14.773 Political Economy of Institutions and Development.
Lectures 10-12: Culture, Norms and Institutions

Daron Acemoglu

MIT

March 2023.

What about Culture? Norms?

- One view is that economic performance and incentives are partly or largely shaped by culture, often equated with religious or national cultural characteristics determining beliefs, preferences and customs (e.g., Landes or the essays in Harrison and Huntington).
- Example:
  - Europe has more growth-enhancing culture than Africa, northern Italy more than southern Italy etc.
- What's the difference between culture and norms?
- Most important challenge: if culture is so important, and very slow-changing, how do economic incentives and performance change sharply as exemplified by China?

Culture-Institutions Interactions

- An alternative, more sophisticated view, is that culture and institutions interact, and "culture" is more about **social norms** than national cultural characteristics.
- Simple examples:
  - China had an "authority-respecting" culture under Mao and now has a more "individualistic" culture.
  - Or North Korea and South Korea now have very different "cultures".
- A related perspective is that culture and attitudes play some of the same role as "de facto power" in what has been described so far.
  - What is the difference between culture and beliefs?
- If so, "culture" could be a mechanism for persistence or even an autonomous force affecting how society is organized.
- Equally importantly, if so, we should really study the determinants of cultural change (or better, changes in social norms).

Culture-Institutions Interactions

- Perhaps

$$\text{culture}_t \Longrightarrow \text{economic performance}_t \Longrightarrow \text{culture}_{t+1}$$

- Or

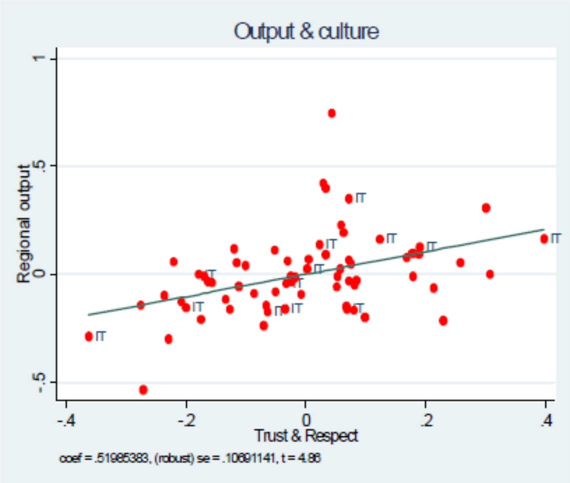
$$\left.\begin{array}{c}\text{political} \\ \text{inst's}_t \\ \\ \text{culture}_t\end{array}\right\} \Longrightarrow \begin{array}{l}\text{econ.} \\ \text{inst's}_t\end{array} \Longrightarrow \left\{\begin{array}{c}\text{econ. perf}_t \\ \\ \text{culture}_{t+1} \\ \\ \text{political inst's}_{t+1}\end{array}\right.$$

Evidence on Effect of Culture

- Plenty of correlation evidence and several studies showing historical effects, but sometimes difficult to interpret.
- Tabellini (2008): use differences in constraints on executive across a subnational units (within current nations) in the past, as an instrument for contemporary culture variables (in particular, generalized vs. clan or family-based trust, as in Banfield).
- Identification assumption: constraints on the executive in the distant past to not have a direct effect. Is this plausible? Tabellini argues yes because there are country fixed effects. But potentially, no if institutions persist at the local level.
- Guiso, Sapienza and Zingales (2009) and Fernandez and Fogli (2009) show that certain behavioral patterns (e.g., fertility and women's labor supply) of second or third generation immigrants are highly correlated with these variables in their country of origin. They argue that the channel must be cultural.

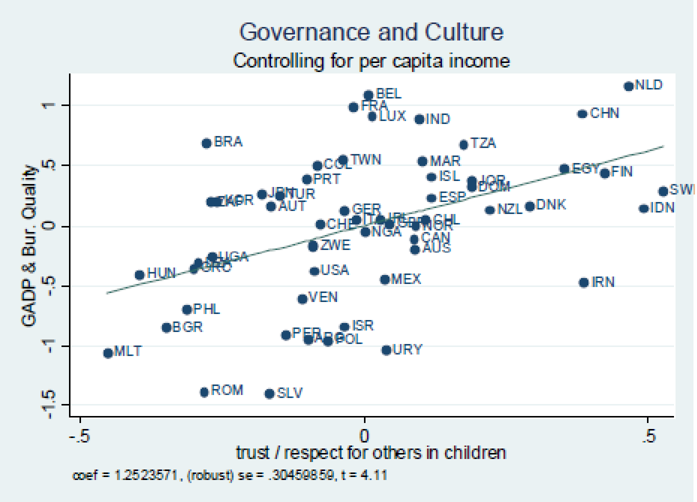
Evidence on Effect of Culture (continued)

- Tabellini (2008): relationship between generalized trust and income per capita



Evidence on Effect of Culture (continued)

- Potentially working through institutions



Evidence on Effect of Culture (continued)

	Trust & Respect	Per capita output 1995-00	Trust & Respect	Growth 1977-2000
Trust & respect		0.69 (0.29)**		1.40 (0.59)**
Past Constr. on executive	0.32 (0.04)***		0.29 (0.07)***	
Log per capita output 1977			0.23 (0.10)**	-1.32 (0.36)***
Estimation	OLS	2SLS	OLS	2SLS
Observations	67	67	67	67

Interpretation

- What could be interpretation be?
- Threat to validity:
  - is what persists culture?
  - is culture adapting or driving?

Evidence on Effect of Culture (continued)

- Fernandez and Fogli (2009) persistent effects of culture on values and behavior

	Dependent variable is hours worked					Dependent variable is children			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Female LFP 1950	0.047***	0.041**	0.072***	0.045***	0.053***				−0.010
	(0.012)	(0.016)	(0.015)	(0.014)	(0.016)				(0.008)
TFR 1950					−0.225**	0.250***	0.219***	0.219***	0.194***
					(0.103)	(0.056)	(0.041)	(0.041)	(0.051)
High school		0.490	2.136***	2.114***	2.059***		−0.415**	−0.393***	−0.378**
		(0.520)	(0.575)	(0.511)	(0.572)		(0.181)	(0.151)	(0.147)
Some college		−0.147	3.205***	3.336***	3.160***		−0.503**	−0.485***	−0.457**
		(1.078)	(1.034)	(0.963)	(1.024)		(0.213)	(0.185)	(0.179)
College +		0.815*	6.032***	6.744***	5.968***		−0.869***	−0.865***	−0.838***
		(0.492)	(0.494)	(0.448)	(0.480)		(0.214)	(0.204)	(0.195)

Interpretation

- What persists from home country to the United States?
- Threats to validity:
  - Do second-generation immigrants from different countries live under similar circumstances? Are they fully integrated?
  - In the same or different neighborhoods?
  - Are they subject to the same labor market opportunities (because of discrimination or externalities)?
- Is this culture, norms or something else? Social organization? Neighborhoods?

Evidence on Effect of Culture (continued)

- Guiso, Sapienza and Zingales: bilateral trust as a function of religious and genetic overlap (why?)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Common language	0.05	0.09*	0.11*	0.09*	0.08	0.02	0.04	0.05	0.08
	(0.07)	(0.05)	(0.06)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)
Log (distance)	−0.11***	−0.04*	−0.05*	−0.04	−0.01	−0.01	0.00	0.01	−0.01
	(0.03)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
Common border	−0.01	−0.05	−0.01	−0.05	−0.04	−0.04	−0.05	−0.05	−0.03
	(0.05)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
Fraction of years at war		−1.16***	−1.07***	−1.16***	−1.07***	−1.16***	−1.26***	−1.26***	−1.07***
(1000–1970)		(0.29)	(0.39)	(0.29)	(0.29)	(0.29)	(0.39)	(0.39)	(0.39)
Religious similarity		0.15***	0.24***	0.15***	0.15***	0.11**	0.13***	0.13***	0.15***
		(0.04)	(0.05)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)
Somatic distance		−0.06***		−0.06***	−0.05***	−0.04***	−0.04***	−0.04***	−0.03***
		(0.01)		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Genetic distance			−10.00*	0.06					
			(5.94)	(5.07)					
Differences in GDP per					−0.14***	−0.14***	−0.13***	−0.13***	−0.09**
capita (percentage)					(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
Same legal origin						0.07**	0.05	0.05	0.05
						(0.03)	(0.03)	(0.03)	(0.04)
Linguistic common roots							0.20*	0.20*	0.21*
							(0.11)	(0.11)	(0.11)
Transportation costs*1,000								−0.58	−1.05
								(1.00)	(0.96)
Press coverage									−0.73**
									(0.34)
Observations	207	207	207	207	207	207	180	180	154
R^2	.772	.840	.806	.840	.854	.858	.832	.832	.837

Evidence on Effect of Culture (continued)

- Export from country i to j on trust of country j to i

	OLS (1)	OLS (2)	OLS (3)	OLS (4)	IVGMM (5)	OLS (6)
Mean trust of people in importing country to people in exporting country	0.36**	0.29*	0.25	0.34**	1.20***	0.19
	(0.17)	(0.17)	(0.19)	(0.16)	(0.20)	(0.22)
Interaction between trust and diversified good						0.83***
						(0.05)
Common language	0.58***	0.32**	0.37**	0.82***	0.94***	1.04***
	(0.22)	(0.16)	(0.16)	(0.21)	(0.14)	(0.27)
Log (distance)	−0.31***	−0.43***	−0.43***	−0.57***	−0.61***	−0.73***
	(0.09)	(0.09)	(0.09)	(0.10)	(0.07)	(0.12)
Common border	0.49***	0.43***	0.41***	0.41***	0.36***	0.35***
	(0.11)	(0.10)	(0.11)	(0.10)	(0.06)	(0.13)
Press coverage	0.45	−0.03	−0.09	−1.34	−0.89	−2.83**
	(1.05)	(0.93)	(0.94)	(1.0)	(0.60)	(1.12)
Transportation costs	−1.81**	−0.33	−0.28	0.10	0.63	−1.83
	(0.79)	(0.74)	(0.76)	(0.73)	(0.52)	(1.17)
Same legal origin		0.45***	0.43***	0.36***	0.24***	0.57***
		(0.10)	(0.10)	(0.11)	(0.07)	(0.15)
Linguistic common roots			0.09			
			(0.28)			
Correlation of consumption between the two countries				−0.95	−1.05***	−1.82**
				(0.68)	(0.37)	(0.89)

Evidence on Effect of Culture (continued)

- FDI from country i to j on trust of country i to j

	OLS (1)	OLS (2)	OLS (3)	OLS (4)	IVGMM (5)
Mean trust toward people in destination country	1.35***	0.94*	0.70	0.84*	6.65***
	(0.51)	(0.51)	(0.48)	(0.49)	(1.24)
Common language	0.12	0.17	−0.57*	−0.75**	−2.05***
	(0.31)	(0.29)	(0.30)	(0.38)	(0.43)
Log (distance)	−0.46*	−0.22	−0.48**	−0.56**	−0.70***
	(0.26)	(0.27)	(0.23)	(0.24)	(0.26)
Common border	0.47**	0.44**	0.26	0.34	0.26
	(0.20)	(0.20)	(0.20)	(0.21)	(0.21)
Press coverage	2.65	1.67	0.76	1.00	8.97***
	(2.29)	(2.18)	(2.04)	(2.24)	(2.69)
Transportation costs		−4.55**	−0.23	−0.32	5.13**
		(1.76)	(1.66)	(1.80)	(2.31)
Common law			1.28***	1.36***	1.38***
			(0.27)	(0.31)	(0.26)
Linguistic common roots				−0.86	−2.41***
				(0.55)	(0.66)
Investing-country fixed effects*years	YES	YES	YES	YES	YES
Destination-country fixed effects*years	YES	YES	YES	YES	YES
Observations	445	445	445	419	419
R^2	.854	.860	.879	.880	
Hansen J-statistic					0.031
χ^2 p-value					.859
Test of excluded instruments in first stage					$F(2,328) =$ 24.34

Interpretation

- History vs. culture/trust.
- Reverse causality.
- What does genetic and religious distance capture?

Overall Interpretation

- While there are potential issues with any single study on institutions and on culture, in both cases the body of evidence probably suggests that these regressions are capturing something real.
- Interpretation is in general open and needs to be done in light of some theoretical and historical ideas.

Introduction

- How do we model the effects of culture and values on social and political outcomes—and through which mechanisms?
- Why do these values persist?
- How they interact with institutions?
- In this lecture, an overview of some related research.

Intergenerational Transmission: Basic Models

- Cavalli-Sforza and Feldman (1981) and to Boyd and Richerson (1985), based on models of evolutionary biology applied to the transmission of cultural traits.
- Suppose that there is a dichotomous cultural trait in the population, $\{a, b\}$. Let the fraction of individuals with trait $i \in \{a, b\}$ be q^i.
- Focus on a continuous time model with "a-sexual" reproduction where each parent has one child at the rate λ and is replaced by the child.
- Two types of cultural transmission:
  1. *direct/vertical* (parental) socialization and
  2. *horizontal/socialization* by the society at large.

Intergenerational Transmission (continued)

- Suppose that direct vertical socialization of the parent's trait, say i, occurs with probability d^i.
- Then, if a child from a family with trait i is not directly socialized, which occurs with probability $1 - d^i$, he/she is horizontally/obliquely socialized by picking the trait of a role model chosen randomly in the population (i.e., he/she picks trait i with probability q^i and trait $j \neq i$ with probability $q^j = 1 - q^i$).
- Therefore, the probability that a child from family with trait i is socialized to have trait j, P^{ij}, is:

$$\begin{aligned} P^{ii} &= d^i + (1 - d^i)q^i \\ P^{ij} &= (1 - d^i)(1 - q^i). \end{aligned} \tag{1}$$

Intergenerational Transmission (continued)

- Now noting that each child replaces their parent in the population (at the rate λ), we have that

$$\dot{q}^i = \lambda \left[\left(d^i + (1-d^i) q^i \right) q^i + (1-d^j) q^i \left(1 - q^i \right) \right] - \lambda q^i.$$

- Simplifying this equation, we obtain:

$$\dot{q}^i = \lambda q^i (1-q^i) \left(d^i - d^j \right). \tag{2}$$

- This is a version of the replicator dynamics in evolutionary biology for a two-trait population dynamic model—i.e., a logistic differential equation.
- If $\left(d^i - d^j \right) > 0$ cultural transmission represents a selection mechanism in favor of trait i, due to its differential vertical socialization.
- However, this selection mechanism implies that there will not be cultural heterogeneity, i.e., a steady-state with $0 < q^{i*} < 1$.

Intergenerational Transmission (continued)

- The following result is now immediate.
- Let $q^i(t, q_0^i)$ denotes the fraction with trait q^i at time t starting with initial condition q_0^i. Then:

Proposition

Suppose $d^i > d^j$. Then, steady states are culturally homogeneous. Moreover, for any $q_0^i \in (0, 1]$, $q^i(t, q_0^i) \to 1$. If instead $d^i = d^j$, then $q^i(t, q_0^i) = q_0^i$, for any $t \geq 0$.

Intergenerational Transmission: Bisin-Verdier Model

- Bisin and Verdier (2000, 2001) introduce "imperfect empathy" into this framework, whereby parents look at the world with their own preferences and thus wants to socialize their offspring according to their preferences.
- Formally, suppose that individuals choose an action $x \in X$ to maximize a utility function $u^i(x)$, which is a function of the cultural trait $i \in \{a, b\}$. Suppose that this utility function is strictly quasi-concave.

Intergenerational Transmission (continued)

- Let V^{ij} denote the utility of a type i parent of a type j child, $i, j \in \{a, b\}$. Then clearly, we have

$$V^{ij} = u^i(x^j)$$

And

$$x^j = \arg\max_{x \in X} \ u^j(x)$$

- This implies the "imperfect empathy" feature:

$$V^{ii} \geq V^{ij}$$

holding with $>$ for generic preferences (i.e., in particular when the maximizers for the two types are different).

Intergenerational Transmission (continued)

- Suppose also that parents have to exert costly effort in order to socialize their children. In particular, parents of type i choose some variable τ^i, which determines

$$d^i = D\left(q^i, \tau^i\right).$$

  The dependence on q captures other sources of direct transmission working from the distribution of traits in the population.
- The cost of τ^i is assumed to be $C\left(\tau^i\right)$.
- Suppose that D is continuous, strictly increasing and strictly concave in τ^i, and satisfies $D\left(q^i, 0\right) = 0$, and C is also continuous, strictly increasing and convex. Moreover, suppose also that $D\left(q^i, \tau^i\right)$ is nonincreasing in q^i.
- Parents of type i will solve the following problem:

$$\max_{\tau^i} -C(\tau^i) + P^{ii}V^{ii} + P^{ij}V^{ij},$$

  where P^{ii} and P^{ij} depend on τ^i via d^i.

Intergenerational Transmission (continued)

- Let us say that the *cultural substitution property* holds if the solution to this problem d^{i*} is is strictly decreasing function of q^i and takes a value $d^{i*} = 0$ at $q^i = 1$. Intuitively, this implies that parents have less incentives to socialize their children when their trait is more popular/dominant in the population.
- This cultural substitution property is satisfied in this model.
- Then, the dynamics of cultural transmission can be more generally written as

$$\dot{q}^i = \lambda q^i (1 - q^i) \left(d^i \left(q^i \right) - d^j \left(1 - q^i \right) \right). \tag{3}$$

- We can also verify that this differential equation has a unique interior steady state, q^{i*}, and moreover,

Proposition

The steady states are now culturally heterogeneous. In particular, $q^i(t, q_0^i) \to q^{i*}$, for any $q_0^i \in (0, 1)$.

Intergenerational Transmission (continued)

- Intuition: the cultural substitution property implies that parents put more effort in socializing their children, i.e., passing on their traits, when their traits are less common in the cooperation.
- The proof of this result follows from the following observations:
  1. Clearly, an interior steady state satisfies

$$d^i\left(q^i\right) - d^j\left(1 - q^i\right) = 0,$$

     and since both d^i and d^j are strictly decreasing, there can at most be one such steady state q^{i*}.
  2. Moreover, since $d^i(1) = 0$, existence is guaranteed.
  3. Global stability then follows from the fact that this pattern implies that $\dot{q}^i > 0$ whenever $q^i \in \left(0, q^{i*}\right)$ and at $\dot{q}^i < 0$ whenever $q^i \in \left(q^{i*}, 1\right)$.

Culture, Values and Cooperation

- Tabellini (2009) considers the following variation on the static prisoners' dilemma game.
- Individuals incur a negative disutility from defecting, but the extent of this disutility depends on how far their partner is according to some distance metric.
    - The most interesting interpretations of this distance are related to "cultural distance" or "kinship distance". For example, some individuals may not receive any disutility from defecting on strangers, but not on cousins.
    - This captures notions related to "generalized trust".

Model

- A continuum of one-period lived individuals, with measure normalized to 1, is uniformly distributed on the circumference of a circle of size $2S$, so that the maximum distance between two individuals is S.
- A higher S implies a more "heterogeneous" society—in geography, ethnicity, religion or other cultural traits.
- Each individual is (uniformly) randomly matched with another located at distance y with probability $g(y) > 0$, and naturally

$$\int_0^S g(y) = 1.$$

Model (continued)

- A matched pair play the following prisoners' dilemma:

	C	D
C	c, c	$h - l, c + w$
D	$c + w, h - l$	h, h

- Naturally, $c > h$ and $l, w > 0$. Let us also suppose that $l \geq w$, so that the loss of being defected when playing cooperate is no less than the reverse benefit.

Model (continued)

- In addition, each individual enjoys a non-economic (psychological or moral) benefit

$$de^{-\theta y}$$

  whenever she plays "cooperate" (regardless of what her opponent plays) but as a function of the distance between herself and the other player, y, with the benefit declining exponentially in distance.
- Let us assume that

$$d > \max\{l, w\},$$

  which ensures that this benefit is sufficient to induce cooperation with people very close.

Model (continued)

- Finally, suppose that there are two types of player indexed by $k = 0, 1$, "good" and "bad," modeled as having different rates at which the benefit from cooperation declines. In particular,

$$\theta^0 > \theta^1.$$

- This captures the idea that what varies across individuals (and perhaps across societies) is the level of "generalized trust".
- The fraction of good $(k = 1)$ types in the population is the same at any point in the circle is $1 > n > 0$.

Equilibrium

- Consider a player in a match of distance y.
- Let $\pi(y)$ denote the probability that her opponent will play C.
- We can express the player's net expected *material* gain from defecting instead of laying C as:

$$T(\pi(y)) = [l - \pi(y)(l - w)] > 0 \tag{4}$$

- This is strictly positive, as it is always better not to cooperate given the prisoners' dilemma nature of the game.
- Note also that cooperation decisions are strategic complements, since, given the assumption that $l \geq w$, the function $T(\pi(y))$ is non-increasing in $\pi(y)$

Equilibrium (continued)

- The temptation to defect will be potentially balanced by the non-economic benefit of cooperation, $de^{-\theta^k y}$, as a function of a player's type.
- To simplify the analysis, let us suppose that

$$\frac{\theta^0}{\theta^1} > \frac{\ln(l/d)}{\ln(w/d)} \tag{A0}$$

and also focus on "best" (Pareto superior) and symmetric (independent of location on the circle) equilibria.

- Then a player of type $k = 0, 1$ will be indifferent between cooperating and not cooperating with a partner of distance $\tilde{y}^k$ defined as

$$T(\pi(\tilde{y}^k)) = de^{-\theta^k \tilde{y}^k}, \tag{5}$$

Or as

$$\tilde{y}^k = \left\{ \ln d - \ln \left[(w - l)\, \pi(\tilde{y}^k) + l \right] \right\} / \theta^k. \tag{6}$$

Equilibrium (continued)

- Thus given the equilibrium probability of cooperation $\pi(y)$ (for all y), each individual will cooperate with players closer than $\tilde{y}^k$ ($y < \tilde{y}^k$) and defect against those farther than $\tilde{y}^k$ as a function of her type k.
- Note that if $l > w$, then the right hand side of (6) is increasing in $\pi(y)$, and there are multiple equilibria, though we are ignoring this by focusing on best equilibria.
- Now consider a bad player, $k = 0$, and suppose that she/he expects the opponent always to cooperate, so that $\pi(y) = 1$ (which will be true, since both types of players will cooperate whenever this player is choosing to operate along the equilibrium path).
- Then (6) reduces to:

$$Y^0 = \left[\ln d - \ln w\right] / \theta^0, \tag{7}$$

and player $k = 0$ will cooperate up to distance $y \leq Y^0$.

Equilibrium (continued)

- The problem of a good player is a little more complicated.
- She will necessarily cooperate up to distance $y \leq Y^0$. But beyond that, she recognizes that only other good players will cooperate, and thus $\pi(y) = n$.
- Using this with (6)

$$Y^1 = \left[\ln d - \ln\left[\left(w - l\right) n + l\right]\right] / \theta^1. \tag{8}$$

- And with players cooperate up to Y^1 (which is strictly greater than Y^0 given the assumption above).

Equilibrium (continued)

- Thus summarizing:

Proposition

In the Pareto superior symmetric equilibrium, a player of type k cooperates in a match of distance $y \leq Y^k$ and does not cooperate if $y > Y^k$, where Y^k is given (7)-(8), for $k = 0, 1$.

- This proposition captures, in a simple way, the role of "generalized trust" in society.
- It also highlights the strategic complementarity in trust, as Y^1 is increasing in n: thus good players trust others more when there are more good players. Interestingly, this does not affect bad types, given the simple structure of the prisoners' dilemma game coupled with the assumption that $l \geq w$.

Endogenous Values

- Values can now be endogenized using the same approach as Bisin and Verdier.
- Parents choose socialization effort τ at cost

$$\frac{1}{2\varphi}\tau^{2},$$

  and as a result, their offspring will be over the "good type," i.e., $\theta^k = \theta^1$, with probability $\delta + \tau$.
- As in Bisin and Verdier, they evaluate this with their own preferences, i.e., there is "imperfect empathy".

Endogenous Values (continued)

- Let V_t^{pk} denote the parent of type p's evaluation of their kid of type k's overall expected utility in the equilibrium of the matching game.
- Since the probability of a match with someone located at distance z is denoted $g(z)$, we have

$$V_t^{pk} = U_t^k + d \int_0^{Y_t^k} e^{-\theta^p z} g(z) dz, \quad (9)$$

  where $U_t^k = U(\theta^k, n_t)$ denotes the expected equilibrium material payoffs of a kid of type k, in a game with a fraction n_t of good players. The integral gives the parent's evaluation of their kid's expected non-economic benefit from their offspring's cooperating in matches of distance smaller than Y_t^k.
- This is where imperfect empathy comes in, as this integral term uses the parent's value parameter, θ^p, rather than with the kid's value.

Endogenous Values (continued)

- With the same argument as in Bisin and Verdier, we have that whenever $k \neq p$, then
$$V_t^{pp} > V_t^{pk}$$
where recall that, given the assumptions, $Y^1 > Y^0$.
- The fact that parents of bad type, according to their values, have nothing to gain from exerting effort to socialize their children to be good (as they do not internalize the "moral" benefit from cooperation with farther away partners), and the fact that the marginal cost of exerting effort at zero is zero, implies the following simple result:

Proposition

A "good" parent $(p = 1)$ *exerts strictly positive effort* $\tau_t > 0$*. A "bad" parent* $(p = 0)$ *exerts no effort.*

Endogenous Values (continued)

- Therefore, the law of motion of types in the population follows the following difference equation:

$$n_t = n_{t-1}(\delta + \tau_t) + (1 - n_{t-1})\delta = \delta + n_{t-1}\tau_t. \tag{10}$$

- It can also be shown that the optimal level of effort for with type parents is $\tau_t = F\left(Y_t^1\right)$, where

$$F\left(Y_t^1\right) \equiv \varphi d[-e^{-\theta^1 Y_t^1} + E[e^{-\theta^1 y} \mid Y_t^1 \geq y \geq Y^0]] \Pr(Y_t^1 \geq y \geq Y^0), \tag{11}$$

where intuitively the benefit to good parents depends on the likelihood that their children will play against an opponent of good type, again highlighting the strategic complementarities. The right-hand side of (11), $F\left(Y_t\right)$, is as a result strictly increasing in Y_t^1.

Endogenous Values (continued)

- This means that (10) can be written as

$$n_t = \delta + n_{t-1} F(Y_t^1) \equiv N(Y_t^1, n_{t-1}), \tag{12}$$

with the date t equilibria value of Y_t^1 being defined as:

$$Y_t^1 = \left[\ln d - \ln\left[(w - l)\, n_t + l\right]\right] / \theta^1 \equiv Y(n_t).$$

- Now using the fact that n_t itself is a function of n_{t-1} and Y_t^1 from (10), we can express endogenous value dynamics as in two equations system:

$$Y_t^{1*} = G^Y(n_{t-1}) \tag{13}$$

$$n_t^* = G^n(n_{t-1}) \tag{14}$$

- Strategic complementarities now imply multiple steady state are possible.

Endogenous Values (continued)

- Naturally, additional conditions ensure uniqueness. One such condition would be

$$\frac{1}{\varphi} > I - w \tag{A1}$$

  which ensures that the marginal cost of effort, $1/\varphi$, rises sufficiently rapidly, relative to the strategic complementarity captured by $(I - w)$.
- Given uniqueness, global stability of dynamics can also be ensured. The following proposition gives one sufficient condition

Proposition

Suppose (A1) holds and $\varphi > 0$ is sufficiently small. Then the equilibrium is unique and is globally stable, i.e., it asymptotically reaches the unique steady state (Y_s^{1*}, n_s^*). Moreover, adjustment to steady state is monotone, i.e., the fraction of what types, n_t^*, and the cooperation threshold, Y_t^{1*}, andmonotonically increase or decrease along the adjustment path.

Effects of Institutions

- Let us introduce institutional enforcement of cooperation simply by assuming that there is a probability $\chi(y)$ that defection gets detected when it takes place in a match of distance y and it gets punished.
- We can think of different types of shifts up the schedule $\chi(y)$ as corresponding to different types of changes in institutions.
- In particular, we can imagine that χ increases for high y. This will encourage more broad-based cooperation and it will also incentivize parents to socialize their children to be of the “good” type. As a result, both n_t^* and Y_t^{1*} will increase.
- At the other extreme, we can think of an improvement in local enforcement, with no change in enforcement for faraway matches. This could be considered as a family- or clan-based enforcement, or what the Mafia achieves in southern Italy. This would increase Y^0, so its static effect is good. However, it would also reduce the parental efforts for good socialization, so ultimately it would reduce n_t^* and Y_t^{1*}.

Endogenous Institutions

- One could also endogenize enforcement through a voting or political economy process.
- In this case, one can obtain richer dynamics, where parental socialization interacts with political economy. For example, more with types today leads to greater enforcement, which then encourages more would socialization.
- Multiple steady states are again possible, this time resulting from the interaction of culture and institutions.

An Alternative Framework

- But perhaps this way of thinking about culture—as largely unchanging, and good vs. bad—is not very helpful.
- Let's think of a different approach (based on Acemoglu and Robinson, 2021) centered on the interplay between **social equilibria** (representing the ensemble of political, economic and social arrangements) and **cultural configurations** (representing cultural constraints as we will describe next).
- Social equilibria and cultural configurations are *jointly determined*.
- But in some instances it will be political, economic and other environmental factors that shape cultural configurations, and yet in others, it will be cultural configurations that fundamentally constrain social equilibria.

Culture Set, Attributes and Cultural Configurations

- We view culture as a set of beliefs, relationships, rituals and obligations that provide a framework to people to interpret the world, coordinate expectations and constrain behaviors.
- In contrast to the approaches that assume that cultures correspond to stable values, we argue that a culture is defined by a **culture set**, containing a number of distinct (cultural) **attributes**.
- Example attributes include: definitions of distinct groups in society; types of social hierarchy; various social responsibilities; family structures; specific rituals (such as witchcraft or ancestor worship); interlocking set of obligations (from parents to children or from regular people to elites); and certain higher ideals (such as different types of virtue or honor, importance of order or in society, or values of equality or hierarchy and status).
- A cultural configuration is created from combinations of subsets of attributes in a culture set.

Multitudes of Cultural Configurations

- Crucially, the attributes in the culture set (except in uniquely rigid cultures) can be combined to support a multitude of feasible cultural configurations.
- For example, social hierarchy could be one in which rulers or certain privileged groups have to be obeyed all the time, or it may include the notion that, as in Confucian culture, virtuous leaders should be obeyed.
- The former, especially if it is rigidly specified, may take the form of a caste-based society (literally and figuratively).
- The latter can on the other hand allow different types of social and political organizations.

Social Equilibria and Cultural Configurations

- So what determines which cultural configuration is realized?
- Our framework is predicated on the notion that this is jointly determined with the social equilibrium.
- If political power lies in the hands of a narrow group that can wield this power in order to shape the social equilibrium, a cultural configuration that justifies this type of social equilibrium may arise (if it is feasible).
- But under other circumstances, a very different cultural configuration may arise as part of the social equilibrium. In fact, cultural change can be discontinuous: **Saltational Transformations.**
- Our discussion below of how Confucian culture has supported Imperial institutions, Communist Party rule and democracy in Taiwan and Hong Kong (and South Korea) will illustrate these ideas.

Cultural Constraints

- However, our framework does not imply that cultures are perfectly malleable, adapting to political power, economic conditions or other environmental factors.
- The set of cultural configurations that a culture allows will be, by its nature, finite.
- This means that certain social equilibria will be ruled out (unless the culture can fundamentally change—more on this below).
- This will be all the more so when a culture is **hardwired**.
- For example, as we discuss next, the Indian caste system and the culture set that it is associated with severely constrain economic and political arrangements if they conflict with the pre-ordained hierarchy of groups.

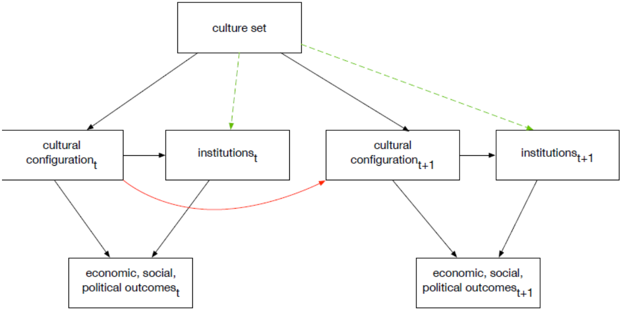
Fluid and Hardwired Cultures

- We say that a culture is **more fluid** than another if it has attributes that can be combined in more distinct ways and thus allow for a *larger set* of cultural configurations.
- By this definition, a fluid culture will allow the formation of more diverse cultural configurations and as a result will be less constraining for social equilibria than a hardwired culture.
- Put differently, a hardwired culture is more likely to be one where cultural configuration act as **hard constraints** on political, social and economic arrangements.

Culture and Social Equilibria

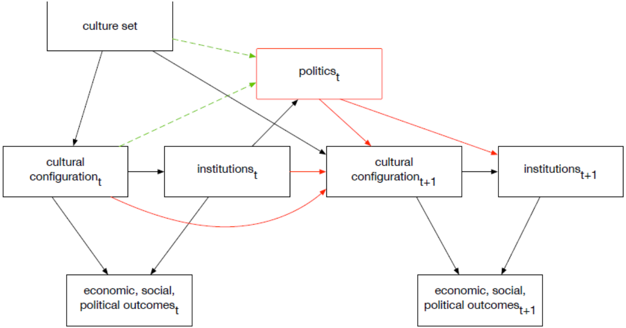
- In principle, cultural configurations and social equilibria are jointly determined—meaning that politics, economic outcomes and social relations shape as much as being shaped by cultural configurations.
- The exception to this is the polar case of extremely hardwired cultures that allow only very limited cultural configurations.
- This gives an extreme special case of our framework corresponding to "vulgar culturalism", where everything is shaped by an unchanging (hugely persistent) culture.

Cultural Configurations and Social Equilibria under an Extreme Hardwired Culture



- Institutions do not affect culture and **politics** is absent.
- We have not dwelt on the green arrows which may also be present (the culture set directly impacting institutions, e.g., via language etc.).

Cultural Configurations and Social Equilibria in General



- Institutions affect culture and politics (and other environmental factors) now play a central role. Culture still plays a role, but in a more limited way, since cultural configurations are adaptable to social conditions and political factors.

Understanding Fluid and Hardwired Cultures

- So what makes a culture more fluid or more hardwired?
- In practice, many factors play a role, but our analysis focuses on two features of its attributes as represented by the next table:

	abstract	specific
free standing	**fluid culture**	intermediate culture
entangled	intermediate culture	**hardwired culture**

- Diagonals are most common and easier to study, but there are some interesting phenomena in off-diagonals as well, as we will discuss.
- Note that there should be no expectation that hardwired cultures will automatically lead to worse social equilibria. The hardwired cultural configurations may be those that favor economic or political development (at least under some circumstances).

Abstract versus Specific Attributes

- We say that an attribute is **abstract** if it can be combined with different attributes in different ways.
- On the other hand, it is **specific** if it narrowly specifies the prescribed roles and/or values and behaviors.
- Because abstract attributes can be combined with others in many different ways, they can be reinterpreted when the social equilibrium or other pressures demand such a change.
- This implies that when attributes are more specific, the set of feasible cultural configurations will be more limited.
- A common attribute, present in many culture sets, is a type of in-group social hierarchy: the in-group can or should be treated better than the out-group.
- But this can be either abstract or specific as we next discuss.

Abstract vs. Specific Attributes in Islamic and British Cultures

- In Islam this social hierarchy attribute is **specific**: Muslims are the in-group.
- The Quran decrees (9.29):
  Fight those of the People of the Book who do not [truly] believe in God and the Last Day, who do not forbid what God and his Messenger have forbidden, who do not obey the rule of justice, until they pay the tax and agree to submit.
- "The tax" (*jizya*) was interpreted as a poll tax on non-Muslims (the People of the Book—Jews and Christians) later specified to be 48 dirhams for the rich, 24 for the middle incomes and 12 for the poorer.
- Compare this to Britain. Before the Reformation being Catholic was one identifier of the in-group, heretics were the persecuted out-group. By Elizabethan times and for 250 years until Catholic Emancipation in 1829 Catholics became the out-group. They paid higher taxes, their lands could be forfeit, and they could not hold public office.

Entangled versus Free-Standing Attributes

- We say that an attribute is **entangled** if its form or function is tightly linked to other attributes in the culture set, limiting how it can be combined with different combinations of other attributes.
- It is **free standing** if it is independent from other attributes and can be more easily combined with others or sidelined.
- Entangled attributes can be combined with others in more limited ways, thus reducing the range of cultural configurations that are feasible.

Entangled vs. Free-Standing Attributes in Islam and Britain

- Returning to the issue of social hierarchy, another key difference between the Islamic and British case is that in Islam, discrimination against non-Muslims was part of the Sharia.
- The Sharia is the law of God revealed by the Archangel Gabriel to Mohammed and as such cannot be changed by legislation (like the Catholic Emancipation Act).
- Thus the distinction between the in-group and out-group in Islamic culture is not just more specific, it is also **entangled** with other attributes, here religious beliefs.
- In Britain, the identities of in-group and out-group were **free standing** and thus relatively amenable to change.

Cultural Persistence

- How do we think about cultural persistence in our framework?
- Several new ideas arise from what we have presented already.
- First, many aspects of cultures will tend to persist precisely because they do not constrain economic, social or political behavior (for example cuisine or non-economic, non-political rituals).
- Second, more fluid cultures will tend to persist more because they can adapt to changing circumstances by generating new cultural configurations—but paradoxically, these are also the cultures that are changing more in the sense of generating distinct cultural configurations. But this is despite the fact that they are using the same attributes and hence appear to rely on the same "values".
- Third, cultures will persist and tend to be defining for the nature of social equilibria when they are hardwired and can withstand forces towards cultural collapse.

Endogenous Norms of Cooperation

- Different focus: how does “cooperation” (or “solution to collective action problem”) emerge, and why does “history” affect the outcome of such cooperation games? How and why do norms of corporation change?
  - Why does a history of distrust leads to distrust? How do we understand “social norms” and why do they persist?
  - Why does a society sometime break out of a history of distrust and change social norms?
  - Why does “collective action” differ across societies and why does it seem to change abruptly from time to time?
  - What is the role of leadership and “prominence”?
- Simple model based on Acemoglu and Jackson (2011).

Model

- Consider an overlapping-generations model where agents live for two periods. We suppose for simplicity that there is a single agent in each period (generation), and each agent's payoffs are determined by his interaction with agents from the two neighboring generations (older and younger agents).
- The action played by the agent born in period t is denoted $A_t \in \{H, L\}$, corresponding to "High" and "Low" actions (also can be interpreted as "honest" and "dishonest" actions).
- An agent chooses an action only once, in the first period of his or her life and that is played in both periods. This can be thought of as a proxy for a case where there is discretion, but also a high cost of changing behavior later in life.

Model

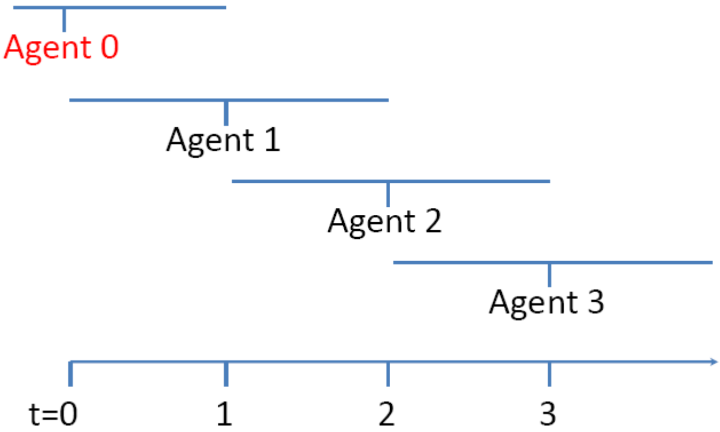
- The stage payoff to an agent playing A when another agent plays A' is denoted $u(A, A')$. The total payoff to the agent born at time t is

$$(1-\lambda)\, u(A_t, A_{t-1}) + \lambda u(A_t, A_{t+1}), \tag{15}$$

where A_{t-1} designates the action of the agent in the previous generation and A_{t+1} is the action of the agent in the next generation.

- Implicit assumption: choose single "pattern of behavior" A_t against both generations
  - $\lambda \in [0, 1]$ is a measure of how much an agent weighs the play with the next generation compared to the previous generation.
  - When $\lambda = 1$ an agent cares only about the next generation's behavior, while when $\lambda = 0$ an agent cares only about the previous generation's actions. We do not explicitly include a discount factor, since it is subsumed by λ.

Demographics



Model (continued)

- The stage payoff function $u(A, A')$ is given by the following matrix:

	H	L
H	β, β	$-\alpha, 0$
L	$0, -\alpha$	$0, 0$

  where β and α are both positive.
- This payoff matrix captures the notion that, from the static point of view, both honesty and dishonesty could arise as social norms—i.e., both (H, H) and (L, L) are static equilibria given this payoff matrix. (H, H) is clearly the Pareto optimal equilibrium, and depending on the values of β and α, it may be the risk dominant equilibrium as well.

Endogenous and Exogenous Agents

- There are four types of agents in this society.
- First, agents are distinguished by whether they choose an action to maximize the utility function given in (15). We refer to those who do so as "endogenous" agents.
- In addition to these endogenous agents, who will choose their behavior given their information and expectations, there will also be some committed or "exogenous" agents, who will choose an exogenously given action.
  - This might be due to some irrationality, or because some agents have a different utility function.

Endogenous and Exogenous Agents (continued)

- Any given agent is an "exogenous type" with probability 2π (independently of all past events), exogenously committed to playing each of the two actions, H and D, with probability $\pi \in (0, \frac{1}{2})$, and think of π as small.
- With the complementary probability, $1 - 2\pi > 0$, the agent is "endogenous" and chooses whether to play H or D, when young and is stuck with the same decision when old.
- Any given agent is also "prominent" with probability q (again independent). Information about prominent agents will be different. Thus:

	non-prominent	prominent
endogenous	$(1 - 2\pi)(1 - q)$	$(1 - 2\pi) q$
exogenous	$2\pi (1 - q)$	$2\pi q$

- Let us refer to endogenous non-prominent agents as *regular.*

Signals, Information and Prominent Agents

- A noisy signal of an action taken by a non-prominent agent of generation t is observed by the agent in generation $t+1$.
- No other agent receives any information about this action.
- In contrast, the actions taken by prominent agents are perfectly observed by all future generations.

Information Structure

- Let h^{t-1} denotes the public history at time t, which includes a list of past prominent agents and their actions up to and including time $t-1$. We denote the set of h^{t-1} histories by $\mathcal{H}^{t-1}$.
- We write $h_t = (T, a)$ if at time t the agent is both prominence type $T \in \{P, N\}$ and has taken action $a \in \{H, L\}$ if $T = P$ (if $T = N$, his action is not part of the public history).
- In addition to observing $h^{t-1} \in \mathcal{H}^{t-1}$, an agent of generation t, when born, receives a signal $s_t \in [0, 1]$ about the behavior of the agent of the previous generation (where the restriction to $[0, 1]$ is without loss of any generality). This signal has a continuous and distribution described by a density function $f_H(s)$ if $A_{t-1} = H$ and $f_L(s)$ if $A_{t-1} = L$.

Information Structure (continued)

- Without loss of generality, we order signals such that higher s has a higher likelihood ratio for H; i.e., so that
$$g(s) \equiv \frac{f_L(s)}{f_H(s)}$$
is nonincreasing in s.
- Suppose also that it is strictly decreasing, so that we have *strict Monotone Likelihood Ratio Property (MLRP)* everywhere.
- Suppose, without loss of any generality, that $s \in [0, 1]$, so that 0 is the worst signal for past H and 1 best signal for past H.
- Let $\Phi(x, s)$ denote the posterior probability that $A_{t-1} = H$ given $s_t = s$ under the belief that an endogenous agent of generation $t-1$ plays H with probability x. This is:
$$\Phi(x, s) \equiv \frac{f_H(s)\, x}{f_H(s)\, x + f_L(s)(1-x)} = \frac{1}{1 + g(s)\frac{1-x}{x}}. \tag{16}$$
The game begins with a prominent agent at time $t = 0$ playing

Strategies

- Let us use N to denote regular agents and P to denote prominent agents.
- With this notation, we can write the strategy of an endogenous agent of generation t (who may or may not be regular) as:

$$\sigma_t : \mathcal{H}^{t-1} \times [0, 1] \times \{P, N\} \rightarrow [0, 1],$$

written as $\sigma_t(h^{t-1}, s, T)$ where $h^{t-1} \in \mathcal{H}^{t-1}$ is the public history of play, $s \in [0, 1]$ is the signal of the previous generation's action, and $T \in \{P, N\}$ denotes whether or not the current agent is prominent.

- The number $\sigma_t(s, h^t, T)$ corresponds to the probability that the agent of generation t plays H.

We denote the strategy profile of all agents by the sequence

$$\sigma = \left(\sigma_1(h^0, s, T), \sigma_2(h^1, s, T), ..., \sigma_t(h^t, s, T), ...\right).$$

Semi-Markovian Strategies

- For the focus here, the most relevant equilibria involve agents ignoring histories that come before the last prominent agent (in particular, it will be apparent that these histories are not payoff-relevant provided others are following similar strategies).
- Let us refer to these as *semi-Markovian* strategies.
- Semi-Markovian strategies are specified for endogenous agents as functions $\sigma_{\tau}^{SM} : \{H, D\} \times [0, 1] \times \{P, N\} \longrightarrow [0, 1]$, written as $\sigma_{\tau}^{SM}(a, s, T)$ where $\tau \in \{1, 2, \ldots\}$ is the number of periods since the last prominent agent, $a \in \{H, D\}$ is the action of the last prominent agent, $s \in [0, 1]$ is the signal of the previous generation's action, and again $T \in \{P, N\}$ is whether or not the current agent is prominent.
- Let us denote a semi-Markovian by the sequence $\sigma^{SM} = \left(\sigma_1^{SM}(a, s, T), \sigma_2^{SM}(a, s, T), ..., \sigma_t^{SM}(a, s, T), ...\right)$.
- With some abuse of notation, write $\sigma_t = H$ or D to denote a strategy (or a semi-Markovian strategy) that corresponds to playing honest (dishonest) with probability one.

Equilibrium Definition

- Perfect Bayesian Equilibrium or Sequential Equilibrium.
- Only need to be careful when $q = 0$.
- Define *greatest* and *least* equilibria, and focus on greatest equilibria.

Towards Equilibrium Behavior

- Let ϕ_{t-1}^{t} be the the probability that the agent of generation t assigns to the agent from generation $t-1$ choosing $A=H$
- Let ϕ_{t+1}^{t} be the probability that the agent of generation t assigns, conditional on herself playing $A=H$, to the agent from generation $t+1$ choosing $A=H$.
- Payoff from L: 0
- Payoff from H:
  $(1-\lambda)\left[\phi_{t-1}^{t}\beta-(1-\phi_{t-1}^{t})\alpha\right]+\lambda\left[\phi_{t+1}^{t}\beta-(1-\phi_{t+1}^{t})\alpha\right].$
- Then an endogenous agent of generation t will prefer to play $A=H$ only if

$$(1-\lambda)\,\phi_{t-1}^{t}+\lambda\phi_{t+1}^{t}\geq\frac{\alpha}{\beta+\alpha}\equiv\gamma. \tag{17}$$

- Parameter γ a convenient way of summarizing relative payoffs (and also "basin of attraction" of L; so the greater is γ, the more attractive it is $A=L$).

Cutoff Strategies

- We say that a strategy σ is a *cutoff strategy* if for each t, h^{t-1} such that $h_{t-1} = N$ and $T_t \in \{P, N\}$, there exists s_t^* such that $\sigma_t(h^t, s, T_t) = 1$ if $s > s_t^*$ and $\sigma_t(h^t, s, T_t) = 0$ if $s < s_t^*$.
    - Clearly, setting $\sigma_t(h^t, s, T) = 1$ (or 0) for all s is a special case of a cutoff strategy.
- Cutoff strategy profile can be represented by the sequence of cutoffs

$$c = \left(c_1^N(h_0), c_1^P(h_0), \ldots c_t^N(h_{t-1}), c_t^P(h_{t-1}), \ldots \right).$$

- Given strict MLRP, all equilibria will be in cutoff strategies.
- Define greatest equilibria using the Euclidean distance on cutoffs.

Equilibrium Characterization

Proposition

1. *All equilibria are in cutoff strategies.*
2. *There exists an equilibrium in semi-Makovian cutoff strategies.*
3. *The set of equilibria and the set of semi-Markovian equilibria form complete lattices, and the greatest (and least) equilibria of the two lattices coincide.*

Understanding History-Driven Behavior

- Look for a unique equilibrium given by history:
  - When following prominent H, will all endogenous agents play H?
  - When following prominent L, will all endogenous agents play L?
- In such an equilibrium, social norms of High and Low emerge and persist, but not forever, since there might be switches because of exogenous prominent agents.
- Related question: when is this the greatest equilibrium?

Understanding History-Driven Behavior (continued)

- Recall that an endogenous agent of generation t will prefer to play $A = H$ only if

$$(1-\lambda)\,\phi^t_{t-1} + \lambda\phi^t_{t+1} \geq \frac{\alpha}{\beta+\alpha} \equiv \gamma. \tag{18}$$

- H is a *unique* best response for all if

$$\begin{aligned} (1-\lambda)\,\phi^t_{t-1} + \lambda\phi^t_{t+1} &\geq \gamma \\ \gamma^*_H \equiv (1-\lambda)\,\Phi\,(1-\pi, 0) + \lambda\pi \geq & \quad \gamma. \end{aligned}$$

- L is a *unique* best response for all if

$$\begin{aligned} (1-\lambda)\,\phi^t_{t-1} + \lambda\phi^t_{t+1} &< \gamma \\ \gamma^*_L \equiv (1-\lambda)\,\Phi\,(\pi, 1) + \lambda(1-\pi) < & \quad \gamma. \end{aligned}$$

Understanding History-Driven Behavior (continued)

Proposition

1. *If $\gamma < \gamma_H^*$, then following $a = H$ at date $t = 0$, the unique continuation equilibrium involves all (prominent and non-prominent) endogenous agents playing H.*
2. *If $\gamma > \gamma_L^*$, then following $a = L$ at date $t = 0$, the unique continuation equilibrium involves all endogenous agents playing L.*
3. *If $\gamma_L^* < \gamma < \gamma_H^*$, then there is a unique equilibrium driven by the starting condition: all endogenous agents take the same action as the action of the prominent agent at date $t = 0$.*

- Interpretation: persistent, but not everlasting, social norms.

Understanding History-Driven Behavior (continued)

- The condition that $\gamma_L^* < \gamma < \gamma_H^*$ boils down to

$$\lambda(1-2\pi) < (1-\lambda)\left[\Phi(1-\pi,0) - \Phi(\pi,1)\right]. \tag{19}$$

- It requires that λ be sufficiently small, so that sufficient weight is placed on the past. Without this, behavior would coordinate with future play, which naturally leads to a multiplicity.
- It also requires that signals are not too strong (so $\Phi(1-\pi,0) - \Phi(\pi,1) > 0$), as otherwise players would react to information about the most recent past generation and could change to High behavior if they had a strong enough signal regarding the past play and would also expect the next generation to have good information.

Understanding History-Driven Behavior (continued)

- Focusing on the greatest equilibrium:
- Let

$$\overline{\gamma}_H \equiv (1-\lambda)\,\Phi(1-\pi,0)+\lambda\,(1-\pi)\,. \tag{20}$$

- Thus relative to γ_H^*, more "optimistic" expectations about the future.

Proposition

The greatest equilibrium is such that:

(i) *following a prominent play of L, there is a low social norm and all endogenous agents play L if and only if* $\overline{\gamma}_L < \gamma$*; and*

(ii) *following a prominent play of H, there is a high social norm and all endogenous agents play H if and only if* $\gamma \leq \overline{\gamma}_H$*.*

Thus, endogenous players always follow the play of the most recent prominent player in the greatest equilibrium if and only if $\overline{\gamma}_L < \gamma \leq \overline{\gamma}_H$*.*

General Characterization of Greatest Equilibrium

- Let
$$\widehat{\gamma}_H \equiv (1-\lambda)\,\Phi(1-\pi, 1) + \lambda\,(1-\pi)\,.$$
- This is the expectation of $(1-\lambda)\phi^t_{t-1} + \lambda\phi^t_{t+1}$ for an agent who believes that any regular agent preceding him or her played H and sees the most optimistic signal, and believes that all subsequent endogenous agents will play H.
- Above, this threshold, no regular agent would ever play H.

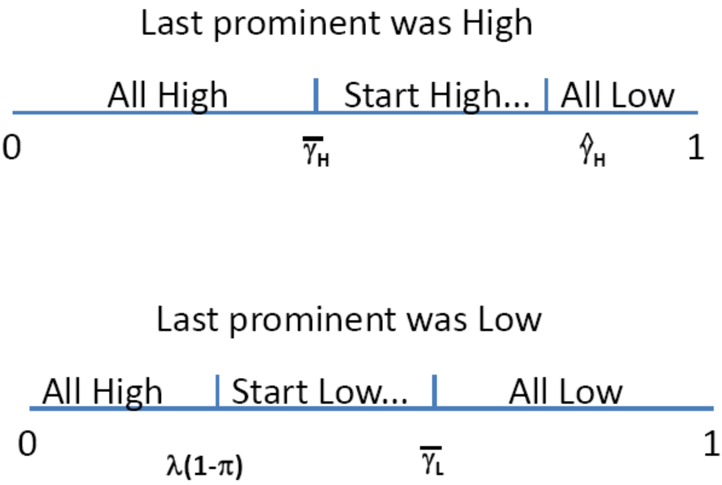
General Characterization (continued)

Proposition

In the greatest equilibrium:

1. *If $\gamma \leq \lambda(1-\pi)$, then all endogenous agents play H in all circumstances, and thus society has a stable high behavioral pattern.*
2. *If $\lambda(1-\pi) < \gamma \leq \overline{\gamma}_H$, then following a prominent play of H (but not following the prominent play of L) all endogenous agents play H.*
3. *If $\overline{\gamma}_L < \gamma \leq \overline{\gamma}_H$, then following a prominent play of L, all endogenous agents play L, and so all endogenous players follow the play of the most recent exogenous prominent player.*
4. *If $\overline{\gamma}_H < \gamma$, then endogenous agents play L for at least some signals, periods, and types even following a prominent play of H.*
5. *If $\widehat{\gamma}_H < \gamma$, then all endogenous agents who do not immediately follow a prominent H play L regardless of signals or types.*

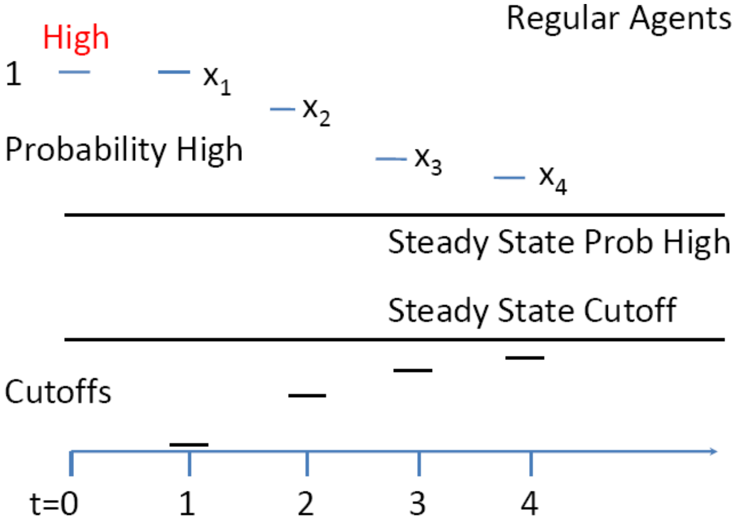
General Characterization of Greatest Equilibrium (continued)



Reversion of Play

- What happens when all High and all Low are not stable social norms?
- *Answer:* play reverts from an *extreme* (started by a prominent agent) to a steady-state distribution.
    - Start with H
    - Next player knows previous is H with probability 1
    - Next player knows previous endogenous played H, but this has probability $1-\pi$, so *action made depend on signal*
        - In fact, even stronger, because she knows that her signals will be interpreted is not necessarily coming from H.
    - Next player knows previous play was H with probability $< 1-\pi$.

Reversion of Play (continued)



Reversion of Play (continued)

- Let us denote the cutoffs used by prominent and non-prominent agents τ periods after the last prominent agent by c_τ^P and c_τ^N respectively.
- We say that high play *decreases* over time if $(c_\tau^P, c_\tau^N) \leq (c_{\tau+1}^P, c_{\tau+1}^N)$ for each τ.
- We say that high play *strictly decreases* over time, if in addition, we have that when $(c_\tau^P, c_\tau^N) \neq (0, 0)$, $(c_\tau^P, c_\tau^N) \neq (c_{\tau+1}^P, c_{\tau+1}^N)$.
- The concepts of high play increasing and strictly increasing are defined analogously.

Reversion of Play (continued)

Proposition

1. *In the greatest and least equilibria, cutoff sequences* $\left(c_{\tau}^{P}, c_{\tau}^{N}\right)$ *are monotone. Thus, following a prominent agent choosing* H, $\left(c_{\tau}^{P}, c_{\tau}^{N}\right)$ *are nondecreasing and following a prominent agent choosing* L, *they are non-increasing.*
2. *If* $\gamma > \overline{\gamma}_H$, *then in the greatest equilibrium, high play strictly decreases over time following high play by a prominent agent.*
3. *If* $\gamma < \overline{\gamma}_L$, *then in the greatest equilibrium, high play strictly increases over time following low play by a prominent agent.*

- But important asymmetry from switching from L to H vs from H to L
    - As we will see next, endogenous prominent agents would not like to the latter, but would prefer to do the former, so the first type of switches are driven by both exogenous and endogenous prominent agents, while the second only by exogenous prominent agents.

Reversion of Play (continued)

- The following is an immediate corollary of Proposition 10.

Corollary

As the distance from the last prominent agent grows ($\tau \to \infty$), cutoffs in the greatest equilibrium converge and the corresponding distributions of play converge to a stationary distribution. Following a choice of H by the last prominent agent, this limiting distribution involves only H by all endogenous agents if and only if $\gamma \leq \overline{\gamma}_H$. Similarly, following a choice of L by the last prominent agent, this limiting distribution involves L by all endogenous agents if and only if $\gamma \geq \overline{\gamma}_L$.

Leadership: Breaking the Low Social Norm

- Can promise breaker social norm of L?
  - Regular agents may be stuck in L for reasons analyzed so far.
  - But prominence, greater visibility in the future, can enable “leadership”
- Idea:
  - endogenous prominent agents can always break the social norm
  - when “all L” is not the unique equilibrium after prominent L, endogenous prominent agents will like to break the social norm of L and start a switch to H

Leadership: Breaking the Low Social Norm (continued)

Proposition

1. *Suppose that the last prominent agent played L and*

$$\widetilde{\gamma}_L \leq \gamma < \tilde{\gamma}_H \equiv (1-\lambda)\Phi(\pi, 0) + \lambda(1-\pi). \quad (21)$$

   Then there exists a fixed cutoff below 1 (after at least one period) such that an endogenous prominent agent chooses High and breaks the Low social norm if the signal is above the cutoff.

2. *Suppose that $\gamma < \widetilde{\gamma}_L$ and $\gamma < \tilde{\gamma}_H$. Suppose that the last prominent agent has played L. Endogenous prominent agents have cutoffs below 1 that decrease with time such that if the signal is above the cutoff then in a greatest equilibrium the endogenous prominent agent will choose H and break a low social norm.*

3. *Moreover, in either case if $\gamma < \gamma_H^*$, the above are the unique continuation equilibrium.*

Role of Prominence

- Prominence provides greater visibility and thus coordinates future actions.
- Crucially: common knowledge of visibility.
    - Without this, prominence is less effective.

Role of Prominence (continued)

- Suppose there is a starting non-prominent agent at time 0 who plays *High* with probability $x_0 \in (0, 1)$, where x_0 is known to all agents who follow, and generates a signal for the first agent in the usual way.
- All agents after time 1 are not prominent.
- In every case all agents (including time 1 agents) are endogenous with probability $(1 - 2\pi)$.

Scenario 1: The agent at time 1 is not prominent and his or her action is observed with the usual signal structure.
Scenario 2: The agent at time 1's action is observed perfectly by the period 2 agent, but not by future agents.
Scenario 2′: The agent at time 1 is only observed by the next agent according to a signal, but then is subsequently perfectly observed by all agents who follow from time 3 onwards.
Scenario 3: The agent at time 1 is prominent, and all later agents are viewed with the usual signal structure.

Role of Prominence (continued)

- Clearly, as we move from Scenario 1 to Scenario 2 (or $2'$) to Scenario 3, we are moving from a non-prominent agent to a prominent one
- Let us focus again on the greatest equilibrium and let $c^k(\lambda, \gamma, f_H, f_L, q, \pi)$ denote the cutoff signal above which the first agent (if endogenous) plays *High* under scenario k as a function of the underlying setting.

Proposition

The cutoffs satisfy $c^2(\cdot) \geq c^3(\cdot)$ *and* $c^1(\cdot) \geq c^{2'}(\cdot) \geq c^3(\cdot)$, *and there are settings* $(\lambda, \gamma, f_H, f_L, q, \pi)$ *for which all of the inequalities are strict.*

Multiple Agents

- Now suppose n agents within each generation, and random matching; unless there is a prominent agent, in which case all those from previous and next generations match with the prominent agent.
- If no prominent agent, then observe a signal generated by the action of a randomly generated agent from the previous generation.
- Results generalize, except but now we can do comparative statics with respect to n.

Multiple Agents (continued)

Proposition

In the model with n agents within each generation, there exist greatest and least equilibria. In the greatest equilibrium:

1. *following a prominent play of Low, there is a Low social norm and all endogenous agents play Low (i.e., $\overline{\sigma}_{\tau}^{SM}(a = Low, s, T) = Low$ for all s, T and all $\tau > 0$) if and only if $\overline{\gamma}_{L}^{n} < \gamma$; and*
2. *following a prominent play of High, there is a High social norm and all endogenous agents play High (i.e., $\overline{\sigma}_{\tau}^{SM}(a = High, s, T) = High$ for all s, T and all $\tau > 0$) if and only if $\gamma \leq \overline{\gamma}_{H}^{n}$.*

The threshold $\overline{\gamma}_{H}^{n}$ is increasing in n and and the threshold $\overline{\gamma}_{L}^{n}$ is decreasing in n, so that both High and Low social norms following, respectively, High and Low prominent play, emerge for a larger set of parameter values.

- Intuition: signals less informative, thus history matters more.

Laws and Norms

- How norms influence institutions? How are they influenced by institutions?
- A specific context is the interplay between laws and norms:
  - norms may make laws less effective (e.g., nobody obeys them because they conflict with the social norms);
  - laws may change social norms.
- A simple model to study these issues: Acemoglu and Jackson (2015).

Static Setup

- Finite population of agents, $N = \{1, \ldots, n\}$, with $n \geq 2$ taken to be even, and much pairwise match pairwise.
- Agent i has type $\theta_i \in [0, 1]$, distributed according to a cumulative distribution function F.
- Agent i chooses a *base behavior* $b_i \in [0, 1]$
- The agent's *actual behavior* may be forced away from this level ex post to some B_i because of the enforcement of a law.

Laws

- A law, $L \in [0, 1]$, is an *upper bound* on the behaviors of agents, meaning that any behavior above L is prohibited.
- The government detects law-breaking in any pair with probability $\eta \in [0, 1)$, but can also find out about law-breaking because of whistle-blowing within the partnership.

- If a law-breaker is caught, her behavior is brought down to $B_i = L$ and she pays a fine ϕ.
- In particular, an agent i can whistle-blower on her partner $m(i)$ if he is breaking the law, i.e., $b_{m(i)} > L$, and she is herself law-abiding, i.e., $b_i \leq L$.
- This last requirement captures the fact that if a law-breaker whistle-blows, then she may get caught herself and be subject to a fine.

Actual Behaviors and Payoffs

- The actual behavior of agent i is

$$B_i\left(w_{m(i)}, b_i\right) = \begin{cases} L & \text{if } b_i > L, \text{ and } w_{m(i)} = 1 \text{ or if there is public enf} \\ b_i & \text{otherwise.} \end{cases}$$

- Agent i's payoff is given by

$$u_i(\theta_i, B_i) = -a\left(B_i - \theta_i\right)^2 - (1-a)\left(B_i - B_{m(i)}\right)^2 - \zeta_m B_{m(i)} - \zeta_o \sum_{j \neq i, m(i)} B_j - \left(\eta + (1-\eta)\, w_{m(i)} \mathbf{I}_{\{b_i > L\}}\right)\phi. \tag{22}$$

- The parameters $\zeta_m, \zeta_o \geq 0$ capture negative externalities from the behaviors of others.

Equilibrium

- It is a strictly dominant strategy for an agent with $b_i \leq L$ and $b_{m(i)} > L$ to whistle-blow (because this reduces both the externality and the mismatch with the partner's behavior).
- Thus just focus on first-stage choice of base behavior and define an equilibrium as a pure-strategy symmetric Bayesian equilibrium described by $\beta : [0,1] \longrightarrow [0,1]$, with $\beta(\theta_i)$ indicating the action taken by type θ_i.

Proposition

An equilibrium exists, and all equilibria are in monotone strategies and are characterized by a threshold θ^* above which all types break the law and below which they obey the law.

Characterization of Equilibrium

Proposition

All equilibria are of the following form. There exists $\theta^* \in [L, 1]$ *such that*

$$\beta(\theta_i) = \beta_{abiding}(\theta_i) \equiv \min[a\theta_i + (1-a)x, L] \text{ if } \theta_i < \theta^* \tag{23}$$

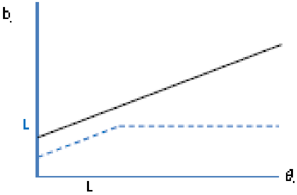
and

$$\beta(\theta_i) = \beta_{breaking}(\theta_i) \equiv a\theta_i + (1-a)\mathbb{E}[\theta|\theta > \theta^*] \text{ if } \theta_i > \theta^*, \tag{24}$$

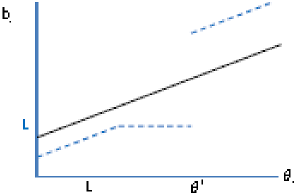
where x *is the unique solution to* $x = \mathbb{E}[\min \beta(\theta), L]$.

Representation of Equilibrium

- Full compliance and partial compliance equilibria:



(a) Law fully obeyed



(b) Law partially obeyed

Intuition

- Without a law, an agent chose a convex combination of her preferred behavior given by her type, θ_i, and expected behavior in society.
- For law-abiding agents, the calculus is still similar, except that as shown by the expression, $a\theta_i + (1 - a)x$, the expected behavior is replaced by x, because she can whistle-blow on a law-breaking partner and bring his behavior down to L.
- On the other hand, a law-breaker knows that his partner, if she is law-abiding, will whistle-blow on him, setting their behavior down to L. Thus, when choosing their behavior, he will need to reason *conditionally* — conditional on matching with another law-breaking agent and not being subject to public enforcement. This is the reason why the term $\mathbb{E}[\theta|\theta > \theta^*]$ appears.

Multiple Equilibria

- In general there can be multiple equilibria. In particular

Proposition

For any $L \in (0,1)$, there exists $\overline{\phi} \geq 0$, such that

- *if $\phi > \overline{\phi}$, then there is a unique equilibrium, which involves full compliance (all types obey the law);*
- *if $\phi < \overline{\phi}$, then there are multiple equilibria: one with full compliance and (at least two) other equilibria ordered by the threshold above which all types break the law.*

- Intuitively, if others are expected to break the law, they cannot whistle-blow and breaking the law becomes more attractive.
- This captures the fact that the effectiveness of laws is interwoven with social norms, here corresponding to expected behavior of others.

Comparative Statics

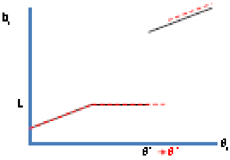
Proposition

Consider the lowest compliance equilibrium. Then:

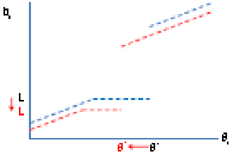
1. *A small increase in* ϕ *(higher fine),* ζ_m *(greater within-match externality), and/or* η *(higher likelihood of public enforcement): increases* θ^* *and so lowers the fraction of agents breaking the law; leaves behavior by each agent who was obeying the law unchanged; but leads to higher behavior among those still breaking the law.*
2. *There exists* $\bar{\zeta}_m > 0$ *such that if* $\zeta_m < \bar{\zeta}_m$*, a small decrease in* L *(a stricter law): decreases* θ^**, increasing the fraction of agents breaking the law; leads to lower behavior by each agent who was already breaking the law; and leads to lower average behavior by those obeying the law.*

- Note the non-monotonicity in behavior.

Comparative Statics in a Diagram



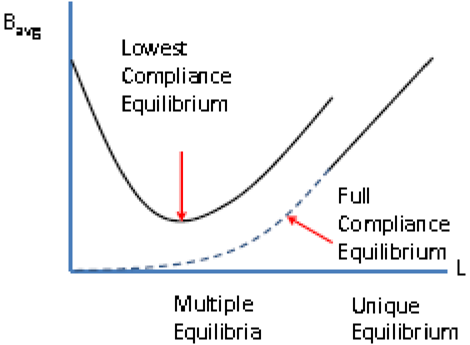
(a) An increase in fine, ϕ; externality, $\hat{\xi}_{m}$; or public enforcement, η.



(b) A decrease in the (tightening of the) law L

- Nonmonotone behavior the population — greater ϕ less law-breaking, but greater behavior among law-breakers.

Nonmonotonicity in Average Behavior



- Nonmonotonicity in response to tighter laws both because of multiple equilibria and competing effects.

Dynamic Model

- Now consider this same set up in the context of an overlapping generations model, where each agent plays with a random partner from the previous and next generations (with weights $1 - \lambda$ and λ on payoffs from previous and next generations).
- The rest of the setup is unchanged.

Dynamic Equilibria

- Dynamic equilibria are similar (a little more involved).
- But steady-state equilibria are identical. In particular:

Proposition

Let $\mathcal{B}$ be the set equilibria of the static game, and $\mathcal{B}^*$ be the set of steady-state behaviors from the equilibria of the dynamic game (with the same parameter values as the static game and with $L_t = L$ and $\phi_t = \phi$ for all t). Then $\mathcal{B} = \mathcal{B}^*$, and every steady-state equilibrium is described by a strategy of the form

$$\beta^*(\theta_i) = \begin{cases} \beta_{abiding}(\theta_i) & \text{if } \theta_i < \theta^* \\ \beta_{breaking}(\theta_i) & \text{if } \theta_i > \theta^* \end{cases}$$

for some threshold θ^*, where $\beta_{abiding}(\theta_i)$ and $\beta_{breaking}(\theta_i)$ are as defined in (23) and (24) in the static proposition, and then i whistle-blows occurs if and only if $\theta_i < \theta^*$ and a match breaks the law.

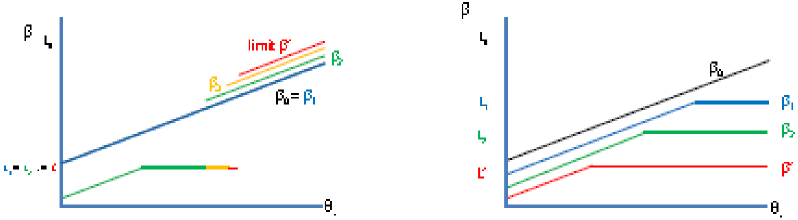
Social Norms and Effectiveness of Laws

Proposition

Fix F, $\phi > 0$ and $\zeta \geq 0$, let $\lambda = 0$, take η and a small, and start with initial non-binding law (i.e., $L \geq (1-a)\mathbb{E}[\theta] + a$). Suppose that society starts at time $t = 0$ in the unique steady-state equilibrium corresponding to L, and consider new law $L' < (1-a)\mathbb{E}[\theta] - a$ (so that there is a less than full compliance steady-state equilibrium under L').

- *(Abrupt tightening of law) If there is an unanticipated and permanent change to L' in period 1, then all agents break the law in period 1, and behavior converges to the lowest compliance equilibrium associated with L'.*
- *(Gradual tightening of law) However, for any such L' there exists a (finite) decreasing sequence of laws $\{L_t\}_{t=1}^{T}$ with $L_T = L'$ such that following a switch to this sequence of laws, all agents comply with the law at their birth and play converges to the full compliance steady-state equilibrium associated with L'.*

Gradual vs. Abrupt Tightening of Laws



(a) An abrupt tightening of the law leading to a switch from full compliance to the lowest compliance equilibrium

(b) A gradual tightening of the law while maintaining the full compliance equilibrium

Interpretation

- A significant tightening of the law creates a conflict between the prevailing norms and the law, leading to an immediate and significant increase in law-breaking.
- This then makes it impossible for society to achieve the full compliance steady-state equilibrium.
- In contrast, a series of gradual laws converging to L' can be much more effective in containing law-breaking and can achieve full compliance.
- In particular, each gradual tightening of the law will have a small impact on behavior, and the next generation will be willing to abide by the law, because this enables both coordination with the current generation and avoids the costs imposed by public law enforcement.
- This gradual sequence of tighter laws slowly changes the prevailing norms, and as norms change, these tighter laws become more and more powerful in restricting behavior.

Multiple Behaviors

- Now suppose that there are two unrelated types of behaviors, $\left(b_i^1, b_i^2\right)$.
- Suppose that types are also two-dimensional, (θ_i^1, θ_i^2), and are independently drawn.
- Thus the only interaction between the two behaviors is that law-breaking on one dimension precludes whistle-blowing.

Equilibria with Multiple Behaviors

Proposition

Consider the model with multiple laws described above. Suppose that there is a non-trivial law in the first dimension ($L^1 < 1$) but no law in the second dimension ($L^2 = 1$), and an associated equilibrium, $\left(\beta^1(\theta), \beta^2(\theta)\right)$ with a law-breaking threshold $\theta^{1*} < 1$ on the first dimension. Then:

- *There exist $\bar{\delta}$ and $\underline{\delta}$ such that if a law $\tilde{L}^2 \in (L^1 - \underline{\delta}, L^1 + \bar{\delta})$ is imposed on the second dimension, then there is a new equilibrium $\left(\tilde{\beta}^1(\theta), \tilde{\beta}^2(\theta)\right)$ that involves a law-breaking threshold $\tilde{\theta}^{1*} > \theta^{1*}$ (i.e., there is less law-breaking on the first dimension).*
- *There exists $\underline{L} > 0$ such that if a law $\tilde{L}^2 < \underline{L}$ is imposed on the second dimension, then the new equilibrium $\left(\tilde{\beta}^1(\theta), \tilde{\beta}^2(\theta)\right)$ involves a law-breaking threshold $\tilde{\theta}^{1*} < \theta^{1*}$ (i.e., there will be more law-breaking on the first dimension).*

Interpretation

- Bad laws in some dimensions create negative spillovers on law-abiding on all dimensions.
- Implications for the "broken windows theory" — the problem may not be one of enforcement but one of bad laws.
- But also at the same time good laws increased law-abiding behavior in other dimensions because individuals now have an incentive to provide with all laws to benefit from the implementation of (reasonable) laws.