

「実証分析における因果性と再現性：Pitt and Khandker (1998)の マイクロファイナンス論文を巡る一連の議論から」

1. はじめに

経済学の実証分析では、複数の事象の間にどのような因果関係が存在するか厳密に精査することが求められる。因果関係とは、事象 X の変化が事象 Y の変化をもたらす場合の、 X と Y の関係についてである。よく似た概念に相関関係がある。これは事象 X の変化に応じて事象 Y も変化する時に使う。 Y の変化が X によってもたらされたのか、 X の変化が Y の変化によってもたらされたのか、或いは両者の変化が X と Y の変化に影響を与える第三の事象 Z の変化によってもたらされたのか、様々なケースが考えられるが、相関関係においては、原因と結果の方向性は基本的に問われない。

経済学の実証分析において、因果性の立証が求められる背景の一つには、経済分析が誤った政策提言につながらないように、配慮していることが挙げられる。例えば、農作物の近代品種を採り入れた家計(X)ほど所得が高い(Y)という事象は相関関係でしかなく、この観察結果だけをもとに、農村所得改善の方策として近代品種の普及を提言することは危険である。なぜなら、近代品種を採り入れる家計は、種子を毎年購入できる豊かな農家であったり(Y が原因となり X が生じている側面がある)、教育水準が高く新技術への対応が容易な家計である(第三の要素 Z によって X が生じており、 X がなくても Y が高い可能性がある)などの場合が考えられるからである。そうした可能性を適切に統御しなければ、近代品種が家計所得上昇にどれだけ貢献しているのか、そして近代品種普及が所得向上という目標に対して望ましい政策なのか、実のところよくわからない。

因果性を立証するためには、他の条件を全く同一にした上で、外部から意図的に X を変化させたら Y が変化するかどうかを観察することが基本となる。例えば、土壌・気温・雨量などあらゆる外部環境が同じ圃場で、同じだけの施肥・除草・水管理を行い、伝統品種と近代品種の栽培比較実験を行う。その結果、近代品種の方が収量が高くなれば、近代品種は高収量をもたらす

という因果関係が確立される。私たちがこれまで触れてきている「科学的知見」とは、多くの場合、このようによく制御された環境の中で生みだされ、同一の条件で再度実験を行えば同じ結果が得られる再現性の高い観察結果に基づいている。

一方、経済データのように環境を制御するのが難しい場合、因果性を立証する方法に苦慮してきたのが事実である。経済学では、最近になりようやくランダム化比較実験(**Randomized Controlled Trials: RCT**)と呼ばれる方法が発展してきたが、それまでは計量経済学手法の工夫を施し、因果性を立証する方法が採られることが多かった。

RCTとは、**X**の変化を無作為に与えた多数の個人を観察すれば、制御すべき要因が、異なる**X**の集団間で平均的に同じになっていくという、統計的法則性をベースに因果性を検証するものである。例えば、「マイクロクレジット(**MC**:小口融資)は貧困層の生計改善に寄与する」という命題を考えたとしよう。この場合、個々人は**MC**融資を受けるか受けないかの二択があるが、各個人はどちらか一つしか選べないため、同一の個人が**MC**を借りた場合と借りない場合の比較は不可能である。また、**MC**借入が個人の自由な意思決定によるものであるのならば、先の近代品種の例の通り、**MC**借入に影響を与える様々な事象を(通常目に見えない努力水準、意欲なども含め)制御しなければ因果性の検証はできない。**RCT**では、まず個人を無作為にグループ分けし、片方のグループは**MC**借入を行い残りのグループは借入を行わない(或いは借入機会が片方のグループのみに与えられる)状況を意図的に作り出し、グループ間の結果指標(生活水準など)の差を比較する。各個人はどちらのグループに所属するかランダムに決められるため、十分な標本サイズの下では、**MC**融資前の生活水準はグループ間で平均的に同じになるはずである。また、それ以外に、制御すべき要因も全て平均的には同じになる。そのため、**MC**融資後の生活水準に有意な差が生じているのだとしたら、それは**MC**借入の差がもたらした帰結と判断できるとするのが**RCT**の考え方である。

MC研究でも近年、**RCT**に基づき、**MC**融資と生活状況の間の因果性を検証した論文はいくつか存在する。他方、そうした手法が流行する以前に、高度な統計的手法を用いて因果関係を確立したと考えられていた論文に、**Pitt and Khandker (1998)**、以下**PK**が挙げられる。**PK**論文は**MC**が盛んに行われていたバングラデシュを研究対象とし、「**MC**は貧困層の生計改善に貢献する」、

「生計改善効果は女性に貸した方が大きい」という重要な二つのファインディングを含んでいたため、公刊されてすぐに大きな反響を呼んだ。その中の一つに **Morduch (1998)**の反論がある。**Morduch** は異なる仮定の下で因果関係を検証すると「MC が貧困層の生計改善に貢献する」とは言えなくなることを示している。その後、**Pitt (1999)**による **Morduch** 論文へのさらなる反論を経た後、**Khandker (2005)**により、PK 論文の拡張が行われ、PK 論文のファインディングは異なる推計方法でも概ね支持されることが示された。

PK、Morduch、Khandker の三つの論文の分析結果がそれぞれ果たして本当に因果関係を確立したと言えるのか、実証結果は再現されるのかどうか、という課題に挑んだのが **Roodman and Morduch (2009、以下 RM)**である。RM は三論文で使われたデータで同じような推計を試したところ、三論文ともに因果性を立証できるほどの十分な証拠が見つからなかったとして話題を呼んだ。

しばらく RM 論文の結果は経済学者の間でも受け入れられてきたが、2011 年になり、未公刊ながら、Pitt が RM の誤りを指摘する二本の論文(Pitt 2011a, 2011b)をウェブ掲載してから、PK 論文の妥当性について経済学者の間で議論が再燃している。本稿は、PK、RM、Pitt (2011a, 2011b)に特に着目し、各論文を整理するとともに、PK 論文を巡る最近の議論の動向についてまとめるものである。なお、PK から RM に至るまでの一連の議論の流れは高野・高橋 (2011) も参照されたい (各論文の詳細には踏み込んでないが、Karnofsky (2011)も最近の議論を手際よくまとめているので興味ある方はそちらも参照のこと)。

2. PK 論文以降の一連の議論

- Pitt and Khandker (1998)のおさらい

PK 論文は 1998 年に **Journal of Political Economy (JPE)**という経済学のトップジャーナルに掲載された。ⁱ この論文ではバングラデシュで 1991 年から 1992 年に三度にわたって実施された家計調査に基づく分析を行っている。1992 年当時、バングラデシュで MC 融資をしている主な団体として、(1) Grameen Bank, (2) Bangladesh Rural Advancement Committee [BRAC], (3) Bangladesh Rural Development Board's [BRDB] Rural Development RD-12 program の三つがあった。標本家計は

29 の Thana(タナ)から無作為に選ばれた 1798 家計である。このうち、25 タナに属する 1538 家計が MC 実施地域に住んでおり、そのうち 905 家計が実際に MC の借り入れを行っている。

PK 論文はこの三団体による 1986 年 12 月からの MC 借入総額 (1992 年価格に実質化した総額) が家計の消費水準や児童の就学などにどのような影響をもたらしたのか検証したものである。これを簡単な式で表すと、

$$y_{ij} = X_{ij}\beta_y + C_{ij}\delta + \varepsilon_{ij}^y \quad (1)$$

となる。ここで y_{ij} は最終的に見たい結果の指標 (家計消費水準や児童教育) ; X_{ij} は y_{ij} に影響を与える家計の特徴 (制御すべき要因) ; C_{ij} は MC 借入額 ; ε_{ij}^y は誤差項である。 β_y と δ はそれぞれ推計パラメーターであり、 δ の値が統計的に有意であるなら、MC 融資と生活水準の向上には相関関係があるという主張がなされる。

しかし、この式だけでは MC によって生計が改善されるかという因果関係までは十分に検証できない。よく知られる通り、 ε_{ij}^y は X_{ij} や C_{ij} で制御できず、かつ y_{ij} の変動に影響を与える全ての要因が含まれている。また、よく親しまれている最小二乗法 (OLS) では ε_{ij}^y の平均値が 0 で正規分布と仮定される。しかし、仮に、 ε_{ij}^y と C_{ij} の間に何らかの相関関係がある場合には、その仮定が崩れ、 δ の推計量は一致性をもたず、バイアスが生じることになる。

大きく分けるとここには 2 つの潜在的な問題がある。第一に、MC 実施地域の選別が恣意的に行われうることから生じる問題である。例えば、MC 実施担当者は MC のインパクトがより大きく見込める豊かな地域に対して事業を実施するかもしれない。この場合、 δ の値は村の特徴を制御した時と比較して大きくなる。逆に、MC 実施担当者は貧困対策の観点からより貧しい地域で MC を貸し出そうとするかもしれない。この場合、 δ の値は村の特徴を制御した時と比較して小さくなる。村の特徴を適切に制御せず、MC 実施地域の選別が恣意的に行われうることから生じるこうした推計のバイアスは **Placement Bias** と呼ばれている。

第二に、MC からの借入を行うか否か、どのくらい借入するかを、家計が自主的に選択でき

ることから生じる問題が挙げられる。家計の特徴はある程度 X_{ij} で制御されているが、 X_{ij} は調査者がデータを取りやすいものしか反映しきれない。仮に、商才など目に見えない要因が C_{ij} に影響を与えていたら、それを明示的に X_{ij} に取り込まない限り、 ε_{ij}^y と C_{ij} の間に相関関係が生じ、 δ は歪んだ推計値となる。MC の場合、商才に長けている人ほど、小さな資金から大きな利益を生み出すことが可能であるため、商才と借入の間に正の相関関係があるというのは、簡単に思いつくストーリーである。その他、商才以外にも、借金に対する家計の選好など、目に見えない数々の要因が C_{ij} と関連しうる。家計が自主的に MC に参加し、その MC 需要の背後にある家計の特徴を全て制御しきれないことから生じるこうした推計のバイアスは **Self-selection bias** と呼ばれている。

Self-selection bias を除去するために一般的に採られる方法は、 C_{ij} に影響を与え、 y_{ij} (及び ε_{ij}^y) と相関しない第三の変数(識別変数: Z)を用い、 C_{ij} の変動から ε_{ij}^y と相関する部分を除く、操作変数(Instrumental Variable: IV)法と呼ばれる推計である。

$$C_{ij} = X_{ij}\beta_c + Z_{ij}\pi + \varepsilon_{ij}^c \quad (2a)$$

$$y_{ij} = X_{ij}\beta_y + \hat{C}_{ij}\delta + \varepsilon_{ij}^y \quad (2b)$$

IV 法は上式の通り二段階最小二乗法 (2SLS)として捉えるとわかりやすい。ここでは、第一段階で変数 Z を用い、予測値 \hat{C}_{ij} を得る。そして第二段階でその予測値 ($\hat{C}_{ij} = X_{ij}\hat{\beta}_c + Z_{ij}\hat{\pi}$) と y_{ij} の関係を検証する。 C_{ij} の変動のうち ε_{ij}^y と相関する、観察されない要因の部分は ε_{ij}^c で表され、それは予測値から除去されている。IV 法を使うと、 Z によって外生的に決定される C_{ij} の変動分が y_{ij} に与えている影響が δ となり、 δ は一致性を持つことが知られている。

PK 論文では、Placement や Self-selection から生じる潜在的な推計バイアスを除去するために、まず村のダミー変数を推計式に入れた。これにより、村ごとの平均的な特徴の違いや、それがもたらす Placement Bias の軽減が期待できる。また Self-selection に対応するため、IV 法に

言及しているものの、有効な操作変数 Z は明確に特定できなかった。そのため、IV を用いた 2SLS ではなく、家計の外部で定められた規則によって、外生的に C_{ij} の変動が生みだされている準実験的状况を考えた。彼らが注目したのは、当時、バングラデシュでは土地所有面積 0.5 エーカー以下の零細農家や土地なし層にしか MC が貸し出されないという受給資格ルールである。ルール自体は個々の家計では変えられないので、受給資格基準である 0.5 エーカー付近で、受給資格を得るためにわざわざ土地を転売するなどが無い限りにおいては、MC 受給資格を持つかどうか(土地面積が 0.5 エーカー以下であるかどうか)は家計にとって外生変数であろう。そして、その条件を制御すれば、 C_{ij} の影響を計測できるだろうというのが PK 論文の識別戦略である。ii

$$y_{ij} = X_{ij} \beta_y + C_{ij} \delta + v_j + \varepsilon_{ij}^y \quad (3)$$

(3)式において、 v_j は村の固定効果を表す。簡単化のために X_{ij} は受給資格があるかないか (=1 受給資格なし、=0 受給資格あり) だけを表すとすると、 j 村で MC が実施されている場合 ($j=2$) と MC が実施されていない場合 ($j=1$) の期待値はそれぞれ、

$$E(y_{ij} | j=1, X_{ij}=0) = v_1^y \quad (4a)$$

$$E(y_{ij} | j=1, X_{ij}=1) = \beta_y + v_1^y \quad (4b)$$

$$E(y_{ij} | j=2, X_{ij}=1) = \beta_y + v_2^y \quad (4c)$$

$$E(y_{ij} | j=2, X_{ij}=0) = p\delta + v_2^y \quad (4d)$$

となる。ここで p は、MC 実施村で MC 受給資格がある家計の中で実際に参加した人の割合である。この仮定のもとで、 C_{ij} の効果は δ として推計できるだろうというのが PK 論文の骨子である。

この方法は、回帰不連続デザイン (RDD) の応用とも考えられる。RDD とは、ある閾値を境に MC 受給資格が急激に変化する場合に、その閾値近辺の家計を比較する方法である。例えば土地保有面積 0.5 エーカーにより完全に受給資格が決められている場合、0.51 エーカーの人は受給資格がないが、0.49 エーカーの人は受給資格を持つ。彼らの資金需要や観察不可能な特徴は、例えば土地を 1 エーカー持っている人と全く持っていない人よりもおそらく小さく、「似た者同士」と言えるだろう。そのため、制御しなければいけない多くの要因の差は無視できるほど小さ

く、違いは MC 受給資格の差と、それにより生じる生活水準の差と考えることも可能である。PK 論文では、土地保有面積 0.5 エーカー近辺の標本家計だけを推計に使っているわけではないが、こうした考えにより、PK の推計は RDD に基づいているという本人たちの主張も見られる。

PK 論文の実際の推計は以上の説明より、さらに複雑である。実際に PK が採用した方法を列挙すると

- (a) まず C_{ij} には打ち切りポイントがあり、潜在需要 C^*_{ij} が $\log(1000)$ タカ以下の場合に $C_{ij} = 0$ であり、 C^*_{ij} が $\log(1000)$ タカ以上の場合に観察データの通り C_{ij} の値をとるという想定がなされている。つまり観察値 C_{ij} は以下の性質を持つ。

$$\begin{aligned} C^*_{ij} &= X_{ij} \beta_c + \nu_j + \varepsilon_{ij}^c \\ C_{ij} &= C^*_{ij} \text{ if } C^*_{ij} > \log 1000 \\ C_{ij} &= 0 \text{ otherwise} \end{aligned}$$

- (b) C_{ij} は男女別に分けられ、さらに各村で Grameen Bank, BRAC, BRDB がそれぞれ男性向け、女性向けに MC を実施しているかというダミー変数と掛け合わせ、最終推計式には 6 種類 (3 つの MC 機関*男女別) の C_{ij} が入れられている。

$$C^*_{ijf} = X_{ij} \beta_{cf} + \nu_{jf} + \varepsilon_{ijf}^c \quad (5a)$$

$$C^*_{ijm} = X_{ij} \beta_{cm} + \nu_{jm} + \varepsilon_{ijm}^c \quad (5b)$$

$$y_{ij} = X_{ij} \beta_y + \nu_j^y + \sum_h C_{ijf} D_{jfh} \delta_{fh} + \sum_h C_{ijm} D_{jmh} \delta_{mh} + \varepsilon_{ij}^y \quad (5c)$$

ここで、添え字 f は女性グループ、添え字 m は男性グループ、 Df は MC プログラム h ($h = \text{Grameen Bank BRAC, BRDB}$) が女性向けに事業を実施している場合に 1 をとるダミー、 Dm は同様に MC プログラム h が男性向けに事業を実施している場合に 1 をとるダミーである。

- (c) 明確に排除制約(exclusion restriction)を満たす識別変数 Z は(5a)(5b)になく、その替わり MC 受給資格がないというダミー変数を作り、それを(5c)の推計の中に取り入れているが、受給資格は土地の大きさ(閾値である 0.5 エーカー以上か以下か)によってのみ決められて

いるのではなく、MC 融資が実際にあったかも考慮している。そのため、例えばプログラムが実施されている村で土地の大きさが 0.5 エーカー以上であっても、観察期間に MC 融資があった場合には、受給資格ありと見なされている。

- (d) また MC 参加者を母集団の分布よりも多めにサンプルに含めるように調査したことから、Weighted exogenous sampling maximum likelihood(WESML)を用いて、サンプルウェイトの調整している。
- (e) (5a, 5b, 5c) 式を 2SLS ではなく、Limited information maximum likelihood (LIML)を用いた同時推計をしている。さらに、村の固定効果(Fixed Effect:FE)を入れている。
- (f) 最終的な推計式はこれらを組み合わせた WESML-LIML-FE であるが、仮に Placement Bias も Self-selection bias も発生していないのであれば、サンプルウェイトのみを調整すれば、精度のよい推計が得られるため WESML のみを試したり、Placement Bias 発生していないのだとすると、FE を入れた推計はより inefficient になるため、WESML-LIML (+村の外生的特徴をだけを入れたもの)に基づくものも試している。iii

これらの推計式に従い、PK は 4 つの推計結果を並べている。

TABLE 2
ALTERNATIVE ESTIMATES OF THE IMPACT OF CREDIT ON PER CAPITA EXPENDITURE AND WOMEN'S NONLAND ASSETS

EXPLANATORY VARIABLES	LOG OF WEEKLY TOTAL PER CAPITA EXPENDITURE				LOG OF WOMEN'S NONLAND ASSETS				
	Unweighted OLS	WESML-OLS	WESML-LIML	WESML-LIML-FE	Unweighted Tobit	WESML-Tobit	WESML-LIML	WESML-FE	WESML-LIML-FE
Amount borrowed by female from BRAC	.007 (3.048)	.007 (2.847)	.0340 (2.291)	.0394 (4.237)	.277 (4.359)	.182 (2.834)	.0425 (2.302)	.1151 (2.003)	.0318 (.356)
Amount borrowed by male from BRAC	.010 (2.906)	.010 (2.835)	-.0161 (-1.658)	-.0192 (1.593)	.141 (1.615)	.110 (1.214)	.2589 (2.367)	.0878 (1.007)	.1005 (.468)
Amount borrowed by female from BRDB	.002 (.573)	.003 (.906)	.0258 (1.723)	.0402 (3.813)	.078 (1.040)	-.096 (-.949)	.0473 (.300)	.2172 (2.408)	.1257 (1.043)
Amount borrowed by male from BRDB	.007 (3.118)	.007 (2.253)	-.0155 (-1.788)	.0233 (1.936)	.234 (3.934)	.138 (1.608)	3.8329 (3.340)	.0244 (.426)	.0334 (.141)
Amount borrowed by female from Grameen Bank	.003 (1.400)	.004 (1.765)	.0371 (2.174)	.0432 (4.249)	.232 (4.402)	.195 (3.318)	1.3484 (1.452)	.1989 (3.950)	.1131 (1.317)
Amount borrowed by male from Grameen Bank	.001 (.252)	.001 (.325)	-.0225 (-2.291)	.0179 (1.431)	.125 (1.676)	.096 (1.029)	.3377 (2.386)	-.0603 (-.878)	-.0457 (-.200)
p (women)			-.3897 (-2.056)	-.4809 (-4.657)			-.0168 (-.198)		.1136 (1.325)
p (men)			.2999 (2.998)	-.2060 (-1.432)			-.7656 (-36.311)		-.0148 (-.053)
Observations	4,567	4,567	5,218	5,218	1,517	1,517	1,757	1,757	1,757

NOTE.—Figures in parentheses are asymptotic t-ratios.

(出所 : Pitt and Khandker, 1998: Table 2)

左側の 4 つの欄が家計の消費水準への影響を見たものである。最初の 2 つは受給資格があるとされた人だけを推計の対象とし、(1) OLS と(2) WESML を用いた結果。次の 2 つの式は全ての

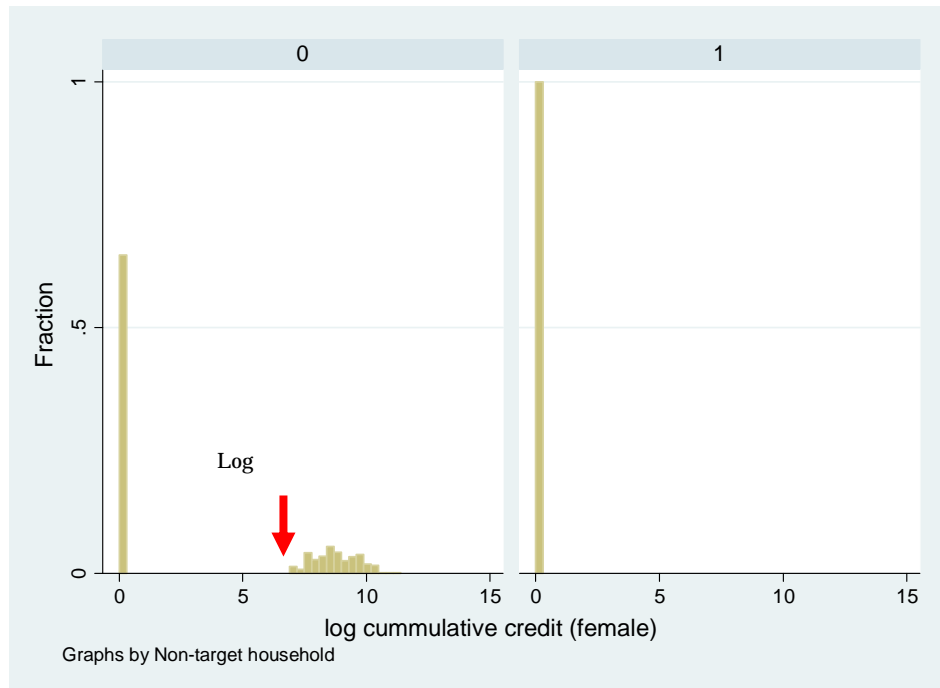
家計を推計の対象とし、(3) 受給資格があるかないかのダミーを加え、(5a)(5b)(5c)式を LIML で同時推計したものと、(4) さらに村固定効果を加えた結果である (括弧内は t 値。ρ は(5a)(5b)式と(5c)式の誤差項の相関度合いを見たもの)。

WESML-LIML-FE (第 4 欄) によれば、BRAC, BRDB, Grameen Bank とも女性への融資の係数は正で統計的に有意な一方、男性については係数は正だが女性よりも値が小さく、また統計的には有意ではない。この推計値をもとに BRAC, BRDB, Grameen Bank を合算すると、1 タカ of MC 融資が女性に対して行われると約 0.18 タカの家計消費上昇が見込まれるのに対し、男性に対して行われると 0.11 タカの上昇にしかない。また ρ が女性も男性も負の数値をとっていることから、MC の借入では self-selection 問題が発生しているものの、もともと村内でより貧しい家計が MC に参加しているという結論を導いた。

WESML-LIML 推計 (第 3 欄) については、女性に対する融資効果が減少すること、男性の場合、融資効果が負で有意になることもあること、また男性の ρ が正に変換することなどをもって、固定効果を入れない (Placement bias を考慮しない) 場合のペナルティについて論じている。

- Roodman and Moduch (2009)のおさらい

RM 論文は上記 PK の結果の妥当性を調べたものである。当時の論文作成に使われたデータがそのままの形では現存していなかったため、素データからデータの再構築を図り、推計結果の再現性を試みた。まずデータ再構築の面では、当初 PK 論文のサマリー統計に記載されていた数値と若干異なる点があったが、主要変数である家計消費額、各 MC 団体からの融資額、その他家計の特徴に関する変数など概ね PK 論文と同様の平均値と分散をとることが判明した。また、PK の一人である Mark Pitt から一部データを譲り受け (Pitt 自身も PK で利用したデータを完全な形式で保有していなかったため、そのデータが PK 論文に使われたものかどうかは確実とは言えないというコメントをしている)、その変数と再構築データの相関係数を調べたところ、ほぼ全ての主要変数で 0.95 以上と高い数値を示した。



(出所 : PKexp.dta から筆者作成)

RM は再構築されたデータを用い、まず(5a)(5b)式で表わされる MC 資金需要の閾値 (censoring threshold)をいくりに設定するか検討した。観察データでは、MC 融資額 C_{ij} が 1000 タカ以下の家計はいないため、受給資格のあるグループでは、全く借入をしなかった ($0 = \log 1$ の値をとる) 家計と、 $\log 1000$ タカ以上の借入をした家計がおり、受給資格がないグループは全員の借入額が $\log 1$ となる上図のような分布となっている (図の左側が受給資格のあるグループ、右側がないグループ)。RM は推計で使う MC の閾値を $\log (1000)$ と当面定め、潜在需要 C_{ij}^* が $\log (1000)$ タカ以上の場合に観察データの通り C_{ij} の値をとり、それ以下の場合には C_{ij} が閾値の $\log (1000)$ をとると想定し、 $\log 1$ を閾値とした場合に結果が変化するか頑強テストを実施している。つまり、彼らの想定は $\log 1000$ を閾値とした場合には下記の通り、

$$\begin{aligned}
 C_{ij}^* &= X_{ij} \beta_c + v_j + \varepsilon_{ij}^c \\
 C_{ij} &= C_{ij}^* \text{ if } C_{ij}^* > \log 1000 \\
 C_{ij} &= \log 1000 \text{ otherwise}
 \end{aligned}$$

となり、 $\log 1$ を閾値とした場合は、

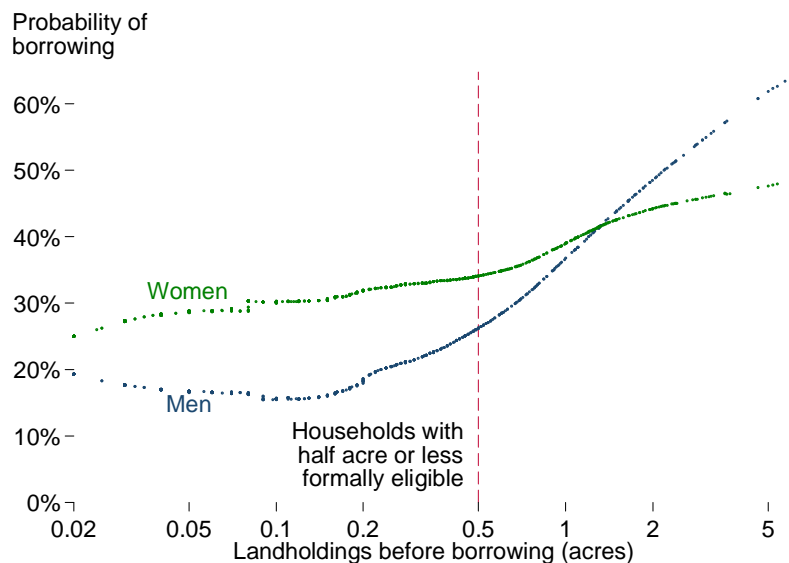
$$C^*_{ij} = X_{ij}\beta_c + v_j + \varepsilon_{ij}^c$$

$$C_{ij} = C^*_{ij} \text{ if } C^*_{ij} > \log 1$$

$$C_{ij} = \log 1 \text{ otherwise}$$

となる。

また、受給資格がないというダミー変数は MC が実施されている村に居住しているかに関わらず、土地所有面積 0.5 エーカー以上であり、それ以外は 1 をとることが基本になっているものの、受給資格ルールが MC 機関や村内で厳格に守られておらず、土地面積が 0.5 エーカー以上でも、借入ができた場合には、0 をとる（受給資格ありと看做される）よう PK 論文では調整がなされている。そのため、0.5 エーカーを境に急激に MC の借入の可能性が変化することがないことを RM は指摘している。この問題は Morduch (1998)による PK 論文への反論の重要な根拠の一つであったが、RM ではそれほど大きく問題視されなかった。



(出所 : Roodman and Morduch, 2009: Figure1 の Replication.)

RM は、再構築されたデータから、PK が使用したものと全く同じと考えられる推計方法で、PK の結果の再現性を試みた。プログラムとしては Roodman が作成した `cmp` と呼ばれるパッケージを使っている。(5a)(5b)式の閾値を $\log 1000$ とし、(5c)式では受給資格があるとされた人だけを推計の対象とし、(1) OLS、(2)LIML(FE の代わりに村の特徴を変数として入れたもの)、

(3)LIML -FE を用いて推計した結果と、全ての家計を推計の対象とし、(4) OLS, (5) LIML(FE の代わりに村の特徴を変数として入れたもの)、(6)LIML -FE (サンプルウェイトは調整済み) を推計した。(6)が PK 論文でいうところの WESML-LIML-FE である。

Table 3. Estimates of impact of cumulative borrowing on log per capita household consumption, PK estimators, target households only

	Target households only			All households		
	LIML, controlling for HH characteristics and...			LIML, controlling for HH characteristics and...		
	OLS	Village characteristics	Village fixed effects	OLS	Village characteristics	Village fixed effects
Log female borrowing from BRAC	0.034 (2.468)**	0.008 (0.156)	-0.021 (0.398)	0.016 (1.163)	-0.107 (2.586)***	-0.103 (2.696)***
Log male borrowing from BRAC	0.042 (2.131)**	-0.010 (0.219)	-0.007 (0.165)	0.024 (1.194)	-0.022 (0.437)	-0.000 (0.007)
Log female borrowing from BRDB	0.016 (0.928)	-0.011 (0.187)	-0.029 (0.433)	-0.007 (0.397)	-0.142 (2.926)***	-0.146 (2.940)***
Log male borrowing from BRDB	0.036 (3.139)***	-0.016 (0.501)	0.032 (0.915)	0.022 (1.962)**	-0.035 (0.871)	0.005 (0.100)
Log female borrowing from Grameen	0.017 (2.324)**	-0.006 (0.151)	-0.015 (0.356)	0.001 (0.101)	-0.099 (3.200)***	-0.087 (3.116)***
Log male borrowing from Grameen	0.000 (0.017)	-0.041 (1.384)	-0.008 (0.240)	-0.017 (1.452)	-0.052 (1.491)	-0.012 (0.313)
Observations	4,567	4,567	4,567	5,218	5,218	5,218
Log pseudolikelihood	-2054.73	-6261.57	-5842.03	-2683.09	-7227.65	-6711.62

(出所 : Roodman and Morduch, 2009: Table3)

結果として、ほぼ同じデータ、ほぼ同じ推計式である第 6 欄において、驚くべきことに、BRAC や BRDB、Grameen Bank からの女性に対する融資は、家計消費水準に有意かつマイナスの影響をもたらすことが判明した。

RM は PK との結果の違いについて

“The sharp contradiction of PK’s headline result poses a mystery

(PK の主要結果と明らかに正反対する結果は不可解である)”

と自ら述べ、(5a)(5b)式の閾値を $\log 1$ にしたり、Pitt から借りたデータを使って再推計してみたが、やはり結果は変わらず、女性に対する融資は消費にマイナスの影響となっていると報告している。

また、PK 論文が使った実証戦略が因果関係を構築するのに有効であったか調べるために、(5a) (5b)(5c)式を LIML の代わりに 2SLS に置き換えて検証した。RM の狙いは、MC 受給資格

の有無が真に外生変数であり、受給資格とその他説明変数の交差項の行列が消費決定式の誤差項と直交（無相関）であるならば、それを **2SLS** の枠組みの中で検定できるだろうということである。**RM** は識別変数の妥当性を検証するために、**Sargan** 検定と **Hansen-J** 検定を行った。**Sargan** 検定は誤差項が互いに独立で均一の分散に従う場合に有効な検定方法であるが、**RM** は **PK** の推計式が閾値のある **Tobit** モデルに準じるものであるため、誤差項の均一分散が暗に仮定されているのではないかと、また、各調査ラウンド間での家庭内の誤差項の系列相関はありうるものの、一貫性のある推計を行うためには各調査ラウンドにおいて家計間の誤差項が互いに独立であると仮定されていなければいけないのではないかと考えた。そのため、**Sargan** 検定は、調査ラウンド間の系列相関がないと想定される、各調査ラウンドの識別テストのみに用い、**Hansen-J** 検定は各調査ラウンド及び三つの調査ラウンドのデータをプールしたものに用いている。

結果から、**Sargan** 検定においては、各調査ラウンドの操作変数行列は誤差項と直交という帰無仮説は強く棄却された。また、全ての操作変数行列を用いた **Hansen-J** 検定においても結果は同様であった。そこで、**Hansen-J** 検定を通った操作変数行列のみを用いて、再度 **Sargan** 検定を試してみたところ、やはり多くの場合において、操作変数行列が誤差項と直交という帰無仮説は棄却された。この結果から、**RM** は誤差項が互いに独立で均一の分散という仮定が満たされていないか、或いはそれが満たされていたのだとしたら、識別戦略が誤っていたかのいずれかであり、いずれの立場をとっても **PK** 論文が **MC** と消費水準の上昇の因果関係を構築したことにはならないと主張した。

RM はデータ構築プロセス、推計コード、推計に使った **cmp** パッケージの全てを一般公開し、第三者がそれを検証できるようにした。

3. Pitt の反論

RM 論文が出された 2009 年には、**RCT** を用いた論文でも相次いで **MC** は生活改善にそれほど大きな貢献をしていないという結果が出されていたこともあり（例えば **Banerjee et al. 2009**, **Karlan and Zinman 2009**）、**RM** の論文は信憑性を持って受け止められた。開発経済学の教科書とも言える **Handbook of Development Economics** でも引用されている。**RM** は論文の中で、**MC**

が本当に負の影響を及ぼすような害のあるものであるとは述べてはおらず、PK の論文によって正の因果性を確立することは困難であると述べているにとどまっているが、貧困削減のためのツールとして MC に過大な期待を寄せすぎてはいけないというムードがこの頃から経済学界の一部に出来上がり始めた。一方、2010 年インドのアンドラ・プラデシュ州で、MC による多重債務者が自殺に追い込まれるケースが多くなったという噂から、州政府が MC の運営を大幅に規制し、債務不履行が広がっていった。このインドのマイクロファイナンス危機において、開発経済学の著名学者達が共同で” **Microfinance is not the enemy** (マイクロファイナンスは敵ではない (Banerjee et al.2010))” という記事を執筆したことにも象徴されるように、MC 自体を悪と考える学者は少ない。MC はある特定の目的には有効であるが、生活全般において有効なわけではない、というのが大方の意見であろう。

そうした中、RM 論文が発表されて 2 年経過した 2011 年、Pitt が自身のウェブページ上において、RM 論文の欠陥を指摘する 2 つの論文を掲載した。以下、その中身を吟味したい。

- Pitt (2011a)

“Response to Roodman and Morduch’s “The Impact of Microcredit on the Poor in Bangladesh: Revisiting the Evidence.”” (Pitt, 2011a)で Pitt が問題にしたのは主に次の点である。RM の再推計は PK を完全に踏襲しておらず、(1) MC 需要の閾値の設定が間違っており、かつ(2) 重要な説明が(5c)式の中から抜けている替りに誤った変数が入れている。(3) これらを是正すると PK の結果は Roodman が作成した `cmp` パッケージでも再現される。

Pitt (2011a)は、(個人的な印象であるが)かなり怒りに満ちている論文である。RM 論文が発表されて以来、PK やその弟子達が書いた論文は信憑性に欠けるという烙印を押され、MC の効果が疑問視されることが当然となってきた。しかし、間違っているのはそもそも RM だったのではないかという問題提起の仕方である。

MC の閾値については、既述の通り、PK の設定はやや特殊である。MC 潜在需要の変数 C_{ij}^* が閾値である $\log 1000$ 以上であるならば観察値 C_{ij} となるが、 C_{ij}^* が $\log 1000$ 以下であるならば

log 1000 ではなく log 1 をとると想定されている。これは通常の Tobit モデルの打ち切り (left censoring) と異なる。また、Pitt (2011a)によれば、PK は各家計が MC を借り入れる可能性があったかという事実を重視し、MC 受給資格のない家計や MC が展開されていない村の家計が正の C_{ij} 値をとる可能性を省いている。これらを見逃し、潜在変数 C_{ij}^* が閾値である log 1000 以下の全ての家計に $C_{ij} = \log 1000$ を与えてしまった RM の計算方法は、資金需要がなかったり、借入機会がそもそもない家計にも 1000 タカを与えた上で、MC の効果を測定しようとしている時点で大きな誤りであると指摘した。

また、(5c)式の中には、MC 受給資格がないというダミー変数が入れているが、RM ではこの変数が使われておらず、代わりに MC 実施村に住み MC 受給資格がありながらも借入を行わなかった家計が 1 をとるようなダミー変数が入れていた。これは PK の識別戦略と全く異なるということである (Pitt (2011a)では詳しく触れられていないが、(3)(4)式に倣うと、プログラムに参加したかしないかはそれ自体内生変数のため、誤差項と相関が生まれる可能性が高く、 C_{ij} の効果をバイアスなく推計できないということなのだろう)。

Pitt (2011a)では RM が公開した cmp パッケージを使った推計式に、MC 閾値とダミー変数の 2 つの修正を加え、Pitt が Roodman に貸し出したデータと RM が再構築したデータを用いて再推計を試みている。その結果、いずれのデータを使っても、女性への貸出は家計の消費水準に有意な効果をもたらすが、男性への貸出は有意な効果をもたらさないという PK オリジナルの結果が再現されることを示した。RM 論文にはこの 2 点の致命的誤りがあったために結果が再現されなかったのであって、

“The “sharp contradiction of PK’s headline result” (RM, p.23) is no longer a
mystery

(PK の主要結果との明らか不一致はもはや不可解ではない) ”

と Pitt (2011a: p12)は述べている。

- Pitt (2011b)

続いて”Overidentification Tests and Causality: A Second Response to Roodman and Morduch” (Pitt, 2011b)と題された論文で、Pitt が議論しているのは、RM の過剰識別検定の妥当性についてである。WESML-LIML-FE 推計を 2SLS に置き換え、識別変数である MC 受給資格と MC 実施村ダミー（村固定効果）を他の説明変数と交差させることで、一致推計が得られることは PK 論文脚注 16 でも述べられている。Pitt (2011b)の主な主張は、RM が議論の拠り所とした Sargan 検定は明確に誤りであり、かつ Hansen J 検定の妥当性も疑わしいというものである。

Pitt(2011b)が Sargan 検定に対して触れているポイントは次の 2 点に集約される。第一に、Sargan 検定は誤差項が互いに独立で均一の分散に従う場合に有効な検定方法となる。しかし、PK 論文で用いたデータは MC 受益者を意図的に多くサンプルに含む調査デザインであったため、その調整としてサンプルウェイトを考慮した WESML を使っている。これは、既に誤差項の不均一分散が仮定されていることになり、Sargan 検定は意味のある検定とはならない、というものである。第二に、PK では MC 受給資格のない家計や MC が展開されていない村の家計が正の C_{ij} 値をとる可能性を意図的に省いており、彼らの C_{ij} の分散は 0 になる。RM の cmp プログラムでもそれが考慮されているが、2SLS では一部の家計の C_{ij} の分散が確定的に(deterministically) 0 になることを許容して第 1 段階の MC 需要式を推計することできない。その結果、2SLS の誤差項は i.i.d ではなくなり、Sargan 検定は決して有効とはならない、というものである。

これらの理由から Pitt(2011b)は、過剰識別制約の検定としては Hansen J 検定しか有効にはなりえないと主張している。ただし Hansen J 検定においても、一部の家計の C_{ij} の分散が確定的に 0 である場合に正しい検出力を持つのかどうか、また、例え Hansen J 検定が有効だったとしても RM が行った 2SLS は男女別に参加資格があるかどうかを考慮していなかったため、厳密な意味で PK の識別戦略が正しかったかどうかを試していない、と批判した。

その上で、Pitt (2011b)は、Pitt (1999)で示した通り、MC 受給資格の代用となる土地保有面積、村ダミーとその他全ての説明変数の交差項を加え、識別変数行列が消費決定式の誤差項の分散と相関しているのか検定した。cmp パッケージに基づく、識別制約に対する F 検定による p

値が **0.794** になり、識別変数行列が消費決定式の誤差項に対して説明力を持たないという帰無仮説を棄却することはできなかった。この結果をもとに、**Pitt (2011b)**は **PK** の識別戦略は妥当であり、「観察データから因果性を確立することは困難である」という **RM** に反し、**PK** は「観察データから因果性を立証したと考えられる」と主張した。

Pitt は、**RM** に倣い、**Pitt (2011a, 2011b)**で利用したデータ、推計コードを一般公開し、第三者が結果を再検証できるようにした。

4. その後の動向

Pitt の反論以降、**PK** 論文と **RM** 論文の妥当性をめぐり、様々な意見がウェブ上で交わされるようになった。その拠点となったのが、世銀が最近開設したインパクト評価に関するブログと **Roodman** が個人的に開設していたブログである。

この問題を世銀のブログで最初にとりあげたのは、**Pitt** とも数本の共著論文を執筆したことのある **Martin Ravallion** である。彼は“The Microfinance Mystery”(Ravallion, 2011)と題された記事において、**PK** 論文と **RM** 論文の概略を整理し、**PK** と **RM** が同じデータで同じ推計方法を用いて、異なる結果になるというのはいかなる理由もないため、どちらかが間違っているとしか考えられず、その間違いは **RM** にあったとして、**Pitt (2011a)**を紹介した。

Ravallion の記事にいち早く反応を示したのが **Roodman** である。**Roodman** は記事へのコメント欄に、**Pitt (2011a)**の全てを許容することはできないが、確かに **RM** 論文には明確な誤りがあったらと書き込んだ。また、**RM** が目指したものは、**MC** が負の影響を及ぼしていることを立証することではなく、**PK** の識別戦略が正しかったか、それによって因果性が確立できたと言えるのかという点であることを改めて表明した。

これに対し、**Ravallion** は、問題は **PK** の識別戦略が妥当であったかという点にあるのではなく、**RM** が提示した **PK** と反対の結果が広く受け入れられている点にあるとし、嫌悪感を示している。

Pitt (2011a)に対する **Roodman** の最初の公式な反応は彼のブログ上で行われた (**Roodman, 2011a**)。この記事では、**Pitt (2011a)**に従い、**MC** 閾値の変更と **MC** 受給資格の有無を(5c)式に入

れ、PK と RM の推計の再検証をしている。その結果、次の表の通り、係数の推計値や符号はもととも PK で報告されたものとほぼ同様になり、PK の結果が再現され、RM の結果は再現されないことが確かめられた。

Estimates of impact of credit on log per capita household consumption, following Pitt & Khandker 1998, using Pitt consumption variable and "current students=0" education variables

Explanatory variables	OLS		Weighted OLS		WESML-LIML		WESML-LIML-FE		2SLS
	PK	New	PK	New	PK	New	PK	New	
Log amount borrowed by female from BRAC	0.007 (3.048)***	0.007 (3.018)***	0.007 (2.847)***	0.007 (2.642)***	0.034 (2.291)**	0.035 (2.547)**	0.039 (4.237)***	0.031 (2.355)**	0.009 (0.530)
Log amount borrowed by male from BRAC	0.010 (2.906)***	0.010 (2.862)***	0.010 (2.835)***	0.010 (2.684)***	-0.016 (1.658)*	-0.002 (0.148)	0.019 (1.593)	0.020 (1.547)	0.034 (1.452)
Log amount borrowed by female from BRDB	0.002 (0.573)	0.001 (0.557)	0.003 (0.906)	0.003 (1.067)	0.026 (1.723)*	0.028 (2.238)**	0.040 (3.813)***	0.031 (2.147)**	-0.024 (0.765)
Log amount borrowed by male from BRDB	0.007 (3.118)***	0.007 (3.066)***	0.007 (2.253)**	0.007 (2.802)***	0.016 (1.788)*	-0.004 (0.370)	0.023 (1.936)*	0.022 (1.810)*	0.001 (0.060)
Log amount borrowed by female from Grameen Bank	0.003 (1.400)	0.003 (1.355)	0.004 (1.765)*	0.004 (1.887)*	0.037 (2.174)**	0.034 (2.202)**	0.043 (4.249)***	0.035 (2.393)**	0.010 (0.753)
Log amount borrowed by male from Grameen Bank	0.001 (0.252)	0.001 (0.197)	0.001 (0.325)	0.001 (0.281)	-0.023 (2.291)**	-0.014 (1.093)	0.018 (1.431)	0.017 (1.317)	-0.002 (0.065)
Log amount borrowed by female from all programs									
Log amount borrowed by male from all programs									
ρ (women)					-0.390 (2.056)**	-0.396 (2.289)**	-0.481 (4.657)***	-0.382 (2.260)**	
ρ (men)					0.300 (2.998)***	0.187 (1.230)	-0.206 (1.432)	-0.193 (1.206)	
Hansen, all instruments (p value)									0.022
Observations	4,567	4,567	4,567	4,567	5,218	5,218	5,218	5,218	5,218

Absolute t statistics in parenthesis, clustered by household except in OLS. *significant at 10%. **significant at 5%. ***significant at 1%.

(出所：Roodman(2011a))

しかし、上表の最後の欄に見られるように、WESML-LIML-FE を 2SLS に置き換えた場合、MC の係数は小さくなり統計的に有意ではなくなる、また Hansen-J 検定では識別変数行列が誤差項と直交(無相関)という仮説は棄却される。そこで RM で行った様に、再び調査ラウンド毎の識別テストに Sargan 検定を用い、各調査ラウンド及び三つの調査ラウンドのデータをプールしたものに Hansen-J 検定を用いてみると、やはり多くの場合で識別変数行列が誤差項と直交という帰無仮説は棄却された。このことから、Roodman(2011a)は Pitt (2011a)によって示された指摘は確かに正しく、PK の結果は再現されるが、それによってむしろ、PK が因果関係を構築するに

至っていないという **RM** の主張が強化された面もある、としている（当初はこうした主張の展開であったが、サンプルウェイトを考慮した推計では **Sargan** 検定が意味を持たず、それに基づいた判断は誤りであることが後に追記されている）。

また、**Roodman** のブログから明らかになっていったことは、**PK** の結果を再検証する際に、データと推計コードを公開することがいかに重要であったか、という点である。**RM** は当初から **Pitt** 及び **Khandker** に **PK** 論文の基となったデータとコードをそのままの形で使わせてもらえるか打診をしていた。**Pitt** はそれに対応し、データを貸したが、注釈付きで、**PK** で使われたものと全く同じであるかはわからないと言ったようである。データ構築のクレームは国際食糧政策研究所(IFPRI)のシニア・リサーチャーである **Maren Duvendack** からなされ、彼女が **PK** 論文を別の角度から検討しようとした時にもデータが完全な形で公開されていなかったことから苦労したことが述べられている。

ここからデータと推計コード公開を巡り、いくつかの議論がわき起こった。**Ravallion** が「**RM** の間違いによって **PK** が汚名を着させられ、**MC** の権威が失墜した。**RM** は結果を公表する前にもっと自らの推計を吟味しなければならなかったのではないか」という指摘を行うと、数人の投稿者から「**RM** が間違いを犯してしまったのは、そもそも **PK** がデータとコードを公開していなかったからであり、その責任の一部は **PK** にもあったのではないか」「もし **PK** がデータと推計コードを公開していれば、**RM** の間違いはもっと早く気付かれたのではないか」という疑問が寄せられた。**Pitt** からは当時のやり取りを克明に綴った投稿がなされ、**RM** が結果を再検討できるようにできる限りの協力をしてきたし、**Duvendack** についても同様の協力を行った様子が伝えられた。

そうした中、**Roodman** から、さらなる検証結果が報告された。これは **Pitt (2011b)** に対応するものである。**Roodman(2011b)** は記事の中で、**Pitt (2011b)** の指摘通り **Sargan** 検定は確かに妥当ではないことを認めながらも、**Pitt (2011b)** で示された **F** 検定は、**MC** 受給資格ダミー変数を識別変数として取り扱ったものではなく、土地面積を使っているものなので、**PK** オリジナルの推計とは意味が異なることを指摘した。そして、**MC** 受給資格ダミー変数を識別変数とすると、識別変数行列が誤差項に対する説明力を持っており、**PK** の識別戦略は妥当ではなかった可能性が

高いと述べている。また、Roodman自身が再推計したRoodman(2011a)の結果で、2SLSで頑強な結果が導けなかった原因として、識別変数行列がそもそも C_{ij} と弱相関にあった可能性について触れた。さらに、受給資格がないというダミー変数を変更し、土地保有面積0.5エーカー以上の家計の全てが1をとるようにデータを作り（繰り返しになるがPKでは土地保有面積0.5エーカー以下の家計+土地保有面積に関係なくMC融資があった家計を受給資格ありとみなしていた）、それを入れて推計をやり直すと、Morduch(1998)の指摘通り、MCと消費の関係は負になりうる可能性を指摘した。

5. まとめ

Roodmanの再度の反論以降、今のところ大きな動きはなく、この問題は完全決着にまでは至っていない。特にRoodmanが主張するように、PKが因果性を立証できたといえるかどうかについてはまだ議論の余地がありそうである。

経済学のジャーナルでは、因果性を厳密に立証することがますます求められており、因果性の特定の根拠が弱いと査読に通らなくなっている。今回の一連の議論は、観察データを用いて、「科学的なエビデンス」を提供することがいかに難しいことであることを示している。しかしこれは、分析上の工夫をそれほど必要としない、もっと単純でシンプルなRCTのみに依拠すればよい、ということでは決してない。Ravallionがブログの中のコメントで述べていたように、RCTには向かないが、現実に解決しなければいけない重要な研究課題はいくつもある。手法の優劣が研究課題を決めるのではなくて、そうした問題意識が研究課題を決めていかなければいけない。

では、現実に問題があり、それを検証する観察データが揃えば、何を言ってもいいのかといえば、もちろんそれも違おうだろう。再検証不可能な言説を、専門家本人の信念に基づいて、あたかも本当のように広めてしまうのは、時として有害ですらある。この点に関し、疫学の分野では科学的エビデンスのグレードが定められており、専門家個人の意見は一番低いグレードに分類されている (<http://www.jsh.or.jp/medical/liver/level%20classification.htm>)。

科学的エビデンスの質を第三者が再検証できるようにするためには、データを公開すること

も有意義だろう。例えば世銀や IFPRI では家計データの収集後、一定期間がたったら、全ての素データを公開する方針を持っている。日本のアカデミック機関では、政策研究大学院大学(GRIPS)の大塚啓二郎教授の研究グループが、やはり「収集データは公共財である」という意識のもと、数々の家計データを公開している。一次データを公開することは研究者にとって時として痛みを伴う。それは、第三者が自分の実証結果を検証しやすくなるからではない。時間と労力を割いて一生懸命集めたデータを無料で公開するのは、他人のフリーライドを許してしまうために、苦痛なのである。しかし、世銀の **Living Standard Measurement Survey** や USAID の **Demographic and Health Surveys** などの家計データに多くの研究者がフリーライドできることで、開発経済学におけるマイクロ実証分析が急速に進展してきたことも忘れてならない。今回、PK を巡り、PK の推計結果が再現されうることや、RM の推計がどの点で間違っていたかを、第三者でも検証できるように、Pitt や Roodman が互いにデータと推計コードを公開したことは、経済学の進展、MC への理解促進の上でも意義は大きい。

一方、今回の議論では推計の技術的な問題にのみ焦点があてられていたが、観察データを使った経済学の実証研究では、データが扱っている地域の現実をきちんと反映しているのか、代替的に想起される仮説があったとすればそれが成り立たない十分な証拠があるかどうかによって、論拠の妥当性を判断することも重要な作業である。例えば MC の議論で言えば受給資格ルールが厳密に守られていたのかというのは、依拠する仮定が地域の実態に沿っていたか判断する上で重要な情報であった。さらには、どれだけ厳密な方法であっても、一つの実証結果が他の文脈に当てはまるとは必ずしも言えないということを忘れてはならない。どのような文脈で何がどのようにうまくいくのか、あるいはいかないのかを判断をする際には、地域の様々な定性的な情報を集めることや、異なる文脈での研究結果を蓄積することが大事であろう。iv

経済学が科学たりうるためには、客観的かつ再現性の高いものでなければいけない。また、ジャーナルに掲載されるためには、新規かつ厳密でなければならない。しかし、ジャーナルに掲載されることを求めるがあまり、理論や仮説検証方法の妥当性、新奇性のみに注意が払われ、現場の状況が無視されることがあってはならない。いかに再検証可能な形で現場の実態に沿ったエビデンスを構築していくか。そうしたことを考えられる議論であった。

参考文献

高野久紀・高橋和志[2011]「マイクロファイナンスの現状と課題－貧困層へのインパクトとプログラム・デザイン」『アジア経済』近刊予定。

Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan. [2009]. “The Miracle of Microfinance? Evidence from a Randomised Evaluation.” Poverty Action Lab Working Paper 101.

Banerjee, Abhijit, Pranab Bardhan, Esther Duflo, Erica Field, Dean Karlan, Asim Khwaja, Dilip Mookherjee, Rohini Pande and Raghuram Rajan. [2010]. “Microfinance is not the enemy.” Financial Times Dec.13.

(<http://www.ft.com/cms/s/0/53e4724c-06f3-11e0-8c29-00144feabdc0.html#axzz1MMSrYkB6>)

Coleman, Brett E. [1999]. “The Impact of Group Lending in Northeast Thailand.” *Journal of Development Economics* 60 (1): pp.105–141.

Karnofsky, Holden (2011). “Microlending debate: an example of why academic research should be used with caution.”

<http://blog.givewell.org/2011/05/13/microlending-debate-an-example-of-why-academic-research-should-be-used-with-caution/>)

Karlan, Dean and Jonathan Zinman. [2009]. “Expanding Microenterprise Credit Access: Using Randomized Supply Decisions to Estimate the Impacts in Manila.” BREAD Working Paper No. 235.

Khandker, Shahid. [2005]. “Microfinance and Poverty: Evidence Using Panel Data from Bangladesh.” *World Bank Economic Review*, 19 (2): pp.263–286.

Morduch, Jonathan. [1998]. “Does Microfinance Really Help the Poor? New Evidence from Flagship Programs in Bangladesh.” Working Papers 198, Princeton University.

Pitt, Mark M. [1999]. “Reply to Jonathan Morduch's “Does Microfinance Really Help the Poor? New Evidence from Flagship Programs in Bangladesh” Mimeograph. Brown University.

Pitt, Mark M. [2011a]. “Response to Roodman and Morduch's “The Impact of Microcredit on the Poor in Bangladesh: Revisiting the Evidence.”” Mimeograph. Brown University. 推計データやコードは(<http://www.pstc.brown.edu/~mp/>)から入手可。

Pitt, Mark M. [2011b]. “Overidentification Tests and Causality: A Second Response to Roodman and Morduch.” Mimeograph. Brown University. 推計データやコードは(<http://www.pstc.brown.edu/~mp/>)から入手可。

Pitt, Mark M. and Shahidur R. Khandker. [1998]. “The Impact of Group-Based Credit

Programs on Poor Households in Bangladesh: Does the Gender of Participants Matter?"

Journal of Political Economy, 106 (5): pp.958–996.

Ravallion, Martin. [2011]. "The Microfinance Mystery."

(<http://blogs.worldbank.org/developmenttalk/the-microfinance-mystery#comment-219>)

Roodman, David and Jonathan Morduch [2009] "The Impact of Microcredit on the Poor in Bangladesh: Revisiting the Evidence." CGIAR Working Paper, 174.推計データやコードは

(<http://www.cgdev.org/content/publications/detail/1422302>)から入手可。

Roodman, David. [2011a]. "Response to Pitt's Response to Roodman and Morduch's Replication of...

etc."(http://blogs.cgdev.org/open_book/2011/03/response-to-pitts-response-to-roodman-and-morduchs-replication-of-etc.php)

Roodman, David. [2011b]. "A Somewhat Less Provisional Analysis of Pitt & Khandker."

(http://blogs.cgdev.org/open_book/2011/04/a-somewhat-less-provisional-analysis-of-pitt-khandker.php)

ⁱ 今でこそ JPE は、分析結果の質と再現性が担保されるよう、推計に利用したデータおよび推計コードを公開することを義務付けているが、当時はまだそうした体制は整えられていなかった。

ⁱⁱ C_{ij} に影響を与える観察不可能な要因を取り出し、それを直接的に制御するのではなく、MC 借入の機会があり、実際に参加しているかどうかを推計式に含めるやり方は Coleman (1999)に近いと言える。しかし Coleman (1999)の場合、MC が既に実施されている村、MC がこれから実施される村において、それぞれ MC 借入を行いたいと表明した人をダミー変数として推計式に含める Difference in Difference に基づいており、Self-selection についてより直感的にわかりやすく、かつクリーンな識別方法を用いていると思われる。

ⁱⁱⁱ WESML-LIML-FE 推計をより詳しく見ると

(5a)(5b)式は性別資金需要に分けられ、

- 被説明変数が男女毎の log MC 借入額 (借入額がない場合は 0)
- 説明変数が [1] HH Head の親族 (両親、兄弟、姉妹) と Spouse の親族 (両親、兄弟、姉妹) で土地を所有している人数、[2] log 家計が保有する土地の価値、[3] HH Head の教育年数、

性別、年齢、[4] 家計内男性成人メンバーの最高教育年数と女性成人メンバーの最高教育年数、[5]成人男子、成人女子、Spouse が一人でも家計内にいるかのダミー、[6](5a)の被説明変数が男性資金需要の場合、男性対象に MC が展開されているかの村ダミー、女性資金需要の場合、女性対象に MC が展開されているかの村ダミー

(5c)式は

- 被説明変数が \log 家計一人当たり消費
- 説明変数が[1] (5a)(5b)式から導かれる、Grameen Bank, BRAC, BRDB の男女別の \log 資金需要 (正の下限は $\log 1000$, それ以下は $\log 0$)、[2] 受給資格が無いというダミー、[3] HH Head の親族 (両親、兄弟、姉妹) と Spouse の親族 (両親、兄弟、姉妹) で土地を所有している人数、[4] \log 家計が保有する土地の価値、[5] HH Head の教育年数、性別、年齢、[6] 家計内男性成人メンバーの最高教育年数と女性成人メンバーの最高教育年数、[6]成人男子、成人女子、Spouse が家計内にいないダミー、[7] 調査ラウンドダミー、[8] 村ダミー

^{iv} これらの論点については、有本寛一橋大学准教授や同僚の高野久紀氏とのインフォーマルな会話から触発されている部分が多い。記して感謝したい。ただし、有りうる誤りの全ての責任は筆者個人に帰される。