

公的統計データを用いた 社会学的研究の可能性とその教育

藤原 翔

(東京大学社会科学研究所)

sho.fujihara@iss.u-tokyo.ac.jp

共同研究集会「官民オープンデータ利活用の動向及び人材育成の取組」

(2019-ISMCRP-5004) プログラム

日時：令和元(2019)年11月15日(金) 9:45～17:50

会場：統計数理研究所 3階セミナー室 1 (D305号室)

社会学における公的統計データの 活用

- 社会学における主要な関心のひとつ
 - 格差・不平等の再生産
 - 親と子どもの社会経済的な地位の関連
- 親の世代の格差・不平等が子の世代の格差・不平等と関連
 - 佐藤俊樹. 2000. 『不平等社会日本』中央公論新社.
 - 吉川徹. 2009. 『学歴分断社会』ちくま書房.
 - 松岡亮二. 2019. 『教育格差』ちくま書房.

どのようにして世代間の格差・不平等を明らかにするのか

- 社会調査
 - 対象者に本人だけではなく親のことについても聞く
 - 社会階層と社会移動調査（SSM調査）
- 対象者
 - 職業， 職歴， 学歴， 収入， 意識・態度， 社会・文化的活動
- 対象者の親の情報
 - 職業（対象者が15歳時， 主な職業）， 学歴， 暮らし向き
 - 収入の情報を得ることは難しい（公的統計を用いた推定は可能。「就業構造基本調査」「全国消費実態調査」）

2

公的統計は社会学者の関心に応えることができるのか

- 対象者の職業はわかっても， 対象者が子どもの頃の親の職業や学歴はわからない
 - 行政データを活用できれば話は別
 - Chetty, R., Hendren, N., Kline, P. and Saez, E., 2014. Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics*, 129(4), pp.1553-1623.
- ただし， **同居**しておりかつ**現在の状況**を分析するのであれば可能
- 世帯員を調査する公的統計の特徴

3

親と子どもの関連に注目

- 子どもが高校卒業くらいまでであれば、同居しており、親と子どもの両方のデータが得られる場合が多い
- 成人した子どもについては無理だが、18歳以下の子どもの格差・不平等について、様々な分析が可能となる
- 教育社会学的関心

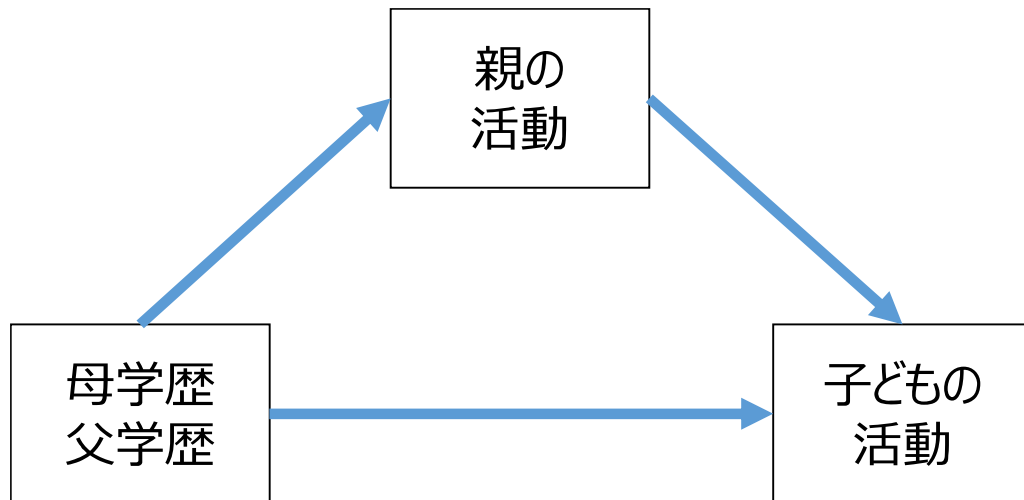
4

アウトカム

- 特に「社会生活基本調査」と「国民生活基礎調査」からは、様々なアウトカムを考えることが可能
 - 子どもの生活時間のパターンなど（胡中 2019）

5

分析戦略



6

説明変数（親）

- 親学歴
- 親職業
- 親収入
- ひとり親
- 祖父母同居
- 子どもの数

• 例えば、「就業構造基本調査」「社会生活基本調査」「国民生活基礎調査」

• 社会階層のBIG3（職業，学歴，収入）を把握可能

7

被説明変数（子ども）

- 子どもの在学状況（中卒・高校中退／高校在学中）
 - 「就業構造基本調査」
- 子どもの主観的健康，メンタルヘルス，ストレスなどの有無
 - 「国民生活基礎調査」
- 子どもの社会・文化的活動，生活時間
 - 「社会生活基本調査」
- 親についても同様の変数を用いることも可能

8

これまでの分析プロジェクト

- 個票データを活用
- 誰が中卒になるのか
 - データ：就業構造基本調査
 - 説明変数：親の学歴，職業，収入
 - 被説明変数：中卒・高校中退／高校在学中
- ひとり親家庭の子どものウェルビーイング
 - データ：国民生活基礎調査
 - 説明変数：ひとり親（死別，離別）
 - 被説明変数：主観的健康，ストレス，メンタルヘルス，中卒・高校中退，喫煙，高校在学中仕事

9

匿名データの活用

・子どもの社会・文化的活動

	男性	女性
美術鑑賞（テレビ・DVDなどは除く）	8.6%	20.2%
音楽会などによるクラシック音楽鑑賞	44.2%	60.3%
CD・テープ・レコードなどによる音楽鑑賞	66.7%	77.4%
楽器の演奏	20.3%	44.1%
趣味としての読書	36.5%	47.5%
遊園地，動植物園，水族館などの見学	39.3%	34.0%
ボランティア活動	24.0%	29.5%

出典) 平成18年社会生活基本調査.

小学・中学・高校在学中で親と同居する子ども. $n = 11,408$. 乗率を使用.

10

親学歴と子どもの社会・文化的活動の関連

- ・被説明変数：子ども美術鑑賞
- ・統制変数：子どもの性別，学年，収入，父職業，母職業，10歳以上人員，10歳未満人員

	Odds Ratio	SE	p	95%CI	
母学歴					
中学	1.025	0.214	0.906	0.681	1.542
高校（基準）	1.000				
短大	1.329	0.110	0.001	1.131	1.562
大学・院	1.425	0.176	0.004	1.119	1.815
父学歴					
中学	0.904	0.140	0.517	0.668	1.225
高校（基準）	1.000				
短大	1.056	0.141	0.685	0.813	1.371
大学・院	1.307	0.121	0.004	1.090	1.567

11

子どもと親の社会・文化的活動の関連

- 先程のモデルに父親美術鑑賞と母親美術鑑賞を追加
- 父親美術鑑賞も母親美術鑑賞も子どもの美術鑑賞と強く関連

	Odds Ratio	SE	p	95%CI	
母美術鑑賞	7.303	0.587	0.000	6.239	8.548
父美術鑑賞	2.886	0.312	0.000	2.335	3.566

12

親の活動は，親学歴の影響をどの程度説明するのか

- KHB法を用いた媒介効果の推定（Karlson et al. 2012）

	Conf_ratio	Conf_Pct	Resc_Fact
母学歴			
中学	0.581	-72.1	6.672
高校（基準）	.	.	.
短大	1.796	44.3	0.885
大学・院	-4.142	124.1	0.754
父学歴			
中学	1.710	41.5	0.824
高校（基準）	.	.	.
短大	-3.528	128.3	0.822
大学・院	1.399	28.5	1.084

13

公的統計データのインパクト

- 巨大なサンプルサイズ
- 世帯を調査
- (狭い意味での) 家族についての分析が可能
 - 親と子だけではなく、夫婦、きょうだいについての分析も可能
- 成人だけではなく子どもの格差・不平等とその生成メカニズムについての分析が可能に

14

匿名データを用いた授業の実践

- 授業スケジュール
- 合計13回（1回105分）、参加者5名程度
- 第1回目：
 - ガイダンス、利用できる匿名データの紹介、先行研究の紹介
 - 課題：調査票と符号表を検討し、研究テーマと使用するデータを定める
- 第2回目：
 - 各自研究関心の報告。先行研究の整理、仮説、変数、モデルの検討。
- 第3回目：
 - 各自研究関心の報告。先行研究の整理、仮説、変数、モデルの検討。

15

- 第4回目：
 - 変数の事前処理. ケースの選択, 変数のリコードなど
- 第5回目：
 - 変数の事前処理. ケースの選択, 変数のリコードなど
- 第6回目～第9回目
 - データの分析とレポートの執筆
 - 課題：分析の進捗を共有
- 第10回目
 - 中間報告
- 第11回目, 第12回目
 - データの分析とレポートの執筆
- 第13回目
 - 最終報告

16

難しさ

- RやStataでデータ分析を行ったことがある院生でないと難しい
 - 分析手法の説明をしている時間はない
- 院生は大きなデータを使い慣れていない
 - マシンのパワーが多少は必要
- 文字列の処理
 - △などの処理 (destring)
 - 普段Rなどを使用していないとつまづく
- データの申請および届くタイミングを考えて授業を計画
- 金銭的負担
 - 運営費

17

利用した匿名データ（平成29年度）

労働力調査	平成元年，14年，24年	各1
住宅・土地統計調査	平成15年	1
全国消費実態調査	平成16年	1
社会生活基本調査（調査票A・生活時間編）	平成18年	1
社会生活基本調査（調査票A・生活行動編）	平成18年	1
就業構造基本調査	平成9年，平成19年	各1

• 学会報告

- 胡中孟徳. 2018. 「子どもの生活時間と家族の関連の実証研究：母親の職業・就業に着目して」第65回数理社会学会大会，2018年3月15日.
- 吉田航. 2018. 「転職後の雇用形態に学歴が与える効果についての分析：「就業構造基本調査」匿名データを用いて」第65回数理社会学会大会，2018年3月14日

18

- 匿名データという制限の中で，仮説を考え結果を示すことの重要性
- 若手研究者に公的統計データに関心を持ってもらうきっかけに
 - 最終的には個票データにチャレンジ
- 社会調査でできること，公的統計でできることを考える
 - 公的統計で分かることはできるだけ公的統計のデータを使用し，それでもわからないことを社会調査で明らかにする

19

公的統計データを社会調査に活かす

- 「就業構造基本調査」を用いた231の職業についての職業的地位尺度の作成（Fujihara 2019）
- 「就業構造基本調査」を用いた職業， 学歴情報を用いた親の収入の推定
- 「全国消費実態調査」を用いた所有財による収入の推定

20

文献

- Karlson, K. B., A. Holm, and R. Breen. 2012. Comparing regression coefficients between same-sample nested models using logit and probit: A new method. *Sociological Methodology* 42: 286-313.
- 胡中孟徳. 2019. 「学校完全5日制による土曜日の生活時間の変化」『教育社会学研究』104:259–278.
- Fujihara, Sho. 2019. “Socio-Economic Standing and Social Status in Contemporary Japan: Scale Constructions and Their Applications” Unpublished manuscript (Under Review).

21

付記

- 本研究はJSPS科研費JP18H05204の成果である.
- 本研究は, 統計法に基づいて、独立行政法人統計センターから「社会生活基本調査」(総務省)の匿名データの提供を受け, 独自に作成・加工した統計であり, 総務省が作成・公表している統計等とは異なる.