Situations and Norms

Formal Foundations of Sociology

Tore Ellingsen

Erik Mohlin

October 2015

A Research Program

A Research Program

Elements from these two intellectual traditions [sociology and economics] cannot be brought together in a pastiche. It is necessary to begin with a conceptually coherent framework from one and introduce elements from the other without destroying that coherence.

A Research Program

Elements from these two intellectual traditions [sociology and economics] cannot be brought together in a pastiche. It is necessary to begin with a conceptually coherent framework from one and introduce elements from the other without destroying that coherence.

James Coleman, 1988, p. S97

Ultimate Goal: A Unified Framework for Social Analysis

Ultimate Goal: A Unified Framework for Social Analysis

- Propose a formal framework to integrate sociology and economics

Ultimate Goal: A Unified Framework for Social Analysis

- Propose a formal framework to integrate sociology and economics
  - comprising natural extensions of familiar game theoretic analysis

Ultimate Goal: A Unified Framework for Social Analysis

- Propose a formal framework to integrate sociology and economics
  - comprising natural extensions of familiar game theoretic analysis
  - yielding a theory of social situations – and thereby distinguishing natural (immutable) and social (partly malleable) aspects of reality.

Ultimate Goal: A Unified Framework for Social Analysis

- Propose a formal framework to integrate sociology and economics
  - comprising natural extensions of familiar game theoretic analysis
  - yielding a theory of social situations – and thereby distinguishing natural (immutable) and social (partly malleable) aspects of reality.
- Use this framework to rationalize some otherwise puzzling findings:

Ultimate Goal: A Unified Framework for Social Analysis

- Propose a formal framework to integrate sociology and economics
  - comprising natural extensions of familiar game theoretic analysis
  - yielding a theory of social situations – and thereby distinguishing natural (immutable) and social (partly malleable) aspects of reality.
- Use this framework to rationalize some otherwise puzzling findings:
  1. Sorting into and out of situations (Dana, Cain, Dawes, 2006 etc).

Ultimate Goal: A Unified Framework for Social Analysis

- Propose a formal framework to integrate sociology and economics
  - comprising natural extensions of familiar game theoretic analysis
  - yielding a theory of social situations – and thereby distinguishing natural (immutable) and social (partly malleable) aspects of reality.
- Use this framework to rationalize some otherwise puzzling findings:
  1. Sorting into and out of situations (Dana, Cain, Dawes, 2006 etc).
  2. Avoidance of situational information (Dana, Weber, Kuang, 2007, etc).

Ultimate Goal: A Unified Framework for Social Analysis

- Propose a formal framework to integrate sociology and economics
  - comprising natural extensions of familiar game theoretic analysis
  - yielding a theory of social situations – and thereby distinguishing natural (immutable) and social (partly malleable) aspects of reality.
- Use this framework to rationalize some otherwise puzzling findings:
  1. Sorting into and out of situations (Dana, Cain, Dawes, 2006 etc).
  2. Avoidance of situational information (Dana, Weber, Kuang, 2007, etc).
  3. Detrimental effects of material incentives (Gneezy, Rustichini, 2000 etc).

Ultimate Goal: A Unified Framework for Social Analysis

- Propose a formal framework to integrate sociology and economics
  - comprising natural extensions of familiar game theoretic analysis
  - yielding a theory of social situations – and thereby distinguishing natural (immutable) and social (partly malleable) aspects of reality.
- Use this framework to rationalize some otherwise puzzling findings:
  1. Sorting into and out of situations (Dana, Cain, Dawes, 2006 etc).
  2. Avoidance of situational information (Dana, Weber, Kuang, 2007, etc).
  3. Detrimental effects of material incentives (Gneezy, Rustichini, 2000 etc).
  4. Blame shifting (Bartling, Fischbacher, 2012, etc)

Ultimate Goal: A Unified Framework for Social Analysis

- Propose a formal framework to integrate sociology and economics
  - comprising natural extensions of familiar game theoretic analysis
  - yielding a theory of social situations – and thereby distinguishing natural (immutable) and social (partly malleable) aspects of reality.
- Use this framework to rationalize some otherwise puzzling findings:
  1. Sorting into and out of situations (Dana, Cain, Dawes, 2006 etc).
  2. Avoidance of situational information (Dana, Weber, Kuang, 2007, etc).
  3. Detrimental effects of material incentives (Gneezy, Rustichini, 2000 etc).
  4. Blame shifting (Bartling, Fischbacher, 2012, etc)
- Key: Behavior is affected by group-level understandings and values.

Example 1a: Unilateral Helping Situation

Example 1a: Unilateral Helping Situation

Suppose Rowena and Colin live in a society that values efficiency and equality. Specifically, both a high sum and an even distribution of material payoffs are desirable.

Example 1a: Unilateral Helping Situation

Suppose Rowena and Colin live in a society that values efficiency and equality. Specifically, both a high sum and an even distribution of material payoffs are desirable.

Rowena can help Colin (play C) or not (play D). Material payoffs are:

C	$-1, 2$
D	$0, 0$

Example 1a: Unilateral Helping Situation

Suppose Rowena and Colin live in a society that values efficiency and equality. Specifically, both a high sum and an even distribution of material payoffs are desirable.

Rowena can help Colin (play C) or not (play D). Material payoffs are:

C	$-1, 2$
D	$0, 0$

Q: Should Rowena help? What would the norm be?

Example 1a: Unilateral Helping Situation

Suppose Rowena and Colin live in a society that values efficiency and equality. Specifically, both a high sum and an even distribution of material payoffs are desirable.

Rowena can help Colin (play C) or not (play D). Material payoffs are:

C	$-1, 2$
D	$0, 0$

Q: Should Rowena help? What would the norm be?

A: It depends on the relative weights attached to efficiency and equality.

Example 1b: Bilateral Helping Situation

Example 1b: Bilateral Helping Situation

Suppose now that Colin can also help Rowena:

C	$-1, 2$
D	$0, 0$

Rowena's choice

C	D
$2, -1$	$0, 0$

Colin's choice

Example 1b: Bilateral Helping Situation

Suppose now that Colin can also help Rowena:

C	$-1, 2$
D	$0, 0$

Rowena's choice

C	D
$2, -1$	$0, 0$

Colin's choice

	C	D
C	$1, 1$	$0, 2$
D	$2, 0$	$0, 0$

Merged situation

Example 1b: Bilateral Helping Situation

Suppose now that Colin can also help Rowena:

C	$-1, 2$
D	$0, 0$

Rowena's choice

C	D
$2, -1$	$0, 0$

Colin's choice

	C	D
C	$1, 1$	$0, 2$
D	$2, 0$	$0, 0$

Merged situation

Example 1b: Bilateral Helping Situation

Suppose now that Colin can also help Rowena:

C	$-1, 2$
D	$0, 0$

Rowena's choice

C	D
$2, -1$	$0, 0$

Colin's choice

	C	D
C	$1, 1$	$0, 2$
D	$2, 0$	$0, 0$

Merged situation

Q: What would the norm be?

Example 1b: Bilateral Helping Situation

Suppose now that Colin can also help Rowena:

C	$-1, 2$
D	$0, 0$

Rowena's choice

C	D
$2, -1$	$0, 0$

Colin's choice

	C	D
C	$1, 1$	$0, 2$
D	$2, 0$	$0, 0$

Merged situation

Q: What would the norm be? A: (C,C). Efficiency and equality now aligned.

Example 1b: Bilateral Helping Situation

Suppose now that Colin can also help Rowena:

C	$-1, 2$
D	$0, 0$

Rowena's choice

C	D
$2, -1$	$0, 0$

Colin's choice

	C	D
C	$1, 1$	$0, 2$
D	$2, 0$	$0, 0$

Merged situation

Q: What would the norm be? A: (C,C). Efficiency and equality now aligned.

- Crucial whether the two choices – Rowena's helping and Colin's helping – are seen as separate situations or part of the same situation.

Example 1b: Bilateral Helping Situation

Suppose now that Colin can also help Rowena:

C	$-1, 2$
D	$0, 0$

Rowena's choice

C	D
$2, -1$	$0, 0$

Colin's choice

	C	D
C	$1, 1$	$0, 2$
D	$2, 0$	$0, 0$

Merged situation

Q: What would the norm be? A: (C,C). Efficiency and equality now aligned.

- Crucial whether the two choices – Rowena's helping and Colin's helping – are seen as separate situations or part of the same situation.
- The "norm of reciprocity" *derived* from more fundamental values.

Example 1c: What is a Social Situation?

Example 1c: What is a Social Situation?

Example 1c: What is a Social Situation?

Rowena can help Colin (play C) or not (play D). Material payoffs are:

C	$-1, 2$
D	$0, 0$

Example 1c: What is a Social Situation?

Rowena can help Colin (play C) or not (play D). Material payoffs are:

C	$-1, 2$
D	$0, 0$

- Suppose only Rowena is aware that action C is available? Unless she takes action C, nobody else will ever know she even had the choice to do so. – To what degree is this a social situation?

Example 1c: What is a Social Situation?

Rowena can help Colin (play C) or not (play D). Material payoffs are:

C	$-1, 2$
D	$0, 0$

- Suppose only Rowena is aware that action C is available? Unless she takes action C, nobody else will ever know she even had the choice to do so. – To what degree is this a social situation?
- Suppose Rowena has a prior choice that either enables C or not. (Walk past the beggar or not.) If it is enabled others will know that C is available, otherwise not. All others are unaware of the enabling choice. – To what extent is the enabling choice a social situation?

Example 1c: What is a Social Situation?

Rowena can help Colin (play C) or not (play D). Material payoffs are:

C	$-1, 2$
D	$0, 0$

- Suppose only Rowena is aware that action C is available? Unless she takes action C, nobody else will ever know she even had the choice to do so. – To what degree is this a social situation?
- Suppose Rowena has a prior choice that either enables C or not. (Walk past the beggar or not.) If it is enabled others will know that C is available, otherwise not. All others are unaware of the enabling choice. – To what extent is the enabling choice a social situation?
- There is an unmanned fruit-stand by the road. Fruit, a cash-box, a sign stating the price. You are alone. To what extent is this a social situation?

Precursors

Precursors

- Related formal work on laws (endogenous social cognition)
  - Mailath, Morris, Postlewaite (2003wp)
  - Basu (2015wp)
- Related formal work on awareness and endogenous bracketing
  - Heifetz, Meier, Schipper (2006)
  - Myerson and Weibull (2015)
- Related formal work on norms
  - Lopez-Perez (2008)
  - Krupka and Weber (2013)
- But primarily lots of informal theorizing, case studies, and experiments
- Caveat: Only complete information environments

Model: I. Players, Actions, Outcomes

Model: I. Players, Actions, Outcomes

Physical Reality (Simple, Static Version)

$N = \{1, \cdots, n\}$	Set of players
$a_i \in A_i$	An action by player i
$a = (a_1, \ldots, a_n) \in A$	An action profile
$x_i(a) \in X_i$	A feasible outcome for i
$x(a) \in X$	A feasible outcome

Model: I. Players, Actions, Outcomes

Physical Reality (Simple, Static Version)

$N = \{1, \cdots, n\}$	Set of players
$a_i \in A_i$	An action by player i
$a = (a_1, \ldots, a_n) \in A$	An action profile
$x_i(a) \in X_i$	A feasible outcome for i
$x(a) \in X$	A feasible outcome

Definition

Reality is separable if, for all players $i \in N$, $x_i(a_i, a_{-i}) - x_i(a_i', a_{-i})$ *is independent of* a_{-i} *for all* $a_i, a_i' \in A_i$ *and all* $a_{-i} \in A_{-i}$.

Model: II. Strategies and Situations

Considered Reality

$s_i \in S_i \subseteq A_i$	Player i's strategy
$S_i^{max} = A_i$	Player i's maximal strategy set
$S^{max} = \times_{j \in N} S_j$	Set of feasible strategy profiles
$G^{max} = \langle N, S, x \rangle$	"Maximal Social Reality" – unbounded cognition
$G = \langle N_G, S_G, x \rangle$	A situation
$\mathcal{G}$	Set of possible situations

Considered Reality

$s_i \in S_i \subseteq A_i$	Player i's strategy
$S_i^{max} = A_i$	Player i's maximal strategy set
$S^{max} = \times_{j \in N} S_j$	Set of feasible strategy profiles
$G^{max} = \langle N, S, x \rangle$	"Maximal Social Reality" – unbounded cognition
$G = \langle N_G, S_G, x \rangle$	A situation
$\mathcal{G}$	Set of possible situations

- A situation could be simple or complex, depending on the sizes of $N_G \subseteq N$ and $S_G \subseteq S$.
- A player could be involved in many situations simultaneously, as one action could be relevant to several situations.
- A situation G is a *game form*.

Model III: Cognition and Beliefs

Individual and Social Realities

$G_i = \langle N_i, S_i, x \rangle$	Player i's understanding of a situation
$B_{ij} \in \mathcal{G}$	Player i's belief about j's understanding
$B_{iji} \in \mathcal{G}$	Player i's belief about j's und. of i's und.

Individual and Social Realities

$G_i = \langle N_i, S_i, x \rangle$	Player i's understanding of a situation
$B_{ij} \in \mathcal{G}$	Player i's belief about j's understanding
$B_{iji} \in \mathcal{G}$	Player i's belief about j's und. of i's und.

We say that i and j have *shared understanding* if

$$B_{iji} = B_{ji} = G_i = B_{jij} = B_{ij} = G_j.$$

Shared understanding is a group-level concept – an element of culture.

Model III: Cognition and Beliefs

Individual and Social Realities

$G_i = \langle N_i, S_i, x \rangle$	Player *i*'s understanding of a situation
$B_{ij} \in \mathcal{G}$	Player *i*'s belief about *j*'s understanding
$B_{iji} \in \mathcal{G}$	Player *i*'s belief about *j*'s und. of *i*'s und.

We say that i and j have *shared understanding* if

$$B_{iji} = B_{ji} = G_i = B_{jij} = B_{ij} = G_j.$$

Shared understanding is a group-level concept – an element of culture.

(Players are said to have common beliefs if we can iterate this criterion indefinitely.

Model III: Cognition and Beliefs

Individual and Social Realities

$G_i = \langle N_i, S_i, x \rangle$	Player i's understanding of a situation
$B_{ij} \in \mathcal{G}$	Player i's belief about j's understanding
$B_{iji} \in \mathcal{G}$	Player i's belief about j's und. of i's und.

We say that i and j have *shared understanding* if

$$B_{iji} = B_{ji} = G_i = B_{jij} = B_{ij} = G_j.$$

Shared understanding is a group-level concept – an element of culture.

(Players are said to have common beliefs if we can iterate this criterion indefinitely. That's the standard assumption in game theory.)

Model IV: Social Outcome Evaluations

A social outcome evaluation orders the set of outcome profiles, providing the value w_G that a society attaches to various outcomes x_G in game G (or, preferably, some large class of situations to which G belongs).

$$w_G : X_G \to \mathbb{R}.$$

Model IV: Social Outcome Evaluations

A social outcome evaluation orders the set of outcome profiles, providing the value w_G that a society attaches to various outcomes x_G in game G (or, preferably, some large class of situations to which G belongs).

$$w_G : X_G \to \mathbb{R}.$$

For example, x might be a profile of monetary payoffs and w might be a weighted summation.

Model V: Social Action Evaluations

Model V: Social Action Evaluations

A social action evaluation orders the set of action profiles, detailing the value v_G that a society attaches to various action profiles in a situation G (or some class of situations to which G belongs).

$$v_G : S_G \to \mathbb{R}.$$

Model V: Social Action Evaluations

A social action evaluation orders the set of action profiles, detailing the value v_G that a society attaches to various action profiles in a situation G (or some class of situations to which G belongs).

$$v_G : S_G \to \mathbb{R}.$$

Sometimes, the evaluation of actions are derived exclusively from the outcomes that they produce. In this case,

$$v_G(a) = w_G(x(a)).$$

We will focus on this case here.

Model VI: Rules – Laws, norms etc

Rules

A rule is a statement about action profiles in a situation. Many rules are binary. For example, binary injunctive rules say whether actions are recommended (permitted) or not (forbidden):

$$r_G : S_G \to \{p, f\} .$$

whereas binary descriptive rules say whether actions are common or not.

Model VI: Rules – Laws, norms etc

Rules

A rule is a statement about action profiles in a situation. Many rules are binary. For example, binary injunctive rules say whether actions are recommended (permitted) or not (forbidden):

$$r_G : S_G \to \{p, f\} .$$

whereas binary descriptive rules say whether actions are common or not.

We will here focus on injunctive rules that recommend

$$s^* \in \arg\max w_G(x(s)).$$

Model VII: Preferences and Games

Model VII: Preferences and Games

Model VII: Preferences and Games

Players have vNM utility functions, but we allow preferences to depend not only on the outcome profile x, but also on:

- Rule compliance – both with injunctive and descriptive rules.

Model VII: Preferences and Games

Players have vNM utility functions, but we allow preferences to depend not only on the outcome profile x, but also on:

- Rule compliance – both with injunctive and descriptive rules.

That is, player i chooses s_i to maximize

$$u_i(s) = \mathbb{E}\, U_i(x(s), r_{G_i}(s)).$$

Let $u = u_1, \ldots, u_n$.
If shared understanding of G, we obtain the *game* $\Gamma = \langle G, u \rangle$.

Varieties of Preferences

Varieties of Preferences

1. **Selfish preferences** – only own outcome matters:

$$U_i(x_i)$$

Varieties of Preferences

1. **Selfish preferences** – only own outcome matters:

$$U_i(x_i)$$

2. **Other-regarding preferences** – everyone's outcome matters:

$$U_i(x)$$

Varieties of Preferences

1. **Selfish preferences** – only own outcome matters:

$$U_i(x_i)$$

2. **Other-regarding preferences** – everyone's outcome matters:

$$U_i(x)$$

3. **Internalised social norms** – own outcome and rules matter:

$$U_i(x_i, r_G)$$

Varieties of Preferences

1. **Selfish preferences** – only own outcome matters:
$$U_i(x_i)$$
2. **Other-regarding preferences** – everyone's outcome matters:
$$U_i(x)$$
3. **Internalised social norms** – own outcome and rules matter:
$$U_i(x_i, r_G)$$

Varieties of Preferences

1. **Selfish preferences** – only own outcome matters:

$$U_i(x_i)$$

2. **Other-regarding preferences** – everyone's outcome matters:

$$U_i(x)$$

3. **Internalised social norms** – own outcome and rules matter:

$$U_i(x_i, r_G)$$

Note: Second-order beliefs may matter in the definition of relevant G. Players care about others' *understandings*, not about their *intentions*. Distinct from (Rabin, 1993).

Varieties of Preferences

1. **Selfish preferences** – only own outcome matters:

$$U_i(x_i)$$

2. **Other-regarding preferences** – everyone's outcome matters:

$$U_i(x)$$

3. **Internalised social norms** – own outcome and rules matter:

$$U_i(x_i, r_G)$$

Note: Second-order beliefs may matter in the definition of relevant G. Players care about others' *understandings*, not about their *intentions*. Distinct from (Rabin, 1993).

Key questions: (i) How is this situation commonly perceived? (Paying for fruit at unmanned roadside stand.) (ii) How do others perceive the situation. (Does the beggar see you or not?)

Other-regarding Preferences (no news)

Other-regarding Preferences (no news)

For example, people may care about the own material outcome as well as about efficiency and distribution. Let $X_i = M_i \subseteq \mathbb{R}$ (monetary payoffs). Let $m = (m_1, \ldots, m_n)$ be a distribution of material outcomes among n players, and let

$$\begin{aligned} m_+ &= m_1 + \cdots + m_n; \\ \underline{m} &= \min(m); \\ \overline{m} &= \max(m). \end{aligned}$$

Other-regarding Preferences (no news)

For example, people may care about the own material outcome as well as about efficiency and distribution. Let $X_i = M_i \subseteq \mathbb{R}$ (monetary payoffs). Let $m = (m_1, \ldots, m_n)$ be a distribution of material outcomes among n players, and let

$$\begin{aligned} m_+ &= m_1 + \cdots + m_n; \\ \underline{m} &= \min(m); \\ \overline{m} &= \max(m). \end{aligned}$$

Player i's preferences might be, for example

$$U_i(m) = \underbrace{\mu_i m_i}_{\text{ego}} + \underbrace{(1-\delta_i)\, m_+}_{\text{efficiency}} + \underbrace{\delta_i(\underline{m} - \overline{m})}_{\text{equality}}.$$

Other-regarding Preferences (no news)

For example, people may care about the own material outcome as well as about efficiency and distribution. Let $X_i = M_i \subseteq \mathbb{R}$ (monetary payoffs). Let $m = (m_1, \ldots, m_n)$ be a distribution of material outcomes among n players, and let

$$\begin{aligned} m_+ &= m_1 + \cdots + m_n; \\ \underline{m} &= \min(m); \\ \overline{m} &= \max(m). \end{aligned}$$

Player i's preferences might be, for example

$$U_i(m) = \underbrace{\mu_i m_i}_{\text{ego}} + \underbrace{(1-\delta_i)\, m_+}_{\text{efficiency}} + \underbrace{\delta_i(\underline{m} - \overline{m})}_{\text{equality}}.$$

Conjecture: μ is usually large; when data sometimes suggest otherwise, it is because of other forces (social norms).

Social Values and Norms (little news)

Social Values and Norms (little news)

Social values and norms are different from other-regarding preferences: They are shared, and they apply to situations.

Social Values and Norms (little news)

Social values and norms are different from other-regarding preferences: They are shared, and they apply to situations.

For example, suppose the *society* values efficiency and equality. Distributions of material payoffs that are due to luck (manna from heaven/experimenter) get a social rank according to the social value:

$$w(m) = (1 - \delta)m_+ + \delta(\underline{m} - \overline{m}).$$

Social Values and Norms (little news)

Social values and norms are different from other-regarding preferences: They are shared, and they apply to situations.

For example, suppose the *society* values efficiency and equality. Distributions of material payoffs that are due to luck (manna from heaven/experimenter) get a social rank according to the social value:

$$w(m) = (1-\delta)m_+ + \delta(\underline{m} - \overline{m}).$$

Social Values and Norms (little news)

Social values and norms are different from other-regarding preferences: They are shared, and they apply to situations.

For example, suppose the *society* values efficiency and equality. Distributions of material payoffs that are due to luck (manna from heaven/experimenter) get a social rank according to the social value:

$$w(m) = (1-\delta)m_+ + \delta(\underline{m} - \overline{m}).$$

Natural candidates for social norms are then strategy profiles s^* that maximize $w(m(s))$.

Social values and norms are different from other-regarding preferences: They are shared, and they apply to situations.

For example, suppose the *society* values efficiency and equality. Distributions of material payoffs that are due to luck (manna from heaven/experimenter) get a social rank according to the social value:

$$w(m) = (1 - \delta)m_+ + \delta(\underline{m} - \overline{m}).$$

Natural candidates for social norms are then strategy profiles s^* that maximize $w(m(s))$.

Questions in need of consideration: (i) What if these norms are too ambitious to be realistic? (ii) How does this play out dynamically – e.g., what is desired behavior after a norm violation? (iii) Who defines the situation? (iv) How do players get information about values?

Internalized Social Norms (similar to Lopez-Perez, 2008)

Internalized Social Norms (similar to Lopez-Perez, 2008)

If selfish people internalize norms, their preferences are represented by a vNM-utility function that depend both on the material payoff and the norm:

$$u_i(s, s^*) = m_i(s) - \nu_i n_i(s, s^*).$$

If selfish people internalize norms, their preferences are represented by a vNM-utility function that depend both on the material payoff and the norm:

$$u_i(s, s^*) = m_i(s) - \nu_i n_i(s, s^*).$$

Suppose players dislike violating a norm for personal benefit at norm compliers' expense (guilt). Thus, define the indicator variable

$$I_i = \begin{cases} 1 & \text{if } m_i(s_i, s_{-i}) > m_i(s_i^*, s_{-i}); \\ 0 & \text{otherwise.} \end{cases}$$

Let the disutility from norm-breaking be linear in externality:

$$n_i(s, s^*) = \overbrace{\sum_{j \neq i : s_j = s_j^*} I_i \left(m_j\left(s_i^*, s_{-i}\right) - m_j\left(s_i, s_{-i}\right)\right)}^{\text{externality on compliers}}.$$

Problem: Asymmetric Situations

Problem: Asymmetric Situations

How to treat opponents who will not take any action, or for whom the normative action does not involve a sacrifice?

Two possible answers:

- Treat such opponents as if they comply. Benefit of the doubt. (Focus on this simple case today.)

Problem: Asymmetric Situations

How to treat opponents who will not take any action, or for whom the normative action does not involve a sacrifice?

Two possible answers:

- Treat such opponents as if they comply. Benefit of the doubt. (Focus on this simple case today.)
- Get disutility only from imposing costs on *those who would have complied with the norm "in my shoes"*. This becomes an equilibrium calculation, defining a threshold ν_i and involving i's prior belief about ν_j.

Example 1a (again): Unilateral Helping Situation

Example 1a (again): Unilateral Helping Situation

In 2-player case, welfare and utility are

$$\begin{aligned} w(m) &= (1-\delta)m_+ + \delta(\underline{m} - \overline{m}) \\ u_i(s, s^*) &= m_i(s) - \nu_i I_i(m_j(s_i^*, s_{-i}^*) - m_j(s_i, s_{-i}^*)). \end{aligned}$$

Example 1a (again): Unilateral Helping Situation

In 2-player case, welfare and utility are

$$\begin{aligned} w(m) &= (1-\delta)m_+ + \delta(\underline{m} - \overline{m}) \\ u_i(s, s^*) &= m_i(s) - \nu_i I_i(m_j(s_i^*, s_{-i}^*) - m_j(s_i, s_{-i}^*)). \end{aligned}$$

Suppose Rowena has the following choice of material outcomes, affecting also Colin's payoff:

C	$0, 3$
D	$1, 0$

Example 1a (again): Unilateral Helping Situation

In 2-player case, welfare and utility are

$$\begin{aligned} w(m) &= (1-\delta)m_{+} + \delta(\underline{m} - \overline{m}) \\ u_i(s, s^*) &= m_i(s) - \nu_i l_i(m_j(s_i^*, s_{-i}^*) - m_j(s_i, s_{-i}^*)). \end{aligned}$$

Suppose Rowena has the following choice of material outcomes, affecting also Colin's payoff:

C	$0, 3$
D	$1, 0$

Norm: $s^* = (C)$ iff $(1-\delta)(3) - \delta(3) \geq (1-\delta)(1) - \delta(1)$, i.e., iff $\delta \leq 1/2$.

Example 1a (again): Unilateral Helping Situation

In 2-player case, welfare and utility are

$$\begin{aligned} w(m) &= (1-\delta)m_{+} + \delta(\underline{m} - \overline{m}) \\ u_i(s, s^*) &= m_i(s) - \nu_i I_i(m_j(s_i^*, s_{-i}^*) - m_j(s_i, s_{-i}^*)). \end{aligned}$$

Suppose Rowena has the following choice of material outcomes, affecting also Colin's payoff:

C	$0, 3$
D	$1, 0$

Norm: $s^* = (C)$ iff $(1-\delta)(3) - \delta(3) \geq (1-\delta)(1) - \delta(1)$, i.e., iff $\delta \leq 1/2$.

Rowena's utility in Helping Game if C is the norm:
$u_R(C) = 0$
$u_R(D) = 1 - 3\nu_R$

Prediction: Rowena complies with norm C iff $\nu_R \geq 1/3$.

Example 1b: Bilateral Helping Situation

Example 1b: Bilateral Helping Situation

$$w(m) = (1-\delta)m_+ + \delta(\underline{m} - \overline{m})$$
$$u_i(s, s^*) = m_i(s) - v_i l_i(m_j(s_i^*, s_{-i}^*) - m_j(s_i, s_{-i}^*)).$$

Suppose both Rowena and Colin move, and do so simultaneously:

C	0, 3
D	1, 0

Rowena

C	D
3, 0	0, 1

Colin

Simultaneous Moves

Example 1b: Bilateral Helping Situation

$$w(m) = (1-\delta)m_+ + \delta(\underline{m} - \overline{m})$$
$$u_i(s, s^*) = m_i(s) - v_i l_i(m_j(s_i^*, s_{-i}^*) - m_j(s_i, s_{-i}^*)).$$

Suppose both Rowena and Colin move, and do so simultaneously:

C	0, 3
D	1, 0

Rowena

C	D
3, 0	0, 1

Colin

Simultaneous Moves

Norm: $a^* = (C, C)$ is the norm (regardless of δ)

Example 1b: Bilateral Helping Situation

$$
\begin{aligned}
w(m) &= (1-\delta)m_+ + \delta(\underline{m} - \overline{m}) \\
u_i(s, s^*) &= m_i(s) - v_i I_i(m_j(s_i^*, s_{-i}^*) - m_j(s_i, s_{-i}^*)).
\end{aligned}
$$

Suppose both Rowena and Colin move, and do so simultaneously:

Rowena

C	$0, 3$
D	$1, 0$

Colin

C	D
$3, 0$	$0, 1$

Simultaneous Moves

Norm: $a^* = (C, C)$ is the norm (regardless of δ)

Utilities in the Helping Game are then:

$$u(C, C) = 3; u(D, D) = 1; u(C, D) = 0; u(D, C) = 4 - 3v_i.$$

$$w(m) = (1-\delta)m_+ + \delta(\underline{m} - \overline{m})$$
$$u_i(s, s^*) = m_i(s) - \nu_i I_i(m_j(s_i^*, s_{-i}^*) - m_j(s_i, s_{-i}^*)).$$

Suppose both Rowena and Colin move, and do so simultaneously:

C	0, 3
D	1, 0

Rowena

C	D
3, 0	0, 1

Colin

Simultaneous Moves

Norm: $a^* = (C, C)$ is the norm (regardless of δ)

Utilities in the Helping Game are then:

$$u(C,C) = 3; u(D,D) = 1; u(C,D) = 0; u(D,C) = 4 - 3\nu_i.$$

Let p_{ij} be i's belief that j plays C.

$$w(m) = (1-\delta)m_+ + \delta(\underline{m} - \overline{m})$$
$$u_i(s, s^*) = m_i(s) - \nu_i I_i(m_j(s_i^*, s_{-i}^*) - m_j(s_i, s_{-i}^*)).$$

Suppose both Rowena and Colin move, and do so simultaneously:

C	0, 3
D	1, 0

Rowena

C	D
3, 0	0, 1

Colin

Simultaneous Moves

Norm: $a^* = (C, C)$ is the norm (regardless of δ)

Utilities in the Helping Game are then:

$$u(C, C) = 3; u(D, D) = 1; u(C, D) = 0; u(D, C) = 4 - 3\nu_i.$$

Let p_{ij} be i's belief that j plays C.

Prediction: Player i chooses C iff $p_{ij}\nu_i \geq 1/3$. (Check!)

$$w(m) = (1-\delta)m_+ + \delta(\underline{m} - \overline{m})$$
$$u_i(s, s^*) = m_i(s) - \nu_i I_i(m_j(s_i^*, s_{-i}^*) - m_j(s_i, s_{-i}^*)).$$

Suppose both Rowena and Colin move, and do so simultaneously:

C	0, 3
D	1, 0

Rowena

C	D
3, 0	0, 1

Colin

Simultaneous Moves

Norm: $a^* = (C, C)$ is the norm (regardless of δ)

Utilities in the Helping Game are then:

$$u(C,C) = 3; u(D,D) = 1; u(C,D) = 0; u(D,C) = 4 - 3\nu_i.$$

Let p_{ij} be i's belief that j plays C.

Prediction: Player i chooses C iff $p_{ij}\nu_i \geq 1/3$. (Check!)

Remark: If $\nu_i = \nu_j < 1/3$ this PD. If $\nu_i = \nu_j > 1/3$, this is Stag Hunt.

Example 2: The importance of others' knowledge

Example 2: The importance of others' knowledge

Example 2: The importance of others' knowledge

Recall the USD 10 Dictator experiment with an unexpected exit option for USD 9, by Dana, Cain, and Dawes (2006). Player i is dictator, j is recipient.

Example 2: The importance of others' knowledge

Recall the USD 10 Dictator experiment with an unexpected exit option for USD 9, by Dana, Cain, and Dawes (2006). Player i is dictator, j is recipient. Exit option changes B_{iji} – and hence enables i to avoid a situation and the norm attached to it.

Example 2: The importance of others' knowledge

Recall the USD 10 Dictator experiment with an unexpected exit option for USD 9, by Dana, Cain, and Dawes (2006). Player i is dictator, j is recipient. Exit option changes B_{iji} – and hence enables i to avoid a situation and the norm attached to it.

- The norm is 50:50. The model predicts that the player will give 5 dollars if $\nu \geq 1$ and nothing otherwise.

Example 2: The importance of others' knowledge

Recall the USD 10 Dictator experiment with an unexpected exit option for USD 9, by Dana, Cain, and Dawes (2006). Player i is dictator, j is recipient. Exit option changes B_{iji} – and hence enables i to avoid a situation and the norm attached to it.

- The norm is 50:50. The model predicts that the player will give 5 dollars if $\nu \geq 1$ and nothing otherwise.
  - These are almost always the two modes.

Example 2: The importance of others' knowledge

Recall the USD 10 Dictator experiment with an unexpected exit option for USD 9, by Dana, Cain, and Dawes (2006). Player i is dictator, j is recipient. Exit option changes B_{iji} – and hence enables i to avoid a situation and the norm attached to it.

- The norm is 50:50. The model predicts that the player will give 5 dollars if $\nu \geq 1$ and nothing otherwise.
  - These are almost always the two modes.
- When the player is offered the opportunity to exit, the model predicts that the exit reservation value is 5 dollars for those that would give 5, and between 5 and 10 for others. Why? Choosing situation!

Example 2: The importance of others' knowledge

Recall the USD 10 Dictator experiment with an unexpected exit option for USD 9, by Dana, Cain, and Dawes (2006). Player i is dictator, j is recipient. Exit option changes B_{iji} – and hence enables i to avoid a situation and the norm attached to it.

- The norm is 50:50. The model predicts that the player will give 5 dollars if $\nu \geq 1$ and nothing otherwise.
  - These are almost always the two modes.
- When the player is offered the opportunity to exit, the model predicts that the exit reservation value is 5 dollars for those that would give 5, and between 5 and 10 for others. Why? Choosing situation!
  - There is indeed a negative empirical relationship between the initial gift and the exit reservation price; Broberg, Ellingsen, and Johannesson (2007); Lazear, Malmendier, Weber (2012).

Example 3: The importance of own knowledge

Example 3: The importance of own knowledge

Suppose norms are defined for *actions in the situation*, not for *information acquisition about the situation*.

Example 3: The importance of own knowledge

Suppose norms are defined for *actions in the situation*, not for *information acquisition about the situation*.

Dana, Weber, Kuang (2007). Dictator chooses between two options, A and B. There are two states of the world.

- State 1 (non-aligned) payoffs: $A = (6, 1), B = (5, 5)$.
- State 2 (aligned) payoffs: $A = (6, 5), B = (5, 1)$.

Example 3: The importance of own knowledge

Suppose norms are defined for *actions in the situation*, not for *information acquisition about the situation*.

Dana, Weber, Kuang (2007). Dictator chooses between two options, A and B. There are two states of the world.

- State 1 (non-aligned) payoffs: $A = (6, 1), B = (5, 5)$.
- State 2 (aligned) payoffs: $A = (6, 5), B = (5, 1)$.

Case 1. State is known.

Example 3: The importance of own knowledge

Suppose norms are defined for *actions in the situation*, not for *information acquisition about the situation*.

Dana, Weber, Kuang (2007). Dictator chooses between two options, A and B. There are two states of the world.

- State 1 (non-aligned) payoffs: $A = (6, 1), B = (5, 5)$.
- State 2 (aligned) payoffs: $A = (6, 5), B = (5, 1)$.

Case 1. State is known.

Case 2. Both states equally likely. You can find out state before choosing, or choose without knowing the state (remaining in the uninformed state).

Example 3: The importance of own knowledge

Suppose norms are defined for *actions in the situation*, not for *information acquisition about the situation*.

Dana, Weber, Kuang (2007). Dictator chooses between two options, A and B. There are two states of the world.

- State 1 (non-aligned) payoffs: $A = (6, 1), B = (5, 5)$.
- State 2 (aligned) payoffs: $A = (6, 5), B = (5, 1)$.

Case 1. State is known.

Case 2. Both states equally likely. You can find out state before choosing, or choose without knowing the state (remaining in the uninformed state).

Prediction: (i) Low ν doesn't care, chooses A. (ii) High ν reveals and chooses B in State 1 and A in state 2. (iii) Intermediate ν does not reveal and chooses A.

Example 3: The importance of own knowledge

Suppose norms are defined for *actions in the situation*, not for *information acquisition about the situation*.

Dana, Weber, Kuang (2007). Dictator chooses between two options, A and B. There are two states of the world.

- State 1 (non-aligned) payoffs: $A = (6, 1), B = (5, 5)$.
- State 2 (aligned) payoffs: $A = (6, 5), B = (5, 1)$.

Case 1. State is known.

Case 2. Both states equally likely. You can find out state before choosing, or choose without knowing the state (remaining in the uninformed state).

Prediction: (i) Low ν doesn't care, chooses A. (ii) High ν reveals and chooses B in State 1 and A in state 2. (iii) Intermediate ν does not reveal and chooses A. Fits evidence that 44% don't reveal, and then usually choose A. Also, more revelation if larger probability of non-alignment (Feiler, 2006).

Example 4: Changing the Game

Suppose Rowena and Colin are selfish and do not care about norms.

At date 1, they play Γ_1, the "Helping Game"

C	0, 3
D	1, 0

Example 4: Changing the Game

Suppose Rowena and Colin are selfish and do not care about norms.

At date 1, they play Γ_1, the "Helping Game"

C	$0, 3$
D	$1, 0$

At date 2, they play Γ_2, the Battle of the Sexes Game

	L	R
T	$0, 0$	$3, 1$
B	$1, 3$	$0, 0$

Suppose the social convention is (B, L).

Example 4: Changing the Game

Suppose Rowena and Colin are selfish and do not care about norms.

At date 1, they play Γ_1, the "Helping Game"

C	$0, 3$
D	$1, 0$

At date 2, they play Γ_2, the Battle of the Sexes Game

	L	R
T	$0, 0$	$3, 1$
B	$1, 3$	$0, 0$

Suppose the social convention is (B, L). Could they do better, and if so, how?

A Combined Game Γ_3

A Combined Game Γ_3

Rowena's 8 strategies		Colin's 4 strategies
$\langle C, TT \rangle$	$\langle C, TB \rangle$	$\langle LL \rangle$
$\langle C, BT \rangle$	$\langle C, BB \rangle$	$\langle RR \rangle$
$\langle D, TT \rangle$	$\langle D, TB \rangle$	$\langle LR \rangle$
$\langle D, BT \rangle$	$\langle D, BB \rangle$	$\langle RL \rangle$

BT means Rowena plays B is she played C at date 1 and T if she played D at date 1, and so on.

A Combined Game Γ_3

Rowena's 8 strategies		Colin's 4 strategies
$\langle C, TT \rangle$	$\langle C, TB \rangle$	$\langle LL \rangle$
$\langle C, BT \rangle$	$\langle C, BB \rangle$	$\langle RR \rangle$
$\langle D, TT \rangle$	$\langle D, TB \rangle$	$\langle LR \rangle$
$\langle D, BT \rangle$	$\langle D, BB \rangle$	$\langle RL \rangle$

BT means Rowena plays B is she played C at date 1 and T if she played D at date 1, and so on.

$(\langle C, BT \rangle, \langle LR \rangle)$ is an equilibrium of Γ_3. Both players' payoff have increased by 1.

A Combined Game Γ_3

Rowena's 8 strategies		Colin's 4 strategies
$\langle C, TT \rangle$	$\langle C, TB \rangle$	$\langle LL \rangle$
$\langle C, BT \rangle$	$\langle C, BB \rangle$	$\langle RR \rangle$
$\langle D, TT \rangle$	$\langle D, TB \rangle$	$\langle LR \rangle$
$\langle D, BT \rangle$	$\langle D, BB \rangle$	$\langle RL \rangle$

BT means Rowena plays B is she played C at date 1 and T if she played D at date 1, and so on.

$(\langle C, BT \rangle, \langle LR \rangle)$ is an equilibrium of Γ_3. Both players' payoff have increased by 1.

Note: Here, we did not need norms to improve the situation. Enough that players view the situation differently!

Bonus example: Crowding Out

Bonus example: Crowding Out

Suppose Choice D is penalized by $p < 1/2$ money units.

$$
\begin{aligned}
w(m) &= (1-\delta) m_+ + \delta(\underline{m} - \overline{m}) \\
u_i(a, a^*) &= m_i(a) - \nu_i I_{ij}(m_j(a_i^*, a_{-i}^*) - m_j(a_i, a_{-i}^*)).
\end{aligned}
$$

Suppose Choice D is penalized by $p < 1/2$ money units.

$$\begin{aligned} w(m) &= (1-\delta)m_+ + \delta(\underline{m} - \overline{m}) \\ u_i(a, a^*) &= m_i(a) - v_i I_{ij}(m_j(a_i^*, a_{-i}^*) - m_j(a_i, a_{-i}^*)). \end{aligned}$$

C	$0, 3$
D	$1 - p, p$

Bonus example: Crowding Out

Suppose Choice D is penalized by $p < 1/2$ money units.

$$w(m) = (1-\delta)m_+ + \delta(\underline{m} - \overline{m})$$
$$u_i(a, a^*) = m_i(a) - v_i I_{ij}(m_j(a_i^*, a_{-i}^*) - m_j(a_i, a_{-i}^*)).$$

C	$0, 3$
D	$1 - p, p$

Norm: $a^* = (C)$ iff $(1-\delta)(3) - \delta(3) \geq (1-\delta)(1) - \delta(1-2p)$, i.e., iff $\delta \leq (1-p)/2$. Stronger: C less likely to be the norm (general).

Bonus example: Crowding Out

Suppose Choice D is penalized by $p < 1/2$ money units.

$$\begin{aligned} w(m) &= (1-\delta)m_{+} + \delta(\underline{m} - \overline{m}) \\ u_i(a, a^*) &= m_i(a) - v_i I_{ij}(m_j(a_i^*, a_{-i}^*) - m_j(a_i, a_{-i}^*)). \end{aligned}$$

C	$0, 3$
D	$1 - p, p$

Norm: $a^* = (C)$ iff $(1-\delta)(3) - \delta(3) \geq (1-\delta)(1) - \delta(1-2p)$, i.e., iff $\delta \leq (1-p)/2$. Stronger: C less likely to be the norm (general).

Since Rowena always plays D if D is the norm (no trade-off), there is a set of parameters such that penalties destroy voluntary cooperation.

Bonus example: Crowding Out

Suppose Choice D is penalized by $p < 1/2$ money units.

$$\begin{aligned} w(m) &= (1-\delta)m_+ + \delta(\underline{m} - \overline{m}) \\ u_i(a, a^*) &= m_i(a) - \nu_i I_{ij}(m_j(a_i^*, a_{-i}^*) - m_j(a_i, a_{-i}^*)). \end{aligned}$$

C	$0, 3$
D	$1 - p, p$

Norm: $a^* = (C)$ iff $(1-\delta)(3) - \delta(3) \geq (1-\delta)(1) - \delta(1-2p)$, i.e., iff $\delta \leq (1-p)/2$. Stronger: C less likely to be the norm (general).

Since Rowena always plays D if D is the norm (no trade-off), there is a set of parameters such that penalties destroy voluntary cooperation.

Simple intuition: The offending action isn't offensive anymore.

Interpreting Gneezy-Rustichini (2000)

Interpreting Gneezy-Rustichini (2000)

The fine for picking up late fundamentally changes the situation.

- Before, the choice is between a clearly selfish action (be late) and an unselfish one (be on time). Being late is a norm violation.

Key difference between:

In order to work as a symbol, it may be important that the rule is communicated together with the fine.

Interpreting Gneezy-Rustichini (2000)

The fine for picking up late fundamentally changes the situation.

- Before, the choice is between a clearly selfish action (be late) and an unselfish one (be on time). Being late is a norm violation.
- After, the choice is less clearly between selfishness and unselfishness. If the price/fine is seen as full compensation, picking up late isn't a norm violation anymore.

Key difference between:

In order to work as a symbol, it may be important that the rule is communicated together with the fine.

Interpreting Gneezy-Rustichini (2000)

The fine for picking up late fundamentally changes the situation.

- Before, the choice is between a clearly selfish action (be late) and an unselfish one (be on time). Being late is a norm violation.
- After, the choice is less clearly between selfishness and unselfishness. If the price/fine is seen as full compensation, picking up late isn't a norm violation anymore.

Key difference between:

- A fine as a symbol – which ought to strengthen the norm.

In order to work as a symbol, it may be important that the rule is communicated together with the fine.

Interpreting Gneezy-Rustichini (2000)

The fine for picking up late fundamentally changes the situation.

- Before, the choice is between a clearly selfish action (be late) and an unselfish one (be on time). Being late is a norm violation.
- After, the choice is less clearly between selfishness and unselfishness. If the price/fine is seen as full compensation, picking up late isn't a norm violation anymore.

Key difference between:

- A fine as a symbol – which ought to strengthen the norm.
- A fine as a price – which weakens the norm, but may have a deterrent effect on its own.

In order to work as a symbol, it may be important that the rule is communicated together with the fine.