

社会的分配をめぐる合意形成の経験は平等原理とマキシミン原理の区別を促すか：
二者間での相互作用場面を用いた実証研究

上島淳史^a 亀田達也^b

要約

希少な資源をいかに分けるかは重要な社会的意思決定である。このような分配の意思決定は、しばしば他者との合意形成のうえで下される。本研究では、社会的分配について他者との合意形成を経験することが、各人の分配選好にどのような影響を与えるかを行動実験で検討した。実験参加者は、第三者の立場から他者3人にお金を分ける課題を行った。分析の結果、他者との合意形成を経験した参加者は、その後個人的に分配の決定を下す際に、平等主義的な選択を減少させ、最も恵まれない受け手の利益を考慮したマキシミン的な選択を増加させることが明らかになった。効用モデルを用いた分析から、合意形成を経験した後、参加者は最不遇の受け手の利益と社会効率（他者3人が受け取る金額の総和）のトレードオフを考慮している可能性が示された。以上の結果は、合意形成の過程が平等原理とマキシミン原理の区別を促進し、参加者が特に不遇を配慮するようになる可能性を示唆している。

JEL 分類番号：D03, D71

キーワード：分配の意思決定, 合意形成, マキシミン的配慮, 平等主義

^a 東京大学大学院人文社会系研究科, 日本学術振興会 (特別研究員 DC1)
ueshima73@gmail.com

^b 東京大学大学院人文社会系研究科 tkameda@l.u-tokyo.ac.jp

1. イントロダクション

希少な資源をいかにして社会の成員に分けるかは重要な意思決定である。このような分配の意思決定は、しばしば他者との合意形成に基づき行われる。本研究では、合意形成の経験が、個人の分配の選好に与える影響を、行動実験を用いて検討した。

協力行動や分配の選好に関するこれまでの行動経済学、社会心理学、発達科学、神経科学、社会調査の研究から、人が極めて平等主義的であることが示されている (Fehr et al., 2008; Tricomi et al., 2010; Norton and Ariely, 2011; Blake et al., 2015)。しかしながら、より最近の研究では人々が平等と呼ぶものには他の価値が混ざっており、人が平等それ自体を強く選好するわけではないことが示唆されている (Starmans et al., 2017)。例えば Kameda et al. (2016)は、不平等の改善と不遇の改善を課題構造から弁別した分配実験において、人々が他者の不遇を改善することに選択的な関心を示すことを明らかにした。同時に、同研究は、人々が最不遇な他者が受け取る金額と受益者全員が受け取る金額の総和に注目し、2つの要素をトレードオフして意思決定することを報告している。

ただし、上記の知見は第三者的立場から分配を個人的に行う場面を扱っており、合意形成を経験した後どのような分配の選好を人々が支持するかについては十分に明らかにしていない(ただし、自己利益に関係する分配の合意形成については Frohlich and Oppenheimer, 1992 を参照)。本研究では、第三者的立場から行う分配について、他者と合意形成することが、各人の分配選好に与える影響を検討した。特に、合意形成の際に望ましい分配法について議論する過程で、平等原理とマキシミン原理という 2つの概念が弁別されることで、合意形成後人々がより不遇に対して強い関心を向ける可能性について検討した。

2. 方法

2.1. 実験概要

北海道大学の学部生 156 名が実験に参加した (男性 99 名, 女性 57 名, 平均年齢 19.3 歳)。参加者は第三者の立場から、別の実験の参加者 3 人にお金を分ける意思決定を複数回行った。分配法の選択肢には、3 人全体の受け取る総額で優れている功利主義選択肢 (図 1 上) と、3 人の受け取る額の平等性で優れている平等主義選択肢 (図 1 中) と、3 人のうち最も恵まれない人が受け取る額 (最低額) の大きさに優れているマキシミン選択肢 (図 1 下) を設けた (Kameda et al., 2016)。なお、3 つの選択肢の呈示順は設問間でランダム化し、最低額、中間額、最高額の呈示順は参加者間でランダム化した。参加者にはお金を受け取る 3 人の情報を一切知らせず、実験後も 3 人に対して参加者が匿名のままであることを課題開始前に伝えた。

2.2. 実験デザイン

全参加者は、フェーズ 1, 2, 3 の 3 フェーズで分配の意思決定を行った。フェーズ 1, 3 ではそれぞれ 40 問、フェーズ 2 では 5 問の決定を行った (計 85 問)。フェーズ 1, 3 では同一の 40 問を呈示し、フェーズ 2 の 5 問もその 40 問に含まれていた。実験終了後に各フェーズから 1 問がランダムに選ばれ、その設問で参加者が選んでいた選択肢に従って、3 人の受け手に対して実際にお

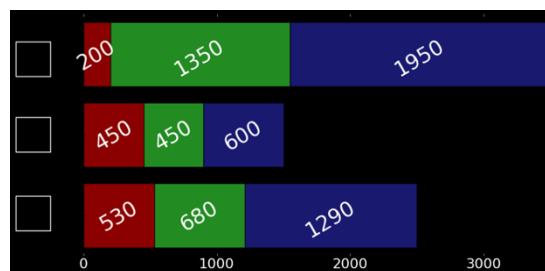


図 1 : 3 つの選択肢の一例。最上段の選択肢 (功利主義選択肢) が選ばれた場合、200 円、1350 円、1950 円がそれぞれ 3 人に分配される。

金を分配することを事前に伝えた。フェーズ 1, 3 ではコンピュータ上に選択肢を呈示し、15 秒以内に意思決定することを求めた。フェーズ 2 では紙面で選択肢を呈示し、5 問について合計 10 分以内に回答することを求めた。なお、分配されたお金を受け取る 3 人は、各フェーズで異なる人物であることが伝えられた。

参加者は個人条件または合意条件のどちらかに割り振られた。個人条件 (56 人、男性 39 人、女性 17 人) に割り振られた参加者は、全てのフェーズで個人的に分け方を決定した。合意条件 (100 人: 男性-男性ペア 24 組、女性-女性ペア 14 組、男性-女性ペア 12 組) に割り振られた参加者は、フェーズ 1, 3 は個人的に決定し、フェーズ 2 は他の実験参加者と 2 人で合意したうえで一つの選択肢を選ぶように指示された。

2.3. 効用モデル

本研究と同一の分配課題を用いた先行研究 (Kameda et al., 2016) では、複数の効用モデルを参加者の行動にフィッティングし、最低額と総額のトレードオフモデルが赤池情報量規準 (AIC) の点で他の効用モデルより優れていることを示した (詳細は Kameda et al., 2016 を参照)。この quasi-maximin model (Charness and Rabin, 2002) では、選択肢 o の効用は、パラメータ α によって最低額と総額のそれぞれが重みづけられた、加重平均として表される。

$$U(o) = \alpha \min(o) + (1 - \alpha)\bar{o}, \quad (1)$$

$\min(o)$ は選択肢 o の分配結果の最低額を表し、 \bar{o} は分配総額を表す。選択確率はソフトマックス関数を用いて推定した。準ニュートン法 (BFGS 法) を用いて、フェーズ 1, 3 におけるマキシミムの配慮の程度 α を参加者ごとに推定した。

3. 結果

本実験では、合意形成を通して平等原理とマキシミン原理という 2 つの概念が弁別されることで、合意形成後、人々がより不遇に対して強い関心を向けるようになるという予測を立てた。

まず、この予測を検証するために、他者との合意形成を経験した参加者が、後の個人決定（フェーズ 3）で平等主義選択を減少させ、マキシミン選択を増加

させることを確認した（図 2）. 独立変数に個人決定のフェーズ（フェーズ 1, フェーズ 3）、条件（個人条件, 合意条件）、

フェーズと条件の交互作用を投入し、マキシミン選択、平等主義選択、功利主義選択を多項ロジスティック回帰した（マキシミン選択を基準カテゴリーとした）. その結果、個人条件と比べ合意条件では、フェーズ 3 で平等主義選択が減少したことが明らかになった（交互作用効果, 95% CI[-0.75, -0.21]). また、同じく合意条件のフェーズ 3 では、功利主義選択が増加することも示された（交互作用効果, 95% CI [0.09,0.56]).

次に、効用モデルを用いてより詳細な解析を行った. 合意形成の経験により平等主義選択が減少し、最低額と総額のトレードオフモデル（quasi-maximin model）が意思決定をより良く説明できるようになると予測した. 合意条件のフェーズ 1, 3 において、参加者の quasi-maximin model の AIC を算出し平均したところ、フェーズ 1 よりもフェーズ 3 で AIC が改善していることが明らかになった（フェーズ 1 の平均 AIC: 42.6, フェーズ 3: 35.83). これは、合意形成を経験した個人の意思決定は、quasi-maximin モデルでより良く予測できることを示している.

最後に、合意形成を行った二者間で、合意形成

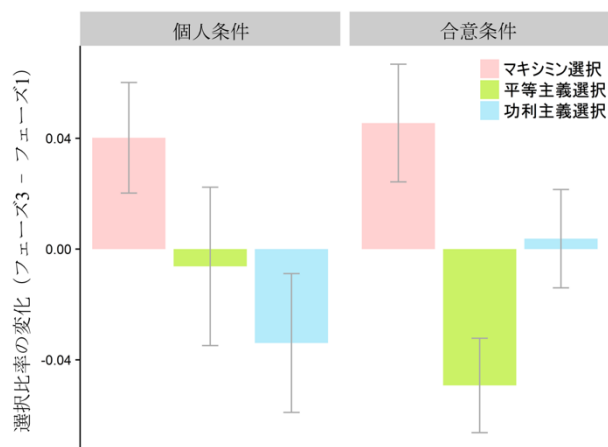


図 2: フェーズ 1, 3 間における選択比率の変化. 合意条件では個人条件に比べ、フェーズ 3 において平等主義選択が減少した.

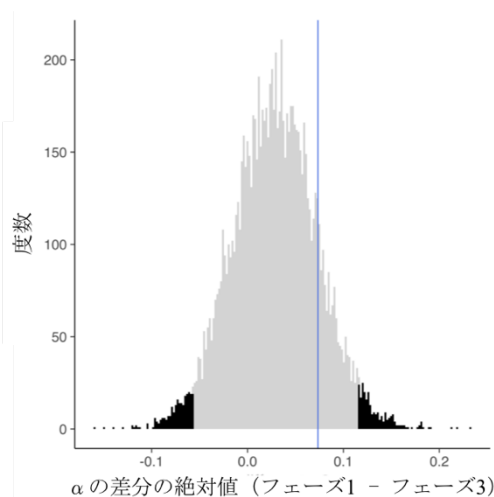


図 3: フェーズ 1, 3 間におけるパラメータ α の収束. 青色の縦線は実際に合意形成した 50 組が平均してどの程度収束したかを示す. 名義ペア 50 組をランダムに生成し平均収束値を算出した結果 (10000 回分) をヒストグラムに示した. 黒色の部分はデータの上下 2.5%を表す.

後に選好の収束が生じていた可能性を検証した。具体的には、合意形成を行なった二者のマキシミンの配慮の度合い (α) が、フェーズ 1 よりもフェーズ 3 で近づいたかどうかを解析した。二者の α はフェーズ 1 よりも合意形成後のフェーズ 3 で 0.07 程度近づいたが (50 組の平均)、個人条件から 50 組の名義ペアを抽出するベンチマーク (10000 回の反復計算に基づく平均) と比べて、5%水準で有意な収束が生じているとは言えないことが明らかになった (図 3)。

3. 考察

本研究では、合意形成を経験することが、社会的分配の選好にどのような影響を与えるかについて検討した。実験の結果、合意形成を経験した後の個人決定において、参加者が平等主義選択を減少させ、マキシミン選択を増加させることが明らかになった。効用モデルを用いた解析からは、合意形成を経験した後の参加者は、最不遇な受け手の利益と社会効率 (3 人の受け取る金額の総和) のトレードオフを考慮して意思決定した可能性が示された。ただし、合意形成した二者間で、最不遇な受け手の利益と社会効率のそれぞれをどの程度重視するか (α) については、有意な収束が見られなかった。

以上の結果は、合意形成を通して平等原理とマキシミン原理という 2 つの概念が弁別され、人々がより不遇に対して強い関心を向けるようになった可能性を示唆している。今後はこのような現象が生じるための境界条件を明らかにする必要がある。例えば、平等主義者同士の合意形成であっても、合意形成の結果、不平等よりも不遇を重視するようになるのか、などは重要な論点となる。

本研究では、最不遇な受け手の利益と社会効率のそれぞれをどの程度重視するか (α) 自体についての収束は見られなかった。しかし、このことは、合意形成の経験によって平等主義選択が減少し、分配問題を最不遇と社会効率のトレードオフで考えるという「メタレベルにおける選好軸の収束可能性」を否定するものではない。このようなメタレベルでの合意は社会的選択において「選好の単峰性」を生むなど、望ましい性質を持つ重要な現象である (List, 2002; List et al., 2012)。それゆえ、最不遇と社会効率のトレードオフが、分配の意思決定について社会的合意を形成するにあたってどの程度頑健にメタレベルでの合意軸となりうるのかについてより詳細に理解することは重要である。

今後眼球運動計測等を用いて、合意形成中にどのような要素 (最低額, 平等性, 総額など) に注目しているのかを直接計測する実験を行い、合意形成の過程で平等原理とマキシミン原理という 2 つの概念が弁別される過程を詳細に検証することを計画している。

引用文献

- Blake, P.R., McAuliffe, K., Corbit, J., Callaghan, T. C., Barry, O., Bowie, A., Kleutsch, L., Kramer, K. L., Ross, E., Vongsachang, H., Wrangham, R. and Warneken, F., 2015. The ontogeny of fairness in seven societies. *Nature*, 528, 258–261.
- Charness, G. and Rabin, M., 2002. Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, 117, 817-869.
- Fehr, E., Bernhard, H. and Rockenbach, B., 2008. Egalitarianism in young children. *Nature*, 454, 1079-1083.
- Frohlich, N. and Oppenheimer, J. A., 1992. *Choosing Justice: An Experimental Approach to Ethical Theory*. University of California Press, Berkeley, CA.
- Kameda, T., Inukai, K., Higuchi, S., Ogawa, A., Kim, H., Matsuda, T. and Sakagami, M., 2016. Rawlsian maximin rule operates as a common cognitive anchor in distributive justice and risky decisions. *Proceedings of the National Academy of Sciences*, 113, 11817-11822.
- List, C., 2002. Two concepts of agreement. *The Good Society*, 11, 72-79.
- List, C., Luskin, R. C., Fishkin, J. S. and McLean, I., 2012. Deliberation, single-peakedness, and the possibility of meaningful democracy: evidence from deliberative polls. *Journal of Politics*, 75, 80–95.
- Norton, M. I. and Ariely, D., 2011. Building a better America—One wealth quintile at a time. *Perspectives on Psychological Science*, 6, 9-12.
- Starmans, C., Sheskin, M. and Bloom, P., 2017. Why people prefer unequal societies. *Nature Human Behavior*, 1, 1–7.
- Tricomi, E., Rangel, A., Camerer, C.F. and O'Doherty, J.P., 2010. Neural evidence for inequality-averse social preferences. *Nature*, 463, 1089-1091.