Frustration, anger, and aggression have important consequences for economic and social behavior.

- Anger arises from the frustration of non-attainment of an expected outcome; as a behavioral consequence, this goal-blockage experience can lead to aggressive behavior.
- Emotions depend on beliefs and then we need belief-dependent preferences to illustrate anger-like motivations.
- We develop a formal framework and a set of models that incorporate frustration and anger in games.
Motivation & Examples

The following is inconsistent with standard social preferences (e.g., inequity aversion), but consistent with our model(s):

- **Psychology**: frustration-aggression hypothesis (Dollard et al., 1939). “Experiences of anger consist of the experience of an event as obstructing one’s goals and as caused by someone else’s blameworthy intent.” (Frijda, 1993).

- **Facts (empirics)**:
  - unexpected losses by home football/soccer teams are associated with increased domestic violence (Card & Dahl, 2011) or violent crime (Munyo & Rossi 2013).
  - firms do not want to “antagonize customers” (Anderson, E. and D. Simester 2010)

- **Facts (experimental)**:
  - Self-reported anger correlates with punishment of free-riders in Public Good Games (Fehr & Gächter, 2002)
  - Rejections in the Ultimatum Game correlate with (manipulated) initially expected offers (Sanfey, 2009; Xiang et al., 2013, with fMRI)
  - Deviations from expectations drive both anger and the destruction of endowments in Power-to-Take Games (van Winden et al. 2002,'05).
  - Agents are blamed by principals for bad outcomes (Gurdal et al. 2014).
Consider **multistage games with observable actions**: 

- Players \( i \in I \)
- Histories/nodes \( h: \emptyset, a^1, (a^1, a^2), \ldots \in \tilde{H} \) (\( \emptyset = \) root of tree \( \tilde{H} \))
- Paths, or terminal nodes \( z = (a^1, \ldots, a^T) \in Z \), where \( a^t = (a^t_i)_{i \in I} \) is a profile of actions
- Monetary payoffs \( \pi(z) = (\pi_i(z))_{i \in I} \)
- Beliefs about paths conditional on each nonterminal history \( h \in H := \tilde{H} \setminus Z \):
  - **First-order beliefs**: \( \alpha_i = (\alpha_i(\cdot|h))_{h \in H} \in \Delta^H_i(Z) \)
  - **Second-order beliefs**: \( \beta_i = (\beta_i(\cdot|h))_{h \in H} \in \Delta^H_i(Z \times \Delta^H_{-i}(Z)) \)
  - **Coherence** of \((\alpha_i, \beta_i)\): \( \alpha_i(\cdot|h) = \text{marg}_Z \beta_i(\cdot|h) \) for each \( h \in H \)
- Main focus: 2-stage games with **complete information** (simplicity)
Frustration

- Anger is anchored in frustration.
- **Frustration** is given by *unavoidable shortfall* in expected material payoff; thus, it depends on his beliefs about others and on own plans.

Player $i$’s frustration, in the second stage given $a^1$, is defined as:

$$F_i(a^1; \alpha_i) = \left[ \mathbb{E}[\pi_i; \alpha_i] - \max_{a_i^2 \in A_i(a^1)} \mathbb{E}[\pi_i(a^1, a_i); \alpha_i] \right]^+$$

where $[x]^+ = \max\{0, x\}$. 
Example: Ultimatum mini-Game

If Bob initially expects $f$ (fair offer), his frustration following $g$ (greedy offer) is

$$F_b(g; \alpha_b) = [2 \cdot (1 - \alpha_b(g)) + 1 \cdot \alpha_b(g) \alpha_b(y|g) - 1]^+.$$ 

Given $g$, Bob is more frustrated (i) the more he expects the fair offer $f$, and (ii) the less he initially plans to reject the greedy offer $g$.

▶ How do players react to frustration?
Blame

A frustrated player may go after his co-players depending on the evaluation of how much his opponents can be blamed for the outcome which frustrates him.

- $B_{ij}$ measures the amount of frustration $i$ attributes to $j$, with:

  $$B_{ij}(a^1; \beta_i) \leq F_i(a^1; \alpha_i)$$

- Player $i$ moving at the second stage chooses $a_i$ to maximize his “decision utility” of the form

  $$u_i(a^1, a_i; \beta_i) = \pi_i(a^1, a_i) - \theta_i \sum_{j \neq i} B_{ij}(a^1; \beta_i) \pi_j(a^1, a_i)$$

  where $\theta_i \geq 0$ is a sensitivity parameter.

- Three approaches to incorporate anger into utility functions according to different levels of cognitive appraisal, reflected by different blame functions.
Simple Anger (SA)

A player’s tendency to hurt others is proportional to his frustration, un-modulated by cognitive appraisal of blame:

\[ B_{ij}(a^1; \alpha_i) = F_i(a^1; \alpha_i) \]

- **Frustration-aggression displacement hypothesis** (Dollard *et al.*, 1939): the existence of frustration leads to some form of aggressive behavior through a displacement effect that directs hostile inclinations at substitute targets.

- **Card & Dahl, 2001**: correlation between an external source of frustration from unexpected loss by local teams and an increasing number of reports of domestic abuse.
Simple Anger: Ultimatum mini-Game

- If Bob initially expects $f$ with certainty, his frustration following $g$ is $F_b(g; \alpha_b) = [2 \cdot 1 + 0 \cdot \alpha_b(y|g) - 1]^+ = 1$. Therefore $u_{SA}^b(g, n; \alpha_b) - u_{SA}^b(g, y; \alpha_b) = 3\theta_b - 1$. Bob rejects $g$ if $\theta_b \geq 1/3$; otherwise, he accepts.

- If Bob initially expects $g$ with certainty, his frustration following $g$ is $F_b(g; \alpha_b) = [2 \cdot 0 + 1 \cdot \alpha_b(y|g) - 1]^+ = 0$. Bob accepts $g$ for every $\theta_b$ and $\alpha_b(y|g)$. 
Anger from Blaming Behavior (ABB)

A player blames whoever causes his frustration.

How much $i$ blames $j$ is determined by a continuous function $B_{ij}(a^1; \alpha_i)$ such that:

$$B_{ij}(a^1; \alpha_i) = \begin{cases} 
0, & \text{if } \{j\} \neq I(\emptyset) \\
F_b(g; \alpha_b), & \text{if } \{j\} = I(\emptyset)
\end{cases}$$

In leader-followers games, SA and ABB are equivalent. For example, in the Ultimatum Minigame the two models yield to the same behavioral prediction.
Example: Hammering One’s Thumb

\[ F_i(B; \alpha_a) = (1 - \varepsilon) \cdot 2 + \varepsilon \alpha_a(N|B) \cdot 1 - 1 > 0 \]

Difference between simple anger and anger from blaming behavior:

- **SA**: given \( B \), Andy chooses \( T \) for \( \theta_a \) sufficiently high;
- **ABB**: given \( B \), Andy chooses \( N \) regardless of \( \theta_a \).
ABB: Could-Have-Been Blame

We propose two specific functional forms for ABB:

1. **Could-have-been blame**
   A frustrated player $i$ considers for each co-player $j$ what $i$ would have obtained at most, in expectation, had $j$ chosen differently:

   $$B_{ij}(a^1, \alpha_i) =$$

   $$\min \left\{ \left[ \max_{a'_j \in A_j(\emptyset)} \mathbb{E}[\pi_i(a^-_j, a'_j); \alpha_i] - \mathbb{E}[\pi_i(a^1, \alpha_i)] \right]^+ ; F_i(a^1; \alpha_i) \right\}$$
Blaming unexpected deviations

A frustrated $i$ assesses for each co-player $j$ how much $i$ would have obtained had $j$ behaved as expected:

$$B_{ij}(a^1, \alpha_i) = \min \left\{ \left[ \sum_{a'_j \in A_j(\varnothing)} \alpha_{ij}(a'_j) \mathbb{E}[\pi_i|(a^1_{-j}, a_j); \alpha_i] - \mathbb{E}[\pi_i|a^1; \alpha_i] \right]^+ ; F_i(a^1; \alpha_i) \right\}$$
Could-Have-Been Blame: Asymmetric Punishment

Consider Penny at \( a^1 = (D, L) \).

\[
B_{pa}((D, L); \alpha_p) = B_{pb}((D, L); \alpha_p) = \min \{ [2 - \mathbb{E}[\pi_p|(D, L); \alpha_p]]^+, [\mathbb{E}[\pi_p; \alpha_p] - 1]^+ \} = [\mathbb{E}[\pi_p; \alpha_p - 1]]^+
\]

- Penny fully blames both Ann and Bob for her frustration, but can and does punish only Bob.
Blaming Unexpected Deviations: Asymmetric Punishment

If Penny is initially certain of \((U, L)\) then \(E[\pi_p; \alpha_p] = 2\). Given \((D, L)\), her frustration is \(F_p((D, L); \alpha_p) = [2 - 1]^+ = 1\).

▶ Penny fully blames Ann, who deviated from \(U\) to \(D\):

\[
B_{pa}((D, L); \alpha_p) = \min \left\{ \left[ 2 - E[\pi_p|a^1; \alpha_p] \right]^+ ; 1 \right\} = 1
\]

▶ Penny does not blame Bob, who played \(L\) as expected.
Anger from Blaming Intentions (ABI)

A player blames whoever he believes *intended* to cause his frustration.

- A frustrated player asks himself whether the co-player intended to give her a low expected payoff:

\[
B_{ij}(a^1_i; \beta_i) = \min \left\{ \mathbb{E} \left[ \max_{a_j} \sum_{a^1_j} \alpha_j, -j(a^1_j) \mathbb{E}[\pi_i | (a^1_j, a^1_i); \alpha_j] - \mathbb{E}[\pi_i; \alpha_j] \right| a^1_i; \beta_i \vphantom{\sum_{a^1_j}} \right], F_i(a^1_i, \alpha_i) \right\}
\]

- Key issue: is observed behavior interpreted as *intentional*? According to the “trembling-hand” story, deviations from equilibrium are unintentional ⇒ no blame, no aggression!
Sequential Equilibrium: Consistent Assessments

Fix psychological utility function \( u_i \) (\( u_i = u_i^{SA}, u_i^{ABB}, u_i^{ABI}, \) or other).

An assessment is a profile of behavioral strategies and coherent beliefs \((\sigma_i, \alpha_i, \beta_i)_{i \in I}\) such that \( \sigma_i \) is the plan \( \alpha_i,i \) entailed by second-order belief \( \beta_i \).

An assessment \((\sigma_i, \beta_i)_{i \in I}\) is consistent if, for all \( i, h \) and \( a = (a_j)_{j \in I} \),

(a) (behav.strat. given by 1\textsuperscript{st}-ord. beliefs) \( \alpha_i(a|h) = \prod_{j \in I(h)} \sigma_j(a_j|h) \),

(b) (correct 2\textsuperscript{nd}-ord. beliefs) \( \text{marg}_{\Delta_{h,i}(Z)} \beta_i(\cdot|h) = \delta_{\alpha_{-i}} \),

where \( \alpha_i \) is derived from \( \beta_i \) (coherence) and \( \delta_{\alpha_{-i}} \) is the Dirac probability measure that assignes probability 1 to the singleton \( \{\alpha_{-i}\} \).
Sequential Equilibrium: Sequential Rationality

An assessment \((σ_i, β_i)_{i ∈ I}\) is a **sequential equilibrium (SE)** if

1. it is consistent and
2. it satisfies the following **sequential rationality** condition (one-shot-deviation property):

\[
∀ i ∈ I, ∀ h ∈ H, \text{Supp}σ_i(⋅|h) \subseteq \arg \max_{a_i ∈ A_i(h)} u_i(h, a_i; β_i).
\]
Equilibrium Analysis

**Theorem** (cf. B&D, 2009)
If $u_i(h, a_i; \cdot)$ is continuous for all $i \in I$, $h \in H$ and $a_i \in A_i(h)$, then there is at least one SE.

- Every game with SA, ABB, or ABI has at least one SE.

**Proposition**
In every perfect-information game form with no chance move and a unique SE of the material-payoff game, this unique material-payoff equilibrium is realization-equivalent to an SE of the psychological game with ABI, ABB, or – with only two players – SA.

- **Intuition:** PI $\Rightarrow$ pure-strat. material-payoff eq. $\Rightarrow$ No on-path surprise $\Rightarrow$ No on-path anger with psy-utility functions $\Rightarrow$ On-path material payoff maximization and angry punishment of deviations makes deviations even less attractive.
Example: Sequential Equilibria of UmG

- Under all models (SA, ABB, ABI), \((g, y)\) is SE for all \(\theta_b\)
  (illustration of previous Proposition):
  - if greedy offer \(g\) is expected no anger \(\Rightarrow\) yes
- Under SA and ABI, also \((f, n)\) is SE if \(\theta_b\) is high enough:
  - off-path offer \(g\) is unexpected \(\Rightarrow\) anger \(\Rightarrow\) no for suff. high \(\theta_b\) (credible threat)
- Under ABI \((f, n)\) is not SE:
  - off-path offer \(g\) interpreted as unintentional mistake (trembling hand) \(\Rightarrow\) no blame \(\Rightarrow\) \(n\) not credible threat
Discussion

- **Dynamic Inconsistency of preferences**: implied by F&A models, but we assume sophistication $\implies$ no change of plan

- **T-period extension**: fast vs. slow play version
  - fast play: benchmark is initial expectation
  - slow play: benchmark is beginning-of-period expectation
  - cooling off effect (experiment by Gneezy and Imas, 2014)

- **Unintuitive implication of SE**: deviations are not blamed and punished, because they are perceived as unintended

- Do not blame (sic) us, we are true to the original concept. Blame instead the trembling-hand notion of consistency built into (original) SE, and look for different solution concepts!

- **Polymorphic SE**: minimal generalization of SE allowing meaningful on-path updating about others’ intentions (even with complete information!)
We make sense of relevant evidence with psychological models where anger is belief-dependent: unavoidable shortfall.

Anger can also depend on regret (e.g., unexpected discounts after purchase, see Anderson & Simester, 2010), which is belief-dependent in a different way ...

... or on perceived unfairness (riots, political unrest), which—however—often is hard to distinguish from deviation from expectations.

F&A vs Negative Reciprocity: both give credibility to threats, but unlike NR, in the F&A models no aggression occurs on a pure-equilibrium path \( \Rightarrow \) no miserable equilibrium.


References (cont.)


