

自然実験環境を用いた留学効果の推定
—Xプログラムを事例にして—

中尾走・樊怡舟・西谷元・村澤昌崇

RIHE Advancement of Higher Education Research

Monograph Series

No. 3

自然実験環境を用いた留学効果の推定

—X プログラムを事例にして—

中尾 走

(広島大学教育学研究科・日本学術振興会特別研究員 (DC2)・広島大学 スーパーグローバル大学創成支援事業データ分析チーム)

樊 怡舟

(広島大学教育学研究科・広島大学 スーパーグローバル大学創成支援事業データ分析チーム)

西谷 元

(広島大学人間社会科学研究科・広島大学 スーパーグローバル大学創成支援事業データ分析チーム)

村澤 昌崇

(広島大学高等教育研究開発センター・広島大学 スーパーグローバル大学創成支援事業データ分析チーム)

1. 序論

本研究の目的は、A 大学で実施されている短期留学プログラムの効果（以下、留学効果）を自然実験（*natural experiment*）¹⁾ の環境を用いて明らかにすることである。

昨今、政策立案全般に対して、根拠に基づく政策立案（*Evidence-based policy making*, 以下 *EBPM*）が求められており、それは教育の分野も例外ではない。もちろん従来からも学者の専門知識を“根拠”に政策が立案されてはいたが、それらと *EBPM* の違いは、“根拠”により厳密な因果効果、すなわち反実仮想のフレームワークに基づき、交絡要因を統制した上で因果効果の推定を根拠とすることを至上としている点にある。

また、高等教育の文脈では *IR* (*Institutional Research*) の重要性が強調されており、大学にも教育効果・アウトプット・アウトカムの測定が今まで以上に求められている。例えば、留学等をはじめとした各種取り組みとその効果検証も *IR* 業務の一環であろう。留学とその効果分析については一部関係者により取り組まれてきたが、特に近年は短期留学が多くの大学で実施されており、その因果推論を用いた効果検証も行われはじめている（日本学生支援機構 2017；川田・西谷 2017）。しかしいずれも交絡要因の統制が困難であるという課題を抱えていた。

なぜなら、教育分野を含む社会科学全般では、交絡要因が多く、調査によっては未観測である場合も多いからである。特に交絡要因が、観測しづらい構成概念の場合は、その影響を取り除くことが難しい（中尾ほか 2021）。例えば、留学による語学力向上の効果を分析する

際に、学生が元来保持している「学習態度」や「学習意欲」などの構成概念が重要な交絡要因として想定される。それらの交絡要因を統制しない限り、欠落変数バイアスによって、純粋な留学効果を推定できない。このような課題への対応策としては、構成概念など統制が困難な交絡要因や未観測要因の交絡要因が統制されることが確率的に期待されるランダム化比較試験（Randomized clinical trial, 以下 RCT）が、EBPM の文脈では推奨されてはいる。けれども、RCT の実施には、「子どもを実験台にするのか」（中澤 2016）という倫理面からの批判や費用面が課題であり、簡単には実施することが出来ないという別の課題がある。

以上を踏まえ、本稿では、自然に発生した実験的状況を用いることで交絡・選択バイアスの問題へ対処可能な自然実験の環境を用いて、留学効果の推定を行うこととする。

2. 方法

（1）自然実験の環境

留学効果を検証してきた先行研究では、様々な手法を用いて交絡要因を統制し、留学効果を推定することが行われてきた。けれども、①未観測の交絡要因を統制できていない点、②学習意欲や学習習慣測定が困難な交絡要因についての統制が不十分であることが予想され、留学効果が過大推定されている可能性が課題として残されている。そこで、本稿は自然実験の環境を用いることで、以上で挙げた二つの限界を克服しつつ、留学効果を明らかにすることとしたい。

以下では、本研究の根幹である自然実験の環境は以下の通りである。A 大学では、毎年選抜された 1 年次の学生が X プログラムという短期留学プログラムへと参加する²⁾。参加者の決定については、TOEIC の点数の他に、面接など計 4 つの評価項目を点数化し、総合的に考慮した上で、参加者が選抜されている。X プログラムは、2010 年度から開始され、毎年 8 月から 9 月にかけての前期派遣と 2 月から 3 月にかけての後期派遣の二回に分けて、それぞれおおよそ百名ずつの学生が約 2 週間の海外留学を行う。2019 年度も同様に、百名程度の学生が、前期と後期にそれぞれ海外留学を行う予定であったが、新型コロナウイルスの影響で、後期のプログラムのみが中止となった。この偶然にも 2019 年度後期の X プログラムが中止になったことで、X プログラムへ参加する予定だった学生の中に、X プログラム参加者（前期に参加した学生）と不参加者（後期に参加する予定だった学生）が混在することとなった。つまり、新型コロナウイルスによる中止という外生的な介入によって、X プログラム参加者と不参加者が分けられる自然実験の環境が生じている。

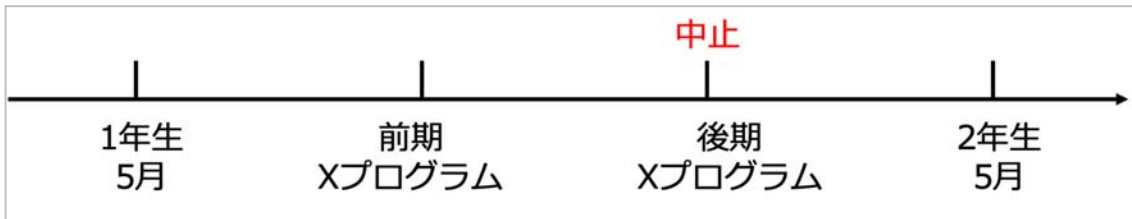


図1 自然実験の環境

(2) 分析手法

A 大学では、学生に対して、1年生の5月に指定受検として全員（本稿で用いるデータの2019年度の受験率は92.6%）にまた2年生の5月には希望受検として、大学が費用負担を行うことで、学生にとっては無料でTOEIC受験が可能な制度を導入している。そのため、1年生の5月と2年生の5月の2回ともTOEICを受験している学生が多く、2時点のパネルデータが得られる。この2回分のTOEICの点数を用いることで、Xプログラムに参加する前後のTOEICの点数の伸びを明らかにすることが出来る。このようなパネルデータを用いた因果効果の推定には、差の差分分析（Difference in Difference, 以下DID）やパネル固定効果モデル（Fixed effect model）が用いられるが、本稿で用いるデータが2時点のパネルデータであるため、DIDを用いることとする。

DIDは、平行トレンド仮定が満たされている場合に、因果効果の推定が可能な手法である。一方で、平行トレンド仮定は、データから確認することは不可能であり、仮定が成立すると思われる証拠をできる限り並べる（伊藤2017:186-7）など、傍証として与えることしかできない³⁾。一般的には、交絡要因やその他の介入によって平行トレンド仮定が満たされないことがあり（Craig et al. 2017）、例えば、川田・西谷（2017）では所属学部と留学前のTOEICの点数を傾向スコア（Propensity Score）として統制することで、これらの条件が同じであれば平行トレンド仮定が満たされるものと見做している。しかしながら、既に指摘した通り、未観測の交絡要因の影響や学習意欲、学習習慣など測定が困難な交絡要因については統制されておらず、平行トレンド仮定が満たされていない可能性が考えられる。実際に、川田ほか（2019）では、Xプログラムに参加が決定することで、学習意欲が増加し、TOEICの点数が上昇することが考察されており、学習意欲など（測定が困難であるが）は統制すべき交絡要因となり得る。

しかしながら、自然実験の環境から得られたデータを用いることで、DIDを行う上で望ましい二つの性質がある。一つ目に、学習意欲や学習習慣、そして未観測の交絡要因がXプログラムに参加したか否かという二つの群で大きく変わらないことが確率的に期待できる。二つ目に、Xプログラム以外の留学プログラムなど、平行トレンド仮定を崩すその他の介入についても、新型コロナウイルスの影響で中止になっていることが考えられるため、その他の介入によって平行トレンド仮定が崩されることがないと見做せる点である。以上の二点を考慮すると、この自然実験の環境を用いることで、平行トレンド仮定が満たされているも

のと想定することが出来る。以上のような利点から、自然実験の環境で得られたデータに DID を用いることで、より純粋な留学効果の推定を行うことが可能となる。

(3) データと変数

データは、A 大学の X プログラムに採択された学生のデータを用いる。そのデータセットの中で、自然実験の環境を用いるために、2019 年度の学生データのみを分析に用いることとする。従属変数は、1 年生 5 月と 2 年生 5 月に受験した TOEIC の点数である。TOEIC は、リスニングとライティングがそれぞれ 495 点ずつの合計 990 点満点のテストであり、リスニング、ライティング、合計の 3 つをそれぞれ従属変数として用いることとする。以下、用いる変数の記述統計量である。

表 1 記述統計量

	受験時期	Xプログラム不参加			受験時期	Xプログラム不参加		
		合計	リスニング	ライティング		合計	リスニング	ライティング
n	1年次	107	107	107	1年次	123	123	123
	2年次	52	52	52	2年次	107	107	107
Missing	1年次	0	0	0	1年次	0	0	0
	2年次	55	55	55	2年次	16	16	16
Mean	1年次	559	298	261	1年次	575	306	270
	2年次	618	324	294	2年次	619	324	295
Median	1年次	565	305	260	1年次	565	300	270
	2年次	613	328	290	2年次	620	330	295
Minimum	1年次	305	130	130	1年次	195	165	5
	2年次	440	230	180	2年次	235	150	85
Maximum	1年次	770	420	395	1年次	915	465	450
	2年次	825	415	430	2年次	870	455	445

3. 分析

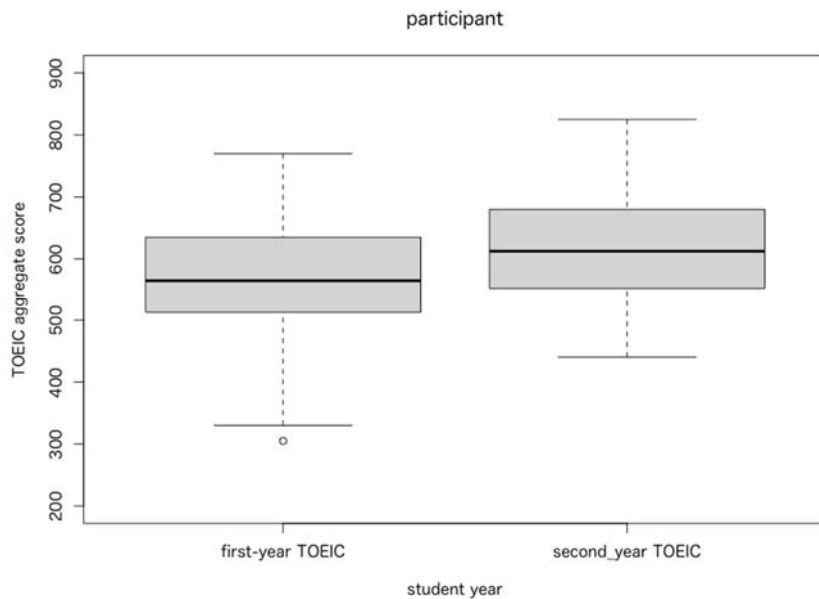


図 2 Xプログラム参加者の TOEIC の点数の変化

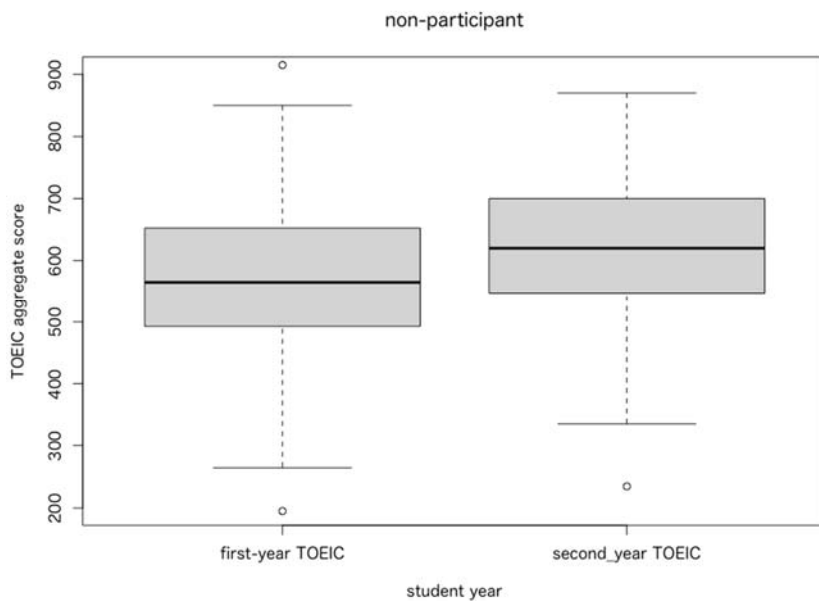


図 3 Xプログラム不参加者の TOEIC の点数の変化

(1) 記述的分析

まず、Xプログラムへの参加・不参加での TOEIC の点数の変化を確認する。以上は、Xプログラムへ参加した群と参加していない群で、1年次と2年次の TOEIC の点数を示した箱ヒゲ図である。この図を見ると、どちらの群も2年次に TOEIC の点数が上昇しているであろうことが確認される。また、Xプログラム参加の有無に関わらず、1年次、2年次の

TOEIC の点数で、満点近くの学生は少なく、天井効果による平行トレンド仮定の未充足 (Ryan et al. 2012) についても考えづらい状況である。

それでは、1年次から2年次への TOEIC の点数の上昇は、X プログラムへ参加したことの効果なのであろうか。この間について次節にて、DID を用いた検証を行う。

(2) 差分の差分分析

本節では、自然実験の環境より得られたデータに、DID を応用し、留学効果の推定を行う。

表 2 DID の分析結果

	合計			リスニング			ライティング		
	Estimate	p. value		Estimate	p. value		Estimate	p. value	
Intercept	559.4 (10.9)	<.001	***	298.04 (5.76)	<.001	***	261.4 (6.24)	<.001	***
time	58.4 (19.1)	0.002	***	25.62 (10.07)	0.011	**	32.83 (10.92)	0.003	***
treat	16 (15)	0.284		7.49 (7.87)	0.342		8.56 (8.54)	0.317	
time:treat	-14.8 (24.3)	0.542		-6.85 (12.78)	0.592		-7.98 (13.86)	0.565	
R ²	0.049			0.035			0.048		
Adj. R ²	0.042			0.027			0.041		
n	346			346			346		

*** <.001; ** <.01; * <.05

表 2 の分析結果より、いずれの従属変数においても X プログラム参加の有無で、TOEIC の点数が上昇しているとは言い難い。本稿では、自然実験の環境を用いたため、学習意欲や学習習慣などの交絡要因が統制されていると考え、川田・西谷 (2017) など先行研究で X プログラムの効果が見出されていたのは、学習意欲や学習習慣などの交絡要因の影響が適切に統制されていなかったことが原因であったと考えられる。

4. まとめ

本稿では、自然実験の環境を用いた DID によって、X プログラムの効果を明らかにした。自然実験の環境を用いることで、先行研究では統制しきれなかった学習意欲や学習習慣などの測定困難な交絡要因や、未観測の交絡要因が統制されていることが確率的に期待できる。このような、偶然生じた自然実験の環境を用いることで、交絡要因を統制した上で、X プログラムの効果が推定可能となった。得られた結果は、TOEIC の合計得点、リスニング、リーディングのいずれも X プログラムによって、伸びているとは言い難い結果であった。本稿では、自然実験の環境を用いたことで、学習意欲や学習習慣などの測定困難な交絡要因

や、未観測の交絡要因が統制されていると考えると、先行研究で X プログラムの効果が見出されたのは、統制されていなかった交絡要因の影響であった可能性がある。

本稿では、交絡要因の統制に重点を置くことで、内的妥当性の高い手法を用いたが、当然ながら課題もある。まず一つ目が、パネルの脱落によるバイアス (attrition bias) が生じている可能性である。そのバイアスの補正のために、周辺構造モデル (Marginal Structure Model) などを用いて再分析することが必要である。二つ目に、効果の異質性 (heterogeneity) の問題である。本稿で明らかになったのは、X プログラム参加者が仮に参加しなかった場合との効果の平均的な差 (Average Treatment effect on Treated : ATT) であり、実際には効果の高い人と低い人が混在しているために、平均的には効果がないという結論が得られた可能性がある。また、効果の異質性があるのであれば、どのような人は効果が高い、または低いのかを明らかにすることで、X プログラムの改善に繋がることも考えられ、今後の課題である。三つ目が外的妥当性の問題である。本稿では自然実験の環境を分析に利用したため、A 大学の一部の学生のみが対象の極めて限られたサンプルの分析結果である。二つ目の課題とも密接に結びつくが⁴⁾、効果の異質性がある場合、他の学生が X プログラムに参加した場合、効果があるという結論が得られる可能性は考えられ、本稿の結論を広く敷衍して解釈することは禁欲的になるべきである。最後に、長期効果の検討である。川田・西谷 (2017) では、1 年次と 3 年生次の指定受検の TOEIC データを用いたため、2 年間における効果を見出している。本研究では期間が 1 年間のため、2 年後や 3 年後など長期的な効果が生じる可能性についても検討が必要であろう。このような課題については、稿を改めて取り組んでいくこととしたい。

【付記】

本研究は、広島大学スーパーグローバル大学創成支援事業の一部として行った、大学施策の立案・実行に役立てるためのデータ分析の成果である。また、本研究推進に際し、JSPS 科研費 JP18K18651・JP 18K18652・JP19H00619・JP19H00621・JP19H01690・JP20H01643・JP20J14673 の資金提供と西谷元教授 (広島大学) より、X 大学の学生データの利用許可を頂いた。ここに記して感謝申し上げる次第である。

【注】

- 1) 自然実験については、伊藤 (2017: 116) や川口 (2019: 423-6) を参照のこと。また、本稿と違いクロスセクションのデータであるが、実際に学生の自然実験の環境を用いた分析については、平尾 (2019) 参照のこと。
- 2) X プログラムについては、以下の A 大学の HP を参照のこと (<https://momiji.hiroshima-u.ac.jp/momiji-top/learning/start.html>)。

- 3) 平行トレンド仮定が満たされない場合，平行トレンド仮定が満たされないと考えられるデータを取り除いたり，平行トレンド仮定の乖離を説明するような変数をモデルに加えたりする方法（安井 2020 : 146-7）や介入前のトレンドが平行であることを示す（伊藤 2017 : 187-8 ; 笹渕 2019 : 65-6）などが方法として挙げられる。
- 4) 効果の異質性と外的妥当性の関係については，Hernan and VanderWeele (2011)，中尾 (2021) 参照のこと。

参考文献

- Craig, P., Katikireddi, S. V., Leyland, A, and Popham, F, 2017, “Natural Experiments: An Overview of Methods, Approaches, and Contributions to Public Health Intervention Research.” *The Annual Review of Public Health*, 38: 39-56.
- Hernan, M. A. and Vanderweele, T. J., 2011, “Compound Treatments and Transportability of Causal Inference.”. *Epidemiology*, 22(3): pp.368-377.
- 平尾智隆, 2019, 「自然実験によるキャリア教育の効果測定」『日本労働研究雑誌』No.707 : 79-92.
- 伊藤公一朗, 2017, 『データ分析の力 因果関係に迫る思考法』光文社新書.
- 川口大司, 2019, 「政策評価モデル」西山慶彦・新谷元嗣・川口大司・奥井亮『計量経済学』有斐閣 : 391-457.
- 川田恵介・西谷元, 2017, 「短期留学プログラムが語学到達度に与える影響について : 広島大学 START プログラムの事例から」 (<https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxrZWlzdWtla2F3YXRhN3xneDoxOTQwNDhhNDNkZGIwZWNm>, 2021年5月20日最終確認日) .
- 川田恵介・西谷元・森田光宏, 2019, 「短期の留学でも語学力は向上するのか？」 (<https://premium.toyokeizai.net/articles/-/20825>, 2021年5月20日最終確認日) .
- 中尾走, 2021, 「EBPMの動向と課題—高等教育政策への影響を見据えて」『大学論集』54 : 39-52.
- 中尾走・樊怡舟・村澤昌崇, 2021, 「交絡要因の測定誤差が合流点バイアスをもたらす影響について」『広島大学高等教育研究開発センター編 ディスカッションペーパーシリーズ』15 : 1-18.
- 中澤渉, 2016, 「教育政策とエビデンス」志水宏吉編『岩波講座 教育 変革への展望 2 社会のなかの教育』: 73-101.
- 日本学生支援機構, 2017, 『平成 27 年度協定等に基づく日本人学生留学状況 調査結果』.

Ryan, A. M., J. Blustein, and L. P. Casalino, 2012, “Medicare’s Flagship Test of Pay-For-Performance Did Not Spur More Rapid Quality Improvement among Low-Performing Hospitals.” *Health Affairs*, 31 (4): 797–805.

Wing, C., Simon, Kell-Gomez, R. A., 2018, “Designing Difference in Difference Studies- Best Practices for Public Health Policy Research.” *The Annual Review of Public of Health*, 39: 453-69.

広島大学高等教育研究開発センター 国際共同研究推進事業
Advancement of Higher Education Research:
RIHE Monograph Series について

『Advancement of Higher Education Research: RIHE Monograph Series (旧ディスカッションペーパーシリーズ)』は、速報性を重視し暫定的にまとめて発信することを目的として企画されたシリーズ (オープン・ジャーナル) です。これまでに、「ディスカッションペーパーシリーズ」として、国際共同研究、公募型研究、戦略的プロジェクト研究、客員研究員による研究等、成果として取りまとめられたものが発信されました。(執筆者の所属は刊行時のものです)

本事業の推進にあたり、以下の資金提供を受けた。記して感謝したい。

- ・JP18K18651 (研究代表者: 村澤昌崇, 「多様な使命と目的を同時最適化する卓越した大学モデルの探索」)
- ・JP 18K18652 (研究代表者: 西谷元, 「英語力向上と教育プログラムとの相関関係: 大規模データによる統計的分析」)
- ・JP19H00619 (研究代表者: 松繁寿和, 「社会的能力の特定化とその育成適正期および教育効果の検証」)
- ・JP19H00621 (研究代表者: 小林信一, 「知のオープン化時代の大学・科学関連システムの再構築」)
- ・JP19H01690 (研究代表者: 西谷元, 「海外留学プログラムの効果検証: 大規模パネルデータによる学生の心理特性の変化の分析」)
- ・JP20H01643 (研究代表者: 村澤昌崇, 「EBPM の批判的検討を通じた高等教育政策・研究の高度化と EIPM への展開可能性」)
- ・JP20J14673 (日本学術振興会特別研究員: 中尾走, 「高等教育政策を事例にした EBPM の批判的検討」)

RIHE Advancement of Higher Education Research
Monograph Series

No. 3

2021 (令和 3) 年 5 月 24 日 発行

自然実験環境を用いた留学効果の推定

—X プログラムを事例にして—

執筆者: 中尾 走 (広島大学教育学研究科・日本学術振興会特別研究員 (DC2)・広島大学 スーパーグローバル大学創成支援事業データ分析チーム)、樊 怡舟 (広島大学教育学研究科・広島大学 スーパーグローバル大学創成支援事業データ分析チーム)、西谷 元 (広島大学人間社会科学研究科・広島大学 スーパーグローバル大学創成支援事業データ分析チーム)、村澤 昌崇 (広島大学高等教育研究開発センター・広島大学 スーパーグローバル大学創成支援事業データ分析チーム)

〒739-8512 広島県東広島市鏡山 1-2-2 電話 (082) 424-6240
<https://rihe.hiroshima-u.ac.jp/>

RIE 広島大学高等教育研究開発センター