

制裁媒介型協調再構成

制度状態、関係損傷および行動バイアスを伴う動学ゲーム

猿星鉄勒

仲地亮二

2026 年 7 月

Sanction-Mediated Reconstruction of Cooperation: A Dynamic Game with Institutional State, Relationship Damage, and Behavioral Biases

要旨

本稿は、継続的關係において、一方の主体が協調から逸脱し、他方の制裁を経た後に、制裁前とは異なる協調状態が形成される過程を定式化する。既存の反復ゲームでは、処罰は主として逸脱後の継続利得を引き下げる装置として扱われる。これに対し、本稿では、制裁を、当期の規律効果と、制度能力および関係損傷を通じた将来状態の変化を同時にもたらす介入として捉える。

完全情報の基準モデルでは、公開状態に制度ストックと関係損傷を含め、制裁主体が制裁強度と制度再構築投資からなるパッケージを選択する。逸脱主体が協調へ復帰するためには、当期の利得差、制裁後の協調価値、および制裁解除の信頼性が同時に十分でなければならない。制裁解除の信頼性は外生的な確率として置かず、制裁主体自身の解除誘因から内生化する。

本稿は、第一に、制裁強度の規律効果が逡減し、関係損傷が凸に増加するという十分条件の下で、協調復帰を実現する制裁強度が下限と上限を持つ有限区間になることを示す。第二に、制裁開始、協調復帰、制裁解除、および制裁後状態の安定性を同時に満たす定常マルコフ完全均衡の存在条件を示す。第三に、制裁主体が選ぶ最適パッケージについて、制裁強度が私的にも社会的にも損傷的である場合、最適制裁は復帰制約を満たす最小有効水準となり、制度再構築が強い制裁を代替することを示す。第四に、ランダム効用を導入することで、理論上の復帰区間を実験で観測される単峰型の復帰確率へ写像する。最後に、現在バイアス、損失回避、および制裁コミットメントに関する不完備情報を拡張として分析し、二段階の実験設計を提示する。

本稿の固有の貢献は、処罰後に元の協調経路へ戻るのではなく、制裁パッケージが制度状態と関係状態を内生的に変え、その新しい状態の下で別の協調均衡が成立することを、状態完備な動学ゲームとして明示する点にある。

キーワード：反復ゲーム、制裁、協調再構成、制度変化、関係損傷、現在バイアス

JEL 分類： C72、C73、D03、D74、D86

Abstract

This paper develops a dynamic model of cooperation reconstructed after a deviation and an intervening sanction episode. Standard repeated-game models typically treat punishment as a device that lowers continuation payoffs following a deviation. In the present model, a sanction package has two distinct effects: it changes the current payoff difference between returning to cooperation and persisting in deviation, and it changes the future state through institutional reconstruction and relationship damage.

The complete-information benchmark uses a payoff-relevant public state that includes institutional capacity and relationship damage. The enforcer chooses a package consisting of sanction intensity and reconstruction investment. The credibility of sanction relief is not imposed as an exogenous probability; it follows from the enforcer's own sequential incentive constraint. Under sufficient conditions in which the disciplinary effect of sanctioning is diminishing while relationship damage is convex, the set of sanction intensities that induce a return to cooperation is a bounded interval with a lower and an upper threshold. The paper then provides conditions for a stationary Markov perfect equilibrium satisfying sanction initiation, return, relief, and

post-sanction stability constraints.

The model further characterizes the enforcer's optimal package. When sanction intensity is costly and destructive, the optimal sanction is the minimum effective intensity consistent with return, while reconstruction investment substitutes for destructive punishment. A random-utility extension maps the deterministic return region into a single-peaked return probability suitable for experimental testing. Behavioral and informational extensions introduce present bias, loss aversion, and Bayesian learning about sanction commitment. Two experimental designs are proposed to identify the return mechanism and the stability mechanism separately.

1 序論

本稿が説明する対象は、次の三点に限定される。

第一に、協調からの逸脱が観察された後、どのような制裁パッケージが逸脱主体を協調へ戻すのか。第二に、なぜ制裁は弱すぎても強すぎても機能しないのか。第三に、制裁後の協調が、なぜ制裁前の協調とは異なる状態として理解されるべきなのか。

本稿は、制裁が常に有効であること、制裁が社会厚生上望ましいこと、または戦争や強制的介入が協調形成の一般的手段であることを主張しない。分析対象は、公開された二者関係における条件付き制裁と協調再構成である。

企業間取引、労使関係、組織内関係、共同資源管理、および国家間関係では、短期的な利益追求によって協調が破られた後、制裁、償い、制度変更、または監視強化を経て、協調が再形成される場合がある。ここで重要なのは、制裁後に成立する関係が、制裁前の関係へ単純に戻るとは限らないことである。制裁は、相手の当期利得を引き下げるだけでなく、制度、能力、信念、信頼、および将来の逸脱可能性を変える。

この違いは、標準的な反復ゲームに対する小さな修正ではない。処罰局面の後に元の協調経路へ戻るモデルでは、協調経路の利得と移行法則はゲーム開始時に与えられている。これに対し、本稿では、制裁主体が選ぶ制裁強度と制度再構築投資が、制裁後の協調状態を内生的に形成する。したがって、制裁は継続利得を一時的に下げだけの装置ではなく、将来ゲームの状態を変える介入である。

本稿では、潜在的な逸脱主体を D 、制裁主体を E と呼ぶ。通常または再構成された協調状態において、 D は協調行動 A または自己利益最大化行動 B を選ぶ。 B を観察した E は、逸脱を受忍するか、制裁パッケージを開始する。制裁パッケージは、制裁強度 c と制度再構築投資 m から構成される。 D はパッケージを観察した後、 A へ復帰するか B を継続する。 D が A へ復帰した場合、 E は制裁を解除するか、制裁を自己目的化して継続する。

本稿の完全情報基準モデルでは、公開状態を単に「通常」「制裁」「再協調」のラベルだけで表さない。制度ストック I と関係損傷 H を状態に含める。制裁パッケージは、制度ストックを増加させる可能性がある一方、関係損傷を増加させる。制裁後の協調価値は、 I から正の影響を受け、 H から負の影響を受ける。このため、制裁後状態は、履歴とパッケージを反映した完全なマルコフ状態となる。

本稿の主結果は五つある。

第一に、制度再構築と関係損傷を含む状態遷移を明示することで、制裁前状態と制裁後状態が一般に異なることを示す。これは単なる名称上の区別ではなく、将来利得と逸脱制

約の違いである。

第二に、協調復帰余剰を定義し、その制裁強度に関する十分条件を示す。当期の規律効果が増加するが逓減し、関係損傷が凸に増加するならば、復帰余剰は厳密な凹関数となり得る。制裁なしと最大制裁の双方で復帰余剰が負であり、内点で正ならば、復帰可能な制裁強度は一意的な下限と上限を持つ有限区間になる。

第三に、制裁解除の信頼性を外生変数として置かない。E が A への復帰後に制裁を解除する誘因を持つ場合にのみ、D は制裁後の協調価値を復帰判断へ織り込む。この解除制約に加え、制裁開始制約、D の復帰制約、および再構成状態の安定制約が同時に成立するとき、定常マルコフ完全均衡が成立する。

第四に、E の最適制裁パッケージを分析する。制裁強度が E 自身にも費用を与え、再構成状態の価値を損なう場合、E は復帰制約を満たす最小有効制裁を選ぶ。制度再構築投資は D の復帰余剰と再構成状態の価値を高めるため、破壊的な制裁強度を代替する。

第五に、決定論的な復帰区間と、実験で観測される復帰確率を区別する。基準モデルが直接示すのは復帰する制裁強度の集合であり、滑らかな逆 U 字型確率ではない。ランダム効用を導入すると、復帰確率は復帰余剰の単調変換となり、同じ内点で最大となる単峰型関数になる。

現在バイアスと損失回避は、基準均衡の存在証明から分離し、行動拡張として導入する。これにより、時間非整合的な主体について独自の均衡概念へ過度に依存することを避ける。現在バイアスは、初期逸脱を増加させるだけでなく、将来の協調価値を軽視させるため復帰を困難にするという二重の作用を持つ。不完備情報拡張では、実際の制裁開始を観察することで、D が E の高コミットメント型に関する信念を更新し、再構成状態での逸脱が抑止されやすくなる。

以下、第 2 節で関連研究と本稿の位置づけを示す。第 3 節で状態、行動順序、利得、および解概念を定義する。第 4 節で協調復帰余剰と制裁区間を導出する。第 5 節で均衡存在と最適パッケージを分析する。第 6 節で確率的選択と行動拡張を扱う。第 7 節で不完備情報と信念更新を導入する。第 8 節で実験設計を提示する。第 9 節で適用範囲と限界を述べ、第 10 節で結論を示す。

2 関連研究と本稿の位置づけ

2.1 反復ゲームと有限処罰

反復ゲームでは、現在の逸脱が将来の相手行動を変えるため、一回ゲームでは実現できない協調を均衡として支持できる。Fudenberg and Maskin (1986) はフォーク定理を体系化し、Abreu (1986, 1988) は、厳しい処罰と協調経路への復帰を組み合わせる最適処罰規則を分析した。Mailath and Samuelson (2006) は、反復関係における評判と長期的誘因を包括的に整理している。

本稿の制裁経路は、外形的には Abreu 型のスティック・アンド・キャロット構造に近い。しかし、本稿では処罰後の協調経路を外生的に与えない。制裁主体が選んだパッケージが制度ストックと関係損傷を変え、その結果として新しい協調利得と逸脱制約が形成される。この点で、処罰は元の均衡を支持するための一時的な脅威ではなく、将来のゲームそのものを変える。

ノイズのある反復囚人のジレンマでは、永久的な報復よりも寛容な戦略が高い成果をも

たらし得る。Fudenberg, Rand, and Dreber (2012) は、協調均衡が存在する環境で、最初の逸脱に直ちに過剰反応せず、処罰後に協調へ戻ろうとする戦略が成功していることを示した。本稿は、赦しの必要性を E の解除制約として形式化し、解除後の状態が内生的に変化する点を追加する。

2.2 費用を伴う処罰、報酬、および関係修復

Fehr and Gächter (2000) は、公共財ゲームにおいて、自ら費用を負担する処罰が非協力を抑え、協調を高め得ることを示した。Andreoni, Harbaugh, and Vesterlund (2003) は、報酬と処罰の組合せを実験的に分析した。Kosfeld and Rustagi (2015) は、同じ処罰権限を持つ指導者でも、効率と公平を重視する処罰と無差別な処罰では、協調成果が異なることを示している。

これらの研究は、処罰の存在だけでなく、その設計と処罰主体の型が重要であることを示す。本稿は、処罰の設計を制裁強度 c と制度再構築 m へ分解し、強度を高めることと、将来の協調可能性を高めることを区別する。

Bottom et al. (2002) は、関係が破られた後の協調再建について、単なる説明や謝罪より、実質的な償いが有効であることを示した。本稿の制度再構築投資 m は、この種の実質的変更を縮約的に表す。ただし、 m は道徳的謝罪そのものではなく、監視、補償、担保、権限配分、および自動執行を含む、将来利得と逸脱可能性を変える制度的介入である。

2.3 関係的契約と関係価値

Levin (2003) は、法的に完全には強制できない約束が、将来の関係価値によって自己執行される関係的契約を分析した。Halac (2012) は、関係価値に関する持続的な私的情報の下で、関係の発展と情報開示を分析している。

本稿は、関係価値を所与の継続余剰として扱わず、制裁によって変化する状態変数の関数とする。制度再構築は関係価値を高め得るが、過大制裁は関係価値を低下させる。したがって、制裁は自己執行制約を強める場合と弱める場合の双方を持つ。

2.4 確率ゲームと内生的状態

Shapley (1953) 以来、確率ゲームは、現在の行動が将来のゲーム状態を変える動学的戦略関係を分析してきた。Hilbe et al. (2018) は、協調と非協調が将来のゲーム状態を変える環境で、状態遷移が協調の進化を促進し得ることを示した。

本稿も状態依存性を用いるが、状態遷移を単純な外生確率としない。制裁主体が選ぶ c と m が、制度ストックと関係損傷を決定する。さらに、状態を制度ストック I と関係損傷 H まで含めて定義するため、異なるパッケージを経た再構成状態を同一の状態として扱わない。

2.5 現在バイアス、損失回避、および互惠性

Laibson (1997) および O'Donoghue and Rabin (1999) は、準双曲割引による現在バイアスと、洗練された主体とナイーブな主体の違いを分析した。Kahneman and Tversky (1979) および Köszegi and Rabin (2006) は、参照点依存と損失回避を理論化した。Rabin (1993) および Fehr and Schmidt (1999) は、公平性、互惠性、および不平等回避を戦略的相互作用へ導入した。

本稿はこれらを基準均衡へ同時に組み込まない。完全情報かつ指数割引の基準モデルで

状態転換と制裁区間を確立した後、現在バイアスと損失回避が復帰制約をどの方向へ変えるかを比較静学として示す。この構成により、本稿固有の状態転換メカニズムと、行動選好の効果を識別する。

2.6 本稿の貢献

本稿の貢献は、次の論理鎖を一つの状態完備な動学ゲームとして提示する点にある。

$$B \longrightarrow z = (c, m) \longrightarrow (I', H') \longrightarrow R(I', H')$$

第一に、制裁後の状態を、制度ストックと関係損傷によって内生的に定義する。

第二に、当期の規律効果と将来の関係損傷を分離し、復帰可能な制裁区間を十分条件から導出する。

第三に、制裁解除の信頼性を E の逐次合理性から内生化する。

第四に、復帰区間、均衡存在、最適パッケージ、および確率の実証予測を明確に区別する。

第五に、現在バイアスと信念更新が、状態転換メカニズムへどのように作用するかを拡張として示す。

3 基準モデル

3.1 プレイヤー、時間、および状態

プレイヤーは、潜在的な逸脱主体 D と制裁主体 E である。時間は離散的で無限に続く。

$$t = 0, 1, 2, \dots$$

プレイヤー i の割引因子を次のように置く。

$$0 < \delta_i < 1, \quad i \in \{D, E\}$$

各期の公開状態は、次の組で表される。

$$x = (s, I, H, z)$$

s は関係局面、I は制度ストック、H は関係損傷、z は制裁節で公開されている制裁パッケージである。局面集合は次のとおりである。

$$s \in \{N, P, R, O\}$$

状態 N は制裁前の協調、状態 P は制裁節、状態 R は再構成された協調、状態 O は関係崩壊後の外部選択肢を表す。状態 O は吸収状態とする。状態 P 以外では z は空集合である。

制度ストックと関係損傷は、次の集合に属する。

$$I \in [0, \bar{I}], \quad H \in [0, \bar{H}]$$

初期状態は次のとおりである。

$$x_0 = (N, I_0, H_0, \emptyset)$$

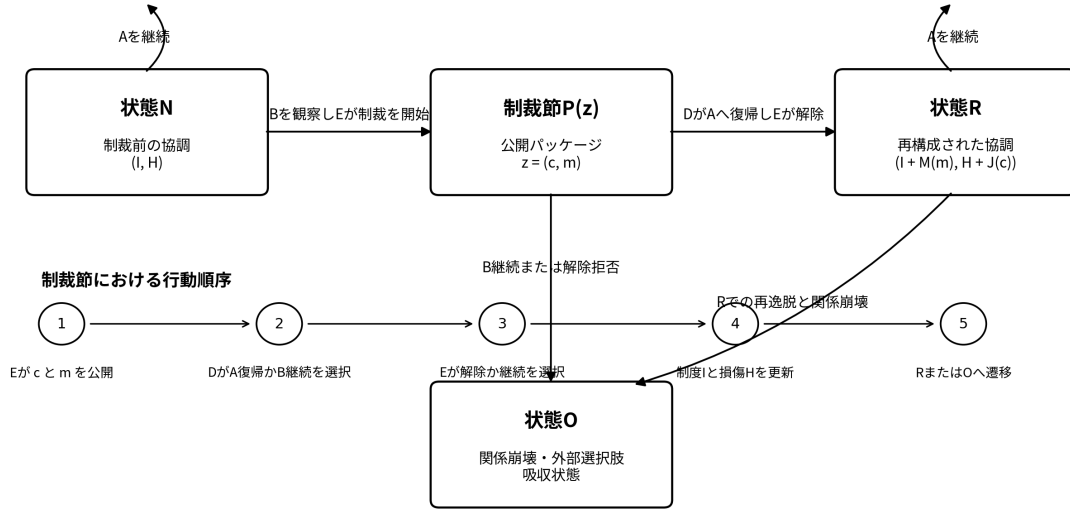


図1 完全な公開状態と制裁節の行動順序

3.2 通常局面と再構成局面

状態 N または状態 R において、D は協調行動 A または自己利益最大化行動 B を選ぶ。

$$a_D \in \{A, B\}$$

D が A を選んだ場合、両者の一期間利得は次のとおりである。

$$r_i(I, H), \quad i \in \{D, E\}$$

利得関数は連続微分可能であり、次を満たす。

$$r_{iI} > 0, \quad r_{iH} < 0$$

制度ストックは協調価値を高め、関係損傷は協調価値を低下させる。

D が B を選んだ場合、D は追加的な当期利益を得る。

$$\Delta_s(I, H) > 0, \quad s \in \{N, R\}$$

このとき、E は逸脱を受忍する行動 T または制裁開始 S を選ぶ。T を選ぶと、次期は状態 O へ移るものとする。これは、E が制裁を行わず、協調契約が崩壊する場合の縮約形である。T の後も関係が継続する一般化は可能であるが、主結果には不要である。

3.3 制裁パッケージと状態遷移

E が S を選ぶ場合、次の制裁パッケージを公開する。

$$z = (c, m) \in \mathcal{Z}$$

ここで、制裁強度 c と制度再構築投資 m の選択集合を次のように置く。

$$\mathcal{Z} = [0, \bar{c}] \times [0, \bar{m}]$$

制裁節 P では、D がパッケージを観察した後、A への復帰または B の継続を選ぶ。D が B を継続した場合、次期は O へ移る。D が A へ復帰した場合、E は制裁解除 L または解除拒否 C を選ぶ。L なら次期は R、C なら次期は O へ移る。

制裁パッケージによる状態更新を次のように定義する。

$$I' = \Gamma_I(I, m) = \min\{\bar{I}, I + M(m)\}$$

$$H' = \Gamma_H(H, c) = \min\{\bar{H}, H + J(c)\}$$

M と J は次を満たす。

$$M(0) = 0, \quad M' > 0, \quad M'' \leq 0$$

$$J(0) = 0, \quad J' > 0, \quad J'' \geq 0$$

制度投資の限界効果は一定または逓減し、関係損傷の限界増加は一定または逓増する。

D が A へ復帰し、E が L を選んだ場合、次期状態は次のとおりである。

$$x' = (R, I', H', \emptyset)$$

したがって、異なる c と m を経た状態 R は、異なる I' と H' を持つ別のマルコフ状態である。

命題 1 制裁後状態の非同一性

少なくとも一方の条件が成立するとする。

$$M(m) > 0$$

$$J(c) > 0$$

さらに、対応する状態変数が少なくとも一方のプレイヤーの利得または逸脱利益へ厳密

に影響するとする。このとき、制裁後状態は制裁前状態と同一ではない。

$$(R, I', H', \emptyset) \neq (N, I, H, \emptyset)$$

証明。 M または J が正ならば、I' または H' が初期値と異なる。対応する状態変数が利得または逸脱利益へ厳密に影響するため、将来の利得関数または逸脱制約が異なる。したがって、両状態は利得関連状態として同一ではない。

命題 1 は、本稿の「再構成」が単なる用語上の区別ではないことを示す。状態 R は、制裁履歴を十分統計量として含む I' と H' によって特徴づけられる。

3.4 制裁節の利得

制裁節における D の当期利得を、A への復帰と B の継続についてそれぞれ次のように置く。

$$p_D^A(c; I, H), \quad p_D^B(c; I, H)$$

当期の規律効果を次のように定義する。

$$g(c; I, H) = p_D^A(c; I, H) - p_D^B(c; I, H)$$

制裁強度は、B を継続する選択を相対的に不利にすると仮定する。

$$g_c > 0$$

E が制裁パッケージを開始する費用を $K(c, m)$ とする。

$$K_c > 0, \quad K_m > 0$$

D が B を継続した場合の状態 O における継続価値を、プレイヤー i について次のように置く。

$$\underline{V}_i$$

状態 O が一定利得 o_i を与える場合、次が成立する。

$$\underline{V}_i = \frac{o_i}{1 - \delta_i}$$

3.5 状態 R の価値

状態 R で A が継続される場合の価値は次のとおりである。

$$V_i^R(I, H) = \frac{r_i(I, H)}{1 - \delta_i}$$

この式は、状態 R で状態変数が追加的な逸脱のない限り固定される基準モデルを表す。制度の減耗または関係損傷の回復を導入する場合には、I と H の遷移式を追加し、 V_i^R を対応するベルマン方程式の解へ置き換えればよい。

3.6 解概念

基準モデルでは、完全情報、完全観測、および指数割引を仮定する。分析対象は、公開状態 x のみに依存する定常戦略である。解概念として定常マルコフ完全均衡を用いる。

本稿の均衡存在定理は、現在バイアスを含まない基準モデルについて示す。現在バイアスと損失回避は第 6 節で局所的な選択条件を変更する行動拡張として扱う。この分離により、時間非整合的な主体に独自の均衡概念を課す必要を避ける。

4 制裁節における協調復帰

4.1 制裁解除の内生的信頼性

D が A へ復帰した後、E が L を選ぶ価値は次のとおりである。

$$W_E^L(c, m; I, H) = \delta_E V_E^R(I', H')$$

E が解除を拒否し、制裁または対立を継続する場合の私的便益を $w(c, m)$ とする。解除拒否後は状態 O へ移るとすれば、その価値は次のとおりである。

$$W_E^C(c, m; I, H) = w(c, m) + \delta_E \underline{V}_E$$

制裁解除が逐次合理的である条件は次のとおりである。

$$\delta_E V_E^R(I', H') \geq w(c, m) + \delta_E \underline{V}_E \quad (1)$$

式 (1) を解除制約と呼ぶ。解除制約を満たすパッケージの集合を次のように定義する。

$$\mathcal{Z}_L(I, H) = \{(c, m) \in \mathcal{Z} : \delta_E V_E^R(I', H') \geq w(c, m) + \delta_E \underline{V}_E\}$$

完全情報の純粋戦略基準モデルでは、解除は信頼されるか、信頼されないかのいずれかである。したがって、解除確率を 0 と 1 の間の外生パラメータとして置かない。中間的な解除確率は、第 7 節の不完備情報または混合戦略拡張でのみ生じる。

解除制約が成立しなければ、D は A へ復帰しても再構成状態へ移れないことを正しく予想する。この場合、制裁後の協調価値は D の復帰判断へ寄与しない。制裁が協調回復の装置として機能するには、B へ制裁する約束だけでなく、A へ戻った後に解除する約束も自己执行的でなければならない。

4.2 協調復帰余剰

以下では、パッケージが解除制約を満たす場合を考える。状態 P で D が A へ復帰する価値と、B を継続して状態 O へ移る価値の差を、協調復帰余剰と呼ぶ。

$$\Phi(c, m; I, H) = g(c; I, H) + \delta_D [V_D^R(I', H') - \underline{V}_D] \quad (2)$$

D は次の条件で A へ復帰する。

$$\Phi(c, m; I, H) \geq 0 \quad (3)$$

式 (2) の第 1 項は当期の規律効果である。第 2 項は、協調へ復帰し制裁が解除された場合と、逸脱を継続して関係が崩壊した場合の将来価値差である。

制裁強度 c は、第 1 項を高める一方、関係損傷 $J(c)$ を通じて第 2 項を低下させる。制度再構築 m は制度ストック $M(m)$ を通じて第 2 項を高める。この二つの経路を分離することが、本稿の制裁区間結果の基礎である。

4.3 復帰余剰の凹性に関する十分条件

復帰余剰が制裁強度に関して厳密な凹関数となるための十分条件を示す。

仮定 A1. g は c に関して二階連続微分可能であり、次を満たす。

$$g_c > 0, \quad g_{cc} \leq 0$$

仮定 A2. 状態 R における D の価値は H に関して減少し、弱く凹である。

$$V_{DH}^R < 0, \quad V_{DHH}^R \leq 0$$

仮定 A3. 関係損傷関数は増加かつ弱く凸である。

$$J' > 0, \quad J'' \geq 0$$

式 (2) を c で二階微分すると、次を得る。

$$\Phi_{cc} = g_{cc} + \delta_D [V_{DHH}^R (J')^2 + V_{DH}^R J''] \quad (4)$$

仮定 A1 から A3 の下では、式 (4) の各項は非正である。少なくとも一つが関連領域で厳密に負ならば、次が成立する。

$$\Phi_{cc} < 0 \quad (5)$$

この条件は、制裁強度の限界的な規律効果が逡減し、関係損傷の限界費用が増加するという経済的仮定から復帰余剰の厳密な凹性が導かれることを示す。厳密な凹性そのものを「一般法則」として仮定するのではなく、明示された十分条件の下で用いる。

4.4 制裁復帰区間

制裁強度の選択集合をコンパクトな区間とする。

$$c \in [0, \bar{c}]$$

さらに、次を仮定する。

仮定 A4. 制裁なしでは D は復帰しない。

$$\Phi(0, m; I, H) < 0$$

仮定 A5。最大制裁では関係損傷が規律効果を上回る。

$$\Phi(\bar{c}, m; I, H) < 0$$

仮定 A6。ある内点で復帰余剰が正である。

$$\max_{c \in [0, \bar{c}]} \Phi(c, m; I, H) > 0$$

命題 2 制裁復帰区間

仮定 A1 から A6 が成立し、式 (5) が関連領域で成立するならば、一意な二つの閾値が存在する。

$$0 < c_L(m; I, H) < c_H(m; I, H) < \bar{c}$$

D が A へ復帰するための必要十分条件は次のとおりである。

$$\Phi(c, m; I, H) \geq 0 \iff c \in [c_L, c_H] \quad (6)$$

証明。付録 A.1 を参照。

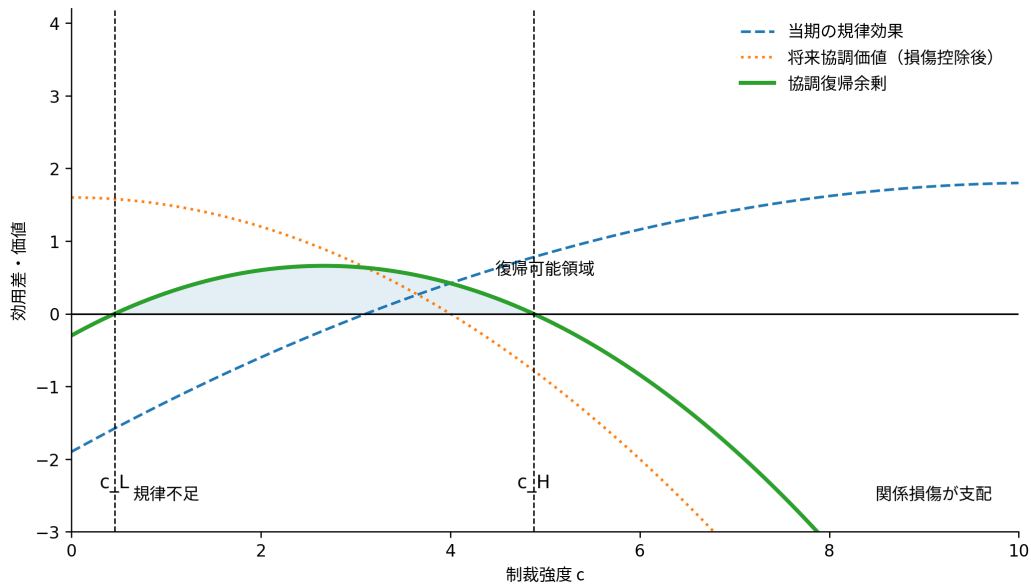


図 2 当期の規律効果、将来の関係損傷、および協調復帰余剰

命題 2 が示すのは、決定論的モデルにおける復帰可能な制裁強度の集合である。制裁が弱すぎる場合、当期の規律効果が不足する。中間領域では、B を継続する当期損失と状態 R の将来価値が復帰を支持する。制裁が過大になると、関係損傷によって状態 R の価値が低下し、D は協調へ戻るより関係崩壊を選ぶ。

この結果は「制裁が強いほど協調が回復する」という単調な見方を否定する。同時に、「制裁は常に逆効果である」という見方も否定する。復帰を支持するのは、状態と制度投資に依存する有限区間である。

4.5 制度再構築の比較静学

制度再構築 m は、状態 R の制度ストックを通じて D の価値を高める。上限制約が結合していない領域では、次が成立する。

$$\Phi_m = \delta_D V_{DI}^R M'(m) > 0 \quad (7)$$

端点は次を満たす。

$$\Phi(c_j, m; I, H) = 0, \quad j \in \{L, H\}$$

陰関数定理により、次を得る。

$$\frac{\partial c_j}{\partial m} = -\frac{\Phi_m}{\Phi_c} \quad (8)$$

厳密な凹性と命題 2 から、下限では正の傾き、上限では負の傾きとなる。

$$\Phi_c(c_L, m) > 0, \quad \Phi_c(c_H, m) < 0$$

したがって、次が成立する。

系 1 制度再構築による復帰区間の拡大

$$\frac{\partial c_L}{\partial m} < 0, \quad \frac{\partial c_H}{\partial m} > 0 \quad (9)$$

制度再構築は、より弱い制裁で復帰を可能にし、同時に関係損傷を吸収して許容可能な上限を引き上げる。したがって、 m は単なる「制裁後の報酬」ではない。制裁の破壊性を抑えながら、再構成状態の自己執行可能性を高める状態投資である。

4.6 関係損傷感応度

関係損傷の強さを表すパラメータを η とし、次のように書く。

$$H' = H + \eta J(c)$$

状態 R の価値が H に関して低下するため、次が成立する。

$$\Phi_\eta = \delta_D V_{DH}^R J(c) < 0 \quad (10)$$

η が上昇すると復帰余剰全体が低下し、復帰区間は縮小する。特に、制裁が物的能力、信頼、または相手の存立可能性を大きく損なう環境では、同じ当期の規律効果を持つ制裁

でも協調再構成に失敗しやすい。

5 均衡存在と最適制裁パッケージ

5.1 制裁開始制約

D の B を観察した時点で、E が逸脱を受忍した場合の継続価値を次のように置く。

$$V_E^T(I, H)$$

E がパッケージ z を開始し、D が A へ復帰し、E が L を選ぶ場合の価値は次のとおりである。

$$W_E^S(c, m; I, H) = -K(c, m) + \delta_E V_E^R(I', H') \quad (11)$$

制裁開始制約は次のとおりである。

$$W_E^S(c, m; I, H) \geq V_E^T(I, H) \quad (12)$$

式 (12) は、制裁が D にとって有効であるだけでは不十分であり、E にとっても費用を負担して開始する価値がなければならないことを示す。

5.2 再構成状態の安定制約

状態 R で D が B を選んだ場合、E は均衡で指定された制裁パッケージ z_R を開始する。逸脱後の制裁節から得る D の均衡価値を次のように置く。

$$W_D^P(I, H; z_R)$$

状態 R で A を選ぶ価値は次のとおりである。

$$r_D(I, H) + \delta_D V_D^R(I, H)$$

B を選ぶ価値は次のとおりである。

$$r_D(I, H) + \Delta_R(I, H) + \delta_D W_D^P(I, H; z_R)$$

したがって、再逸脱を防ぐ条件は次のとおりである。

$$\Delta_R(I, H) \leq \delta_D [V_D^R(I, H) - W_D^P(I, H; z_R)] \quad (13)$$

制度再構築は、左辺の逸脱利益を監視、担保、補償、または自動執行によって低下させる可能性がある。また、状態 R の協調価値を高めることで右辺を増加させる可能性がある。

5.3 実行可能パッケージ

状態 $x = (s, I, H, \emptyset)$ において、次の集合を定義する。

D の復帰制約を満たす集合は次のとおりである。

$$\mathcal{Z}_D(I, H) = \{z \in \mathcal{Z} : \Phi(z; I, H) \geq 0\}$$

E の解除制約を満たす集合は、前節で定義した $\mathcal{Z}_L(I, H)$ である。

E の開始制約を満たす集合は次のとおりである。

$$\mathcal{Z}_S(I, H) = \{z \in \mathcal{Z} : W_E^S(z; I, H) \geq V_E^T(I, H)\}$$

状態 R の安定制約を満たす集合を次のように置く。

$$\mathcal{Z}_R(I, H)$$

実行可能集合は、これらの共通部分である。

$$\mathcal{Z}^*(I, H) = \mathcal{Z}_D(I, H) \cap \mathcal{Z}_L(I, H) \cap \mathcal{Z}_S(I, H) \cap \mathcal{Z}_R(I', H') \quad (14)$$

定理 1 制裁媒介型協調再構成均衡

各関連状態で実行可能集合が空でなく、E が選ぶ定常パッケージ z^* がその集合に属するとする。このとき、次の定常戦略プロファイルを含むマルコフ完全均衡が存在する。

1. 状態 N および状態 R で D は A を選ぶ。
2. D が B を選んだ場合、E は z^* を開始する。
3. 制裁節で D は A へ復帰する。
4. D が A へ復帰した後、E は L を選ぶ。
5. 制裁解除後、両者は更新された状態 R で A を継続する。
6. D が B を継続するか E が解除を拒否した場合、状態 O へ移る。

証明。付録 A.2 を参照。

定理 1 の中心は、単一の不等式ではなく、四つの制約の同時成立である。制裁が D を動かしても、E が開始しなければ経路は成立しない。E が開始しても、A への復帰後に解除しなければ D は将来協調価値を評価しない。解除されても、状態 R で再逸脱が有利なら協調は持続しない。

5.4 E の最適パッケージ

E は、実行可能集合の中から、自らの制裁開始価値を最大化するパッケージを選ぶ。

$$z^*(I, H) \in \arg \max_{(c, m) \in \mathcal{Z}^*(I, H)} \{-K(c, m) + \delta_E V_E^R(I', H')\} \quad (15)$$

この問題は、制裁の強さと制度再構築の配分問題である。c は D の当期行動を変えるが、

K と J を通じて E 自身の価値も低下させる。m は費用を伴うが、制度ストックを高める。

命題 3 最小有効制裁

m を固定し、実行可能な制裁強度の集合が非空の閉区間であるとする。さらに、E の目的関数が実行可能領域で c に関して厳密に減少するとする。

$$-K_c(c, m) + \delta_E V_{EH}^R(I', H') J'(c) < 0 \quad (16)$$

このとき、E が選ぶ最適制裁強度は、すべての制約を満たす最小の制裁強度である。

$$c^*(m) = \min\{c : (c, m) \in \mathcal{Z}^*(I, H)\} \quad (17)$$

証明。付録 A.3 を参照。

命題 3 は、「最大限の制裁」が均衡設計上の最適解ではないことを示す。制裁強度が E にも費用を与え、状態 R の価値を損なうならば、E は復帰を実現する最小水準を選ぶ。

5.5 制度再構築による代替

m が増加すると D の復帰区間が拡大するため、最小有効制裁 $c^*(m)$ は低下し得る。内部解で復帰制約が結合している場合、次が成立する。

$$\Phi(c^*(m), m; I, H) = 0$$

したがって、次を得る。

$$\frac{dc^*}{dm} = -\frac{\Phi_m}{\Phi_c} < 0 \quad (18)$$

ここで、最小有効制裁は下側の根にあるため、 $\Phi_c > 0$ である。

式 (18) は、制度再構築が制裁強度を代替することを示す。制裁と再構築を一つの総量として扱うのではなく、規律と将来状態への異なる作用を持つ二つの政策変数として分離する必要がある。

5.6 制裁の自己目的化

$w(c, m)$ が大きい場合、解除制約は成立しない。例えば、制裁継続が E に政治的支持、資源移転、復讐効用、または組織維持利益を与える場合である。

解除制約の余剰を次のように定義する。

$$\Lambda_E(c, m; I, H) = \delta_E V_E^R(I', H') - w(c, m) - \delta_E V_E \quad (19)$$

$$\Lambda_E < 0$$

であれば、D が A へ復帰しても E は解除しない。この場合、制裁は条件付き協調回復の

装置ではなく、相手を恒常的に低利得状態へ置く装置となる。Dがこの誘因を予想するなら、協調復帰経路は均衡として成立しない。

6 確率的予測と行動拡張

6.1 復帰区間と復帰確率の区別

命題 2 が示すのは、決定論的選択の下で A が最適となる制裁強度の区間である。実験データで観測される滑らかな復帰確率を得るには、選好異質性、判断誤差、または一時的利得ショックを導入する必要がある。

D の復帰判断に、独立なランダム効用ショック ε を導入する。

$$A \iff \Phi(c, m; I, H) + \varepsilon \geq 0$$

ε の累積分布関数を F_ε とし、連続かつ厳密に増加するとする。復帰確率は次のとおりである。

$$P_A(c, m; I, H) = F_\varepsilon(\Phi(c, m; I, H)) \quad (20)$$

命題 4 単峰型の復帰確率

Φ が c に関して厳密な凹関数で、一意な最大点 $c^\dagger$ を持つとする。 F_ε が厳密に増加するならば、復帰確率 P_A も同じ $c^\dagger$ で一意に最大となる。

証明。 F_ε は厳密に増加するため、 $P_A(c_1) > P_A(c_2)$ と $\Phi(c_1) > \Phi(c_2)$ は同値である。したがって、両関数は同じ最大点を持つ。

命題 4 により、理論上の復帰区間と、実験上の単峰型または逆 U 字型の平均復帰率を整合的に結びつけることができる。ただし、復帰確率自体が厳密な凹関数になるとは限らない。実証分析では、二次項の符号だけでなく、処置水準別の平均と事前登録された対比を用いるべきである。

6.2 現在バイアス

D に準双曲割引を導入する。各時点の D は当期利得に重み 1、次期以降の利得に重み $\beta_D \delta_D^k$ を置く。

$$0 < \beta_D \leq 1$$

基準モデルの復帰余剰は次のように変化する。

$$\Phi_\beta(c, m; I, H) = g(c; I, H) + \beta_D \delta_D [V_D^R(I', H') - \underline{V}_D] \quad (21)$$

状態 R の価値が状態 O の価値を上回るならば、次が成立する。

$$\frac{\partial \Phi_\beta}{\partial \beta_D} = \delta_D [V_D^R(I', H') - \underline{V}_D] > 0 \quad (22)$$

現在バイアスが強い、すなわち β_D が低い主体ほど、将来の協調価値を軽視するため A へ復帰しにくい。

一方、状態 N または R で B を選ぶ当期利益を Δ_s 、B によって失う将来価値を L_s とすると、D の逸脱条件は次のとおりである。

$$\Delta_s > \beta_D \delta_D L_s \quad (23)$$

β_D の低下は式 (23) を満たしやすくする。したがって、現在バイアスには二重の作用がある。

1. 即時の逸脱利益を相対的に重くし、B を増加させる。
2. A へ戻った後の将来価値を軽くし、復帰を困難にする。

系 2 現在バイアスと復帰区間

命題 2 の端点が内部にあり、式 (22) が成立するならば、 β_D の上昇は下限を低下させ、上限を上昇させる。

$$\frac{\partial c_L}{\partial \beta_D} < 0, \quad \frac{\partial c_H}{\partial \beta_D} > 0 \quad (24)$$

逆に、現在バイアスが強い主体に対しては、将来の協調便益だけを示す政策は弱い。A への復帰と同時に当期または直近の制裁負担を軽減する、前倒しされた出口が重要となる。

6.3 損失回避

損失回避の効果は、A への復帰が当期の損失をどの程度軽減するか依存する。参照点を ρ とし、制裁節の両利得が参照点を下回る場合を考える。損失領域の傾きを $\lambda > 1$ とする。

物質的利得差を次のように置く。

$$\tilde{g}(c) = y_D^A(c) - y_D^B(c)$$

両選択が損失領域にある場合、行動効用上の当期差は次のとおりである。

$$g(c) = \lambda \tilde{g}(c) \quad (25)$$

したがって、 $\tilde{g}(c) > 0$ の領域では損失回避が A への復帰を促進するが、 $\tilde{g}(c) < 0$ の領域では B の継続を強める。

この結果は、損失回避が制裁の効果を一律に高めるわけではないことを示す。A へ戻っても制裁が軽減されない、または A と B が同程度に罰せられる設計では、損失回避は協調復帰を支えない。

7 不完備情報と制裁コミットメント

7.1 制裁開始能力に関する型

E には高コミットメント型 H と低コミットメント型 L が存在するとする。D の事前信念を次のように置く。

$$q = \Pr(\theta = H)$$

B を観察した後、高コミットメント型が制裁を開始する確率を ρ_H 、低コミットメント型の開始確率を ρ_L とする。

$$1 \geq \rho_H > \rho_L \geq 0$$

制裁開始 S を観察した後の事後信念は、ベイズ則により次のようになる。

$$q' = \frac{q\rho_H}{q\rho_H + (1-q)\rho_L} \quad (26)$$

命題 5 制裁実績によるコミットメント信念の上昇

$0 < q < 1$ かつ $\rho_H > \rho_L$ ならば、次が成立する。

$$q' > q \quad (27)$$

証明。 付録 A.4 を参照。

将来の制裁開始確率は次のとおりである。

$$\pi(q) = q\rho_H + (1-q)\rho_L \quad (28)$$

命題 5 から、次を得る。

$$\pi(q') > \pi(q) \quad (29)$$

状態 R の逸脱制約が、予想制裁確率を含む次の形であるとする。

$$\Delta_R(I, H) \leq \delta_D \pi(q) [V_D^R(I, H) - W_D^P(I, H; z_R)] \quad (30)$$

制裁実績による信念更新は、式 (30) の右辺を高める。このため、同じ制度状態と同じ制裁技術であっても、実際に制裁を実行した履歴を持つ状態 R の方が、再逸脱を抑止しやすい。

7.2 解除コミットメント

制裁開始へのコミットメントと、A への復帰後に解除するコミットメントは異なる型特性である。強硬型が開始には積極的だが解除には消極的である場合、制裁実績は開始信頼性

を高める一方、出口信頼性を低下させる可能性がある。

このため、不完備情報モデルを実証する際には、次の二つの信念を分離して測定する必要がある。

$$q_S = \Pr(S \mid B)$$

$$q_L = \Pr(L \mid A, S)$$

高い q_S だけでは協調再構成に十分ではない。復帰経路には高い q_L も必要である。制裁主体の一貫性とは、「B には制裁する」とことと「A には解除する」とことの組合せである。

8 実験による識別

本稿の実験目的は、すべての心理変数を一つの実験で同時に識別することではない。協調復帰を決める制裁節のメカニズムと、状態 R の安定性を決めるメカニズムを二つの実験へ分離する。

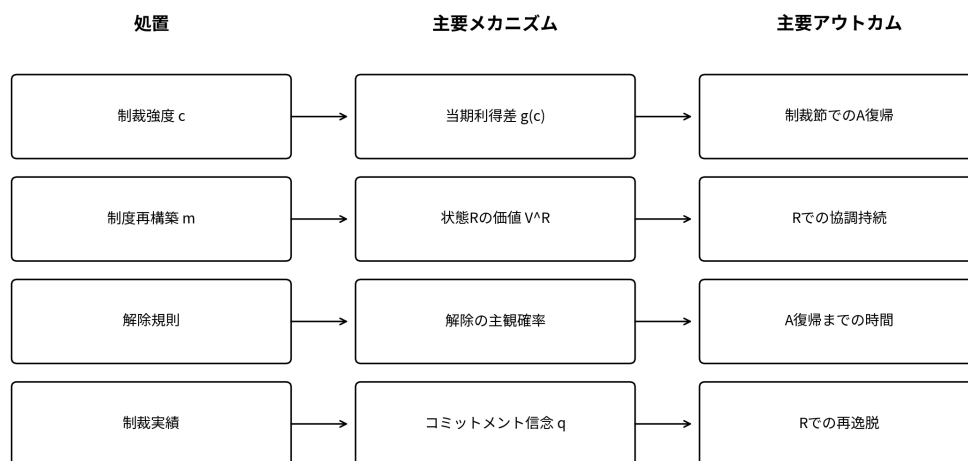


図 3 理論変数、処置、および主要アウトカムの対応

8.1 実験 1：制裁節からの復帰

固定された二人一組に、D と E の役割を割り当てる。通常局面で逸脱が発生した後、E が選ぶ制裁パッケージを実験者が処置として固定する。主要処置は次の三つである。

処置	水準	理論上の対象
制裁強度 c	低・中・高・過大	当期の規律効果と関係損傷
制度再構築 m	無・有	状態 R の制度価値
解除規則	信頼可能・不信頼	E の解除制約に対応する出口

制裁強度処置では、B 継続時の当期利得を段階的に低下させる。同時に、過大制裁群で

は制裁後の双方の協調利得も低下させる。これにより、当期の規律効果と将来の関係損傷を別の画面と利得表で明示する。

制度再構築群では、状態 R での逸脱利益を自動的に減らす担保または補償制度を導入する。制度なし群では、同じ制裁後も当初の逸脱利益が残る。

信頼可能な解除群では、D が A を選ぶと制裁が機械的に解除される。不信頼群では、A を選んでも一定確率で制裁が継続する。この操作は、完全情報基準モデルの解除制約が成立する状態と、不完備情報下で出口信頼性が低い状態を比較する。

主要従属変数は、制裁パッケージ提示後の最初の選択で D が A へ復帰するかどうかである。副次従属変数は、復帰までのラウンド数とする。

主要な主観測定は一つに限定する。D に、A を選んだ場合に次期に制裁が解除される確率を、適切なスコアリング規則を用いて回答させる。この主観確率を解除処置の主要媒介変数とする。

実験 1 の事前登録仮説は次のとおりである。

1. 制裁強度と復帰率の関係は単峰型である。
2. 制度再構築は各制裁強度における復帰率を高める。
3. 信頼可能な解除規則は復帰率を高め、その効果は主観的解除確率によって媒介される。

8.2 実験 2：再構成状態の安定性

実験 2 では、参加者を状態 R から開始させる。制度ストックと関係損傷を固定し、制裁主体の過去履歴だけを操作する。

一方の群では、E が過去の逸脱に対して費用を負担して制裁を開始した履歴を示す。他方の群では、同じ利得状態にあるが、制裁実績を示さない。

主要な主観測定は、B を選んだ場合に E が制裁を開始する確率である。主要従属変数は、D が状態 R で B へ再逸脱するかどうかである。

実験 2 の事前登録仮説は次のとおりである。

1. 制裁実績は主観的な制裁開始確率を高める。
2. 主観的な制裁開始確率は再逸脱率を低下させる。
3. 制裁実績の再逸脱への効果は、制裁開始信念によって媒介される。

8.3 推定方法

実験 1 の主要推定は、処置水準ごとの平均復帰率と事前登録された対比によって行う。ランダム効用拡張に対応する補助推定として、次のロジットモデルを用いる。

$$\Pr(A_i = 1) = \Lambda(\theta_0 + \theta_1 c_i + \theta_2 c_i^2 + \theta_3 m_i + \theta_4 L_i + \theta_5 c_i m_i) \quad (31)$$

L_i は信頼可能な解除処置を表す。単峰型仮説については、 $\theta_1 > 0$ および $\theta_2 < 0$ だけに依存せず、中間制裁群と低・過大制裁群の事前登録対比を主要検定とする。

実験 2 では、制裁実績処置による再逸脱率と主観的制裁確率の差を推定する。媒介分析は各実験で一つの主要媒介変数に限定し、その他の信頼、敵意、損失回避、および現在バイアス尺度は異質性分析として扱う。

標準誤差はペア単位でクラスタ化し、主要仮説と主要アウトカムを事前登録する。サンプル数は、処置水準別の最小検出効果を設定したシミュレーションによって決定する。

9 適用範囲、規範的含意、および限界

9.1 適用範囲

本モデルは、継続的な二者関係で、逸脱が公開され、制裁主体が条件付き制裁と制度再構築を選択できる状況へ適用される。企業間の供給契約、労使関係、プラットフォームと参加者、共同事業、国際的な協定などが候補となる。

国家間関係への適用には特に注意が必要である。国家は単一主体ではなく、国内政治、第三国、軍事力、国際法、および複数の意思決定機関を含む。本稿の二者モデルは、歴史的事例の因果的説明を直接には提供しない。

20 世紀の日米関係のように、対立と強制的介入の後に、制度、安全保障、および経済関係が変化した事例は、制裁後状態が制裁前状態と同一ではないという直感を示すことはできる。しかし、この例示は本モデルの実証的検証ではなく、戦争または占領を協調形成の望ましい手段として評価するものではない。

9.2 規範的含意

均衡として実行可能であることと、社会厚生上望ましいことは異なる。制裁は D と E だけでなく、第三者へ損害を与える可能性がある。人命、権利、資本、制度能力、および長期的信頼への損傷は、単純な二者利得では捉えきれない。

本モデルから導かれる規範的含意は、「強い制裁を選ぶべきである」ことではない。むしろ、次の設計原則である。

1. B を継続する場合と A へ復帰する場合の処遇を明確に区別する。
2. A への復帰後に解除する約束を自己執行可能にする。
3. 復帰を実現する最小有効制裁を選ぶ。
4. 破壊的な強度を制度再構築で代替する。
5. 制裁主体が制裁継続から私的利益を得る構造を制限する。

9.3 限界

第一に、基準モデルは完全観測を仮定する。実装誤差または誤認がある場合、非意図的な B に対する制裁が関係損傷を増加させる。Fudenberg, Rand, and Dreber (2012) が示すように、ノイズの下では寛容性が重要になる。

第二に、状態 O を吸収状態として縮約した。受忍後または解除拒否後にも交渉が継続する場合、追加の状態と交渉ゲームが必要である。

第三に、制度ストックと関係損傷を二つの集約変数で表した。現実には、監視、補償、担保、能力破壊、敵意、および評判損失は異なる遷移を持つ。

第四に、基準モデルでは状態 R に入った後の I と H を固定した。制度の減耗と関係損傷の回復を導入すると、再構成状態の安定性は時間依存になる。

第五に、制裁主体を一人の主体として扱った。制裁開始を支持する主体と解除を決定する主体が異なる場合、解除制約は組織内ゲームとして定式化する必要がある。

第六に、現在バイアスと損失回避は比較静学として扱った。時間非整合的な主体の完全

な動学均衡、内生的参照点、および学習を統合することは今後の課題である。

第七に、本稿の新規性は、制裁、制度、関係損傷、または現在バイアスの各要素を個別に初めて導入することではない。貢献は、制裁後の状態を内生的に形成するパッケージ、復帰区間、解除の内生的信頼性、および最適パッケージを一つの動学ゲームへ統合した点にある。

10 結論

本稿は、協調からの逸脱、制裁節、および制裁後の協調再構成を、制度ストックと関係損傷を含む動学ゲームとして定式化した。

中心的な区別は、制裁後に元の協調へ戻ることと、制裁によって変化した状態の下で新しい協調を形成することの違いである。制裁パッケージは、当期の B 継続を不利にするだけでなく、制度ストック I と関係損傷 H を変える。このため、状態 R の協調利得、逸脱利益、および自己執行制約は状態 N と異なる。

本稿は、復帰余剰の厳密な凹性を、当期の規律効果の逡減と関係損傷の凸性から導く十分条件を示した。制裁なしと過大制裁の双方で復帰余剰が負となり、内点で正となる場合、協調復帰を実現する制裁強度は有限区間になる。

また、制裁開始、D の復帰、E の解除、および状態 R の安定性を同時に満たすパッケージが存在するとき、制裁媒介型協調再構成均衡が成立する。解除の信頼性は外生的な約束ではなく、E 自身の解除制約から生じる。

E の最適政策は、一般に最大制裁ではない。制裁が E にも費用を与え、将来関係を損なうならば、E は復帰を実現する最小有効制裁を選ぶ。制度再構築は復帰余剰と状態 R の価値を高めるため、破壊的な制裁を代替する。

決定論的な復帰区間と、実験で観測される復帰確率も区別した。ランダム効用を導入すると、復帰確率は復帰余剰と同じ最大点を持つ単峰型関数になる。現在バイアスは、初期逸脱を増加させると同時に、将来協調価値を軽視させて復帰を困難にする。制裁実績は制裁コミットメントに関する信念を高め、状態 R の安定性を強化し得る。

以上の結果を一つの論理鎖で表すと、次のようになる。

$$B \longrightarrow z = (c, m) \longrightarrow (I', H') \longrightarrow R(I', H')$$

制裁が協調を再構成するのは、相手を単に損失へ追い込むからではない。B を相対的に不利にし、A への信頼可能な出口を設け、将来の協調を支える制度を形成し、同時に関係損傷を上限内に抑える場合に限られる。

付録 A 証明

A.1 命題 2 の証明

式 (5) により、 Φ は $[0, \bar{c}]$ 上で厳密な凹関数である。仮定 A4 と A6 から、 $\Phi(0) < 0$ であり、ある内点 $c^\dagger$ で $\Phi(c^\dagger) > 0$ となる。連続性により、 $(0, c^\dagger)$ に少なくとも一つの根が存在する。

仮定 A5 と A6 から、 $\Phi(\bar{c}) < 0$ であり、 $c^\dagger$ で正となる。したがって、 $(c^\dagger, \bar{c})$ にも少なくとも一つの根が存在する。

厳密な凹関数の上位集合は凸集合であり、一次元では区間である。また、厳密な凹関数が正となる内点を持つ場合、ゼロ点は二つを超えない。よって、二つの根 c_L と c_H は一意であり、 $\Phi \geq 0$ となる集合は $[c_L, c_H]$ である。

A.2 定理 1 の証明

候補戦略を固定する。状態 N および R で D は A を選び、B が観察された場合に E は z^* を開始する。制裁節で D は A へ復帰し、その後 E は L を選ぶ。D が B を継続するか E が C を選ぶ場合には O へ移る。

まず、 $z^* \in \mathcal{Z}_D$ であるため、制裁節で D は B 継続より A 復帰を弱く選好する。次に、 $z^* \in \mathcal{Z}_L$ であるため、A 復帰後の E は解除拒否より L を弱く選好する。さらに、 $z^* \in \mathcal{Z}_S$ であるため、B 観察後の E は受忍より制裁開始を弱く選好する。

状態 R では、 $z^* \in \mathcal{Z}_R$ により D は B への再逸脱より A を弱く選好する。状態 N についても同様の協調制約を課すか、初期逸脱を外生ショックとして扱う。状態 O は吸収状態であり、追加の選択は存在しない。

各決定節で一段階逸脱が利益を与えないため、割引動学ゲームの一段階逸脱原理により、候補戦略は定常マルコフ完全均衡を構成する。

A.3 命題 3 の証明

m を固定する。実行可能な c の集合を $C^*(m)$ とする。仮定により、E の目的関数は $C^*(m)$ 上で厳密に減少する。したがって、任意の $c_2 > c_1$ について、 c_1 と c_2 がともに実行可能ならば、 c_1 の方が E へ高い価値を与える。よって、最大化解は実行可能集合の最小要素である。

A.4 命題 5 の証明

式 (26) について、 $q' > q$ は次と同値である。

$$\frac{q\rho_H}{q\rho_H + (1-q)\rho_L} > q$$

分母は正であるため、整理すると次を得る。

$$q(1-q)(\rho_H - \rho_L) > 0$$

$0 < q < 1$ かつ $\rho_H > \rho_L$ より、この不等式は成立する。

付録 B パラメトリックな特殊例

閉形式の制裁閾値を示すため、制裁節の物質的利得を次のように置く。

$$y_D^A(c) = p - ac$$

$$y_D^B(c) = p + b - dc, \quad d > a \geq 0, \quad b > 0$$

両方の利得が参照点を下回り、損失回避係数が $\lambda > 1$ である場合、当期の規律効果は次のとおりである。

$$g(c) = \lambda [(d - a)c - b]$$

状態 R の一期間利得を次のように置く。

$$r_D(I', H') = r_0 + \gamma m - \eta c^2$$

状態 O の一期間利得を o_D とする。このとき、復帰余剰は次の二次関数となる。

$$\Phi(c, m) = \lambda [(d - a)c - b] + \frac{\delta_D}{1 - \delta_D} [r_0 - o_D + \gamma m - \eta c^2] \quad (B1)$$

次を定義する。

$$A_0 = \frac{\delta_D \eta}{1 - \delta_D}$$

$$B_0 = \lambda(d - a)$$

$$C_0 = \lambda b - \frac{\delta_D}{1 - \delta_D} (r_0 - o_D + \gamma m)$$

すると、次を得る。

$$\Phi(c, m) = -A_0 c^2 + B_0 c - C_0$$

$A_0 > 0$ 、 $C_0 > 0$ 、および次の判別式条件が成立するとする。

$$B_0^2 - 4A_0 C_0 > 0$$

二つの閾値は次のとおりである。

$$c_L = \frac{B_0 - \sqrt{B_0^2 - 4A_0C_0}}{2A_0}$$

$$c_H = \frac{B_0 + \sqrt{B_0^2 - 4A_0C_0}}{2A_0}$$

この特殊例は命題 2 の直感を閉形式で示すためのものであり、主結果は二次関数の仮定に依存しない。

付録 C 記号表

記号	定義
D	潜在的な逸脱主体
E	制裁主体
A	協調行動または協調への復帰
B	自己利益最大化行動または逸脱継続
N	制裁前の協調局面
P	制裁節
R	再構成された協調局面
O	関係崩壊後の吸収状態
I	制度ストック
H	関係損傷
c	制裁強度
m	制度再構築投資
z	制裁パッケージ、 $z = (c, m)$
M(m)	制度ストックの増加
J(c)	関係損傷の増加
K(c,m)	E の制裁パッケージ費用
g(c)	制裁節における A と B の当期利得差
$\Phi(c, m)$	D の協調復帰余剰
c_L	協調復帰を実現する制裁強度の下限
c_H	協調復帰を実現する制裁強度の上限
V_i^R	状態 R におけるプレイヤー i の価値
$\underline{V}_i$	状態 O におけるプレイヤー i の価値
w(c,m)	E が解除を拒否する私的便益
β_D	D の現在バイアス係数
λ	損失回避係数
q	E が高コミットメント型である事前確率

参考文献

Abreu, D. (1986), “Extremal Equilibria of Oligopolistic Supergames,” *Journal of Economic Theory*, 39(1), 191-225. DOI: 10.1016/0022-0531(86)90025-6.

Abreu, D. (1988), “On the Theory of Infinitely Repeated Games with Discounting,” *Econometrica*, 56(2), 383-396.

Andreoni, J., Harbaugh, W., and Vesterlund, L. (2003), “The Carrot or the Stick: Rewards, Punishments, and Cooperation,” *American Economic Review*, 93(3), 893-902. DOI: 10.1257/000282803322157142.

Bottom, W. P., Gibson, K., Daniels, S. E., and Murnighan, J. K. (2002), “When Talk Is Not Cheap: Substantive Penance and Expressions of Intent in Rebuilding Cooperation,” *Organization Science*, 13(5), 497-513. DOI: 10.1287/orsc.13.5.497.7816.

Fehr, E. and Gächter, S. (2000), “Cooperation and Punishment in Public Goods Experiments,” *American Economic Review*, 90(4), 980-994. DOI: 10.1257/aer.90.4.980.

Fehr, E. and Schmidt, K. M. (1999), “A Theory of Fairness, Competition, and Cooperation,” *Quarterly Journal of Economics*, 114(3), 817-868.

Fudenberg, D. and Maskin, E. (1986), “The Folk Theorem in Repeated Games with Discounting or with Incomplete Information,” *Econometrica*, 54(3), 533-554.

Fudenberg, D., Rand, D. G., and Dreber, A. (2012), “Slow to Anger and Fast to Forgive: Cooperation in an Uncertain World,” *American Economic Review*, 102(2), 720-749. DOI: 10.1257/aer.102.2.720.

Halac, M. (2012), “Relational Contracts and the Value of Relationships,” *American Economic Review*, 102(2), 750-779. DOI: 10.1257/aer.102.2.750.

Hilbe, C., Šimsa, Š., Chatterjee, K., and Nowak, M. A. (2018), “Evolution of Cooperation in Stochastic Games,” *Nature*, 559, 246-249. DOI: 10.1038/s41586-018-0277-x.

Kahneman, D. and Tversky, A. (1979), “Prospect Theory: An Analysis of Decision under Risk,” *Econometrica*, 47(2), 263-291.

Kosfeld, M. and Rustagi, D. (2015), “Leader Punishment and Cooperation in Groups: Experimental Field Evidence from Commons Management in Ethiopia,” *American Economic Review*, 105(2), 747-783. DOI: 10.1257/aer.20120700.

Kőszegi, B. and Rabin, M. (2006), “A Model of Reference-Dependent Preferences,” *Quarterly Journal of Economics*, 121(4), 1133-1165.

Laibson, D. (1997), “Golden Eggs and Hyperbolic Discounting,” *Quarterly Journal of Economics*, 112(2), 443-478.

Levin, J. (2003), “Relational Incentive Contracts,” *American Economic Review*, 93(3), 835-857. DOI: 10.1257/000282803322157115.

Mailath, G. J. and Samuelson, L. (2006), *Repeated Games and Reputations: Long-Run Relationships*, Oxford University Press.

O'Donoghue, T. and Rabin, M. (1999), “Doing It Now or Later,” *American Economic Review*, 89(1), 103-124. DOI: 10.1257/aer.89.1.103.

Rabin, M. (1993), “Incorporating Fairness into Game Theory and Economics,” *American Economic Review*, 83(5), 1281-1302.

Shapley, L. S. (1953), “Stochastic Games,” *Proceedings of the National Academy of Sciences*, 39(10), 1095-1100.