**第13章　労働経済学における実験的手法**

1. **はじめに**

本章では、近年、観察データを使った実証分析と並び、新たな分析手法として注目されている経済実験の手法について解説する。ここでは、今後、経済実験をデザインし、実際に自分で経済実験を行ってみたいと考えている学部上級者や大学院生をターゲットにするので、基本的な実験作法（参加者の募集、プログラムの作成、実験指示書の作成、実験室の準備、謝金の準備と支払い方）、実験デザイン（参加者の振り分け方や実験回数）、そして実験メソッドを中心に説明する。また、読者が理解しやすいように、贈与交換を検証する経済実験の例題を交えて、実践的にわかりやく説明することを心がける。加えて、労働経済学の研究領域における経済実験手法の有効性と可能性について説明する。したがって、本章では、労働経済学の研究領域に関連する経済実験研究のサーベイではないことを留意していただきたい[[1]](#footnote-1)。

　最初に、混同されやすい「実験経済学」と「行動経済学」の区別について簡単に説明したい。「実験経済学」は分析手法に関する分類であり、どのように経済実験を行うことで経済理論や制度設計に示唆を与えられるかを研究する分野である。その一方で、「行動経済学」は経済理論の組み立て方に関する分類であり、個人の意思決定や効用関数により現実に近い人間行動を反映する仮定を採用したものである。確かに行動経済学の検証のために経済実験が使われることが多いが、両者は大きく性質が異なる点には注意が必要である。

　経済実験の目的は、科学実験と同様に、経済理論分析から得られた理論を数量的に検証することである。特に、介入の効果を厳密に推定するのに効果的である。しかし、Friedman (1953)が指摘するように、経済学を含めた社会科学は、科学とは異なり、理論分析から得られる理論や予測を実験から検証することはできないと言われ、多くの研究者はそう信じてきた。Friedman (1953)は、人々を被験者（参加者）としたラボ実験を実施すること自体は可能であるが、ラボ実験で制御された環境は実際の経済社会とは大きく異なると考えていた。実際の社会において直接実験をして、理論や経済的予測を検証できればいいが、それには莫大な費用が掛かり、到底無理な話である。また、理論分析において、経済主体は利己的、合理的な経済人（ホモ・エコノミカス）と仮定するが、分析しやすいためにこのような極端な仮定を設定しているだけで、実際の社会の人々の特性を必ずしも描写しているわけではない。従って、実際の社会の人々を使って実験することで、合理的な経済主体を前提とした理論分析から得られた理論や予測を検証することに意味はないと主張する。

Binmore (1999)は、利己的、合理的な経済人（ホモ・エコノミカス）の選択や行動を探る理論研究と実際の人間を参加者として行う経済実験とでは整合的な結果を得られなかったとしても、それがすぐに経済人を前提とした伝統的経済学の理論研究の否定にはつながらないと主張する。彼は、経済理論を実験で検証するためには、参加者が慎重に考える誘因をもち、じっくり考える時間がある状況を整える必要があると説く。

Rubinstein(2001)は、Friedman(1953)と同様に、経済理論の検証に対する経済実験の有効性に否定的であったが、理論分析によって得られた理論が道理的であるか、筋が通っているか、もっともらしいかを確かめる術として経済実験は有効であると認めた。ただし、実験実施者が、科学的な実験手順に従い、データ分析による結果に都合のいいような解釈をしない場合に限ってと釘を刺している。科学的に正しい実験の手順を踏まないような実験結果は信頼できるものではなく、それだったら、理論家による思考実験で十分だと述べている。近年、経済実験による研究は盛んに行われ、多くの研究結果が蓄積されているのは事実であるが、経済実験を実施する際に、手順だけでなく、解釈についても注意を払う必要がある。

　本章の構成は以下の通りである。次節では2種類の経済実験（ラボ実験とフィールド実験）を紹介し、メリットとデメリットを示す。第3節では、経済実験の例として贈与交換の実験を紹介し、第4節で経済実験を行う際に必要な事務的な手続き実践的に解説する。第5節では、様々な実験のデザインとメソッドを紹介するとともに、それぞれのメリットとデメリットを解説する。第6節では、労働経済学の研究領域においる経済実験という手法の有効性と可能性を示す。最後の第7節は、本章のまとめである。

1. **実験的手法のタイプとメリット・デメリット**

本節では、経済実験の手法と既存の観察データ（２次データ）を利用した実証分析の手法の位置関係を整理すると共に、データの性質に沿ってどの分析手法が適切かを解説する。そして、実験的手法のメリットとデメリットを説明する。

図13-1は、List and Rasul (2011)から抜粋したものである(Figure 2)。横軸の右側は、データが「自然発生的」、または分析者によって観察されるデータであることを意味し、横軸の左側はデータが「人工発生的」、もしくは分析者によって制御された実験データであることを意味する。

データの性質が自然発生的な場合、構造モデル（Structural modeling）を仮定して推定する方法や、「疑似実験(quasi-experiment)」と考えられる場面を探し出して政策の介入効果を推定する方法がある。推定方法として、差の差(difference-in-difference)分析や傾向スコア・マッチング(propensity score matching)のような計量経済学で学習する分析ツールが採用される。また、これらよりも古く、基本的な介入効果の推計手法として、操作変数法(instrument variable)がある（第11章参照）。分析対象の１つの例として、高齢者雇用安定法改正の効果を考えよう。2012年法改正により定年の引き上げ、廃止、または継続雇用の導入を雇用主に義務付けされた。このような法制度の改正は疑似実験として利用することができる。高齢者の雇用確保措置によって高齢者の就業率が上昇したかを、『労働力調査』（総務省）や『雇用動向調査』（厚生労働省）などの大規模なデータを利用して、計量経済学の手法を駆使して検証することができる。

ただ、公表されている大規模データから政策評価を統計的に評価する場合、注意しなければいけないことがある。それは、対照群と処置群が明確に分かれているかどうかである。法改正が適用されるグループと適用外のグループが存在しなければいけない。高齢者雇用安定法を含め、一般的な法改正の場合、企業規模別に応じて移行措置期間が設けられる。その適用時期のずれを利用して対照群と処置群に識別し、法改正による介入効果を評価することは可能である。しかし、もう１つ介入評価の分析に必要は前提条件として、ランダム化比較試験法（Randomized Controlled Trial、略してRCT）の仮定が満たされているかどうかを確認する必要がある。すなわち、法改正がなかった場合、対照群と処置群の平均的アウトカム（高齢者の就業率）が同じでなければいけない。その前提条件を満たすためには、対照群と処置群に高齢者がランダムに振り分けていなければいけない。ところが、高齢者雇用安定法改正による政策評価の場合、処置群は移行措置期間がない大企業に勤める高齢者が属し、対照群は移行措置期間が設けられる中小企業に勤める高齢者が属することになる。そうすると、法改正前の対照群と実験群が平均的に同じ属性をもったグループとは言いにくく、サンプル・セレクション・バイアスが発生する可能性がある。高齢者雇用安定法改正による介入効果を含め多くのケースでは、RCTの仮定を満たしていないことがある。企業規模別に関する変数をコントロールして推計しても、対照群と処置群の異質性を完全に制御できるとはいえない。

既存の観察データを使った実証研究とは異なり、経済実験の場合、対象群と処置群を明確に区別できる環境を実験実施者が作り出すことができる。経済実験の手法は大きく分けて、「ラボ実験（実験室実験）」と「フィールド実験」がある。ラボ実験では、参加者を実験室に集めるので、実験環境や処置群に対する介入を最も厳密に制御することができる。多くの場合、ラボ実験は大学内の実験室で実施され、参加者は大学生である。

大学生を実験室に集めるのではなく、広く一般の人々を参加者にするのがフィールド実験の手法である。List and Rasul (2011)によると、主なフィールド実験手法は３つある[[2]](#footnote-2)。1つ目は、「人工的フィールド実験」(artfactual filed experiment)であり、ラボ実験で実施するような実験内容（独占ゲーム、最後通牒ゲーム、信頼ゲームなど）をそのまま実験室外にいる特定グループの人々（ある地域の漁師、バイクメッセンジャー、村人など）を参加者として実験を行うことである。ラボ実験との違いは、実験対象となる参加者が大学生ではないことだけである。2つ目は、「枠組み型フィールド実験」(framed field experiment)である。実験室で行うような取引を、実際の社会で行われている取引に当てはめて、参加者の選択や行動を観察する。例の１つとして、List (2002)の研究がある。彼は、Smith (1962)が実験室で実施した買い手と売り手による市場実験を、実際に取引が行われている野球カードの交換所で参加者を募り市場実験を実施した。ただ、この場合、参加者は経済実験に参加していることを認識・承諾している。したがって、「参加者である」という認識が自然な振る舞いを阻害する可能性がある（ホーソン効果,またはdemand effect）。3つ目は、「自然型フィールド実験」(natural field experiment)である。参加者が実験に参加していることを認識していないこと以外、枠組み型フィールド実験と同じである。枠組み型フィールド実験よりもより自然な設定であるが、参加者が知らないところでの実験の参加者となっていることに倫理的な問題が生じる可能性がある。自然型フィールド実験を実施する場合は、倫理的な問題に発展しないように注意を払うべきであろう。List and Rasul (2011)がラボ実験と3種類のフィールド実験の違いをまとめた表(Table 1)を表13-1に抜粋する。

次に、経済実験手法のメリットについて簡潔にまとめる（Charness and Kuhn 2011）。以下のメリットはラボ実験、フィールド実験ともにあてはまるが、実験環境を強く制御できるラボ実験において、メリットがより際立つ。1つ目は、実験実施者によって実験環境を完全に制御できることから、介入変数を完全に外生的に設定でき、原因と結果の因果関係を厳密に明らかにすることができることである。既存の観察データによる実証分析の場合、着目する介入変数が完全に外生であることは非常に稀である。観察できない欠落変数との相関性（見せかけの相関）や逆の因果関係によるバイアスが懸念される。2つ目のメリットは、実験環境を制御できるので、理論モデルと同じような環境を設定し、理論分析から得られる理論や知見を忠実に検証することができることである。3つ目は、大規模な観察データを収集したり、または購入したりすることに比べれば、実験費用は比較的に安いことである。もちろん、参加者の人数や謝金の金額に依存する。4つ目は、妨害行為や差別行為などの反道徳的な事象が発生する環境や過程を経済実験から解明できることである。しかし、参加者に対する倫理的な配慮は十分に必要であり、通常、実験者が所属する機関の倫理委員会からの承認を得ることが好ましい。5つ目は、起こり得る全ての状況に対して意思決定を行わせることで実験を通じて取引が成立した結果だけでなく、取引が成立しなかった場合の結果も観察できることである。例えば、最後通牒ゲームにおいて、受け手はゲームが始まる前に、様々な提示額に対して受諾するか否かをあらかじめ回答してもらい、その後、実際にゲームを始める。そうすることで、提示者と受け手の間で取引が成立しなくても、受け手の意思を示すデータを収集することができる[[3]](#footnote-3)。この手法を戦略選択法（strategy method）と呼ぶ。この方法に関しては後で詳細に説明する。

実験のデメリットは、外的妥当性(external validity)を満たさない場合が多いことである。特に、ラボ実験ではこのデメリットが指摘される。ほとんどのラボ実験は大学内で実施され、参加者は大学生である。このように参加者を大学生に限定してしまうと、推定結果にサンプル・セレクションの問題が生じる。あるラボ実験に参加した参加者グループを使って得られた実験結果は、別の参加者グループを使って実施した実験の結果と必ずしも一致しないことがある。外的妥当性を満たしにくいラボ実験において、実験の反復再現性をとることは難しい。したがって、実験結果は慎重に解釈することに留意すべきである。外的妥当性をできるだけ満たすために、フィールド実験では、実験室を飛び出し、大学生ではない一般の人々を参加者とすることで発展してきた。

1. **経済実験の例：贈与交換ゲーム**

　本節では、経済実験を参加した経験の無い読者のために、経済実験に参加するとどのようなことが起きるのかを実験の流れに沿って説明する。視点は参加者である学生とする。例に挙げる実験として、贈与交換（Gift-exchange）ゲームを用いる。これは、労働市場モデルの１つであるAkerlof (1982)を経済実験に落とし込んだもので、Fehr et al.(1993)を初めとして多数の労働市場実験で用いられている。

　まず、参加者は学内の掲示版（Web含む）、チラシ配布、メールなどを通して経済実験の情報を知る。実験情報として、時間と報酬（経済実験では報酬が確定されていないので、平均報酬などが示されている）は知ることができるが、詳しい内容は明らかにされていない。情報を見て参加したいと思った場合、メールやWeb上のシステムで実験に申し込みを行う。申込後、参加日時の確定の連絡がくるので、指定された場所（主に経済実験室）に指定された時間に向かう。複数人で実験を行うことが一般的な経済実験では、遅刻は認められない。

　当日指定された場所に向かい、受付を行う。座席は指定されている。経済実験の開始前に、同意書やその他の事務書類へ記入を行う。実験は任意参加が原則なので、いつでも辞退することができる。

　経済実験の開始後、実験指示書（インストラクション）が配布される。例えば、贈与交換ゲームでは、以下のような指示書が配布される。なお、以下の指示書は全体の抜粋であり、完全版はWebにアップロードする。

実験指示書

* 実験にご参加いただきありがとうございます。あなたは実験参加に対して、1000円が与えられます。さらに、実験中で得られたポイントに応じて報酬を支払います。
* あなたには役割Aまたは役割Bのどちらかが割り当てられます。
* 役割Aと役割Bは1人ずつのペアとなり、意思決定を行います。ペアの相手が誰なのかはわかりません。最後まで匿名性が保証されます。
* 役割A・役割Bはそれぞれ10ポイントを与えられています。
* 実験には2つのステージがあります。第1ステージでは役割Aが役割Bにポイントを渡します。渡した値をXとします。第1ステージの結果が役割Bに伝えられた上で、第2ステージでは役割Bが役割Aにポイントを渡します。渡した値をYとします。
* ポイントは以下のように決まります。
  + 役割A　10-X+5Y
  + 役割B　10+5X-Y
* この実験は1回のみ実施します。
* ポイントは、1ポイント＝50円で換算して終了後にお支払いします。

　指示書はスタッフが音読を行うので、静かに聞く。説明後には質疑応答の時間が設けられる。その後、実験指示書の理解度に関するテストが行われるので回答し、スタッフに回答が合っているかどうかの確認をもらう。

　ここまでの手順を踏んだ上で、ようやく実験開始である。実験はコンピュータで行われる。参加者は役割Aを与えられたので、役割Bに7ポイントを渡した。役割Bからは2ポイントが返ってきたとする。最終的なポイントは10-7+(2×5)=13ポイントである。実験の結果当初の10ポイントからは増えたが、相手の役割Bはさらにポイントを増やしている（43ポイント）ので、少し悔しい。

　実験終了後は、アンケートに回答する。アンケートでは、個人属性を回答し、実験内容に関する質問に回答する。

　全ての作業が終了した後に、謝金が支払われる。個人ごとに謝金が異なるため、支払いは他の人に見えないような形で行われるのが一般的である。今回は13ポイントだったので、参加報酬1000円と、実験での報酬13×50＝650円が合わせて支払われる。

　実験が果たしてどういう目的で行われたのか、参加者としては知りたい気持ちになる。しかし、経済実験では実験内容についての説明はしないのが一般的である。

1. **事務的作業**

　経済実験を実施するには、事前に様々な準備が必要になる。本節では、経済実験に必要な実験場所・参加者・実験指示書・実験プログラム・研究費についてそれぞれ簡単な解説を行い、実験実施への手助けとしたい。

　経済実験では参加者間で意思決定が観察できないような処置を行い、参加者の名前は伏せ匿名で実験が行われるのが通常である。実験者が経済実験を通して得たいのは、参加者個人の意思決定及びその意思決定が合わさったことで生まれる結果である。この参加者個人の意思決定は、実験指示書で設定される実験上の条件にのみ基づくのが理想的であり、それ以外の意思決定に影響を与える要素はできるだけ排除したい。例えば、実験でペアを組む相手が誰かが特定できてしまった場合、実験外での人間関係が意思決定に影響を与えてしまう可能性がある。また、自身の利得に直接関係無い場合であっても、他の人の行動が見えてしまうとそれを参考として意思決定を行ってしまう可能性がある。

　このような理想的状況として整備されているのが、経済実験室である。経済実験室では、参加者はパソコンが配備された席を割り当てられ、席はパーティションで区切られている（図13-2）。パソコンはネットワークを介して意思決定のやり取りを行うため、匿名性を担保することができる。また、パーティションで区切ることで意思決定が観察されることを防いでいる。ただし、専用の経済実験室を持っている大学・学部は限定されている。使用規定は実験室ごとに異なるが、学外者の利用は想定されていないことが多い。内部に共同研究者がいることが望ましいだろう。

　専用の経済実験室を利用できない場合は、主に２つの選択肢がある。１つはコンピュータが配備されている情報処理室での実施である。この場合、実験者が別途パーティションを準備するなど参加者間で意思決定を見えないようにする工夫が必要になる。もう１つは、コンピュータは用いず、「紙とペン」を用いて通常の教室で実施する方法である。この場合も、パーティションなどの利用が望ましい。

　さらに厳格に参加者間の相互作用を防ぐために、部屋を複数使用する場合もある。例えば、贈与交換の例のように役割が2つに分かれる場合、役割ごとに部屋を分ける場合もある。

　参加者を募って実験を始める前に、実験内容が倫理的に問題ないかを倫理委員会で確認することが好ましい。通常、大学内では部局別に倫理委員会が設置されている。ない場合は全学共通の倫理委員会に申請する。実験を通じて参加者の健康を損ねたり、参加者のプライバシーを侵害したりする可能性がなければ、問題なく実験の実施は承諾される[[4]](#footnote-4)。

　参加者は実験場所の大学生に募集をかけるのが一般的である。実験内容によっては、学部や学年を限定する場合がある。例えば、大学で学んだ知識が実験に大きな影響を与えてしまうことが懸念される場合、経済学部生を排除したり、１年生に限定したりする場合がある。

　大学生ばかり（しかも多くの場合いわゆる上位大学で実施される）が経済実験の参加者となってしまい、一般的な人々とは異なるのではないかという批判は、心理学の実験と同様に数多くなされている。もちろんこの批判には真摯に対応する必要があるが、学生を対象とすることで参加者がある程度同質的となり、実験の運営や実験結果の解釈がしやすくなるというメリットがあることは重要な点である。学生ではない一般参加者をランダムに実験室に集めるのは非常にコストがかかる上、学生よりセレクションが強くなってしまう可能性もあることは考慮される必要がある。

　参加者プールの管理は事務的な課題の１つである。実験が１回きりの場合は、実験者がExcelなどを用いて管理すれば良いが、恒常的に多数の実験を行う場合は、大学院生・ポスドクを含めたスタッフを指定し、参加者プールを管理する必要がある。多数の実験を行う場合は、参加者が同種の実験に複数回参加することのないよう確認する必要がある。参加者募集・管理のシステムとして、海外ではORSEEというシステムが使われているが、日本では一般的ではない。

　コンピュータを用いる場合、何らかのプログラムを用いる必要があるが、プログラムに指定があるわけではない。JavaやPythonなどを用いて自分でプログラミングをする研究者もいる。経済実験で最も用いられているのがzTreeというソフトウェアである（Fischbacher 2007）。zTreeは経済実験に特化したソフトウェアであり、許可を得れば無償で使用できる（http://www.ztree.uzh.ch/en.html）。zTreeはWindowsのみで動作するソフトウェアのため、現在はMac, Linux, 各種タブレットやWeb上で動作できるソフトウェアの開発が進められている。例えばoTree（Chen et al. 2016）は次のスタンダードとして有力なソフトウェアである。

　参加者に対する指示（インストラクション）は、文字情報と音声情報を組み合わせて行うのが一般的である。これは、実験上の設定を参加者に確実に理解させるために必要である。文字情報としては、紙の説明書を配布するか、コンピュータのスクリーン上に表示させる。どちらの場合でも、実験中にいつでも設定を確認できるようにするのが望ましい。音声情報としては、実験者が説明を読み上げるか、事前に録音しておいた音声あるいは機械音声を流す[[5]](#footnote-5)。音声情報を使う理由は、参加者が説明を「聞いた」ことを担保するためである。誘導を避けるため、実験の説明は研究者本人ではないスタッフ（直接関係のない大学院生など）が行うことが望ましい。

　実験指示書における注意点を何点か述べる。まず、実験ではできるだけニュートラルな言葉を使うのが好ましい。「経営者」や「労働者」のように、参加者の役割を明確に与えてしまうと、自分の意思で行動を決定するのではなく、与えられた役割を想定して行動を選択する可能性がある（フレーム効果）。ただし、ニュートラルな言葉を使用することで指示が抽象的になりすぎてしまいゲームの理解に影響が出る場合は、多少のコンテクストを加えることもあり得る。また、報酬に関しては、ゲームによって参加者が獲得する金額を直接インストラクションに記載するのではなく、獲得ポイントを示し、そしてポイントと金額の交換比率を記載することが好ましい。実験上のパラメータとポイントの関係を表現する方法として、数式を用いる方法と、利得表を用いる方法がある。数式を用いる場合は簡潔に表現できるが、数式に慣れていない人には難しい可能性がある（特に複雑な式となってしまう場合）。利得表を用いて説明するほうがわかりやすい場合もあるが、選択肢が多い場合は表が大きくなりすぎてしまうなどの別の問題が発生しうる。また、利得表自体も馴染みがあるとは言えない可能性もある。

　実験で得られる報酬金額は、参加者が報酬を獲得したいというインセンティブを抱くように設定する。あまりにも安く報酬設定すると、金銭的なインセンティブが機能しないし、高すぎても費用が高くなるだけである。一般に、参加者が普通にアルバイトをして稼ぐことのできる時給の1.5倍程度で支払うのが適切と考えられている。参加者が一度に複数の実験を行う場合、それぞれの実験で獲得した金額の合計を支払う方法と、複数の実験結果のうち１つを元に報酬を支払う方法がある。金額の合計を支払う場合、実験の最初の段階で好成績を残し、報酬金額が高くなると、後半でやる気を失う可能性があるという問題がある（wealth effect）。これを避けるため、複数の実験結果から無作為に１つの結果を選び、その結果をもとに報酬を決めるほうが好ましいと考えられている。

　実験を実施するにあたって現実的に必要となるのが研究費である。研究費の大半を占めるのが参加者への謝金である。必要額は研究によって大きく異なるが、例えば参加者が300人、謝金を平均2,000円と想定すると、計60万円となり、そこまで安い金額ではないだろう。次に多いのがスタッフ（事務職員または学生）への謝金である。スタッフは事前準備や当日運営に従事し、大学規定の給与を支払う。サイコロなどの物品や紙代などが必要になるが、高額になることは稀である。

　参加者への謝金支払いについては、大学の事務方と十分な折衝が必要である。なぜなら、経済実験における謝金は、大学規定より高い金額を支払い、実験結果によって報酬が変動するという事務的にはイレギュラーな対応が必要だからである。現在は先例が増えてきているので、折衝は比較的やりやすくなっていると思われる。

1. **実験デザインの実験メソッド**

　本節では、一般的な実験デザインや実験メソッドにおいて、注意しなければいけない３つの作法を解説する。まずは、対照群と処置群を設定する方法である参加者間計画法(between-subject method)と参加者内計画法(within-subject method)を紹介し、各方法のメリットとデメリットを説明する。2つ目は、実験回数の設定に関して、1回限りの実験(one shot experiment)にするのか、それとも参加者が複数回の実験を行う繰り返し実験（repeated experiment）にするのかを、各実験設定のメリット・デメリットを踏まえながら説明する。3つ目は、実験メソッドに関することである。ここで取り上げる実験メソッドは、直接選択法(direct response method)と戦略選択法(strategy method)である。２つの手法はそれぞれメリット・デメリットがあり、ここではどちらの手法が優れているかの議論はしない。２つの方法のメリット・デメリットを「経済実験の例」で紹介された贈与交換の実験例を使って説明する。

1つ目の対照群と処置群の設定方法から始める。参加者間計画法では、複数いる参加者を対照群と処置群に振り分け、処置群に属する参加者に対して実験実施者が関心のある介入をする。そして、対照群に属する参加者の選択や行動と比較することで、介入効果を厳密に抽出する。この手法で留意しなければいけない点は、参加者を対照群と処置群に無作為に振り分けることである。そうすることで、対照群と処置群それぞれ属する参加者の属性や観察不可能な要素が平均的に同質となることがわかっている。対照群の選択・行動は、介入を受けた処置群の反事実仮想的な選択・行動とみなすことができる。反対に参加者を無作為に振り分けないと、介入することで観察される処置群と対象群の結果の違いが、介入の効果によるものか、それとも対照群と処置群の属性の違いによるものかが識別できない。参加者間計画法は、ランダム化比較試験法と同様に、適度な数の参加者を対照群と処置群に無作為に振り分けることが可能なら、介入のみの効果を正確に測ることができる。

しかし、実際に、参加者を無作為に２群に振り分けることはそれほど容易ではない。特に、実験施設や設備の制限により、対照群と処置群の実験セッションを別々の時間帯に実施するとなると、参加者の無作為な振り分けは難しくなる。例えば、対照群に対する実験を朝のセッションに実施し、処置群には午後のセッションで実施する場合である。朝早く起きられる参加者は朝のセッションを選択し、夜更かしするような参加者は午後のセッションを選択する。このような分け方では、自己選抜バイアスが発生し、対照群と処置群に属する参加者の属性は平均的に同質ではなくなり、対照群が処置群の反事実仮想的な群に相当しないことになる。もし2室の実験室が設備され、午前中のセッションに参加する参加者を無作為に対照群と処置群と2グループに分けることができるなら、自己選抜バイアスを心配する必要はなくなる。

もう１つの方法である参加者内計画法では、各参加者に対照群に属する参加者として実験をしてもらい、同時に処置群に属する参加者としても実験をしてもらう。参加者は介入のある実験と介入のない実験の両方を受けることになる。この場合、対照群と処置群に属する参加者は同じなので、属性の同質性の条件は満たされる。したがって、実験において、介入することで処置群と対象群の実験結果の違いが、介入の効果によるものだけと判断できる。しかし、この方法の大きな欠点は、順番によるバイアス(order bias or sequential effect)や学習効果（過去の結果）によるバイアスが発生することである。仮に1番目に対照群の実験、2番目に処置群の実験を設定した場合、2番目の実験の選択は、1番目の実験結果に影響されやすいと考えられる。解決方法としては、実験の順番を無作為に並べて、複数回実験をすることで、順番バイアスを軽減することが考えられる。または、1番目の実験の結果を参加者にフィード・バックすることなく、次の実験に進むことで、順番バイアスを軽減する方法も考えられる。その他のメリットとして、参加者間計画法に比べて、参加者数が少なくても対照群と処置群のデータが得られるので、実験費用が安く済むことである。

　参加者を無作為に対照群と処置群に振り分けることができるだけの参加者数と実験設備が整っているのなら、参加者間計画法により厳密に介入効果を明らかにすることができる。参加者内計画法の場合、バイアスが生じる可能性があるが、1人の参加者が複数回実験することになるので、パネル・データとなる。個別効果と実験の順番を示すダミー変数をコントロールした上で、介入の効果を推定することが可能である。実験データはパネル・データなので、様々な統計的手法を応用することも参加者内計画法の利点である[[6]](#footnote-6)。どちらの計画法を採用するかは、実験内容や実験設備の状況に応じて判断すべきであろう。

　2つ目の作法は、実験回数の設定に関することである。設定方法として、1回限りの実験(one shot experiment)と参加者が複数実験を行う繰り返し実験（repeated experiment）が主に考えられる。1回限りの実験を参加者内計画法の実験では適用できないので、必然的に、参加者内計画法の実験は繰り返し実験となる。裏を返せば、1回限りの実験のメリットは、繰り返しの実験に比べて、順番によるバイアス(order bias or sequential effect)や学習効果によるバイアスを気にする必要がないことである。それゆえに、そのバイアスを軽減するための統計的処置も必要としない。デメリットは、1回限りなので、実験内容を理解しないまま実験を終わってしまう参加者がいる可能性がある。そのようなことがないように、実験を始める前に、実験内容を確実に理解しているかどうかを確認するクイズを参加者に課す必要がある。もう１つのデメリットとしては、参加者1人につき１つの観察値しか収集できないので、分析に十分耐えられるだけのデータを収集するためには、多くの参加者を集めなければいけないことである。謝金支払いによる金銭的費用が高くなることを覚悟しなければいけない。

　先ほど述べた通り、参加者内計画法を採用した場合、必然的に繰り返し実験となるので、繰り返し実験のメリット・デメリットは参加者内計画法のメリット・デメリットでもある。デメリットは、順番によるバイアスや学習効果によるバイアスが発生する可能性があることである。対策としては、毎回ごとに実験結果を参加者に知らせるのではなく、次々と実験を繰り返し、実験について学習できない措置を採ることである。繰り返し実験法のメリットは、1人の参加者から複数の観測値を収集できることができるので、費用対効果が高いことが挙げられる。多くの繰り返し実験では、序盤と終盤の観測値を使わずに、真ん中の回から収集した観察値だけを分析に使うことが多い。この理由としては、序盤は参加者が実験に慣れていないから、そして終盤は気持ちが緩み真剣に取り組んでくれないからである。また、贈与交換のような2人一組の実験を繰り返す場合、毎回、異なる参加者とペアを組むような設定をすることに留意する（ストレンジャー・マッチング）。

3つ目は、実験メソッドに関することである。ここでは、直接選択法と戦略選択法を解説する。２つの方法のメリット・デメリットを「経済実験の例」で紹介された贈与交換の実験例を使って説明する。

直接選択法では、参加者が実験実施者の指示のもと、決められたルールに従い自分の効用を最大にするような選択を行う。この方法では、通常、参加者が選択した結果しか実験実施者は観察することができない。贈与交換の例の場合、役割Bの参加者は、役割Aの参加者が提示するXをもとに最適なYを決定する。しかし、実験実施者が観察できるのは、役割Aの参加者が提示したある１つのXと、それに対する役割Bの参加者が選択した１つのYだけである。もし仮にその他のXが提示されたら場合、この役割Bの参加者が選択すると考えられるYは、この実験方法から観察することはできない。直接選択法は、実験手順そのものは単純であり、且つ不自然ではないので、参加者にとって実験内容を理解しやすい。したがって、実験データは信頼できるものと考えられる。しかし、情報量が少ないという欠点がある。情報量を多くするためには参加者数・実験回数を増やさなければいけない。

その欠点を補うのが、戦略選択法である。この方法の場合、実験が始まる前に、参加者にあらゆる状況に応じて、どの選択肢を選択するかを仮想的に回答してもらい、それから実際の実験を始める。実験による事後の選択と仮想的な選択を照らし合わせて、実際に参加者に支払う報酬金額を決める。このような方法の場合、実験実施者は参加者が実際に選択した結果だけではなく、あらゆる状況に応じて選択する結果も観察できる。

贈与交換の例の場合、実験を開始する前、役割Bの参加者に対して、仮想的に役割Aが提示する色々なXに応じてどれ位のYを払うかを回答してもらう。役割Aの参加者に対しては、役割Bが選択するYを予想させることで、役割Aである参加者の信念を探る。そして、実験を開始し、役割Aの参加者が実際に選択したXと役割Bの参加者が事前に回答したYと対応せることで、両参加者の結果と報酬金額が決定する。実際に成立した取引結果だけでなく、実際に提示されなかったXに対して役割Bの参加者が支払うYも仮想的な回答から実験実施者は観察することができる。したがって、戦略選択法から得られる情報量は、直接選択法から得られる情報量よりも多い。しかし、戦略選択の欠点は、まず実験手続きが不自然なところである。事前に仮想的な質問に回答し、その回答をもとに取引結果が決まるような手順は、直接選択法の手順に比べて、不自然なので、参加者が実験内容を理解できなかったり、手順によるバイアスが生じたりする可能性がある。もう１つの欠点は、実験前に課す仮想的な質問で、役割Bの参加者が回答する負担が大きすぎると、認知費用が高くなり、真剣に回答に取り組んでくれない可能性がある。仮想的な質問数に配慮が必要である。

1. **労働経済学における経済実験**

　これまでは、経済実験を実施する上で必要な作法、実験デザイン、分析するためのメソッドを、贈与交換の実験を題材にしながら紹介してきた。本章では、労働経済学の研究における経済実験的なアプローチの適用性と可能性について述べる。

　労働経済学の理論モデルの解析によって得られる知見を、これまでは既存の観察データから統計的に検証してきた。データがアクセスしやすくなったこと、性能の高いコンピュータが安く購入できること、そして、統計ソフトの性能向上により、観察データを使った実証研究は盛んになり、労働経済学分野における研究方法の主流となった。しかし、分析方法として手軽になったとはいえ、やはり実証分析方法だけでは理論分析の検証に限界がある。既存の観察データでは観察できない変数は多くあるし、また、原因と結果の因果関係を明らかにしたくても、他の変数が原因を示す変数に影響する場合、着目する因果効果を厳密に推定することができないという問題がある（内生性の問題）。

　そこで、実証研究の分析方法では検証が難しい領域を経済実験の手法で解き明かすことができる。特に、経済実験の強みが発揮される研究領域として、逐次的サーチ・モデルのような個人の意思決定モデルやプリンシパル＝エージェント・モデル(principal-agent model)を応用した内部労働市場の分析が考えられる。経済実験の手法が適している理由は以下の通りである。

繰り返しになるが、経済実験のメリットは、実験環境を制御できて、通常の観察データでは観察できない変数を観察できるように設計できることである。サーチ・モデルによる求職活動の分析から解明されることとして、「留保賃金の特性」(reservation wage property)がある。その特性は、有期限の逐次的サーチ・モデルの場合、求職期間が長くなるにつれ、求職者の留保賃金は低下することである。また、失業保険給付のような補償金が増えれば、求職者が負う求職するための努力水準は低下し、留保賃金は高くなることがわかっている。では、「留保賃金の特性」を実証的に検証する場合、適切な観察データはあるだろうか。残念ながら、1人の求職者の求職活動開始から終了までの期間を追いかけた長期間のデータ(duration data)はそれほど多くないし、仮にそのようなデータを利用できたとしても、求職の努力水準や留保賃金は通常データとして観察できない。観察できるのは就職できた人の求職期間だけである。したがって、「留保賃金の特性」を実証的に検証することは難しい。

そこで、経済実験の登場である。ラボ実験から「留保賃金の特性」を検証するために、次のような簡単なゲームを参加者に受けてもらう。あるポイント分布から無作為にポイントを引き、受諾すればそのポイントが報酬となり、拒否すれば、次のラウンドに進み、再びポイント分布からポイントを引く。ただ、サーチできるラウンド数は有限とし、最後まで受諾しなかった場合、報酬はゼロとする[[7]](#footnote-7)。実験室内で受けるのは簡単なゲームであり、直接、参加者の求職活動に結びつくものではない。しかし、逐次的サーチ・モデルのエッセンスを取り入れた実験なので、この簡単なゲームでも「留保賃金の特性」を検証することができる。また、次のポイントを引く前に、参加者が払う努力水準（または努力費用）や留保水準を記入するように設定すれば、留保ポイントや努力費用の変遷が明示的に観察でき、「留保賃金の特性」を検証することができる。

　プリンシパル＝エージェント・モデルをベースとした内部労働市場のメカニズムの解明でも経済実験の手法が有効である。この場合もサーチ・モデルと同様に、雇用者が負う努力水準（努力費用）を通常の観察データから観察できないので、実証分析には限界がある。プリンシパル＝エージェント・モデルでの労働契約のタイミングは、雇用主が提示した賃金で労働契約を交わした後に、雇用者が自身の努力水準を決めることになっている。雇用主は雇用者の努力水準を観察できないので、雇用者の最適な選択は努力をしないことになる。つまり、モラル・ハザードの問題が発生する。それを見越して、雇用主は固定給だけでなく、歩合給で構成する賃金体系を提示する。生産量に応じて支払う歩合給を設定することで、生産性ショックに対するリスクを雇用者とシェアすることになる。もし、雇用主が雇用者の努力水準を完全に観察できるなら（雇用主が雇用者の努力水準を決められるなら）、賃金体系は固定給だけとなり、生産性ショックのリスクは雇用主だけが負うことになる。

経済実験では、制御された実験環境のもと、雇用主と雇用者の役割する参加者がそれぞれ理論分析の知見通りに行動するのかを検証する。雇用主が雇用者の努力水準を観察できる確率や生産ショックの頻度や大きさを変えることで、参加者の選択や行動が理論的に整合的であるかを確認する。更に、1人のプリンシパルに対して複数のエージェントがいるモデルのラボ実験では、どのような支払い方が雇用者側である参加者の努力を最大に引き出せるか、そして雇用主側である参加者の利潤を最大にするのかを検証する。例えば、歩合制対トーナメント制、個別ワーク対チーム・ワークが考えられる。その他にマルチ・タスクやダブル・モラル・ハザードに関するラボ実験がある。ラボ実験のサーベイをしたCharness and Kuhn (2011)が紹介した既存研究の多くはプリンシパル＝エージェント・モデルに関するラボ実験である。

プリンシパル＝エージェント・モデルのようにモラル・ハザードが内在するモデルでは、エージェントの「努力費用」をどのように設定するかは重要な問題である。報酬と同じ単位に揃える金銭的費用と、実際に参加者に作業させることで発生する「本当の努力(real effort)」に分けられる。前者の金銭的費用は、報酬と同じ単位なので、報酬と比較しやすいし、明示的に観察できるが、厳密には「努力」と解釈しがたい。その一方で、後者の場合、本当の意味での「努力」を払うことによる費用といえる。実際の作業といっても、計算問題や迷路クイズを解くようなタスクは、実社会のタスクと大いに異なる。また、その努力費用は報酬に対してどの程度なのかは比較できないし、そもそも観察できない。観察できるのは、努力を払うことで完了したタスクの量だけである。

著者の個人的な意見として、逐次的サーチ・モデルのような個人の意思決定モデルやプリンシパル＝エージェント・モデルのような2人組（または1対２～３人）による相互作用を検証するのに経済実験は効果的と考える。その一方で、複数の参加者間で取引を行うような市場実験や一般均衡モデルの実験は実行することが難しいと感じる。2002年にノーベル経済学賞を受賞したSmith (1962)が開発したダブル・オークションのような基本的な市場実験は別として、複雑な一般均衡モデルを経済実験で検証することは難しく、理論から得られる知見をサポートする実験結果は出にくいことがある。理由としては、実験構造が複雑になりすぎて、参加者が実験内容を十分に理解していない可能性がある。経済実験に関しては、できるだけシンプルに、参加者が理解しやすいようにデザインすべきである。したがって、個人の最適意思決定モデルや1対１（または1対2～3人）のプリンシパル＝エージェント・モデルのようなシンプルなモデルを検証するのに経済実験は有効といえる。

　次に、労働経済学の研究領域で活用されるフィールド実験を紹介する。これまで贈与交換の研究を題材にして経済実験のノウハウを紹介してきたので、贈与交換に関するフィールド実験の研究成果をここで取り上げる。まず、Gneezy and List (2006)では、データ打ち込みの作業や各戸を訪れて寄付を集める作業に雇われた参加者に対し、当初は時給12ドルと伝えていたのを当日に20ドルに増やすと伝えた。給与を増やすことで、雇用者（参加者）は雇用主（実験実施者）の善意に感謝し、より頑張って働くかを検証した。作業開始直後は、時給12ドルのままのグループ（対照群）に比べると、時給20ドルに増えたグループ（処置群）の生産性が高かったが、後半になるとその差は無くなった。Kube et al. (2013)は、給与を当日に引き上げる介入だけでなく、引き下げる介入も取り入れて実験を行った。図書館のデータ打ち込みの作業を任された参加者の生産性は、給与増加に影響はなかったが、給与減少に対して明らかに低下した。Kube et al. (2012)は、非金銭的な報酬の贈与交換効果を検証した。時給12ユーロに加えて、7ユーロ相当の水筒を贈与する介入を行った。実験結果によると、現金給与が増加した場合、生産性は微増であったが、水筒を贈与した場合、生産性は大きく増加した。現金よりも、モノのほうが雇用主の善意が伝えやすいことがわかった。

「差別」を含め男女間の賃金格差や雇用格差は労働経済学の重要な研究領域である。経済実験の手法から、心理的態度の違いに基づいて、男女間の賃金や雇用の違いを説明しようとする研究が進んでいる[[8]](#footnote-8)。特に、それらの男女間の違いをリスク回避度、損失回避度、競争回避度の男女間の違いに着目する研究が増えている[[9]](#footnote-9)。

1. **おわりに**

本章では、近年、分析手法として活用頻度が高い経済実験のメリット・デメリット、そして実験実施の作法やメソッドについて解説した。加えて、労働経済学の研究における経済実験手法の有効性や可能性について説明した。

今回紹介したラボ実験やフィールド実験に加えて、新しい実験手法が開発されている。それは、クラウド・コンピューティングを利用した実験である。例えば、実験室で行う3人一組の公共財実験を、クラウド・コンピューティングにより参加者とネットワークでつなげて、ラボ実験と同じような実験を行ってもらう。Amazon Mechanical Turkがクラウド・コンピューティングのサービスを提供しおり、ラボ室を飛び越えた実験が盛んに行われようとしている。ラボ実験の最大の批判は、参加者が大学生であること、すなわち外的妥当性が満たされないことであるが、クラウド・コンピューティングのサービスを活用した実験により、大学生ではない一般の人々を参加者として実験を行うことができる。しかし、これで完全に外的妥当性が満たされたとはいえない。このサービスによる実験に参加するには、ネットワークの環境がある程度整っていなければいけない。参加できる潜在的な参加者が、ネットワークの整備具合によって左右されるということなので、サンプル・セレクション・バイアスは残る。しかし、参加者をより幅広く、手軽に募集できるようになったので、参加者のサンプルに一般性は高まると期待できる。

このようにネットワークやオンラインを活用した実験手法は、経済学の教材として学生の教育に活用されている。林良平氏（東海大学）は、長年、オンライン経済実験教材の開発に取り組み、経済実験を通じた新しい経済学教授法を提案している。彼は、スマホを通じて教室の学生間で相互作用ができる実験プラットファームを完成させた。その教材を使って、ダブル・オークションによる一般均衡市場実験、公共財実験、信頼ゲームなど様々な実験を教室で行うことができるようになった。林氏が管理するサイト(<https://xee.jp/>)にアクセスし、申請書を提出すると実験教材のファイルが提供される。

ブックガイド

　実験手法・作法の改良余地はまだあり、今後の発展が期待できる。より詳しく経済実験の手法について学びたい読者には、Fréchette et al. (2015)やBardsley et al. (2010)を推奨する。日本語の教科書としては、川越(2007)と下村(2015)がある。

**参照文献**

Akerlof, G. A. (1982). “Labor Contracts as Partial Gift Exchange.” *Quarterly Journal of Economics*, *97*(4), pp.543–569.

Bardsley, N., Cubitt, R., Loomes, G., Moffatt, P., Starmer, C., and Sugden, R. (2010) *Experimental Economics: Rethinking the Rules*, Princeton University Press, Princeton

Binmore, K. G (1999) “Why experiment in economics?” *Economic Journal,* 109, F16-F24.

Charness, G., and Kuhn, P. (2011)”Labo Labor: What Can Labor Economists Learn from the Lab?”*Handbook of Labor Economics* 4a, edited by D. Card and O. Ashenfelter, Ch3, pp.229-330, North Holland

Charness, G., Gneezy, U., and Kuhn, M. A. (2012) “Experimental methods: Between-subject and within-subject design,” *Journal of Economic Behavior & Organization*, 81, pp.1-8

Chen, D. L., Schonger, M., & Wickens, C. (2016). “oTree-An open-source platform for laboratory, online, and field experiments.” *Journal of Behavioral and Experimental Finance*, *9*, pp.88–97.

Croson, R., and Gneezy, U. (2009) “Gender differences in preferences,” *Journal of Economic Literature*, 47(2), pp.448-474

Falk, A., & Fehr, E. (2003). “Why Labour Market Experiments?” *Labour Economics*, *10*(4), pp.399–406.

Fehr, E., Kirchsteiger, G., & Riedl, A. (1993). “Does Fairness Prevent Market Clearing? An Experimental Investigation.” *Quarterly Journal of Economics*, *108*(2), pp.437–459.

Fischbacher, U. (2007) “z-Tree: Zurich Toolbox for Ready-made Economic Experiments.” *Experimental Economics*, 10(2), pp.171–178.

Friedman, M. (1953) “The methodology of positive economics,” In *Essays in Positive Economics*, pp.3-43, Chicago University Press, Chicago

Fréchette, G.R., and Schotter, A. (2015) *Handbook of Experimental Economic Methodology*, Oxford University Press, New York

Kube, S., Maréchal, M. A., and Puppe, C. (2012) “The Currency of Reciprocity: Gift Exchange in the Workplace,”*American Economic Review,* 102(4), pp.1644-1662

Kube, S., Maréchal, M. A., and Puppe, C. (2013) “Does Wage Cuts Damage from a Natural Experiment? Journal of European Economic Association, 11(4), pp.853-870

List, J. A. (2002) “Testing neoclassical competitive market theory in the field,” *PNAS、*99(24), pp. 15827-15830.

List, J. A., and Rasul, I. (2011) “Field Experiments in Labor Economics,” Handbook of Labor Economics 4a, edited by D. Card and O. Ashenfelter, Ch2, pp.103-227, North Holland

Greezy, U., and List, J. A. (2006) “Putting behavioral economics to work: testing for gift exchange in labor markets using field experiments,” *Econometrica*, 74, pp.1365-1384

Greezy, U., Niederle, M., and Rustichini, A. (2003) “Performance in Competitive Environments: Gender Differences,” *Quarterly Journal of Economics*, 118(3), pp.1049-1074.

Niederle, M., and Vesterlund, L. (2007) “Do Women Shy Away from Competition? Do Men Compete Too Much?” *Quarterly Journal of Economics*, 122(3), pp.1067-1101.

Rubinstein, A. (2001) “A theorist’s view of experiments” *European Economic Review*, 45, pp.615-628

Smith, V. L. (1962) “An Experimental Study of Competitive Market Behavior,” *Journal of Political Economy*, 70(2), pp.111-137

川越敏司（2007）実験経済学　東京大学出版会

澤田康幸(2016) 「経済における実証分析の進化」『進化する経済学の実証分析』経済セミナー

下村研一(2015) 実験経済学入門　経済学コア・テキスト＆最先端、新世社

高野久紀(2007) 「フィールド実験の歩き方」西條辰義編著『実験経済学への招待』第8章　ＮＴＴ出版

図13-2



大阪大学大学院経済学研究科の情報処理室

1. 参考となるサーベイとしては、Falk and Fehr (2003)やCharness and Kuhn (2011)がある。 [↑](#footnote-ref-1)
2. List and Rasul (2011)は紹介した3種類のフィールド実験の日本語訳は、澤田（2016）や高野(2007)に従う。 [↑](#footnote-ref-2)
3. 最後通牒ゲームは、「提示者」と「受け手」がペアを組んで行うゲームである。まず提示者は、自身の初期保有金から受け手に渡す提示金額を伝える。受け手が承諾すれば、その金額を受け手は獲得し、残りは提示者が受け取る。もし受け手が提示金額を拒否すれば、両者とも獲得金額はゼロとなる。 [↑](#footnote-ref-3)
4. 脳内の血流の変化を観察する装置(MRI)を使うような神経経済学分野の実験の場合、通常の経済実験に比べて、より一層参加者の健康状態に配慮する必要がある。 [↑](#footnote-ref-4)
5. 最近は、実験指示書を読み上げる音声読み上げソフトを利用して参加者に説明するケースが多い。 [↑](#footnote-ref-5)
6. Charness et al. (2012)は、参加者間計画法と参加者内計画法の違いや問題点について解説している。 [↑](#footnote-ref-6)
7. サーチできるラウンド数は無限だが、手元のポイントを拒絶し、次のラウンドに進む時、ある確率で強制的に終了させられる「無期限サーチ・モデル」のラボ実験もある。 [↑](#footnote-ref-7)
8. Croson and Gneezy (2009)は、経済実験を通じて検証された心理的態度における男女間の違いについてサーベイした。 [↑](#footnote-ref-8)
9. 男女間の競争心の違いに関する研究として、Greezy et al. (2003)やNiederle and Vesterlund (2007)がある。 [↑](#footnote-ref-9)