Enrico Mattia Salonia
Game Theory - Spring 2026

Guilt Aversion

A traveler (the *tipper*, Player 2) takes a taxi from the airport to the city center. The driver (Player 1) provides the service first, and then the tipper decides how much to tip. Let the tip be $t \in [0,1]$, measured as the percentage of the ride price. The tipper prefers to keep more money, so tipping is costly. However, the tipper is **guilt-averse**: she dislikes disappointing what she believes the driver expects to receive. Let $\tau \in [0,1]$ be the tip that the tipper **believes** the driver expects. Her utility is

$$U_2(t \mid \tau) = (1-t) - \beta\,[\tau - t]^+,$$

where $[x]^+ = \max\{x, 0\}$ and $\beta > 0$ measures guilt sensitivity.

1. For fixed belief τ, find the tipper's best response $t^*(\tau)$.
2. Show that if $\beta \leq 1$, the unique best response is $t^*(\tau) = 0$ for every τ.
3. Show that if $\beta > 1$, the best response is

$$t^*(\tau) = \tau.$$

4. Briefly interpret the result: when does guilt sustain positive tipping?
5. Now require *belief consistency*: in equilibrium, the driver's expectation must be correct, i.e. $\tau = t^*(\tau)$. Find all equilibria as a function of β.
6. Use your answer to part (e) to explain why tipping norms can differ sharply across countries (e.g. 20% in the US, 0% in Japan) even if individuals have identical preferences.

Solution. **Parts (a)–(c).** For $t \geq \tau$, we have $[\tau - t]^+ = 0$, so

$$U_2(t \mid \tau) = 1 - t,$$

which is decreasing in t. Hence in this region the best choice is $t = \tau$.

For $t < \tau$, we have $[\tau - t]^+ = \tau - t$, so

$$U_2(t \mid \tau) = 1 - t - \beta(\tau - t) = 1 - \beta\tau + (\beta - 1)t.$$

Therefore:

- If $\beta \leq 1$, this is weakly decreasing in t, so the best point in $[0,\tau)$ is $t = 0$. Comparing with $t = \tau$:

$$U_2(0 \mid \tau) = 1 - \beta\tau \geq 1 - \tau = U_2(\tau \mid \tau),$$

  so $t^*(\tau) = 0$.
- If $\beta > 1$, this is increasing in t, so the best point in $[0,\tau)$ is the upper boundary, i.e. approaching τ. In the region $t \geq \tau$, utility decreases from $t = \tau$. Hence $t = \tau$ is optimal, so $t^*(\tau) = \tau$.

Part (d). When guilt is weak ($\beta \le 1$), monetary incentives dominate and the tipper leaves no tip regardless of what she thinks the driver expects. When guilt is strong ($\beta > 1$), the tipper matches what she believes the driver expects, so beliefs can sustain positive tipping.

Part (e). An equilibrium requires belief consistency: the driver's expectation equals the tip he actually receives, i.e. $\tau = t^*(\tau)$.

Case $\beta \le 1$. From part (b), $t^*(\tau) = 0$ for every τ. Consistency requires $\tau = 0$. This is the unique equilibrium: the driver expects nothing and receives nothing.

Case $\beta > 1$. From part (c), $t^*(\tau) = \tau$. The consistency condition $\tau = t^*(\tau) = \tau$ is satisfied for every $\tau \in [0, 1]$. Therefore there is a *continuum* of equilibria: for any $\tau^* \in [0, 1]$, the profile $(\tau, t) = (\tau^*, \tau^*)$ is an equilibrium.

In particular, both $t = 0$ and $t = 1$ (and everything in between) can be sustained in equilibrium. The equilibrium tip level is indeterminate from preferences alone: it depends entirely on which self-fulfilling belief is selected.

Part (f). When $\beta > 1$, the model generates a continuum of equilibria indexed by τ^*. Each equilibrium is sustained by self-fulfilling beliefs: the driver expects τ^*, and the guilt-averse tipper matches that expectation, confirming the driver's belief.

This means that two societies with identical preferences (same $\beta > 1$) can settle on very different tipping norms. A 20% norm in the US and a 0% norm in Japan are both equilibria of the same game. What differs is not guilt sensitivity but the *belief* that coordinates behavior. Tipping norms are social conventions: arbitrary in origin but self-reinforcing once established, because deviating triggers guilt. The exercise thus explains cross-cultural variation without requiring any difference in underlying preferences.

A game where players' beliefs affect their utility over outcomes, like the one in this example, is called a psychological game. Psychological games were introduced by Geanakoplos, Pearce, and Stacchetti (1989) and further studied by Battigalli and Dufwenberg (2009).

References

Battigalli, P., & Dufwenberg, M. (2009). Dynamic psychological games. *Journal of Economic Theory*, *144*(1), 1–35.

Geanakoplos, J., Pearce, D., & Stacchetti, E. (1989). Psychological games and sequential rationality. *Games and economic Behavior*, *1*(1), 60–79.