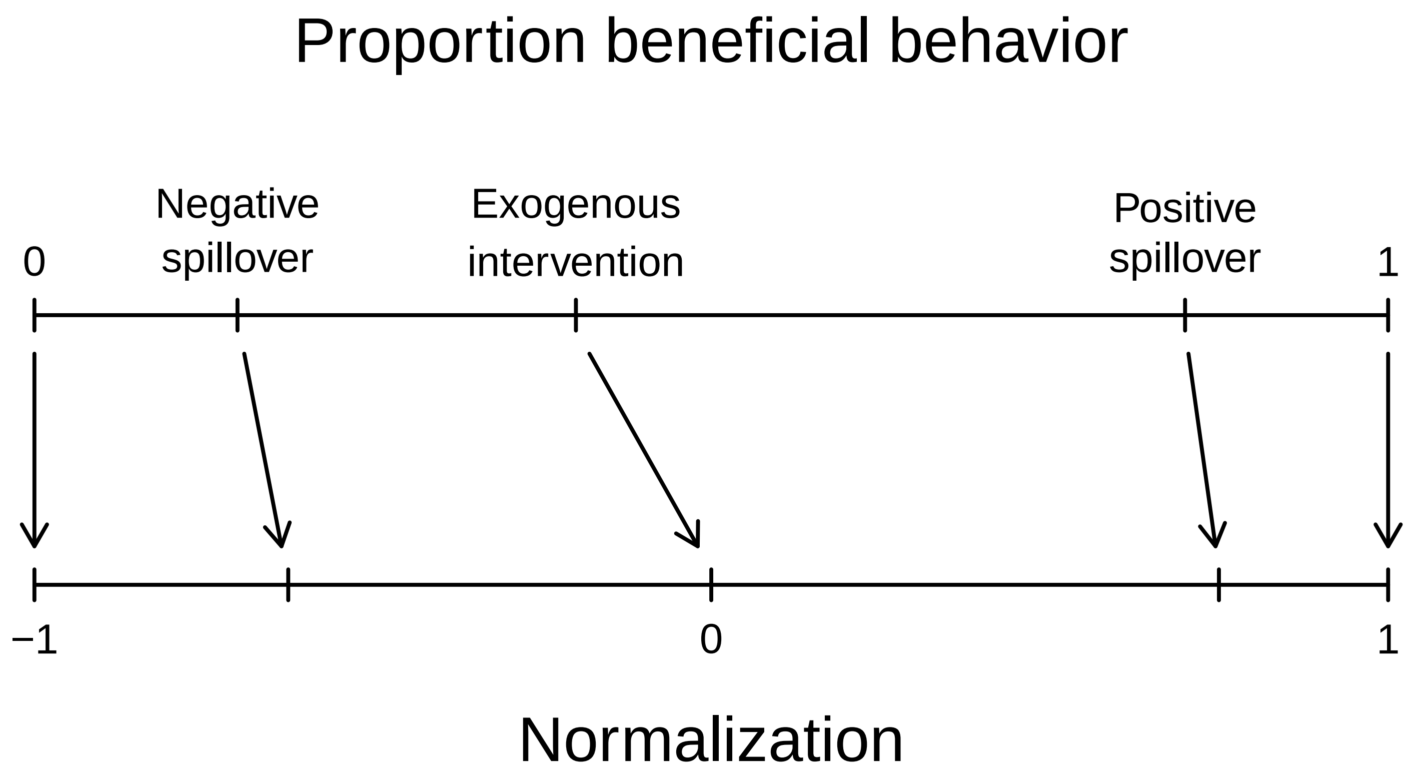
Experimental Design in Practice

- Treatment variation based on interventions that are (i) either small ($1/3 \Rightarrow 4$ players) or large ($2/3 \Rightarrow 8$ players) and target (ii) either the resistant or amenable tail.



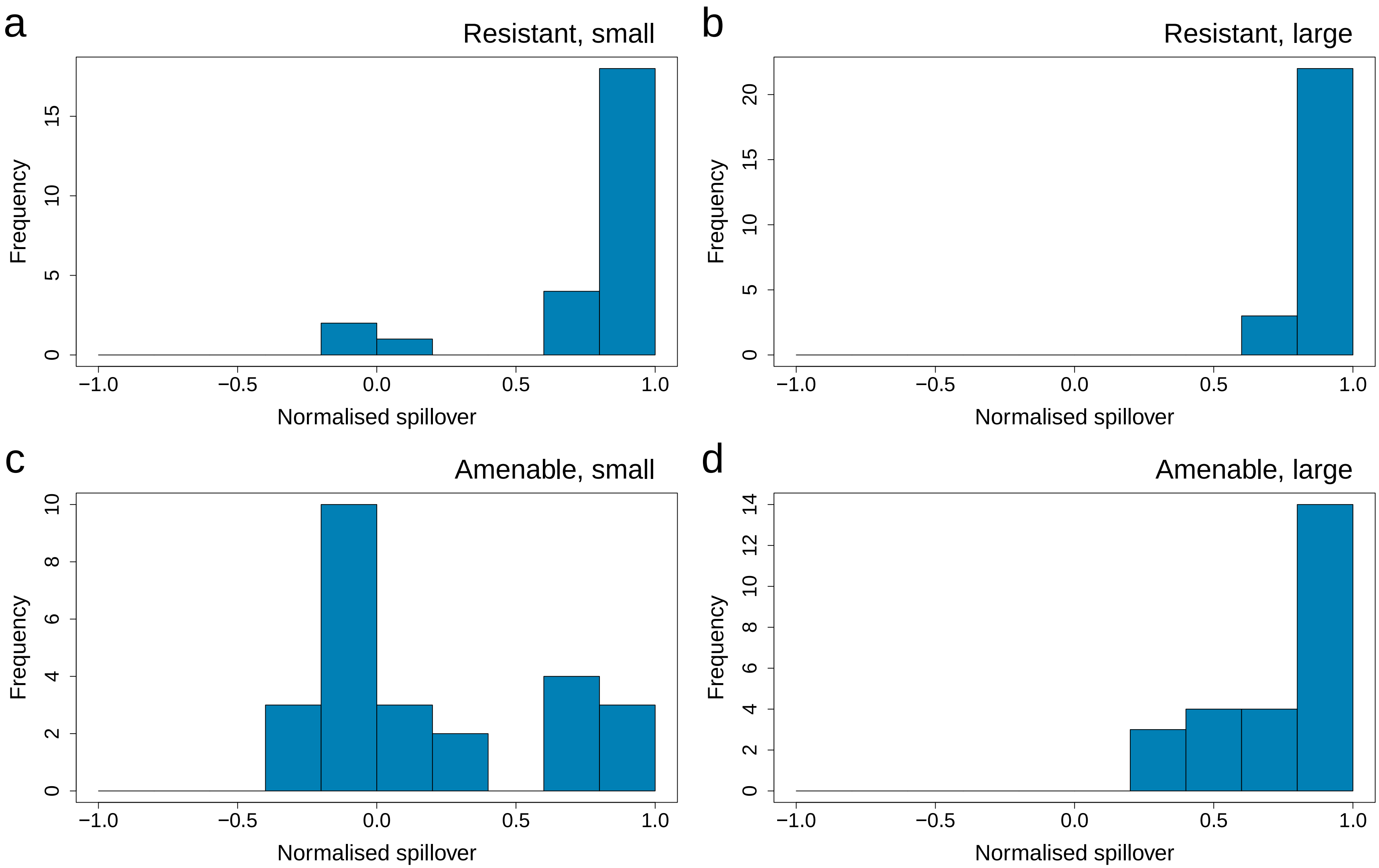
- Even the small intervention is larger than 25% (Centola *et al.*, 2018).

Results: spillovers



$$\Theta_s = \max\left\{0, \frac{[\hat{q}_s > \phi](\hat{q}_s - \phi)}{1-\phi} + \frac{[\phi \geq \hat{q}_s](\hat{q}_s - \phi)}{\phi}\right\}$$

Spillovers by Treatment



Spillovers by Treatment

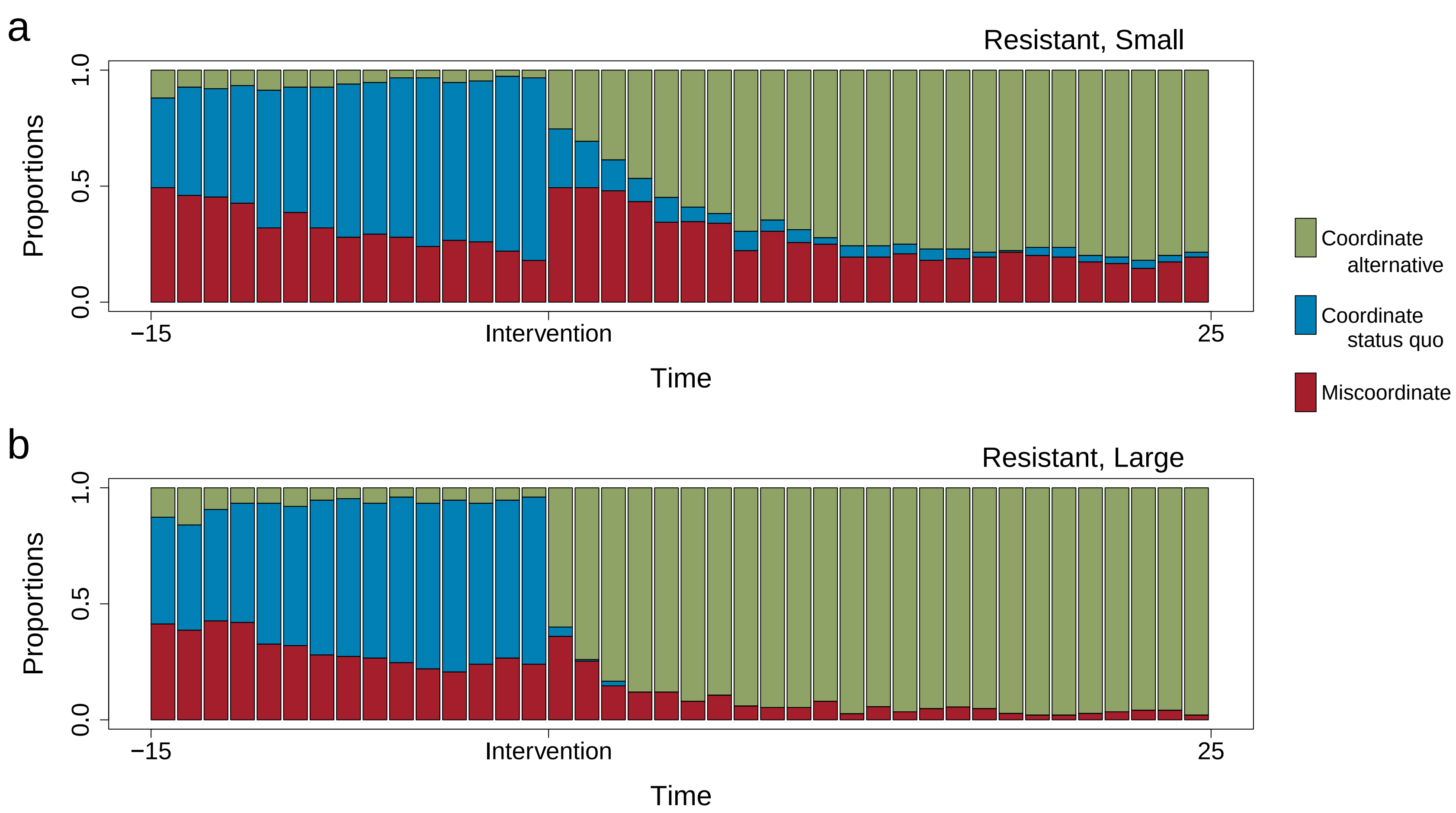
- For any other treatment, spillovers significantly larger than those under a small intervention in the amenable tail.

	Spillovers (Pre-reg)
Intercept	0.24** (0.09)
Resistant, small	0.55*** (0.10)
Resistant, large	0.73*** (0.08)
Amenable, large	0.55*** (0.10)
US sample	-0.01 (0.06)
$N = 100$	(Robust s.e.)

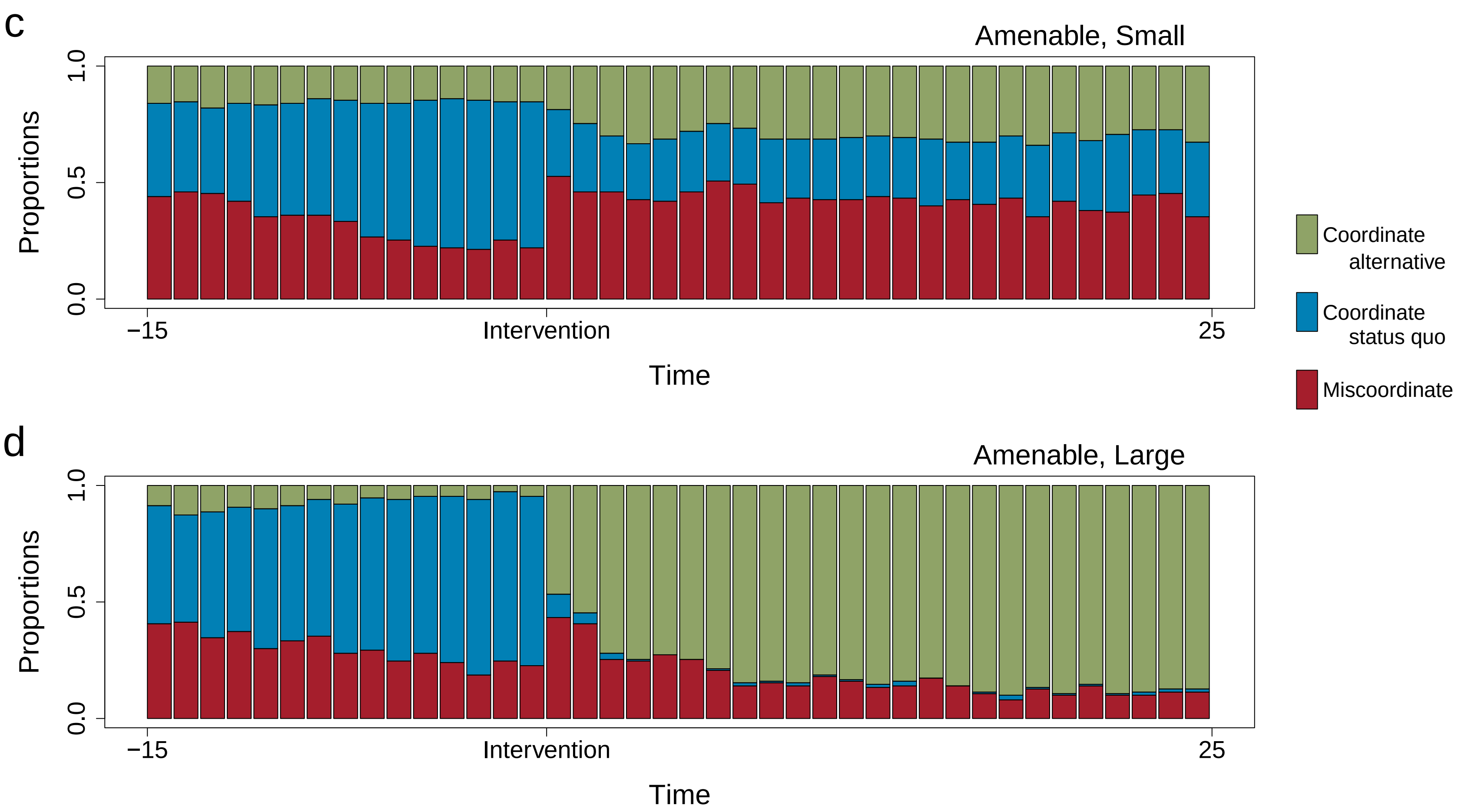
- In addition, $\beta_{\text{Resistant, large}} > \beta_{\text{Resistant, small}}$ ($p = 0.0053$) and $\beta_{\text{Resistant, large}} > \beta_{\text{Amenable, large}}$ ($p = 0.0026$).

Results: choice dynamics

Pooled Outcome Dynamics, Resistant



Pooled Outcome Dynamics, Amenable



Individual Choice, Before and After Intervention

- Let's now turn to individual choice to see who was choosing which behavior and when?
- Linear probability model of choosing Alt as a function of . . .
  - the individual was in a group in which the Amenable tail or Resistant tail was targeted,
  - the individual was targeted (T) or not (NT),
  - the individual was in a group in which the intervention was small (S) or large (L),
  - the choice in question was before intervention (Pre-int) or after (Post-int).
- We did this in three ways: (i) the final pre-intervention and final post-intervention period, (ii) the final five periods each of pre- and post-intervention, and (iii) the final 10 periods each of pre- and post-intervention.

Individual Choice, Before and After Intervention

- As mentioned earlier, we tried to design an intervention that would be equally effective regardless of pre-existing preferences.
- The idea was to hold the *direct* effect of the intervention more or less constant and focus on how different intervention strategies interact with preference heterogeneity to affect the *indirect* effects.
- What were the choices of targeted participants post-intervention?

Extremely similar!

Targeted Individual Choice, Before and After Intervention

	Last period	Last 5 periods (Pre-reg)	Last 10 periods (Pre-reg)
Int (Amenable, NT, S, Pre-int)	0.23***	0.22***	0.23***
	(0.05)	(0.04)	(0.04)
Resistant, NT, S, Pre-Int	-0.07	-0.02	-0.01
	(0.06)	(0.05)	(0.05)
Resistant, T, S, Pre-Int	-0.18**	-0.14**	-0.13**
	(0.06)	(0.05)	(0.05)
Resistant, NT, L, Pre-Int	0.02	0.07	0.06
	(0.07)	(0.06)	(0.06)
Resistant, T, L, Pre-Int	-0.12*	-0.10*	-0.11*
	(0.06)	(0.05)	(0.05)
Resistant, T, L, Post-Int	0.76***	0.77***	0.75***
	(0.05)	(0.04)	(0.04)
Resistant, NT, L, Post-Int	0.74***	0.76***	0.74***
	(0.05)	(0.05)	(0.04)
Resistant, T, S, Post-Int	0.75***	0.77***	0.75***
	(0.05)	(0.04)	(0.04)
Resistant, NT, S, Post-Int	0.57***	0.58***	0.57***
	(0.06)	(0.05)	(0.05)
— Additional effects associated with amenable treatments (see next slide) —			
US sample	0.00	0.00	0.00
	(0.02)	(0.02)	(0.02)
Num.Obs.	3598	13186	25126

(Cluster robust s.e.)

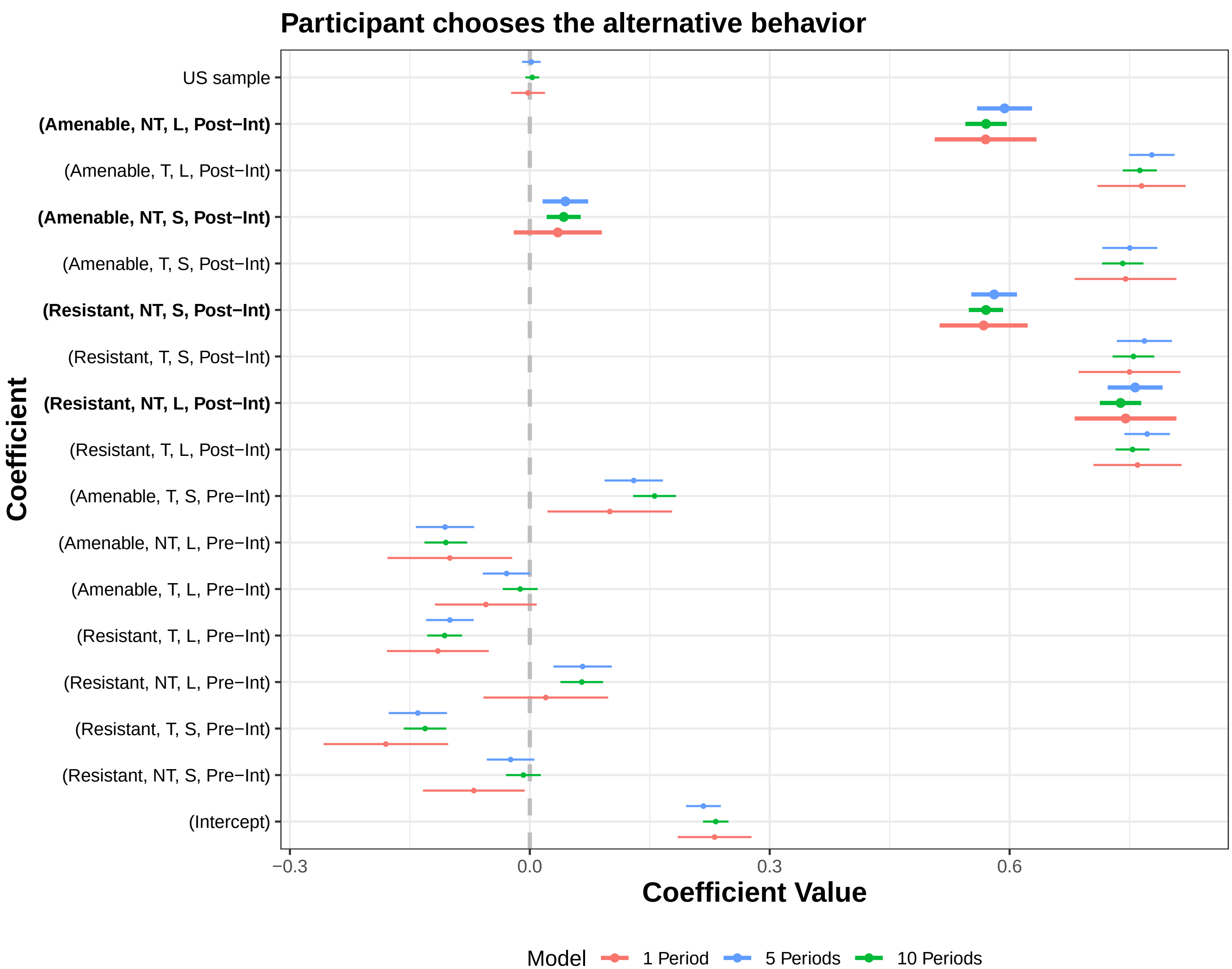
Targeted Individual Choice, Before and After Intervention

	Last period	Last 5 periods (Pre-reg)	Last 10 periods (Pre-reg)
Int (Amenable, NT, S, Pre-int)	0.23***	0.22***	0.23***
	(0.05)	(0.04)	(0.04)
Amenable, T, L, Pre-Int	-0.06	-0.03	-0.01
	(0.06)	(0.05)	(0.05)
Amenable, NT, L, Pre-Int	-0.10	-0.11*	-0.11*
	(0.06)	(0.05)	(0.04)
Amenable, T, S, Pre-Int	0.10*	0.13***	0.16***
	(0.05)	(0.03)	(0.02)
Amenable, T, S, Post-Int	0.74***	0.75***	0.74***
	(0.05)	(0.04)	(0.04)
Amenable, NT, S, Post-Int	0.03	0.04	0.04
	(0.06)	(0.06)	(0.06)
Amenable, T, L, Post-Int	0.76***	0.78***	0.76***
	(0.05)	(0.04)	(0.04)
Amenable, NT, L, Post-Int	0.57***	0.59***	0.57***
	(0.07)	(0.06)	(0.05)
— Additional effects associated with resistant treatments (see previous slide) —			
US sample	0.00	0.00	0.00
	(0.02)	(0.02)	(0.02)
Num.Obs.	3598	13186	25126

(Cluster robust s.e.)

What about non-targeted players post-intervention?

Non-Targeted Individual Choice, Before and After Intervention



Non-Targeted Individual Choice, Before and After Intervention

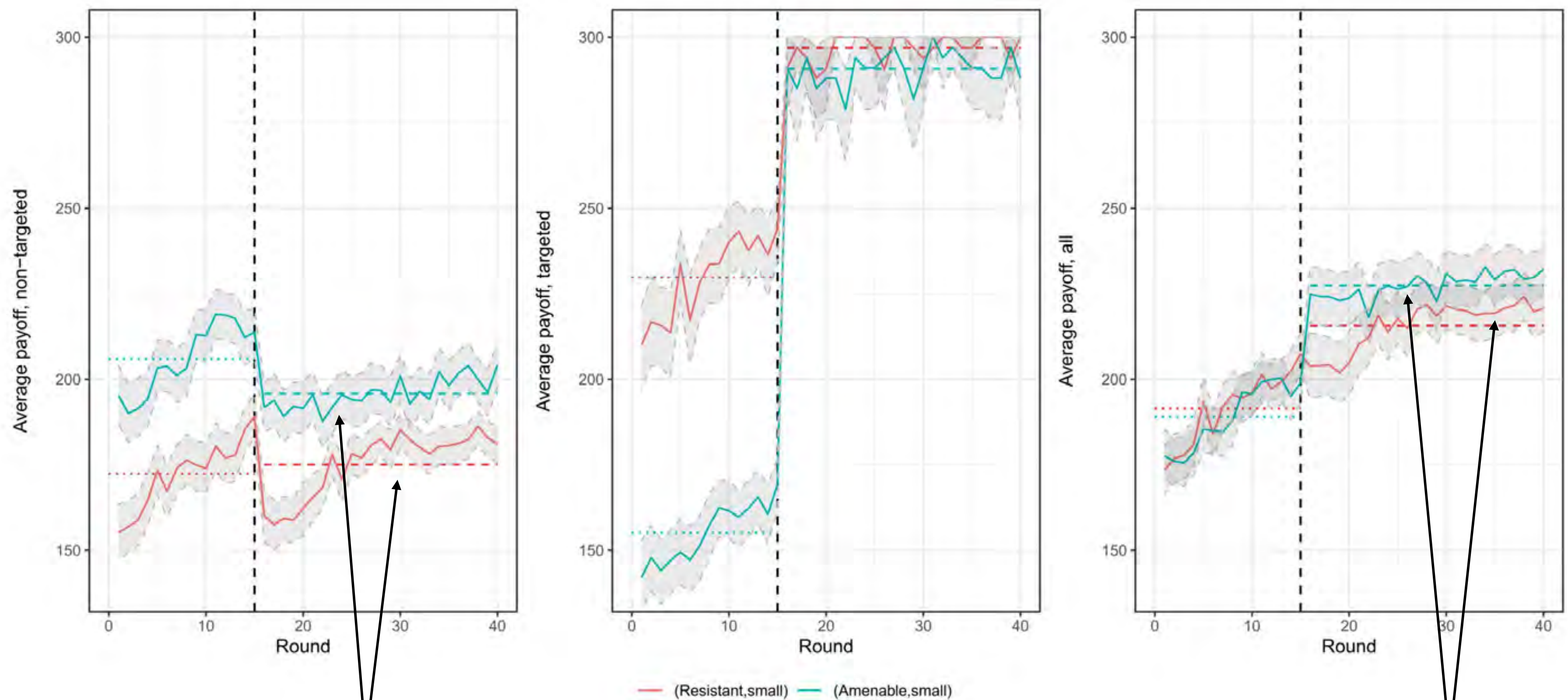
- Holding the targeted tail constant, a large intervention leads to more Alt choices than a small intervention:
  - (Resistant, NT, L, Post-int) vs. (Resistant, NT, S, Post-int) $\rightarrow p < 0.001$,
  - (Amenable, NT, L, Post-int) vs. (Amenable, NT, S, Post-int) $\rightarrow p < 0.001$.
- Holding the size of the intervention constant, targeting resistant leads to more Alt choices than targeting amenable:
  - (Resistant, NT, S, Post-int) vs. (Amenable, NT, S, Post-int) $\rightarrow p < 0.001$,
  - (Resistant, NT, L, Post-int) vs. (Amenable, NT, L, Post-int) $\rightarrow p < 0.001$.
- A small intervention in the resistant tail is no different than a large intervention in the amenable tail, i.e. (Resistant, NT, S, Post-int) vs. (Amenable, NT, L, Post-int) $\rightarrow p = 0.80$.

Results: group welfare

The Welfare Surprise . . .

- We adopt a utilitarian perspective and simply focus on average/total payoffs. This implies the social planner's first objective is to maximize productivity, and she can redistribute ex post to accomplish other social objectives.
- Recall that we had four treatments, and three of them tipped to Alt after intervention.
- The one treatment that did not tip – a small intervention in the amenable tail – persisted in a state of chronic disagreement with miscoordination rates near the maximum as a result.
- Surprisingly, this treatment did *not* have the lowest payoffs post-intervention.

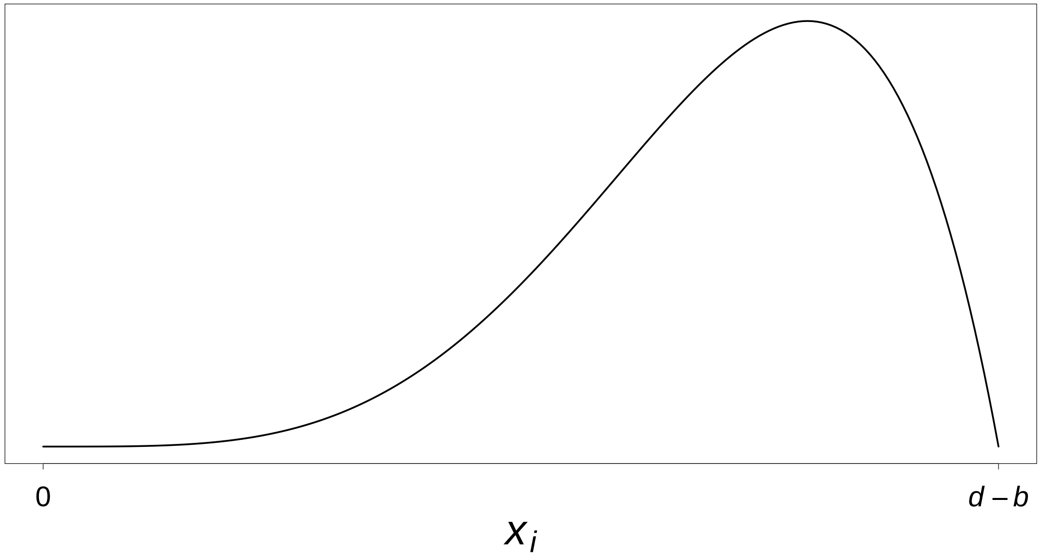
What are the welfare consequences of a *small* intervention?



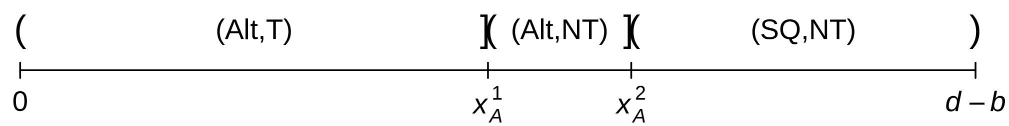
Under an amenable target, miscoordination was persistent, but non-targeted players earned more (10 periods, $p < 0.001$) than under a resistant target, which led to tipping.

Difference is also significant if we average over targeted and non-targeted players (10 periods, $p < 0.001$).

A Generalization (Efferson *et al.*, 2024, *Phil Trans Roy Soc B*)



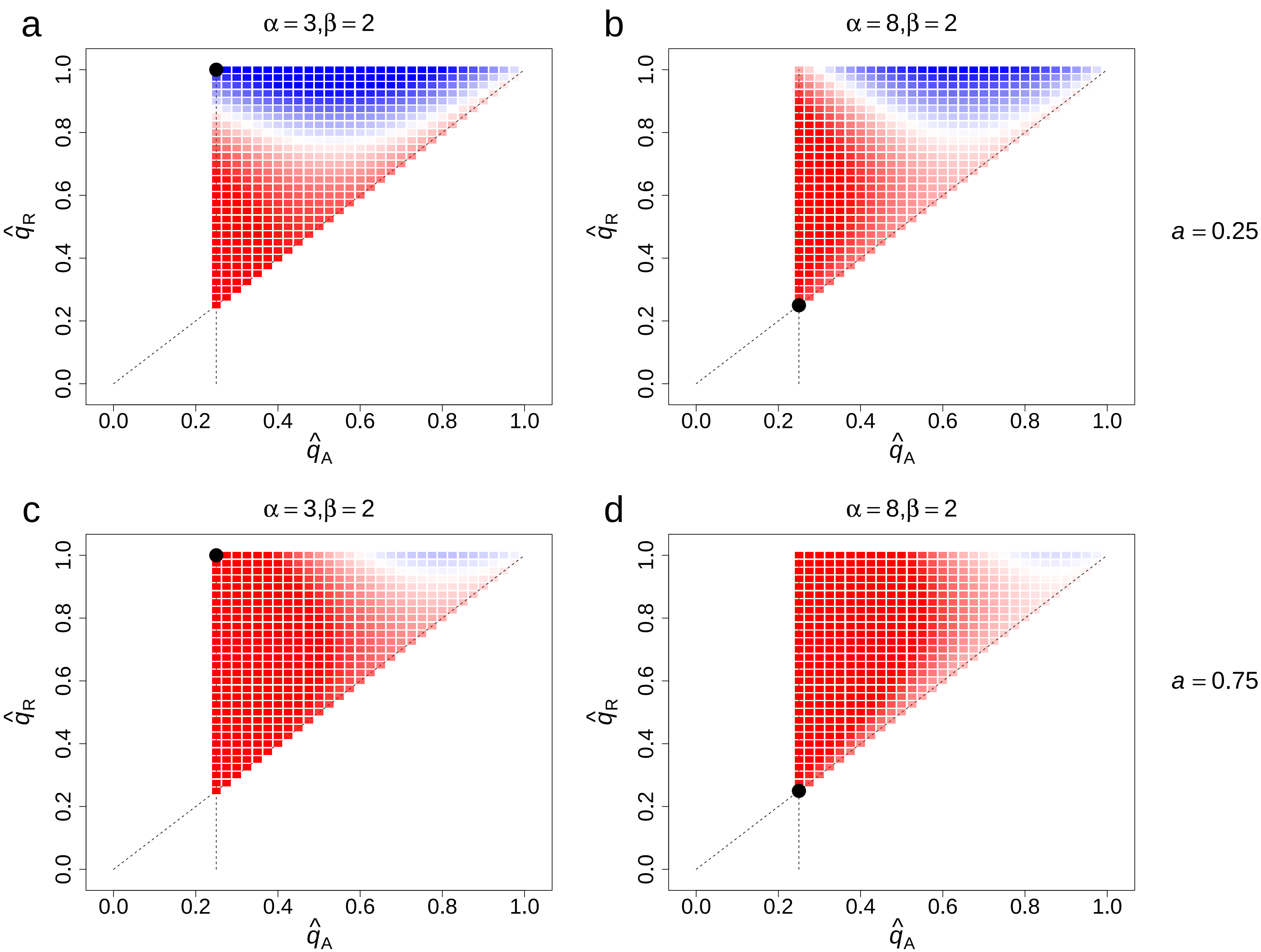
Amenable target:



Resistant target:

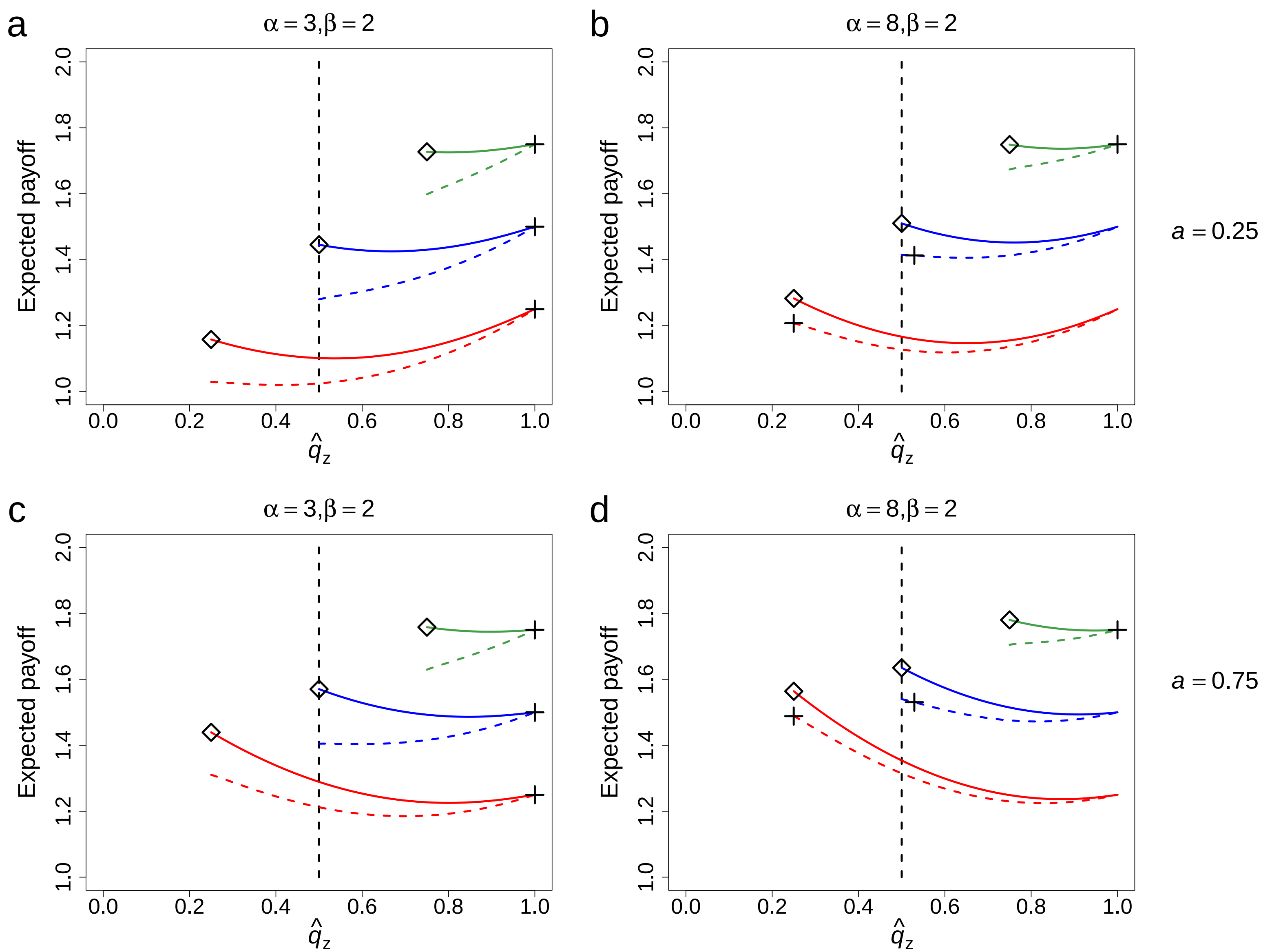
A Generalization (Efferson *et al.*, 2024, *Phil Trans Roy Soc B*)

$$\text{Red} \Rightarrow E_A[\Pi] > E_R[\Pi]$$



A Generalization (Efferson *et al.*, 2024, *Phil Trans Roy Soc B*)

Amenable (S) & Resistant (D)



What about the fundamental trade-off?

Experiment #2

The Weakest Possible Equivalence

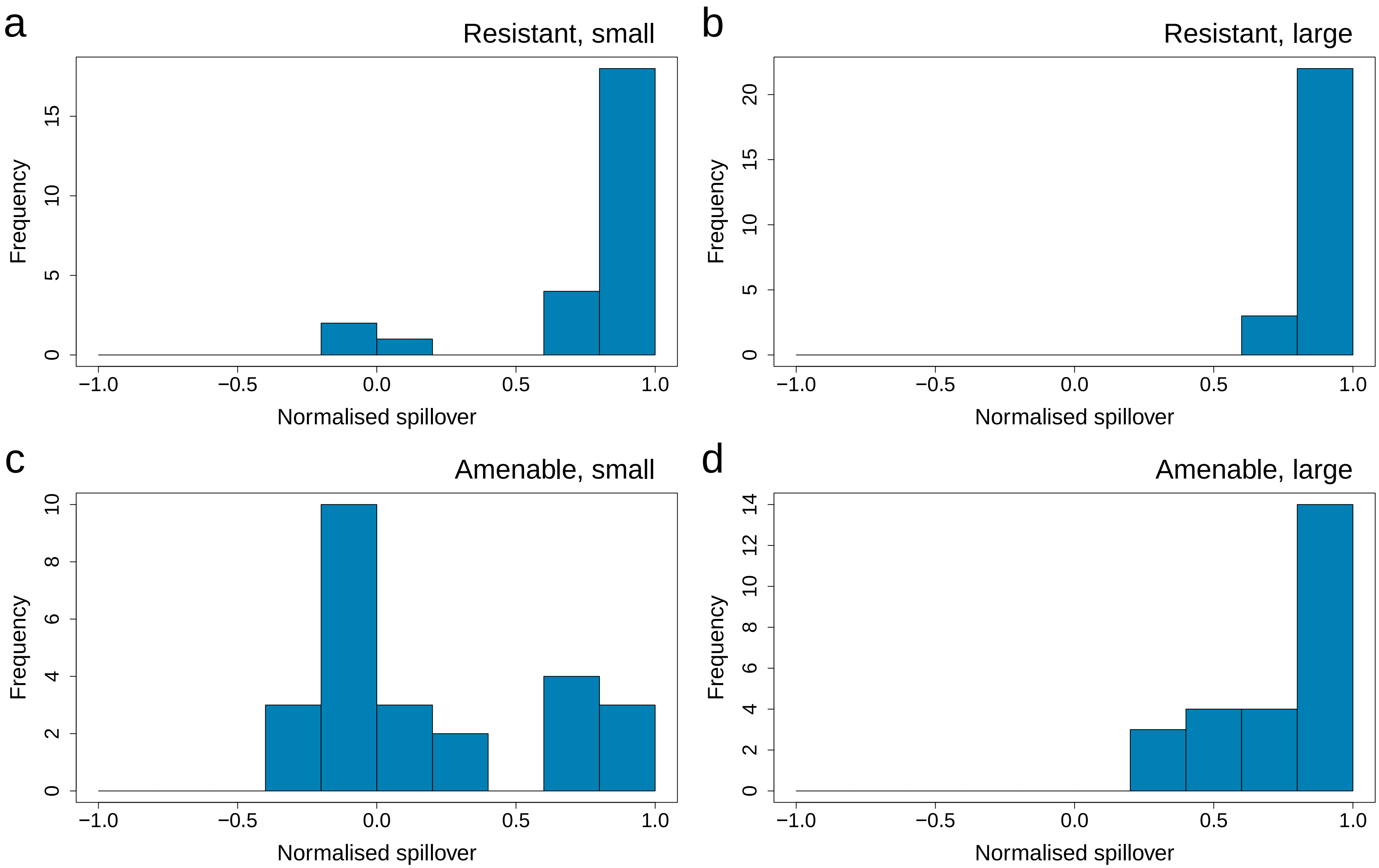
- Before intervention, repeated play with stranger matching and x_i randomly and uniquely assigned to individual players from $\{52, 78, 92, 104, 116, 124, 134, 142, 152, 160, 170, 184\}$:

	E[SQ] (#)	Alt (@)
E[SQ] (#)	$100 + x_i$	x_i
Alt (@)	100	200

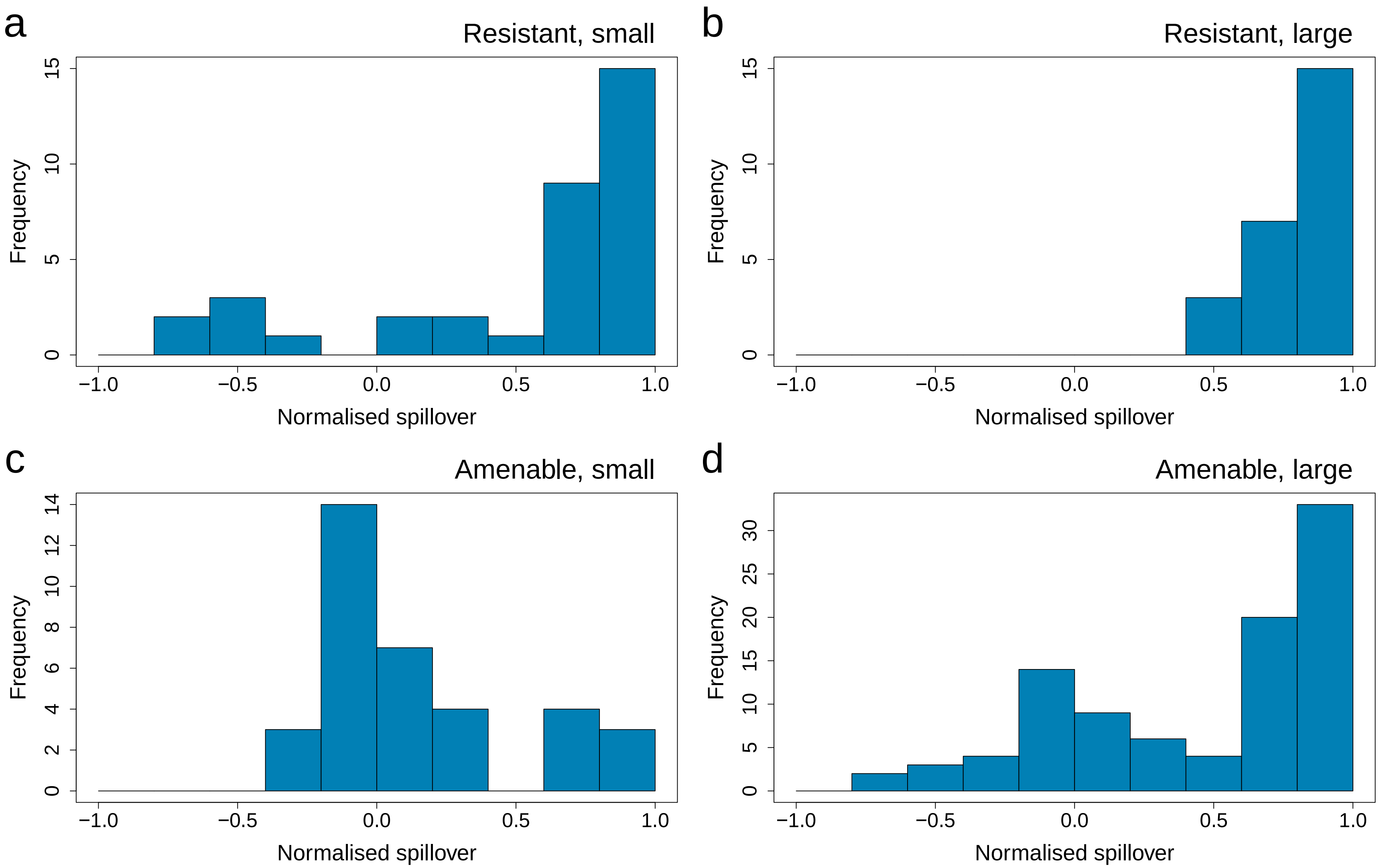
- After 15 periods, which we estimated would be enough for convergence, intervene by targeting some (T) but not all (NT) participants with new incentives ($s = 152$):

	(a) Pre-int (all)		**(b)** Post-int (T)		**(c)** Post-int (NT)	
	E[SQ] (#)	Alt (@)	E[SQ] (#)	Alt (@)	E[SQ] (#)	Alt (@)
E[SQ] (#)	$100 + x_i$	x_i	$100 + x_i$	x_i	$100 + x_i$	x_i
Alt (@)	100	200	$100 + s$	$200 + s$	100	200

Reminder Experiment #1: Spillovers by Treatment



Experiment #2: Spillovers by Treatment



Both Experiments: Spillovers by Treatment

Table 1: **Spillovers by treatment.** Spillovers take values in $[-1, 1]$ and provide a normalized measure of long-run behavior in a population while accounting for the size of the intervention. Results are from OLS regressions that model spillovers as a function of treatment. Composite treatment dummies are defined jointly over (i) intervention target (amenable vs. resistant) and (ii) intervention size (S vs. L). Omitted category is (Amenable, Small). Robust standard errors (parentheses). Models were pre-registered.

	Strong Version	Weak Version
(Intercept)	0.24**	0.25***
	(0.09)	(0.07)
(Resistant, small)	0.55***	0.32**
	(0.10)	(0.11)
(Resistant, large)	0.73***	0.66***
	(0.08)	(0.07)
(Amenable, large)	0.55***	0.40***
	(0.10)	(0.10)
US sample	-0.01	-0.12
	(0.06)	(0.08)
Num.Obs.	100	120

$^{*}\, p \in (0.01, 0.05]$ $^{**}\, p \in (0.001, 0.01]$ $^{***}\, p \leq 0.001$

Both Experiments: Spillovers by Treatment

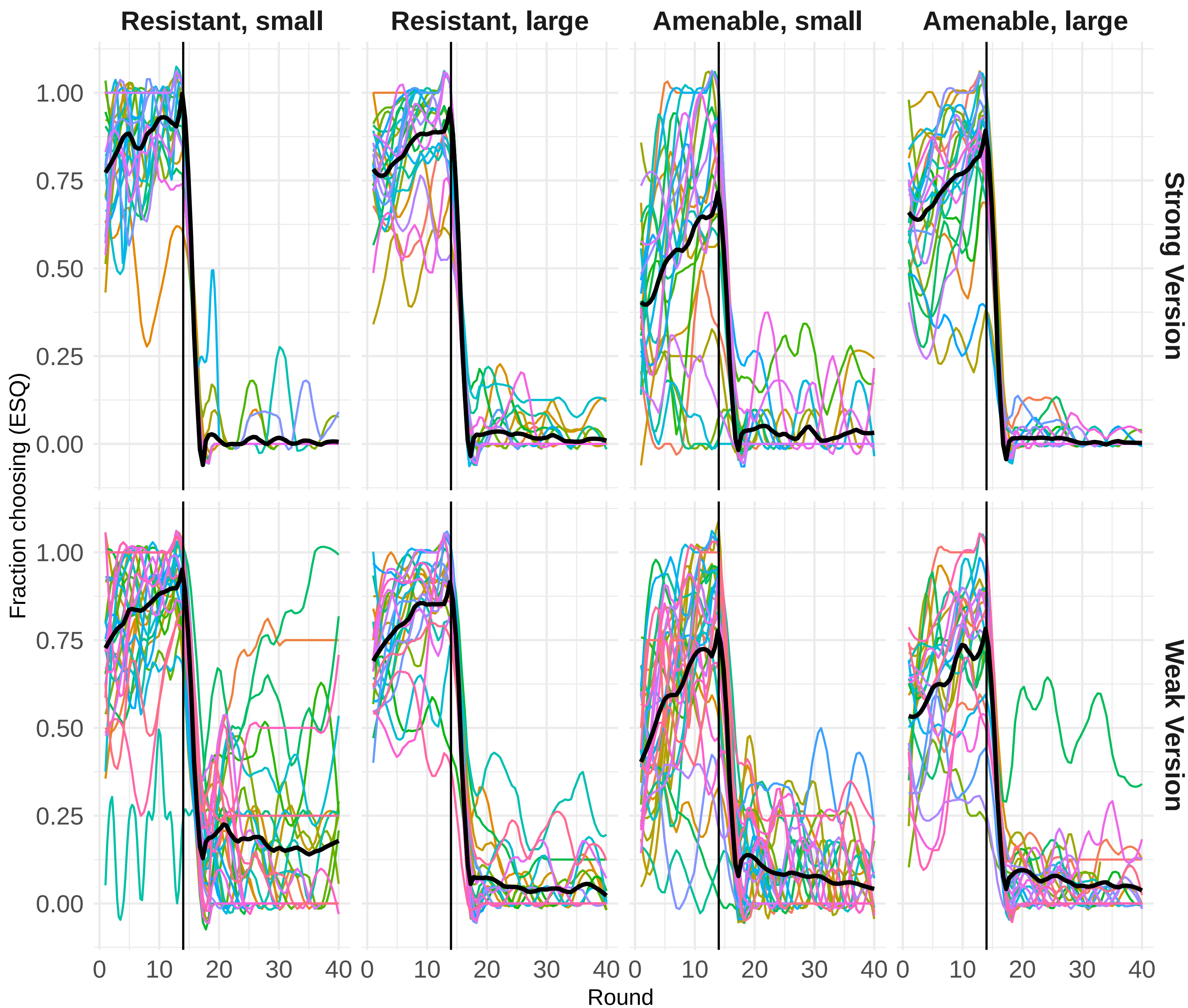
Table 2: **Linear combination tests, spillover model, size and target.**

		Strong Version			Weak Version			
cif.	Linear combination	F	Pr(>F)	95% CI	cif.	F	Pr(>F)	95% CI
1	Resistant Large - Resistant Small	8.13	0.0053	[0.05, 0.31]	4	11.45	0.001	[0.14, 0.53]
2	Resistant Large - Amenable Large	9.60	0.0026	[0.06, 0.30]	5	8.46	0.0044	[0.083, 0.44]
3	Resistant Small - Amenable Large	0	1.0000	[-0.16, 0.16]	6	0.36	0.55	[-0.32, 0.17]

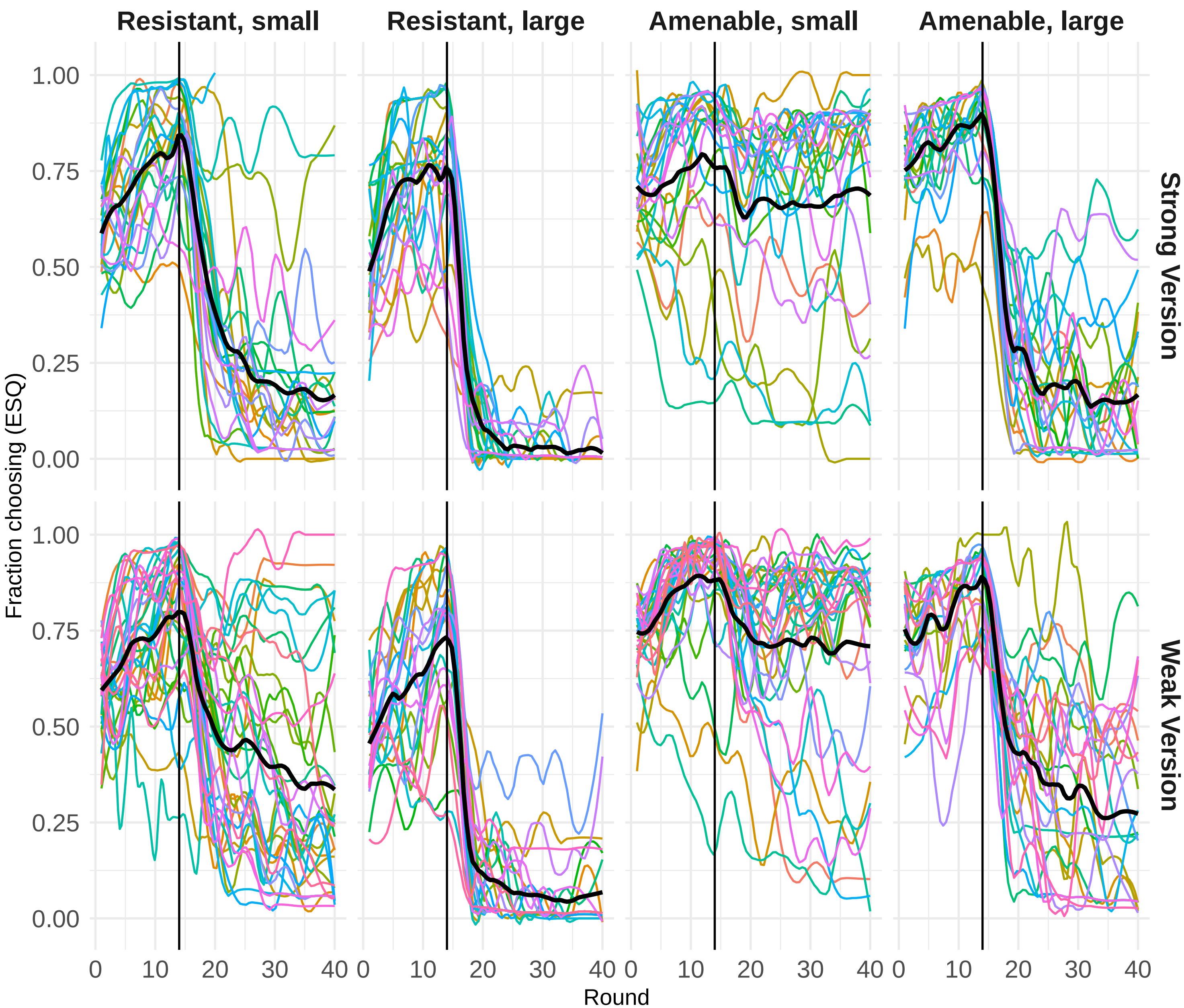
Table 3: **Linear combination tests, spillover model, strong versus weak.**

cif.	Coefficient	Δ Weak Version - Strong Version	Pr(>F)	95% CI
1	Resistant Small	-0.27	0.005	[-0.45,-0.08]
2	Resistant Large	-0.11	0.28	[-0.31,0.09]
3	Amenable Small	-0.04	0.66	[-0.23,0.14]
4	Amenable Large	-0.19	0.06	[-0.39,0.01]

Dynamics by Treatment, Targeted



Dynamics by Treatment, Non-targeted



Conclusions

- In heterogeneous populations, the idea of a critical mass has no meaning on its own, even if everyone is a strong conformist and/or faces coordination incentives.
- For an intervention of a given size, targeting more resistant individuals should be better for tipping and behavior change *so long as* the intervention remains sufficiently effective when the target is resistant.
- This is the case we isolated with the strong intervention, and we found that resistant targets, conditional on intervention size, did in fact lead to more behavior change than amenable targets.

Conclusions

- The "*so long as*" is a strong caveat, with potential for poorly understood trade-offs when violated.
- Our weak intervention undercut behavior change relative to the strong intervention in the (Resistant,S) case.
- Once again, however, conditional on intervention size, resistant targets led to more behavior change than amenable targets.

Conclusions

- We once presented an early version of this thinking to the UN and the European Commission in Brussels, and the consensus in the room seemed to be that development organizations often have a culture of working with people amenable to change, or more generally people who simply support the organization.
- If true, this tendency would imply an important form of selection in practice that could actually undercut behavior change due specifically to endogenous cultural evolution.
- We do not know of any research on this important possibility.

Conclusions

- We were surprised to find that the small amenable intervention, when compared to the small resistant intervention, led to far more miscoordination but higher average payoffs.
- This result illustrates the following:
  - The idea that tipping and high coordination rates after intervention are better than not tipping and low coordination rates is not necessarily true.
  - The social planner should consider both how a specific type of intervention generates behavior change (targeted and non-targeted) and the distribution of residual preferences among those not targeted.

Acknowledgments

- Swiss National Science Foundation Nr. 100018_185417/1
- UNICEF Sudan
- National (Khartoum) and Gezira State (Wad Medani) Councils for Child Welfare
- UNICEF Armenia
- ARMSTAT (Yerevan)
- Women's Resource Council (Yerevan)
- Swiss National Committee of UNICEF
- Amy Elhadi, Hilal El Fadil Ahmed, Nadia Ahmed Mohmmed Zaid, Katelyn Bonner

Thank you!

References

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, **95**(9), 1082–1095.

Andreoni, J., Nikiforakis, N., and Siegenthaler, S. (2021). Predicting social tipping and norm change in controlled experiments. *Proceedings of the National Academy of Sciences*, **118**(16).

Bicchieri, C., Jiang, T., and Lindemans, J. W. (2014, accessed 27 July 2017). A social norms perspective on child marriage: The general framework. `http://repository.upenn.edu/pennsong/13/`.

Camilotti, G. (2016). *Fighting against harmful customs: the case of female genital cutting*. Ph.D. thesis, University of Namur.

Castilla-Rho, J. C., Rojas, R., Andersen, M. S., Holley, C., and Mariethoz, G. (2017). Social tipping points in global groundwater management. *Nature Human Behaviour*, **1**(9), 640–649.

Centola, D., Becker, J., Brackbill, D., and Baronchelli, A. (2018). Experimental evidence for tipping points in social convention. *Science*, **360**(6393), 1116–1119.

Christakis, N. A. and Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England Journal of Medicine*, **2007**(357), 370–379.

Christakis, N. A. and Fowler, J. H. (2008). The collective dynamics of smoking in a large social network. *New England Journal of Medicine*, **358**(21), 2249–2258.

Cloward, K. (2016). *When Norms Collide: Local Responses to Activism Against Female Genital Mutilation and Early Marriage*. Oxford University Press.

Dávila-Fernández, M. J. and Sordi, S. (2020). Attitudes towards climate policies in a macrodynamic model of the economy. *Ecological Economics*, **169**, 106319.

Delneuville, Amy (2012, accessed 27 July 2017). Reaching the tipping point: child marriage in Bangladesh. http://www.sas.upenn.edu/ppe/Events/uniconf_2012/documents/Delneuville.Amy_Final.Paper_000.pdf.

Efferson, C., Vogt, S., Elhadi, A., Ahmed, H. E. F., and Fehr, E. (2015). Female genital cutting is not a social coordination norm. *Science*, **349**(6255), 1446–1447.

Efferson, C., Vogt, S., and Fehr, E. (2020). The promise and the peril of using social influence to reverse harmful traditions. *Nature Human Behaviour*, **4**, 55–68.

Efferson, C., Vogt, S., and von Flüe, L. (2023). Activating cultural evolution for good when people differ from each other. In J. Kendal, R. Kendal, and J. Tehrani, editors, *Oxford Handbook of Cultural Evolution*, chapter TBD. Oxford University Press.

Ehret, S., Constantino, S., Weber, E., Efferson, C., and Vogt, S. (2022). Group identities can undermine social tipping after intervention. *Nature Human Behaviour*, **Forthcoming**, TBD.

Farmer, J. D., Hepburn, C., Ives, M. C., Hale, T., Wetzer, T., Mealy, P., Rafaty, R., Srivastav, S., and Way, R. (2019). Sensitive intervention points in the post-carbon transition. *Science*, **364**(6436), 132–134.

Hallsworth, M., List, J. A., Metcalfe, R. D., and Vlaev, I. (2017). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of Public Economics*, **148**, 14–31.

Koch, C. M. and Nax, H. H. (2017). Rethinking free-riding and tragedy of the commons. https://ssrn.com/abstract=3075935.

Lee-Rife, S., Malhotra, A., Warner, A., and Glinski, A. M. (2012). What works to prevent child marriage: A review of the evidence. *Studies in Family Planning*, **43**(4), 287–303.

Mackie, G. (1996). Ending Footbinding and Infibulation: A Convention Account. *American Sociological Review*, **61**, 999–1017.

Malhotra, Anju and Warner, Ann and McGonagle, Allison and Lee-Rife, Susan (2011, accessed 27 July 2017). Solutions to end child marriage: what the evidence shows. https://www.icrw.org/wp-content/uploads/2016/10/Solutions-to-End-Child-Marriage.pdf.

Nyborg, K., Anderies, J. M., Dannenberg, A., Lindahl, T., Schill, C., Schlüter, M., Adger, W. N., Arrow, K. J., Barrett, S., Carpenter, S., *et al.* (2016). Social norms as solutions. *Science*, **354**(6308), 42–43.

Otto, I. M., Donges, J. F., Cremades, R., Bhowmik, A., Hewitt, R. J., Lucht, W., Rockström, J., Allerberger, F., McCaffrey, M., Doe, S. S., *et al.* (2020). Social tipping dynamics for stabilizing earth's climate by 2050. *Proceedings of the National Academy of Sciences*, **117**(5), 2354–2365.

Paluck, E. L., Shepherd, H., and Aronow, P. M. (2016). Changing climates of conflict: A social network experiment in 56 schools. *Proceedings of the National Academy of Sciences*, **113**(3), 566–571.

Platteau, J.-P., Camilotti, G., and Auriol, E. (2018). Eradicating women-hurting customs. In S. Anderson, L. Beaman, and J. Platteau, editors, *Towards Gender Equity in Development*, pages 319–356. Oxford University Press.

Prentice, D. A. and Miller, D. T. (1993). Pluralistic ignorance and alcohol use on campus: some consequences of misperceiving the social norm. *Journal of Personality and Social Psychology*, **64**(2), 243.

Schief, M., Vogt, S., and Efferson, C. (2021). Investigating the structure of son bias in Armenia with novel measures of individual preferences. *Demography*, **58**, 1737–1764.

Shakya, H. B., Christakis, N. A., and Fowler, J. H. (2015). Social network predictors of latrine ownership. *Social Science & Medicine*, **125**, 129–138.

Shell-Duncan, B. and Hernlund, Y. (2000). Female "Circumcision" in Africa: Dimensions of the Practice and Debates. In B. Shell-Duncan and Y. Hernlund, editors, *Female "Circumcision" in Africa: Culture, Controversy, and Change*, pages 1–40. Boulder, CO: Lynne Rienner.

Travers, H., Walsh, J., Vogt, S., Clements, T., and Milner-Gulland, E. (2021). Delivering behavioural change at scale: What conservation can learn from other fields. *Biological Conservation*, **257**, 109092.

UNFPA-UNICEF (2013, accessed 15 April 2015). Joint evaluation of the UNFPA-UNICEF joint programme on female genital mutilation/cutting: Accelerating change. http://www.unfpa.org/public/home/about/Evaluation/EBIER/TE/pid/10103.

Vivalt, E. (2015). Heterogeneous treatment effects in impact evaluation. *American Economic Review*, **105**(5), 467–70.

Vogt, S., Zaid, N. A. M., Ahmed, H. E. F., Fehr, E., and Efferson, C. (2016). Changing cultural attitudes towards female genital cutting. *Nature*, **538**, 506–509.

World Health Organization (2010, accessed 27 July 2017). Violence prevention, the evidence: changing cultural and social norms that support violence. http://apps.who.int/iris/bitstream/10665/77936/1/9789241500845_eng.pdf.