

The Problem of Business Evil

John Thanassoulis ¹

¹University of Warwick, Warwick Business School
CEPR
U.K. Competition and Markets Authority
<https://sites.google.com/site/thanassoulis/>

August 25, 2025

The problem of evil

The problem of
business evil

Explains why even very good people can knowingly choose to do bad things

- **Rogue trading**
 - ▶ Jérôme Kerviel and Société Générale
 - ▶ Nick Leeson and Barings
- **WorldCom accounting fraud**
 - ▶ Head of Accounting *“there are no words to describe my shame and humiliation,”* Mr Burford Yates Jr.
- **The Sackler Family, OxyContin, opioid addiction**
 - ▶ Donors to the NY Met, Louvre, Taylorian Oxford, etc.
 - ▶ Trained as doctors, worked in NY Psychiatric Centre.
- **Bernie Madoff (financial fraud)**
 - ▶ *“I am so deeply sorry and ashamed [of my crimes] ... I cannot adequately express how sorry I am for what I have done.”*

This paper

- **Model:** Misconduct stopping decision

- ▶ **Ingredients:** Uncertain news, uncertain longevity, fines and ethics

- **Key Forces**

- ▶ Stopping apparently profitable activity \Rightarrow guilt. So fine \uparrow
- ▶ Continuing harmful activity causes pain to others, and inevitably public will find out and have to stop, so fines $\uparrow\uparrow$
- ▶ But game might end before public realise – salvation.

Key Results

- **In-too-deep threshold:**

- ① No matter how ethical, there exists a time such that agent will harm knowingly if discover activity is bad after this point.
- ② Increasing fines makes lock-into-evil sooner for the most ethical.

Key Results

- **In-too-deep threshold:**

- ① No matter how ethical, there exists a time such that agent will harm knowingly if discover activity is bad after this point.
- ② Increasing fines makes lock-into-evil sooner for the most ethical.

- **Constrained game:**

- ① **Patents:** In-too-deep threshold *always* triggered before patent expiry.
- ② **Limited liability:** in-too-deep threshold unless both sufficiently poor & sufficiently ethical.

Key Results

- **In-too-deep threshold:**

- ① No matter how ethical, there exists a time such that agent will harm knowingly if discover activity is bad after this point.
- ② Increasing fines makes lock-into-evil sooner for the most ethical.

- **Constrained game:**

- ① **Patents:** In-too-deep threshold *always* triggered before patent expiry.
- ② **Limited liability:** in-too-deep threshold unless both sufficiently poor & sufficiently ethical.

- **Entry:**

- ① Ethical decision makers unlikely to win an innovation race.
- ② On entry the most ethical entrepreneurs have the safest new products.

Key Results

- **In-too-deep threshold:**

- ① No matter how ethical, there exists a time such that agent will harm knowingly if discover activity is bad after this point.
- ② Increasing fines makes lock-into-evil sooner for the most ethical.

- **Constrained game:**

- ① **Patents:** In-too-deep threshold *always* triggered before patent expiry.
- ② **Limited liability:** in-too-deep threshold unless both sufficiently poor & sufficiently ethical.

- **Entry:**

- ① Ethical decision makers unlikely to win an innovation race.
- ② On entry the most ethical entrepreneurs have the safest new products.

- **Punishment regimes**

- ▶ If extreme, can time-bound in-too-deep lying.

Some key related literature

- **Intuitive why unethical agents (*homo-economicus*) harm others:** Tirole (1996), Sobel (1985), Becker (1968).
- **Bring ethics into Finance:** Arrow (1973), Hausman and McPherson (1993). **Market might flip to misconduct:** Easley and O'Hara (2023), Thanassoulis (2023), Dupont and Karpoff (2025). **Fraud Triangle:** Cressey (1953), Schuchter and Levi (2016).
- **Stopping problem with Poisson news:** Dixit and Pindyck, 1994, Benmelech, Kandel, Veronesi (2010)
- **Different qn: how do good people come to terms with doing bad things: Cognitive dissonance, and directed beliefs –** Bichierri et al. 2023, Akerlof and Dickens, 1982, Gennaioli et al., 2022, Vitell et al., 2011, Bem, 1967, Lowell, 2012.
- **Short-term temptation, want-self and should-self:** Thaler (1980), Benabou and Tirole (2011).

The Model

Decision maker undertakes a probably good activity.

$\Pr(\text{harmful}) = q$.

Profit flow π . If harmful, flow of harm $v > \pi$.

Decision maker ethical. Other-regarding preferences, $\alpha \in \mathbf{R}_+$.

Discount rate r .

Decision maker's involvement ends if death/merger. Poisson probability sdt (salvation).

If activity harmful:

- Private signal activity is bad arrives with Poisson probability bdt .
- Then public signal *will* arrive: Poisson probability cdt

Model Decisions

Private bad news signal fully revealing. Public signal coming.

- Can voluntarily stop activity at any time.
 - ▶ But \Rightarrow admission guilt.
 - ▶ Fine = $\phi \times$ [time activity conducted]
- On public bad signal activity *must* stop:
 - ▶ Fine = $\phi \times$ [time activity conducted]

Model comments

- Public signal (e.g. OxContin, VW, smoking, Boeing).
- Fine rate after private and public signal same (ϕ) – front-running.
 - ▶ Punitive punishments explored.
- Punishment rate ϕ is:
 - 1 Legal liabilities (e.g. Elizabeth Holmes & Theranos).
 - 2 Shame (e.g. Madoff). (Dufwenberg and Dufwenberg, 2018, Abeler et al., 2019).
- No defined end-date allows waiting incentive to be cleanly studied.
- Liability after merger – results remain.

The in-too-deep Threshold

Bellman equation for continuation region, bad news received at t_0 , time t after this:

$$F(t; t_0) = \max \left[-\phi(t + t_0), \frac{(\pi - \alpha v)dt - \phi(t + t_0)c dt + s dt \cdot 0}{1 + r dt} (1 - s dt)(1 - c dt) F(t + dt; t_0) \right]$$

The in-too-deep Threshold

Bellman equation for continuation region, bad news received at t_0 , time t after this:

$$F(t; t_0) = \max \left[\begin{array}{l} \text{Fine if confess} \rightarrow -\phi(t + t_0), \\ \text{Profit \& ethics flow} \rightarrow (\pi - \alpha v)dt - \phi(t + t_0)c dt + s dt \cdot 0 \\ \text{Continuation value} \rightarrow + \frac{1}{1+rdt} (1 - s dt)(1 - c dt) F(t + dt; t_0) \end{array} \right]$$

Fine if caught

salvation

The in-too-deep Threshold

Bellman equation for continuation region, bad news received at t_0 , time t after this:

$$F(t; t_0) = \max \left[\begin{array}{l} \text{Fine if confess} \rightarrow -\phi(t + t_0), \\ \text{Profit \& ethics flow} \rightarrow (\pi - \alpha v)dt - \phi(t + t_0)c dt + s dt \cdot 0 \\ \text{Continuation value} \rightarrow + \frac{1}{1+rdt} (1 - s dt)(1 - c dt) F(t + dt; t_0) \end{array} \right]$$

Fine if caught

salvation

Develop analysis

The business evil problem

Theorem: No matter how ethical the decision maker is ($\alpha \in \mathbb{R}_+$) there exists an *in-too-deep threshold* $t_{0*}(\alpha) \geq 0$ such that if the decision maker discovers the activity is harmful after time $t_{0*}(\alpha)$ since business launch then the decision maker will choose to continue with the known harmful activity until he is either saved by merger or death or the authorities receive the public bad-news signal causing cessation to be enforced.

Why is it possible for any ethical person to do this?

- Admitting immediately means fine incurred.
- Delay means knowingly hurting others, and fine will come.
- But salvation might arrive in the form of merger/death.

Even ethicals ultimately find this gamble appealing.

Why is it possible for any ethical person to do this?

- Admitting immediately means fine incurred.
- Delay means knowingly hurting others, and fine will come.
- But salvation might arrive in the form of merger/death.

Even ethicals ultimately find this gamble appealing.

In the words of Bernie Madoff: *“When I began the Ponzi scheme, I believed it would end shortly and I would be able to extricate myself and my clients from the scheme. However, this proved difficult, and ultimately impossible, and as the years went by I realized that my arrest and this day would inevitably come.”*

Extensions

- Analysis of the in-too-deep threshold
- Patents
- Limited liability
- Entry
- Extra punishment for involuntary cessation

Punishments can backfire

Proposition:

If punishments are increased then the in-too-deep threshold:

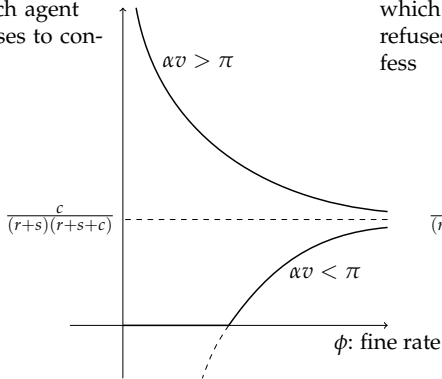
$$\frac{\partial}{\partial \phi} t_{0*} \begin{cases} \leq 0 & \text{if } \alpha v > \pi \text{ (most ethical)} \\ \geq 0 & \text{if } \alpha v < \pi \text{ (least ethical)} \end{cases}$$

Punishment $\uparrow \Rightarrow$ the most ethical agents are trapped into evil sooner.

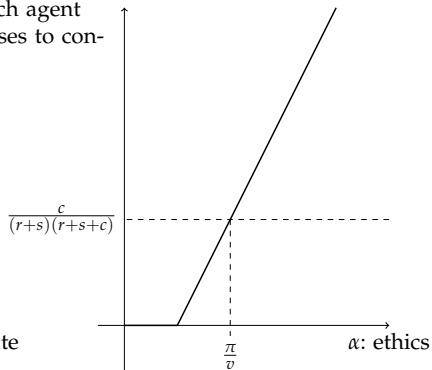
Intuition: punishment only affects part of disutility from continuing.

In-too-deep threshold characteristics

t_{0*} : earliest date after which agent refuses to confess



t_{0*} : earliest date after which agent refuses to confess



Extensions

- Determined end date (e.g. patents)

In-too-deep threshold is always shorter than patent length.

Extensions

- Determined end date (e.g. patents)

In-too-deep threshold is always shorter than patent length.

- Limited liability

Suppose maximum penalty is W then there is a finite in-too-deep threshold if $\pi - \alpha v + (r + s)W > 0$. So applies to all except the poorest and most ethical.

Extensions

- **Determined end date (e.g. patents)**
In-too-deep threshold is always shorter than patent length.
- **Limited liability**
Suppose maximum penalty is W then there is a finite in-too-deep threshold if $\pi - \alpha v + (r + s)W > 0$. So applies to all except the poorest and most ethical.
- **Ethical agents more likely to lose innovation race**
Not due to date of in-too-deep threshold by envelope theorem argument.

Extensions

- **Determined end date (e.g. patents)**
In-too-deep threshold is always shorter than patent length.
- **Limited liability**
Suppose maximum penalty is W then there is a finite in-too-deep threshold if $\pi - \alpha v + (r + s)W > 0$. So applies to all except the poorest and most ethical.
- **Ethical agents more likely to lose innovation race**
Not due to date of in-too-deep threshold by envelope theorem argument.
- **Ethical agents are safest**
Expected probability of harm declines with decision maker ethics. Ethicals need to be more certain product is safe before launching.

Extensions

- **Determined end date (e.g. patents)**
In-too-deep threshold is always shorter than patent length.
- **Limited liability**
Suppose maximum penalty is W then there is a finite in-too-deep threshold if $\pi - \alpha v + (r + s)W > 0$. So applies to all except the poorest and most ethical.
- **Ethical agents more likely to lose innovation race**
Not due to date of in-too-deep threshold by envelope theorem argument.
- **Ethical agents are safest**
Expected probability of harm declines with decision maker ethics. Ethicals need to be more certain product is safe before launching.
- **Extra punishment from forced stop: $(\phi + \phi^{extra})t$**
Extreme extra punishment can time-bound in-too-deep threshold.

Policy Implications and Conclusion

- ① No matter how ethical one is, it is always possible to become trapped into an evil action which causes harm to people.
 - ▶ True for patents & mostly for limited liability.
- ② Unethical agents more likely to win an entry contest – and then more likely to conceal evidence of harm:
 - ▶ AI, self-driving cars ...
- ③ How to choose decision makers with better ethics:
 - ▶ fit-and-proper regime,
 - ▶ ethics training via profession or internally.

Analysis 2

\therefore continuation ode: $(r + s + c)F - \dot{F} = \pi - \alpha v - c\phi(t + t_0)$

Leading to general solution in continuation region:

$$F(t; t_0) = \mathcal{A}e^{(r+s+c)t} + \frac{1}{(r+s+c)} \left[\pi - \alpha v - c\phi(t + t_0) - \frac{c\phi}{r+s+c} \right].$$

Analysis 2

\therefore continuation ode: $(r + s + c)F - \dot{F} = \pi - \alpha v - c\phi(t + t_0)$

Leading to general solution in continuation region:

$$F(t; t_0) = \mathcal{A}e^{(r+s+c)t} + \frac{1}{(r+s+c)} \left[\pi - \alpha v - c\phi(t + t_0) - \frac{c\phi}{r+s+c} \right].$$

General strategy: wait for calendar time τ after bad news received – then confess. Value $\mathcal{F}(t; \tau, t_0)$.

Optimise:

- $\mathcal{F}(t; \tau, t_0)$ in continuation region determined above.
- Constant \mathcal{A} from boundary: $\mathcal{F}(\tau; \tau, t_0) = -\phi(\tau + t_0)$.
- $F(t; t_0) = \sup_{\tau \geq 0} \mathcal{F}(t; \tau, t_0)$

Analysis 3

At bad news arrival, value of waiting for τ to confess:

$$\begin{aligned}\mathcal{F}(0; \tau, t_0) = & -\frac{1}{r+s+c} e^{-(r+s+c)\tau} \left[\pi - \alpha v + (r+s)\phi(t_0 + \tau) - \frac{c\phi}{r+s+c} \right] \\ & + \frac{1}{r+s+c} \left[\pi - \alpha v - c\phi t_0 - \frac{c\phi}{r+s+c} \right]\end{aligned}$$

Quasi-convex in τ .

Analysis 3

At bad news arrival, value of waiting for τ to confess:

$$\mathcal{F}(0; \tau, t_0) = -\frac{1}{r+s+c} e^{-(r+s+c)\tau} \left[\pi - \alpha v + (r+s)\phi(t_0 + \tau) - \frac{c\phi}{r+s+c} \right] \\ + \frac{1}{r+s+c} \left[\pi - \alpha v - c\phi t_0 - \frac{c\phi}{r+s+c} \right]$$

Quasi-convex in τ .

- Define $G := \pi - \alpha v + (r+s)\phi t_0 - \frac{c\phi}{r+s+c}$.
- $G \geq 0 \Rightarrow \sup_{\tau} F(0; \tau, t_0)$ attained as $\tau \rightarrow \infty$. Never confess.
- $G < 0 \Rightarrow \sup_{\tau} F(0; \tau, t_0)$ attained at $\tau = 0$. Confess immediately.

[Return to presentation](#)