

交互最小二乗法を用いた大量欠損の成績表データからの因子抽出

——X大学の留学効果推定への応用の試み——

樊 怡舟*, 中尾 走*
西谷 元**, 村澤 昌崇**

Factor Extraction from Student Test Grade Sheets Using the Alternating Least Squares Method:
An Example of Estimating the Effects of University Students' Studying Abroad

Yizhou FAN*, Ran NAKAO*, Hajime NISHITANI** and Masataka MURASAWA**

In the research area of evaluating the effectiveness of study abroad programs, adopting counterfactual frameworks such as DID, PSM or IV has been considered a valid analytical approach. Previous findings drawn based on these conventional frameworks suggest that even short-term study abroad programs have a significant effect on the improvement of TOEIC scores. However, these studies are often designed to estimate the effects with students' prior TOEIC scores, and only controlling departmental or school affiliations, while confounding factors, particularly students' competency such as learning attitudes as well as learning motivations, remain uncontrolled. This study attempts to extract students' competency from high-dimensional data with a large volume of missing values in student's test score sheets, using the Alternating Least Squares (ALS) method. Injecting the extracted competency in the subsequent regression analysis enables us to accurately estimate the causality between the study abroad experience and the observed outcomes. Our analysis result reveals that, unlike findings from earlier studies, once students' competency is properly controlled, the estimated effect of the study abroad programs becomes negligible with no significance. Therefore, the finding suggests that the causal effect claimed by the previous studies might be due to a bias engendered by students' self-selection. The result also indicates that datasets readily accessible at any university, such as student test score sheets, could effectively be used for project evaluations within an institution, notably because the confounding factors are properly controlled as suggested by the current study.

Key words: Alternating Least Squares, Test Grade Forms, Causal Inference, Effect of Study Abroad Programs

キーワード: 交互最小二乗法, 成績表データ, 因果推論, 留学効果

1. はじめに

本稿は、学生の成績表データのような欠損率の高い多次元データから、測定困難な構成概念の交絡要因を抽出し、留学効果推定に応用することが目的である。

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。EBPM は、反実仮想のフレームワークに基づき、交

*広島大学スーパーグローバル大学創生支援事業データ分析チーム (広島大学大学院教育学研究科)
(Graduate School of Education, Hiroshima University)

連絡先: 〒739-8512 広島県東広島市鏡山 1-2-2

E-mail: fanyizhou921106@yahoo.co.jp

**広島大学スーパーグローバル大学創生支援事業データ分析チーム (広島大学)
(Hiroshima University)

連絡先: 〒739-8512 広島県東広島市鏡山 1-2-2

絡要因を統制した上で因果効果の推定が要請されている。

また、高等教育の文脈では IR (Institutional Research) が強調されており、大学にも教育効果・アウトプット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、そこにおける因果推論を用いた効果検証も行われはじめた(日本学生支援機構, 2017; 川田・西谷, 2017)。例えば川田・西谷(2017)は、傾向スコア+差の差分分析の枠組みから、短期留学と留学直後の英語力(TOEIC 得点)の伸びとの因果効果を調べ、結果的に 39 点の留学効果が有意に推定された。

しかし、それらの研究は交絡要因の統制が不十分である。例えば、川田・西谷(2017)では、傾向スコア算出のために投入した共変量は学生の所属と事前の TOEIC 得点のみであり、さらなる個人属性の追加を課題としている。

留学効果の推定に関しては、実のところ、交絡要因の統制が特に困難といえる。なぜなら、教育分野を含む社会科学全般では、交絡要因が観測しづらい構成概念の場合が多く、これら交絡要因の影響を取り除くことは困難である(中尾・樊・村澤, 2021)。留学に関しては、学生が元来保持している「学習態度」や「学習意欲」¹⁾などの学生の「コンピテンシー」²⁾に相当する構成概念が重要な交絡要因として想定される。それ故に、そのような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。一方、「コンピテンシー」のような観測しづらい構成概念を、正確に且つ留学に参加する前に測定していなければ、留学効果の推定は極めて難しいといえよう。

そこで、測定が困難な交絡要因を統制するために、本研究は、学内の既存データの活用をめざし、成績表データに注目した。実は、交絡要因にあたる学生の「コンピテンシー」が、留学プログラム参加や語学力のほか、他の科目の成績にも影響を及ぼしていると予想される。そうすると、留学プログラム参加前の成績

表 1. 成績表データのイメージ

	科目 1	科目 2	科目 3	科目 J
学生 1	E_{11}	欠損	欠損	E_{1J}
学生 2	欠損	欠損	E_{23}	E_{2J}
学生 3	欠損	E_{32}	欠損	E_{3J}
学生 I	E_{I1}	E_{I2}	E_{I3}	E_{IJ}

表データより「コンピテンシー」に関する構成概念の情報が抽出できるのではないかと考えられる。ゆえに、成績表から抽出したものは代理指標として、交絡要因の部分的な統制 (partial control) となる (Elwert & Winship, 2014)。

成績表データのイメージは表 1 のようになる。各学生の履修科目の期末成績がデータ化されており、未履修の科目はデータの欠損 (missing values) とみなせる。つまり、欠損ありの多変量データである。

ただし、成績表データを交絡要因統制のために活用するには、難点が二点ほど挙げられる。

第一に、成績表データはあまりにも変数が多いことである。X 大学の場合、学部学生が 1 万人以上おり、開設科目数も凡そ 4 千ある。ゆえに主成分分析などの処理で情報の集約を行うことが望まれる。

第二に、成績表データは欠損率が高いことである。普通の主成分分析では、欠損が発生する行を削って計算することとなり、欠損率が高い場合に主成分分析の手法が適用できない(柴山, 1992)³⁾。特に成績表データの場合、全科目を履修できている学生が少ない限り、すべてのデータが削除されることとなるので、別のアプローチを考える必要がある。

そこで、本稿は、このように欠損率の高い成績表データから学生の「コンピテンシー」という測定困難な交絡要因を抽出して、統制変数として回帰モデルへ投入し、より純粋な留学効果の推定を目的とする。

以降、次節では多変量かつ欠損率の高いデータから因子抽出する理論的背景を述べ、3 節では、そのようなデータに適用可能な交互最小二乗法の概要を述べる。4 節では、実際の X 大学の成績表データを用いて、留学効果の推定を行い、最後に 5 節でディスカッションとして、分析結果のまとめと今後の課題を提示する。

¹⁾ その他にも学習能力などが考えられるが、本稿では便宜的に「学習態度」「学習意欲」という用語を用いる。

²⁾ 本稿は「コンピテンシー」について学業成績に影響する学生の属性・特性全般として広義的にとらえている。

³⁾ 欠損を含むデータに主成分分析を適用する手法については、Dray & Josse (2015) 参照のこと。

2. 交互固定効果モデル

本稿のアプローチについて議論する前に、まず成績表データの性質を理解し、本稿の文脈で適用可能な情報集約の方法について、データの生成プロセスに関する仮定から理論的に確認する。成績表の生成プロセスについて、以下のモデルが挙げられる。

まず考えられるのは、個人固定効果モデルと科目固定効果モデルである（以下成績を E 、誤差項を u 、それに準じた係数を β で示す）。

①個人固定効果モデル

下式のように、科目に不変な個人属性 (z 、例えば性別など) を各個体の切片として扱うモデルである。

$$E_{ij} = \beta_0 + \beta_1 x_{ij} + \beta_2 z_i + u_{ij}$$

②科目固定効果モデル

下式のように、個体に不変な科目属性 (s 、例えば講義担当教員の厳しさなど) を各科目の切片として扱うモデルである。

$$E_{ij} = \beta_0 + \beta_1 x_{ij} + \beta_3 s_j + u_{ij}$$

個人固定効果モデルでは、仮定としてどの科目を受験しても高く評価される学生の属性、即ち全科目に通用する能力のようなものが考慮されている。対して科目固定効果モデルでは、どの学生でも得点しやすい科目の特徴、即ち科目の難易度のようなものが考慮されている。

上記の両モデルを合わせると、個人固定効果と科目固定効果を同時に推計できる双方向モデルとなる。

③双方向固定効果モデル

$$E_{ij} = \beta_0 + \beta_1 x_{ij} + \beta_2 z_i + \beta_3 s_j + u_{ij}$$

一方、現実では、学生は科目によって得意・不得意があり、また科目の難易度の捉え方も学生によって異なると考えられる。ゆえに学生の属性と科目の特徴の交互作用を表現するモデルが必要である。そこで本稿は交互固定効果モデルを使用することとした。その仮定は次のようになる。

④交互固定効果モデル

I の学生に対して、一律にそれぞれ J のテストを行い、そして学生 i がテスト j を受けたときの得点を E_{ij} とする。各学生の属性は R 次元のベクトル $\ddot{B}_i = ({}^{(1)}\ddot{B}_i, {}^{(2)}\ddot{B}_i, {}^{(3)}\ddot{B}_i, \dots, {}^{(r)}\ddot{B}_i, \dots, {}^{(R)}\ddot{B}_i)'$ から

なるとし（操作的に学生の「コンピテンシー」と定義する）、そして、テストの特徴も R 次元のベクトル $\ddot{F}_j = ({}^{(1)}\ddot{F}_j, {}^{(2)}\ddot{F}_j, {}^{(3)}\ddot{F}_j, \dots, {}^{(r)}\ddot{F}_j, \dots, {}^{(R)}\ddot{F}_j)'$ からなるとする（操作的に各コンピテンシーの次元に対する「評価ウエイト」と定義する）。そうすると、得点は式 (1) のように各学生の「コンピテンシー」と各科目の「評価ウエイト」という二つのベクトルの内積として定義できる。

$$E_{ij} = \ddot{B}_i \cdot \ddot{F}_j + u_{ij} = \sum_r ({}^{(r)}\ddot{B}_i \cdot {}^{(r)}\ddot{F}_j) + u_{ij} \quad (1)$$

\ddot{B}_i と \ddot{F}_j を並べた行列として B と F を定義すると、成績表 E は式 (2) となる。

$$E = \ddot{B}\ddot{F}' + U \quad (2)$$

交互固定効果モデルの下で、 $I \times J$ のセルを持つ成績表 E は、 R 次元の「コンピテンシー」ベクトル計 I 個と R 次元の「評価ウエイト」ベクトル計 J 個から生成しているとみなすことができ、即ち自由度は $R \times (I+J)$ である。つまり、成績表の情報集約が可能であり、そして、情報集約は即ち潜在因子である「コンピテンシー」と「評価ウエイト」の抽出に相当する。

しかしながら、交互固定効果モデルの因子抽出に関して、識別問題が指摘されている。つまり、任意の正則行列 A をとり、 $B^* = BA$ かつ $F^* = FA^{-1}$ としても、 $B^*F^{*'} = BAA^{-1}F' = BF'$ が常に成立し、観測上同値 (observationally equivalent) のモデルが成立してしまう (奥井, 2014)。ゆえに、推定する際に追加条件が必要となる。

よく使用されるのは、 BB' が単位行列の I (行数) 倍で、かつ $F'F$ が対角行列という識別条件である。 I と J が無限に拡大すると、上記の識別条件の下で行う共通因子の抽出は、主成分分析の結果と整合的であることが証明されている (Stock & Watson 2002; Bai 2009; 奥井 2014)。

けれども、前述したように、成績表は大量に欠損が発生しているため、主成分分析法が直接的に適用できず、別のアプローチを試みる必要がある。そこで本稿が注目したのは交互最小二乗法 (Alternating Least Square, 以下 ALS) である。

3. 交互最小二乗法 (ALS)

I と J が十分に大きい場合、誤差項は式 (3) を満た

すとみなせる.

$$\sum_i u_{ij} = \sum_j u_{ij} = 0 \quad (3)$$

ゆえに、式(2)について、一般的な最小二乗法 (Ordinary Least Square, 以下 OLS) の適用によって、以下のことが成り立つ. 仮に「評価ウエイト」行列 F が既知であれば、「コンピテンシー」行列 B は OLS で不偏に推定できる. 一方、「コンピテンシー」行列 B が既知であれば、「評価ウエイト」行列 F も OLS で不偏に推定できる.

以上を踏まえると、潜在因子の抽出は実質的に、 F から B を推定し、同様に B から F も推定されるという、適当な F と B の組み合わせを見つけることである.

これは即ち、ALS のアルゴリズムであり、欠損ありのデータに対する主成分分析の代替案の一つとして推奨されている (柴山, 1992) ものと実質的に同じである.

このような欠損のあるデータに対する主成分分析については、Dray & Josse (2015) がまとめている. その中で、本稿で用いる推定手法は、欠損値を明示的に考慮しないアルゴリズムの一つである Non-linear iterative partial least squares (以下, Nipals) である. Interactive PCA method (以下, Ipca) (Kiers, 1997) などの推定手法もあるが、Ipca は、真の次元数をあらかじめ既知である必要があり、既知でない場合に過大推定する可能性が指摘されている (Dray & Josse, 2015) ため、相対的に簡便な推定手法である Nipals を用いる. 以下、Nipals の推定手順を本稿の関心に沿って説明していく.

なお、先行研究においては、識別条件を付随することで推定を行ってきたが、本稿では ALS を用いて因子を一つずつ抽出することとする. 理由は、抽出した因子が元の成績表の各主成分の定数倍に収束するためである.

まずは、第一主成分を抽出する. 言い換えると、 F と B それぞれが OLS によって導かれる適当な第一「評価ウエイトベクトル」 ${}^1F = ({}^1f_1, {}^1f_2, \dots, {}^1f_j, {}^1f_j)$ と第一「コンピテンシーベクトル」 ${}^1B = ({}^1b_1, {}^1b_2, \dots, {}^1b_j, \dots, {}^1b_j)$ とを見つける. 両者のベクトル 1F と 1B は以下の式 (4) と (5) を同時に満たしている. (E_i は E の i 行目の行ベクトル, E_j は E の j 列目の列ベクトルとする)

$$\begin{cases} {}^1b_i = E_i {}^1F ({}^1F' {}^1F)^{-1} & (4) \\ {}^1f_j = E_j' {}^1B ({}^1B {}^1B')^{-1} & (5) \end{cases}$$

(1) F と (1) B の推定については、以下のように行う.

① 科目 j にあらかじめランダムに初期値 ${}^1f_j^{(0)}$ を与え、それらの初期値を並べると第一「評価ウエイトベクトル」 1F の初期値 ${}^1F^{(0)}$ となる.

② 式 (4) に基づき、 ${}^1F^{(0)}$ と成績表のそれぞれの行の情報から各個人 i の「コンピテンシー」の一次元目 ${}^1b_i^{(1)}$ が推定できる. その値を並べて第一「コンピテンシーベクトル」 ${}^1b^{(1)}$ とする.

③ 式 (5) に基づき、 ${}^1b^{(1)}$ と成績表のそれぞれの列の情報から各科目 j の「評価ウエイト」の一次元目 ${}^1f_j^{(1)}$ を求められる. そうすると、第一「評価ウエイトベクトル」 1F の値が ${}^1F^{(0)}$ から ${}^1F^{(1)}$ へ更新できる.

④ ②と③を繰り返し、式 (4) と式 (5) を用い、 1F と 1B の値を交互に OLS によって更新していく.

欠損がない場合、上記のような推定を、 n 回繰り返すと、 1F と 1B の値は次のようになる⁴⁾.

$$\begin{cases} {}^1B^{(n)} = (E'E)^{n-1} {}^1B^{(1)} \frac{\prod_{i=1}^{n-1} {}^1F^{(0)'}(E'E)^{2n-1} {}^1F^{(0)}}{\prod_{i=1}^n {}^1F^{(0)'}(E'E)^{2n-2} {}^1F^{(0)}} \\ {}^1F^{(n)} = (E'E)^n {}^1F^{(0)} \frac{\prod_{i=1}^n {}^1F^{(0)'}(E'E)^{2n-2} {}^1F^{(0)}}{\prod_{i=1}^n {}^1F^{(0)'}(E'E)^{2n-1} {}^1F^{(0)}} \end{cases}$$

1F の形は $J \times 1$ のベクトルで、 E は $I \times J$ の行列なので、総乗部分は定数となり、 ${}^1B^{(n)}$ は実質的に、 $(E'E)$ という正方行列に対して、べき乗法による固有値問題の解析を行っており、 $(E'E)$ の絶対値最大の固有値に対応する固有ベクトルの定数倍に収束することが知られている. 同様に、 ${}^1F^{(n)}$ も $(E'E)$ の絶対値最大の固有値に対応する固有ベクトルの定数倍に収束する.

収斂した時の 1F と 1B を掛け合せると、 E の予測値 ${}^1\hat{E}$ 及び残差 ${}^1D = E - {}^1B {}^1F$ が算出される. この残差行列に対して、①~④の同じアルゴリズムを適用し、 1D に対して抽出を行う. 即ち、(4) 式と (5) 式を同時に満たす第二の「評価ウエイトベクトル」 2F と第二の「コンピテンシーベクトル」 2B の適当な組み合わせを推定する. 結果として、残差行列も ${}^2D = {}^1D - {}^2B {}^2F$ に更新される.

このように、残差の更新とともに抽出の手順を繰り返していくと、 E の行と列の各主成分が得られる⁵⁾.

⁴⁾ 数学帰納法によって簡単に証明できる.

⁵⁾ Nipals のアルゴリズムでは、推定値を更新する際に常に共通因子の標準化を行っている. けれども、本稿

なお、上記の手法では、FとBを更新するときに、各学生の「コンピテンシー」と各科目の「評価ウエイト」が一つ一つ並行で推定されている。ゆえに、成績表Eに欠損が発生しているとしても、それぞれの 1f_j と 1b_i を求める際に、すべての列又は行を削除することなく、それぞれの列または行の観測されている部分から推定することができ、データの情報を最大限に活用できると期待される。

ただし、欠損率の高さによっては、上記の手法では安定的な推定が難しい場合がある(Grung & Manne, 1998; Josse, Pagès, & Husson, 2009)。そのために、実際に分析を行う際には、初期値の生成について工夫が必要である。本稿は「多重スタート法」を用いることとする(4.2で詳しく説明する)。

4. 分 析

本節では、実際に成績表のデータにALSを適用することで、共通因子を抽出し、その因子を交絡変数の代理指標として、留学効果の推定に用いる。

4.1. データと変数

本稿で用いるデータは、X大学の2018年度入学者のデータである。X大学では、2011年度以降、毎年200名近い学生が2週間程度の短期の海外留学プログラム(以下、留学プログラム)によって海外留学を行っている。その中で、留学プログラム参加の有無と従属変数である留学プログラム参加後のTOEICの点数(留学後TOEIC得点)、成績表のデータが利用可能な2018年度入学者のデータを用い、留学効果の推定を行うこととする。

成績表のデータについては、交絡要因の統制として用いるため、留学前の成績表データとして2018年度入学者の1年次前期の成績表データを用いた。そのデータから、1つしか講義の成績が得られていない学生と一人の学生しか履修していない講義を削除した。また、成績表データの中で、点数ではなくS, A, B, Cなどのカテゴリーに分けられた成績評価結果の講義については、Sを90点、Aを80点、Bを70点、Cを60点と変換した。このような過程を経て、最終的

において共通因子抽出の目的はあくまで交絡要因の統制であるため、主成分の定数倍(初期値に依存)に収束するだけで充分であると考え、そういった作業をあえて省くことにした。



図 1. 分析に用いたデータ

に2410人の学生、326個の講義からなる成績表データを得た。

留学プログラムは、1年次の後期である2019年2月から3月にかけて留学を行っており、105名の学生が参加している。従属変数は、2年次の前期である2019年5月に実施されたTOEICの点数を用いた。分析に用いたデータの時間関係については、図1に示す通りである。

その他の統制変数として、学生の性別、所属学部、2018年5月に実施されたTOEICの点数(以下、事前TOEIC得点)などを用いた。

4.2. 「コンピテンシー」の抽出

未履修科目の成績をすべて欠損として扱うと、本稿で使用する成績表のデータは欠損率が95.43%となる。本節ではこの成績表のデータから「コンピテンシー」(B)という共通因子を3節で説明した手順に従って抽出する。

ところが、特に成績表のように大量に欠損しているデータに対して、推定値が局所最適に陥る可能性がある。即ち、共通因子として抽出したものが成績表データの観測されている部分さえうまく説明できないということである。

抽出した共通因子の説明力を評価するためによく使われるのは「最小二乗基準」である。つまり、データの観測されている部分に関して、予測値(\widehat{E}_v)と観測値(E_v)の差(即ち残差)の二乗和を最小にするような推定が望ましいとされるものである。本稿では、それに基づいて、式(6)で抽出された共通因子の「説明力」(R)⁶⁾を評価する。そのうえで、多重スタート法(森・黒田・足立, 2017)を適用し、「コンピテンシーベクトル」の初期値を複数回付与したうえでそれぞれ

⁶⁾ 決定係数の発想に似ているが、式(6)では誤って推定された場合に、共通因子の説明力(R)がマイナスになり得る。

表 2. 各抽出試行における第一「コンピテンシーベクトル」の説明力 (R1)

R1		R1	
試行 1	0.829	試行 6	-2.714
試行 2	0.367	試行 7	0.098
試行 3	-0.461	試行 8	0.829
試行 4	0.829	試行 9	0.829
試行 5	0.573	試行 10	-1.046

表 3. 抽出した次元数とその時の説明力

R1	R2	R3	R4	R5
0.829	0.875	0.897	0.913	0.925
R6	R7	R8	R9	R10
0.935	0.944	0.952	0.959	0.965

の初期値に基づいて試行的に抽出を行い、そのうちの説明力の最も高い抽出結果を選んでいく。

$$R = 1 - \frac{\text{var}(\widehat{E}_v - E_v)}{\text{var}(E_v)} \quad (6)$$

以下、第一「コンピテンシーベクトル」(各個人の「コンピテンシー」の一次元目を並べたもの)の抽出を例に説明する。試行的に10回抽出を行い、抽出した第一「コンピテンシーベクトル」の説明力を表2でまとめた。

表2から、試行1, 4, 8, 9のように、共通因子がうまく抽出された場合では、一次元だけで成績表の観測されている部分の約83%の分散を説明できていることが確認できる。一方、初期値によって抽出された共通因子の説明力が低かったり、場合によって説明力がマイナスになったりすることもある。「最小二乗基準」により、試行1の抽出結果を各個人の「コンピテンシー」の一次元目として、この先の分析で用い⁷⁾、二次元目以降の「コンピテンシー」を抽出した。

抽出の結果は表3のようになる。10次元抽出した時点で既に観測されているデータの96%以上説明できており、それ以上抽出を行うと過学習の危険性が大きくなると考え、本稿は10次元から構成される各学

⁷⁾ 試行1と同様に説明力が高い、試行4, 8, 9で分析を行っても同様の分析結果であることは確認済みである。また、試行2のように説明力が低い共通因子であっても、統制しないよりも統制した結果の方が、留学プログラムの係数は小さくなる。

生の「コンピテンシー」(特に説明力の最も高い一次元目のデータ)を統制変数として留学効果の分析に投入することとする。

一方、抽出した「コンピテンシー」が一部の学生しか説明できていない可能性を確認するために、説明力を学生(行)ごとに計算した。例として1次元のみ抽出した時(全体の説明力は0.829)、5次元抽出した時(全体の説明力は0.935)と10次元抽出した時(全体の説明力は0.965)の結果をヒストグラムに示した(図2)。抽出した「コンピテンシー」は、一部の学生のみに対して高い説明力をもっているわけではないことがこの図から確認できる。

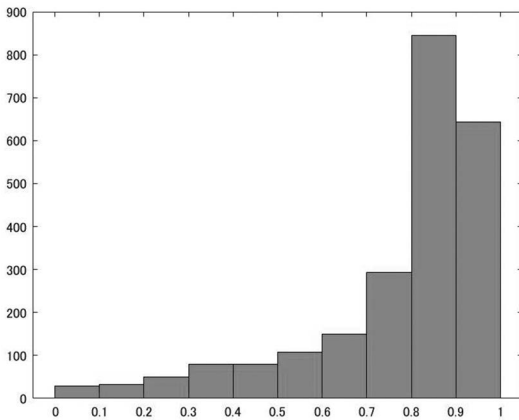
4.3. 留学効果の推定

留学効果の推定として、成績表データから抽出した「コンピテンシー」を統制するか否かで、分析結果にどのような違いが生じるだろうか。分析手法は、全て重回帰分析である。従属変数のTOEICの点数を回帰した結果を図3に示す。

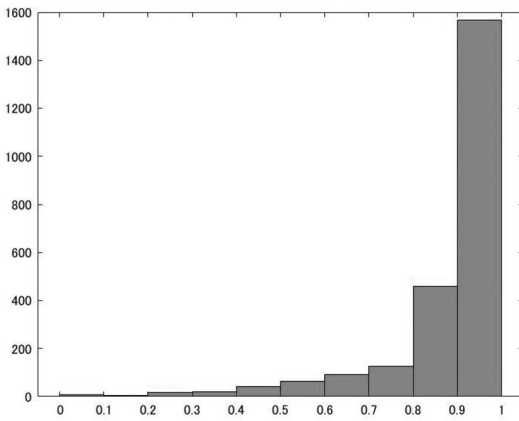
モデル1は、統制変数が投入されておらず、欠落変数バイアスを含む推定結果となる。モデル2は事前のTOEIC得点を統制した。モデル3は事前のTOEIC得点に加え、性別と所属を統制変数として投入した。モデル4は成績表データから得られた「コンピテンシー」の1次元目のみを統制した。そして、モデル6はモデル4に「コンピテンシー」の1次元目のみを統制変数として投入したものである。

また、「コンピテンシー」を複数次元統制した場合の推定結果の頑健性を検証するため、一事例として「コンピテンシー」の10次元をすべて投入したモデル5とモデル7を、それぞれモデル4とモデル6の比較対象として掲載している。

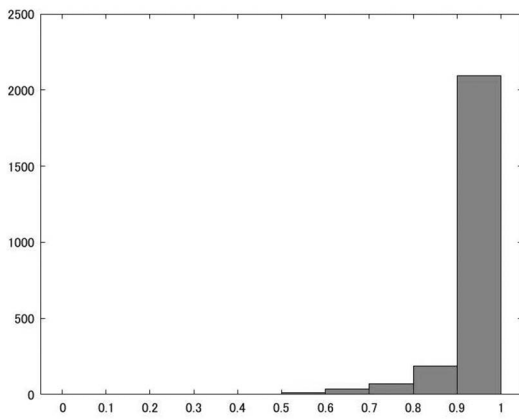
まずモデル1の分析結果より、何も統制していない分析では、留学プログラムの参加の有無によってTOEICの点数が35点ほど上昇すると推定された。しかしながら、この結果は交絡要因を統制していない欠落変数バイアスが起きていると考えられる。そこで、今得られているデータから統制可能な交絡要因を統制した結果がモデル2とモデル3である。性別、所属学部と事前TOEIC得点を統制すると、留学効果は依然として有意に推定されているが、係数が半分ほどまでに小さくなった。モデル1から係数が小さくなったのは、交絡要因を統制したことによる欠落変数バイアスの一部が統制されたためだと考えられる。



①行ごと説明力ヒストグラム (1次元)



②行ごと説明力ヒストグラム (5次元)



③行ごと説明力ヒストグラム (10次元)

図 2.

対して、モデル 4 とモデル 5 で統制したのは成績表データから抽出した「コンピテンシー」のみだが、留学効果の推定値が 11~12 ほどとなり、有意水準も満たさなくなった。モデル 2 とモデル 3 の結果と比較すると、事前の TOEIC 得点を含む他の統制変数よりも、成績表から抽出した「コンピテンシー」を一次元でも統制したほうが、より交絡が取り除かれると推察できる。

モデル 6 とモデル 7 は、性別と所属学部、事前 TOEIC 得点に加えて、「コンピテンシー」を統制変数として投入した。この場合、処置効果の推定がゼロにかなり近い値になり、有意性も完全に失われる。つまり、留学プログラム参加の有無による TOEIC の点数の上昇は有意に検出されないということになる。

モデル 1 からモデル 7 までの推定結果の変遷を見ると、モデル 6 とモデル 7 ではほとんどの交絡が取り除かれているのではないかと推測される。

また、「コンピテンシー」を 1 次元のみ投入したモデル (モデル 4 とモデル 6) と 10 次元まで投入したモデル (モデル 5 とモデル 7) を比較すると、留学効果の推定結果にほとんど変化が無いことが分かる。

最後に、各モデルの決定係数に注目すると、事前 TOEIC 得点を投入した場合、回帰モデルの決定係数が著しく高くなるのに対して、抽出した「コンピテンシー」は、モデルの決定係数に与える影響が相対的に低い。即ち、留学後の TOEIC の点数を予測するという意味では、「コンピテンシー」がそこまで高い説明力を持っているとは言えない。その反面、すでに議論したように「コンピテンシー」を統制するか否かが留学効果の推定に大きく影響している。このことから、本稿が抽出した「コンピテンシー」が、影響力の強い交絡要因の有力な代理指標の一つであると考えられる。

一方、成績表から抽出し統制変数として用いた「コンピテンシー」は、構成概念であり、その中身の議論は簡単にはできない。そのため、投入した「コンピテンシー」の偏回帰係数が有意であっても、本稿はこのことに対して、あえて積極的に解釈しない。結果的に、抽出された「コンピテンシー」は、想定通り統制変数として有効であったという点だけを強調しておきたい。

5. ディスカッション

本稿では、成績表データのような高次元かつ欠損率の高いデータから共通因子「コンピテンシー」の抽出

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
	Estimate	Estimate	Estimate	Estimate	Estimate	Estimate	Estimate
Intercept	524.37*** (2.75)	137.07*** (22.14)	-21.75 (48.82)	-162.16*** (33.38)	-130.24*** (41.87)	-14.86 (50.00)	-153.60 (168.29)
A プログラム参加 (v.s. 不参加)	35.49** (13.18)	22.98* (8.97)	18.41* (8.97)	10.96 (12.17)	12.09 (11.97)	-4.24 (11.03)	-4.23 (11.18)
統制変数	なし	事前 TOEIC 得点	事前 TOEIC 得点・ 性別・所属	コンピテンシー (1次元)	コンピテンシー (10次元)	事前 TOEIC 得点・ 性別・所属+コンピ テンシー (1次元)	事前 TOEIC 得点・ 性別・所属+コンピ テンシー (10次元)
R ²	0.00	0.51	0.54	0.16	0.19	0.54	0.56
Adj. R ²	0.00	0.51	0.52	0.15	0.19	0.52	0.52
Num. obs.	2382	1972	1843	2365	2365	1827	1827

*** p < 0.001; ** < 0.01; * < 0.05

図 3. 重回帰分析の結果の比較

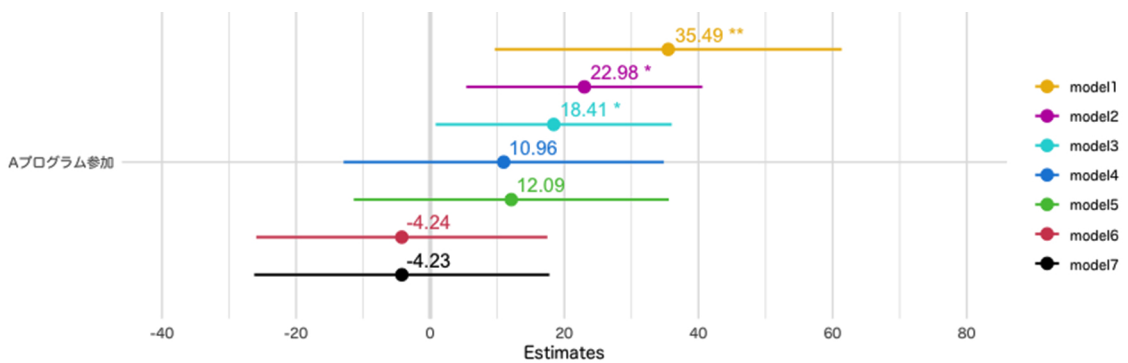


図 4. 分析結果のプロット

を行い、交絡要因の統制変数として用いることで、より厳密な留学効果の推定を試みた。

本稿の関心は、留学などの大学での改革や授業等の介入効果を検討する場合に、どの大学でも保有している成績表データを活用することにより、直接学生の「コンピテンシー」という構成概念を抽出し、これで交絡要因の部分的統制ができることを示すことにあった。実際には、十次元の因子を抽出したが、これらはすべて成績表データをもとにした構成概念であるので、各因子の命名は困難であるとしても、一年次学生の介入前の能力・コンピテンシー等を表していると思われる。

そして、この「コンピテンシー」を投入した留学プログラムの効果を検討したところ、留学プログラムへの参加は、留学後の TOEIC 得点に対する効果が検出されなかった。本稿の知見から、川田・西谷 (2017)

が主張した短期留学の効果は恐らく「コンピテンシー」を統制しなかったことによるバイアスであったと推測される。

ただし、この分析結果から、ただちに（短期の）留学推進が無意味であると結論づけるのは早計ではあろう。国際化の成果の目安として、従来からよく用いられる語学力（実質は英語力）、しかも TOEIC の点数のみに基づいて検証し解釈すること自体が、（短期の）留学プログラムの効果を矮小化するリスクがある。また、本研究の分析結果は、むしろ伝統的な大学教育における積み重ね（=成績）の重要性を傍証しえたとも言えるので、近年の矢継ぎ早に繰り出される改革を再考する契機を与えるものである、とも言える。本稿では、こうした大学における諸改革や取り組みの効果検証を、どの大学にも横たわっているデータから行いう一つの事例を示したと言えるであろう。

ただし、以下のように方法論的課題が残されている。まず、成績表データから因子抽出を行う上でも、今回は Nipals のアルゴリズムで抽出を行なったが、Nipals は収束が不安定であることが指摘されている (Grung & Manne 1998; Josse, Pagès, & Husson 2009)。そのため、今後は Nipals よりも安定したアルゴリズムを試みる必要がある。

また、共通因子の抽出について、多重代入法やランダムフォレストなど、いったん何らかの手法で欠損値補完を行なった上で、主成分分析を行うという代替的アプローチも考えられる。本稿のアプローチと比べて、抽出結果がどのように異なるかについてさらに考察していく必要がある⁸⁾。

最後に、本稿の手法で抽出した各学生の「コンピテンシー」は、傾向スコアなど他の因果推論の手法にも共変量として活用できる。頑健性の検証も含めて、多様な因果推論の手法で評価していくことなど、本稿の課題を改めて論じていく必要もあるだろう。

謝 辞

本研究の改訂にあたり、2名の査読者の先生より大変有益なコメントをいただきました。また、本研究についてアリゾナ州立大学サンダーバードグローバル経営大学院の渡邊聡先生より貴重なコメントをいただきました。この場をお借りして御礼申し上げます。なお、本研究は科研費 JP20H01643, JP18K18651, JP19H00621, JP20J14673, JP19H01690, JP18K18652, JP19H00619 の助成を受けています。

参 考 文 献

- Bai, Jushan (2009). Panel data models with interactive fixed effects. *Econometrica*, **77**, 4, 1229–1279.
- Dray, Stephane & Josse, Julie (2015). Principal component analysis with missing values—a comparative survey of methods. *Plant Ecology*, **216**, 657–667.
- Elwert, Felix & Christopher Winship (2014). Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable. *Annual Review of Sociology*, **40**, 1, 31–53.
- Grung, Bjørn & Manne, Rolf (1998). Missing values in Principal component analysis. *Chemometrics and intelligent laboratory systems*, **42**, 125–139.
- Josse J, Pagès J & Husson F (2009). Gestion des Données Manquantes en Analyse en Composantes Principales. *Journal de la Société Française de Statistique*, **150**, 2, 28–51.
- 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=ZGVmYXVsdGRvbWFpbmnrZWlzdWtla2F3YXRhN3xneDoxOTQwNDhhNDNkZGIwZWNm>, 2021年4月15日最終確認日)。
- 森裕一・黒田正博・足立浩平 (2017). 最小二乗法・交互最小二乗法. 共立出版。
- 中尾走・樊怡舟・村澤昌崇 (2021). 交絡変数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.
- 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。
- 奥井亮 (2014). 因子モデルに関する近年の計量経済学研究の進展. 日本統計学会誌, **43**, 2, 247–273.
- 柴山直 (1992). 欠測値を含む多変量データのための主成分分析的方法. 教育心理学研究, **40**, 257–265.

(2021年4月17日受理, 2021年7月27日採択)

(この間審査 2 回・審査期間合計 72 日)

⁸⁾ 本稿は欠損値補完をせずに主成分を直接抽出するアプローチを使用しているため、欠損値補完後のデータから主成分抽出を行うよりも、主成分抽出に関して、欠損値補完の精度に依存したバイアスは比較的に抑えられると思われる。そのため、一旦他の手法で欠損値補完をしたうえで主成分分析を抽出するというアプローチよりも望ましいと考えている。