

Bootstrapping Sustainability Norms in Cultural–Institutional Coevolutionary Systems

Jenna Bednar* and Scott E Page†

January 15, 2026

Abstract

To better understand the potential and interdependence of complex goal structures like the UN’s Sustainable Development Goals, we construct a framework for studying the co-evolution of culture and institutions in which behaviors and outcomes in one domain influence the returns to behaviors in other domains through **behavioral spillovers** and **domain externalities**. For example, in our framework, economic well-being depends on behaviors in environmental domains. This assumption is consistent with long-standing empirical research showing interdependence across the seventeen United Nations Sustainability Goals. Within this framework, we demonstrate the potential for multiple stable configurations, including some that are far from efficient. Second, we show that increasing the magnitudes of behavioral spillovers and domain externalities has differing effects on the nature of the stable configurations supported. Third, we show how policies that reward particular actions function similarly to domain externalities and not behavioral spillovers.

Keywords: Behavioral Spillovers, Externalities, Cultural-Institutional Coevolution

Introduction

Human well-being cannot be captured by economic variables alone. It depends on a variety of factors—health, security, housing, food, social relationships, and meaningful work—and also on how we manage the natural environment.

Humans rely on natural systems to provide food, water, and air.

*Department of Political Science and Ford School of Public Policy, University of Michigan, and Santa Fe Institute

†Ross School of Business, University of Michigan, Ann Arbor, MI 48104 and Santa Fe Institute, 1399 Hyde Park Rd, Santa Fe, NM 87501

The recognition of the breadth of influences on human well-being was a driving force behind the development of United Nations Sustainable Development Goals (SDGs). The broader aperture induced by the SDGs has resulted in a sustained effort to gather data. That data reveals a rich interdependence between the human well being, the economy and the natural world. With that has come the realization that efforts to improve human well-being that focus only on economic growth will not suffice. A growth mindset can damage the natural environment and hinder efforts to improve well-being.

Social well-being also depends on the natural and built environment (Bednar 2023). What we build in the world and how we treat the natural world influences who we are, the quality of our lives, our relationships to one another, and the strengths of our societies.

The interdependence across domains implies that policies that address domains in isolation will have limited success. To be most effective, policies for growth, sustainability, and education must be coordinated. We need sustainable growth policies linked to education.

Over the past two decades, progress on most of the SDGs has been steady. Sustainability, though, has proven among the most difficult. The efficacy of policies intended to promote sustainable growth has been mixed. On the one hand, top-down technological progress in green energy has been impressive. In many regions of the world, solar, wind, and water produce power at a lower cost than fossil fuels (Way et al 2022). However, policies designed to shift human behavior toward more sustainable actions have proven less successful. Encouraging people to plant a tree or ride a bike to work has not produced large reductions in energy consumption. And the introduction of AI appears to increase demand for energy.

The failure to change behavior through policies should come as no surprise.

Changing behavior is notoriously difficult. Humans are creatures of habit. We do not necessarily respond to information or incentives. Collective behaviors that are supported by norms, whether implicitly through expectations or explicitly through punishments, are even more difficult to change. Policymakers, therefore, must implement bold changes such as large fines or incentives or, in some cases, legal prohibitions, if they wish to change behaviors. Nudges can work, but often existing behaviors and norms claw their way back.

The idea that societies can become stuck in inefficient configurations is not new nor is the idea that institutions and culture co-evolve (Belloc and Bowles 2017). Cultural-institutional co-evolution is documented in case studies, empirical analyses, and experiments. A variety of theoretical models reveal the potential for multiple stable culture-institutional configurations and that the configurations most likely to evolve need not be the most efficient. In fact, the most likely stable configurations can be far from efficient depending on what assumptions one makes about initial conditions, institutional choices, and behavioral adaptations. Even more problematic, norms can reinforce behaviors which hinder bottom-up efforts to escape these inefficient configurations.

A starting point for our analysis is that behaviors across domains are interdependent. Nudges, incentives, and information, for the most part, treat behavior within a domain in isolation. They do not consider the possibility that by changing behavior in one domain, we might induce changes in another domain. That is our focus here, and we consider two distinct mechanisms through which influence might occur.

We refer to the first mechanism as *behavioral spillovers*. In a series of theoretical, computational, and experimental papers, we explore whether people tend to act consistently across domains, i.e. whether they try to avoid dissonance, and what the implication of those spillovers might be (Bednar et al

2007, 2015, 2018, 2025). In brief, those papers provide a framework for exploring how cultures and institutions co-evolve (culture depends on institutions and institutional choices depend on culture), how behavior diversity reduces path dependence, and on why sequencing of institutional choices matters.

The second mechanism, *domain externalities* is, to our knowledge, new to the literature on cultural-institutional co-evolution. A domain externality occurs if behaviors in one domain alter the payoffs to actions in another domain. In the model that follows, we will assume two ecological domains. If people cooperate in those domains, this will raise the value of cooperating in the economic domain. In the terminology of economics, the positive externality depends upon a particular equilibrium being chosen in another game.

Consider a city situated along a river. Actions in the economic domain might be raising taxes to build public libraries and parks. Actions in the ecological domains might affect water quality and air quality. The economic value of cooperating to build parks depends on whether people cooperate to clean the river and the air. Sitting on a park bench beside a clean river, breathing clean air, has higher utility than looking over a polluted river breathing soot. Therein lies the externality.

We can also motivate this assumption more broadly. Assume that a population of agents wish to maximize a broad metric of human flourishing, health, and well-being, such as described by United Nations' Sustainable Development Goal 3 (SDG 3): "To ensure healthy lives and promote well-being for all at all ages." Improvements in SDG 3 depend in part on progress toward sustainability goals: climate action (SDG 13), life on land (SDG 15), and life below water (SDG 14). Climate action, ecosystem preservation, and clean water should all increase our health and well-being. Thus our assumption aligns with evidence showing that progress on one SDG often impacts on the utility or value of progress on other

SDGs (Breuer et al 2019). Here, we take the effect to be positive, which is most often the case, but it could also be negative.

To capture these two effects formally, we do assume that in each domain, an institution determines a game. Individuals receive a benefit (α) if they exhibit similar behavior across the various games. This is the behavioral spillover. The domain externality applies only to an economic domain. Whether it exists depends on the equilibrium that arises in the ecological domains.

Our model assumes a populations of players who interact in environmental and socioeconomic domains. Creating new behaviors and norms to achieve these goals requires policy interventions, often in the form of new institutions. The goal of these policy interventions should be to create pathways to new, more efficient stable configurations through changes in behaviors and norms. This model demonstrates how domain externalities, together with behavioral spillovers, can increase the likelihood of change. This model extends previous work that considers only behavioral spillovers (Bednar and Page 2018).

Moving from an inefficient stable equilibrium to a new, better equilibrium involves escaping a basin of attraction. Here, we encounter opposing conceptions of stability between social and evolutionary systems in which equilibria are peaks on landscapes and in physics in which equilibria are minima of potential functions. In social and evolutionary systems, adaptation and selection keep populations near a peak on a payoff or fitness landscape. In physical systems, gravity keeps a ball at the bottom of a basin.

Peaks and basins can be measured both horizontally and vertically. The horizontal measure gives the value of the equilibrium. The vertical measure gives the size of the basin of attraction.¹ Both contribute to the difficulty of escape. Higher peaks (deeper basins) create more attraction. Wider basins of attraction require more movement away from the local peak (minimum) to

¹Traditionally, *basin of attraction* is used for peaks as well as basins.

escape. In some models, peak height and the width of basins of attraction are perfectly correlated. When that holds, higher peaks will be more stable. That need not be the case. Lower peaks can have larger basins of attraction, and thus be more likely to arise.

Within our theoretical framework, we show that if individuals recognize the interdependence among their goals, then policies that create sustainability norms yield higher payoffs through their indirect effects on well-being. Further, sustainability norms in one domain can spread to other domains, resulting in an equilibrium with better outcomes across multiple goals. Together, these effects bootstrap the system toward a more efficient configuration. If individuals perceive only weak linkages among goals, sustainability norms are less likely to take hold in any domain, and the existing inefficient equilibrium appears robust.

A Three Domain Model

To provide an introduction to our general framework, we begin with a simple model where a population of agents interact in three domains, one socioeconomic and two that are environmental. Within each domain, an institution specifies the set of actions and their payoffs. In other words, each institution creates a game, which we assume to be a variation of a *Stag Hunt* in which each player $n \in \{1, 2\}$ chooses an action

$$a_n \in \{S, H\},$$

where S denotes *Stag* and H denotes *Hare*. Payoffs are symmetric and given by

	S	H
S	(R, R)	(Q, T)
H	(T, Q)	(P, P)

with payoff parameters satisfying

$$R > T \quad \text{and} \quad P > Q.$$

The Stag Hunt game has two pure-strategy Nash equilibria, (S, S) and (H, H) . Mutual cooperation on the (S, S) yields the higher payoff but (H, H) is risk-dominant. In the language of the introduction, this means that (H, H) has a wider basin of attraction, but that (S, S) creates the higher peak. We cannot, therefore, expect the more likely outcome to be the better one. In the language of game theory, the risk-dominant equilibrium need not be the payoff dominant equilibrium.

To create *domain externalities*, we assume that if players choose (S, S) in an environmental domain, they increase the payoffs to choosing S in the social domain. We formalize that using a parameter β . Denote the two environmental domains by $E1$ and $E2$ and $p_{S,Ei}$ denote the proportion of players who choose S in Ei . We then increase the payoff to S in the social domain by $\beta(p_{S,E1} + p_{S,E2})$ as shown in the matrix below.

The resulting payoff matrix for the *social* domain X is given by

	S	H
S	$(R + \beta(p_{S,E1} + p_{S,E2}), R + \beta(p_{S,E1} + p_{S,E2}))$	$(Q + \beta(p_{S,E1} + p_{S,E2}), T)$
H	$(T, Q + \beta(p_{S,E1} + p_{S,E2}))$	(P, P)

To capture *behavioral spillovers*, we assume that individuals receive an additional payoff of α if they play the same strategy in all three games. This additional payoff can represent the social value of conformity. If the behavior is a socially-enforced norm, it is equivalent to assuming that deviating from choosing the same strategy in all three games confers a punishment.

For the purposes of illustration, we assume the following payoffs for the Stag

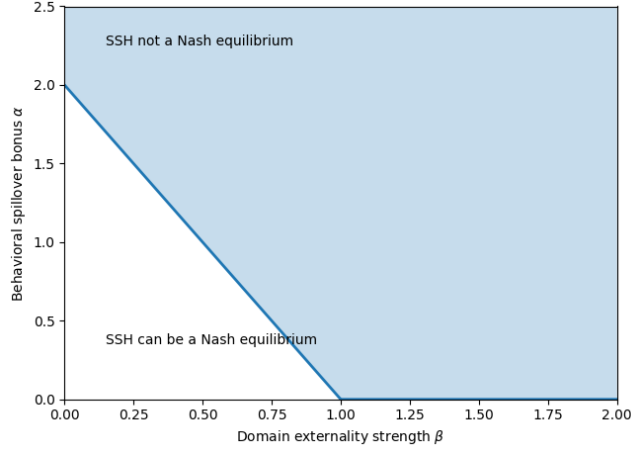


Figure 1: Spillover and Externalities and Support of *SSH*

Hunt game in all three domains (not including the externalities).

	<i>S</i>	<i>H</i>
<i>S</i>	(4, 4)	(0, 3)
<i>H</i>	(3, 0)	(2, 2)

Define a **configuration** to be an equilibrium in each of the three domains $E1, E2, X$. There are eight possible configurations.

As α and β vary, so do the possible equilibrium configurations. When $\alpha = 0$ and $\beta = 1.5$, domain externalities link the social domain to the environmental domains. Coordination on (S, S) in the environmental domains raises the payoff to choosing S in the social domain, making the configuration *SSH* no longer supportable. Thus, the domain externalities can produce behavioral consistency without any behavioral spillovers.

If we make behavioral spillovers large ($\alpha > 2$), then the only configurations that can be supported are *HHH* and *SSS*. Figure 2 below shows the values of α and β at which the configuration *SSH* can no longer be supported.

Whether or not *HHS* is an equilibrium configuration only depends on the

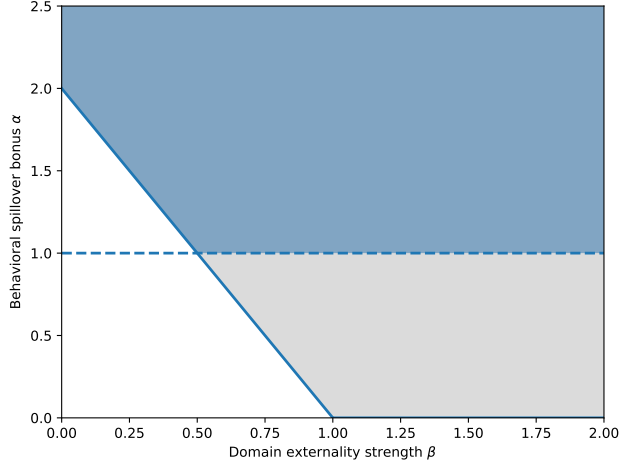


Figure 2: Spillovers and Externalities that Support Only Homogeneous Configurations

size of the behavioral spillover, α . If $\alpha > 1$, then HHS is not supportable because an individual receives a higher payoff by deviating to H in the social domain. Combining this result with the previous result, we can characterize the value of α and β for which only the behaviorally consistent configurations HHH and SSS are equilibria. This is shown in the blue region of Figure 2.

Increases in the domain externality can never eliminate the HHH equilibrium configuration. Neither can increases in the behavioral spillover. Nor can any combination of the two.

Many Domain Model

We can extend the model to allow for an arbitrary number of domains. For convenience, we represent each as a Stag Hunt game. Between each pair of domains, we can assume a domain externality β_{ij} , which describes how much playing (S, S) in domain i increases payoffs in domain j . We can also assume a behavioral spillover $\alpha_{ij} = \alpha j i \geq 0$, which characterizes the benefit from taking

similar actions in domains i and j .

A larger system of domains with externalities and spillovers could support multiple stable configurations. With N domains, there would exist 2^N possible configurations, but the spillovers and externalities would rule out many of these. And, as we saw in the case of three domains, as both effects become larger we would expect pressure toward similar actions in all domains.

In an extension, we seek to explore which domains offer *lever points* to social change.

References

- Acemoglu, Daron and James A Robinson. 2021. “Culture, institutions and social equilibria: A Framework” National Bureau of Economic Research.
- Adler, Paul S. 2001. “Market, Hierarchy, and Trust: The Knowledge Economy and the Future of Capitalism.” *Organization Science* 12(2):215-234.
- Azeem, Muhammad, Munir Ahmed, Sajid Haider, and Muhammad Sajjad, 2021. “Expanding competitive advantage through organizational culture, knowledge sharing and organizational innovation,” *Technology in Society*, 66:101635.
- Bednar, Jenna. 2023. “Governance for Human Social Flourishing.” *Daedalus* 152 (1): 31–45.
- Bednar, Jenna, and Scott E Page. 2007. “Can game (s) theory explain culture? The emergence of cultural behavior within multiple games.” *Rationality and Society* 19(1): 65-97.
- Bednar, Jenna, Andrea Jones-Rooy, and Scott E. Page. 2015. “Choosing a future based on the past: Institutions, behavior, and path dependence.” *European*

Journal of Political Economy 40: 312-332.

Bednar, Jenna and Scott E. Page. 2018. "When Order Affects Performance: Culture, Behavioral Spillovers, and Institutional Path Dependence." *American Political Science Review*. 112(1):82-98.

Bednar, Jenna and Scott E. Page. 2025. "Institutions and cultural capacity: A systems perspective." *Journal of Economic Behavior & Organization* 234: 106990.

Belloc, Marianna, and Samuel Bowles. 2013 "The persistence of inferior cultural-institutional conventions." *American Economic Review* 103(3):93-98.

Belloc, Marianna, and Samuel Bowles. 2017."Persistence and change in culture and institutions under autarchy, trade, and factor mobility." *American Economic Journal: Microeconomics* 9(4):245-276.

Berg, Joyce, John Dickhaut, Kevin McCabe. 1995. "Trust, Reciprocity, and Social History." *Games and Economic Behavior* 10:: 123, 122-142

Bicchieri, Christine, and Xiao, Erte. 2009. "Do the right thing: but only if others do so." *Journal of Behavioral decision-making*, 22(2), 191-208.

Bicchieri, Cristina. 2016 *Norms in the wild: How to diagnose, measure, and change social norms*. Oxford University Press, Oxford, UK.

Bisin, Alberto, and Thierry Verdier. 2021 "On the Joint Evolution of Culture and Institutions: Elites and Civil Society." Working Paper, NYU Department of Economics.

Bowles, Sam, 2011. "Is liberal society a parasite on tradition?" *Philosophy & Public Affairs*, 39(1), pp.46-81.

- Boyd, Robert and Peter J Richerson. 2009. "Culture and the evolution of human cooperation." *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1533), 3281-3288.
- Breuer, A., Janetschek, H. and Malerba, D., 2019. "Translating sustainable development goal (SDG) interdependencies into policy advice." *Sustainability*, 11(7), p.2092.
- Cameron, Kim, and Robert Quinn (eds). 2006. *Diagnosing and Changing Organizational Culture: Based on the Competing Values Framework*. Jossey-Bass, San Francisco, CA.
- Chatman, J. and C.A. O'Reilly. 2016. "Paradigm lost: reinvigorating the study of organizational culture." *Research in Organizational Behavior*, 36 pp. 199-224.
- Chwe, Michael 2003. *Rational Ritual: Culture, Coordination, and Common Knowledge* Princeton University Press. Princeton, NJ.
- Debreu, Gerard 1972. *Theory of Value: An Axiomatic Analysis of Economic Equilibrium*(Cowles Foundation Monographs Series) Yale University Press, New Haven, CT.
- Delhey, Jan, Kenneth Newton, and Christian Welzel. 2011. "How general is trust in "most people"? Solving the radius of trust problem." *American sociological review* 76, no. 5: 786-807.
- DiMaggio, Paul. 1997. "Culture and cognition." *Annual review of sociology* 23.1: 263-287.
- DiMaggio, P. J., and W. W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review*, 48(2), 147-160.

- Dobbin F., Kalev A. 2022. *Getting to diversity: What works and what doesn't*. Harvard University Press.
- Fiske, S.T. and S.E. Taylor. 1991. *Social Cognition*. New York, NY: McGraw-Hill
- Gelfand, Michelle. 2019. *Rule makers, rule breakers: Tight and loose cultures and the secret signals that direct our lives*. Scribner. New York, NY.
- Gelfand, M. J., et. al. 2021. "The relationship between cultural tightness–looseness and COVID-19 cases and deaths: a global analysis." *The Lancet Planetary Health*, 5(3), e135-e144.
- Giddens, Anthony. 1984. *The Constitution of Society: Outline of the Theory of Structuration*. University of California Press, Berkeley, CA.
- Gorodnichenko, Yuriy and Gerard Roland. 2017 "Culture, Institutions, and the Wealth of Nations." *The Review of Economics and Statistics* 99 (3): 402–416.
- Greif, Avner, and David D. Laitin. 2004. "A Theory of Endogenous Institutional Change" *American Political Science Review* 98 (4): 633–52.
- Handy, Charles B. 1994. *The Age of Paradox*, Harvard Business Press, Cambridge, MA
- Harrison, J.R., Z. Lin, G.R. Carroll, K.M. Carley. 2007. "Simulation modeling in organizational and management research" *Academy of Management Review*, 32(4):1229-1245.
- Herrmann, Benedikt. Christian Thoni, and Simon Gächter. 2008. "Antisocial Punishment Across Societies." *Science* 319:1362.
- Hofstede, Gerte, 2002. "Dimensions do not exist: a reply to Brendan Mc-

Sweeney *Human Relations*, 55(11); pp. 1355-1360

Inglehart, Ron (ed.). 2004. *Human beliefs and values: A cross-cultural source-book based on the 1999-2002 values surveys*. Siglo XXI.

Johnson, N. D., and Mislin, A. A. 2011. "Trust games: a meta-analysis." *Journal of Economic Psychology* 32, 865-889.

Jung D, and D.A. Lake. 2011. "Markets, hierarchies, and networks: an agent-based organizational ecology." *American Journal of Political Science* 55(4): 972-990.

Jung, T. J.T. Scott, H.T.O. Davies, P. Bower, D. Whalley, R. McNally, et al. 2009 "Instruments for reviewing organizational culture: A review of the literature." *Public Administration Review*, 69(6): 1087-1096

Kuhn, Kristine. 2009. "Compensation as a signal of organizational culture: the effects of advertising individual or collective incentives," *The International Journal of Human Resource Management*, 20:7, 1634-1648.

Kitayama, Shinobu, and Toshie Imada. 2010. "Implicit independence and interdependence." *The Mind in Context* 174-200.

Kranton, Rachel E. 1996. "Reciprocal Exchange: A Self-Sustaining System," *American Economic Review* 86(4): 830-851.

Laitin, David and Barry Weingast. 2006. "An Equilibrium alternative to the Study of Culture." *The Good Society*, 15(1), pp.15-20.

Mayer, R. C., Davis, J. H., and F.D. Schoorman. 1995. "An Integrative Model of Organizational Trust." *The Academy of Management Review*, 20(3), 709-734.

Mintzberg H. 1979. *The Structuring of Organizations*. Englewood Cliffs, NJ:

Prentice Hall.

Mokyr, Joel. 2014 “Culture, institutions, and modern growth.” *Institutions, property rights, and economic growth: The legacy of Douglass North*: 151-191.

Nadler, David and Michael Tushman.1980. “A Model for Diagnosing Organizational Behavior.” *Organizational Dynamics*, 9(2):35-51.

Nebojsa Janicijevic. 2013. ”The mutual impact of organizational culture and structure.” *Economic Annals* 58(198):35-60.

Pavetic, Suncica, Ana Aleksic, and Najla Podrug. 2023. “Remote work culture: Literature review.” *Economic and Social Development: Book of Proceedings*: 58-67.

Powell, Walter W. 1990. “Neither Market nor Hierarchy: Network Forms of Organization.” *Research in Organizational Behavior* 12: 295–336.

Powell, W. W., and DiMaggio, P. J. 2023. “The Iron Cage Redux: Looking Back and Forward.” *Organization Theory*. 4(4).

Priem, Richard, L. 1994. “Executive Judgment, Organizational Congruence, and Firm Performance.” *Organization Science* 5(3):421-437.

Quinn, Robert E and John Rohrbaugh. 1981 “A Competing Values Approach to Organizational Effectiveness” *Public Productivity Review*. 5(2)122-140.

Scott, W Richard, and Gerald F. Davis. 2000. *Organizations and organizing: Rational, Natural and Open Systems Perspectives* Routledge, New York, NY.

Schein Edward. 2010. *Organizational Culture and Leadership*. 4th ed. Jossey-Bass

Sivadas, E., N. T. Bruvold,, and M.R. Nelson. 2008. “A reduced version of

- the horizontal and vertical individualism and collectivism scale: A four-country assessment.” *Journal of Business Research*, 61(3), 201-210.
- Strang D., Macy M. W. 2001. “In search of excellence: Fads, success stories and adaptive emulation.” *American Journal of Sociology*, 107(1), 147–182
- Triandis, Harry C., and Michele J. Gelfand. 1998. “Converging measurement of horizontal and vertical individualism and collectivism.” *Journal of personality and social psychology* 74.1 118.
- Way, Rupert, Matthew C. Ives, Penny Mealy, and J. Doyne Farmer. 2022 “Empirically grounded technology forecasts and the energy transition.” *Joule* 6(9): 2057-2082.
- Whittington, Richard. 2010. “Giddens, Structuration Theory and Strategy as Practice.” *Cambridge Handbook of Strategy as Practice* 109-126.
- Young, H. Peyton. 1996 “The economics of convention.” *Journal of Economic Perspectives*.10(2): 105-122.
- Zak, Paul J., and Stephen Knack. 2001, “Trust and growth.” *The Economic Journal* 111.470: 295-321.