

総務省四国行政評価支局
平成30年度政策評価に関する統一研修

政策分析を通じた評価手法
～航空行政を中心に～

福井秀樹


愛媛大学法文学部

fukui.hideki.hz@ehime-u.ac.jp

平成30年12月4日

目次

1. 国土交通省航空局への出向：いきさつと業務
2. 政策レビュー評価書作成にたずさわり感じた問題点
3. 政策分析のトレンド：ランダム化比較試験（RCT）
4. 官庁における政策評価の実態は？
5. ランダム化比較試験（RCT）が行えない場合：
 マッチング、傾向スコア分析
6. 政策デザイン、データの問題点：
 米国との比較から見えること
7. 私見：「遅い思考＝熟慮」としての政策評価の重要性
8. 「遅い思考＝熟慮」としての政策評価の実現に必要なのは
9. 主要参考文献



1. 国土交通省航空局への出向

いきさつと業務

出向のいきさつと業務

- 出向のいきさつ

- 2015年2月 打診
- 2015年4月 – 2017年3月 出向
 - 国土交通省 航空局 航空戦略課 政策調査室長

- 業務

- 航空局の所掌事務に関する総合的な政策の企画及び立案に必要な調査
- 外国の航空政策及び航空事情に関する調査
- 政策レビュー(総合評価方式)評価書の作成
 - 『LCCの事業展開の促進』 (2017年)
 - 本文 <http://www.mlit.go.jp/common/001179271.pdf>
 - 概要 <http://www.mlit.go.jp/common/001179269.pdf>

担当した政策レビュー評価書の概要

- LCC事業展開の促進政策の進捗と影響について評価

政策の目的

- 新たな航空需要の喚起、航空サービス多様化の促進

評価の視点

- 施策の進捗や施策が航空需要に与えた影響を評価

評価手法


- LCC就航後の旅客数やシェアの年推移データ、新規参入路線数データ等の現状を分析

評価結果

- LCCは、FSAの需要を奪うことなく、航空旅客数全体の増加に貢献

政策への反映の方法

- 観光や地方創生の観点を重視して、地方空港のゲートウェイ機能強化と併せ、LCC就航をさらに促進



2. 政策レビュー評価書作成に たずさわり感じた問題点

政策評価にたずさわり感じた問題点

- データの制約
 - 例：航空輸送統計年報は2011年分より航空会社別の実績を掲載せず（統計法改正による）。
- より厳密なデータ分析（因果推定など）の必要性



3. 政策分析のトレンド

ランダム化比較試験 (RCT)

政策分析で重視されていること

- より厳密な手法による因果推定
最善の方法はランダム化比較試験（RCT）。

例

- 対象を、ランダム割付により、2群に分ける。
- 処置群に介入（政策）を施し、対照群には何もしない。
- 両群の結果を比較し、介入（政策）の効果を推定する。

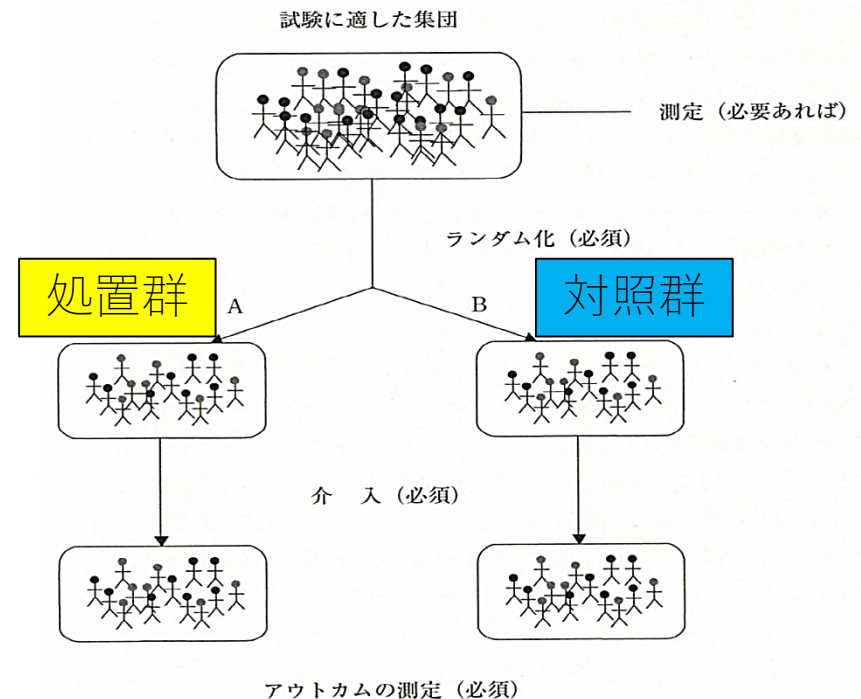


図1.1 ランダム化比較試験の概念図

ランダム化はなぜ重要か

- 介入（政策）の効果をより正確に取り出すことを可能にするため。

ランダム化は

- 介入（政策）の決定に影響を及ぼすその他の要因（共変量、交絡変数）を平均としてバランスさせる。
- そのため、処置群と対照群のアウトカムの平均値に差をもたらすのは「介入（政策）」のみ、ということになる。

ランダム化がなされないと介入効果の推定を誤りうる

• 空港民営化という介入（政策）を例に

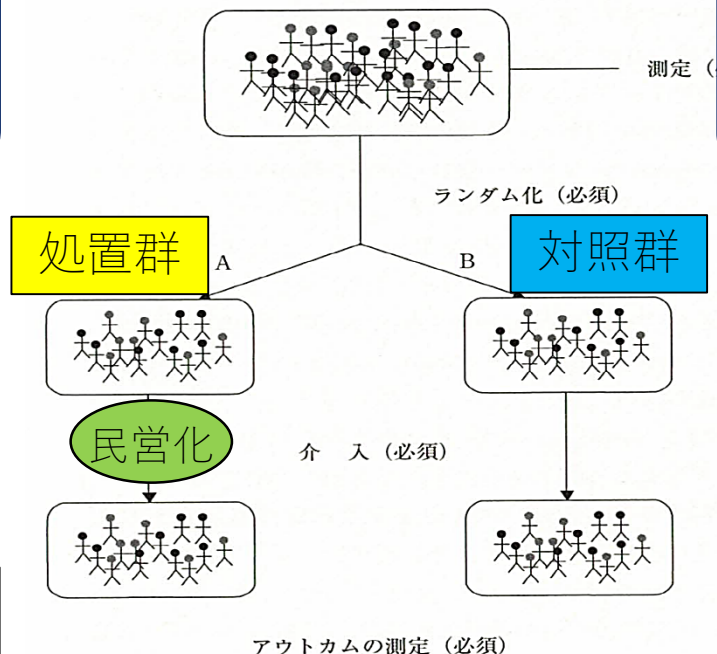
この空港はインバウンドが好調。民営化でさらに伸びるだろう。民営化が失敗と批判される可能性は低いだろう。

民営化前のインバウンド
・ 100万人
・ 3年前の1.5倍

交絡変数（例）
人口 = 200万人

民営化後のインバウンド
・ 200万人

ランダム化がなされない場合



この空港はインバウンドが不調。民営化で伸びることはないだろう。民営化すれば失敗と批判される可能性あり。

民営化前のインバウンド
・ 50万人
・ 5年前とほぼ同じ


交絡変数（例）
人口 = 100万人

民営化後のインバウンド
・ 55万人

図1.1 ランダム化比較試験の概念図

問：民営化は「インバウンド増加を促す」と結論できるか？

答：難しい。介入（政策）の効果とは別に、交絡変数の大きな違いが処置群のインバウンド増加に影響している可能性があるため。




4. 官庁における政策評価の 実態は？

政策評価を難しくする要因にしばしば直面

- 介入（政策）の対象が、関係者のほぼ全てというケースがある。
 - この場合、対照群が存在しない。そのため、介入（政策）の効果が正確に推定できない。
例：LCC関連の施策は実はこのケースに近い。
- 効果が期待できる対象のみに介入（政策）を施すことがある。
 - この場合、介入（政策）がランダム化されていない。
 - そのため、処置群と対照群で、その他の要因がアウトカムに及ぼす影響に差があり、介入（政策）の効果が正確に推定できない。
例：選別による民営化、補助等。

つまり現実の政策は

- ランダム化比較試験（RCT）に適合的な形では行われていないことが多い（と思われる）。
- そのため、現状では、厳密な政策評価はかなり難しい。
- しかし、介入（政策）をランダムに実施できるか？
- これも難しいだろう。



5. ランダム化比較試験 (RCT) が 行えない場合

マッチング、傾向スコア分析

ランダム化比較試験（RCT）が行えない場合

- 観察データを使うしかない。
- 観察データの問題点
 - 処置群、対照群で、交絡変数がバランスしていない（違いが大きい）。
 - そのため、処置群と対照群のアウトカムの平均値に差をもたらすのは「介入（政策）」のみ、とは言えない（交絡変数の影響かもしれない）。
 - 従って、介入（政策）の効果は正確には分からない。

観察データの問題点を緩和する分析手法

- マッチング

- 処置群、対照群それぞれから、交絡変数（例：人口）の値が近いデータ・ユニットを取り出しペアをつくる。
- 交絡変数の影響は、処置群、対照群のいずれにおいても、平均的には同じとなる。
- その結果、介入（政策）の効果をより正確に取り出すことが可能になる。

- マッチングの問題点

- 交絡変数が多くなると、よいマッチングが難しくなる = 次元の呪い

次元の呪いを克服する手法

- 傾向スコア分析(Cf., Guo & Fraser, 2014)
 - 介入（政策）を受ける「傾向」を、交絡変数を用いて推定し、数値化する。
 - それにより得られる「傾向スコア」が近いデータ・ユニットを、処置群、対照群それぞれから取り出し、ペアをつくる。
 - 傾向スコアを利用してデータに重み付けをする手法等もある。
 - その結果、処置群、対照群の間で、交絡変数の影響はほぼ均質化される。つまり、擬似ランダム化が可能になる。

- 限界


- 観察・測定（データ化）できない交絡変数をもたらすバイアス（効果推定の歪み）は取り除けない。
- データの収集が不完全な場合も当然、交絡変数をもたらすバイアスは取り除けない。

- 対処法

- 交絡変数の徹底的な洗い出し。
- データの徹底的な収集。

小括：RCT、傾向スコア、政策評価

- より厳密な政策評価を行うには
 - ランダム化比較試験（RCT）が最善。
- 仮にRCTが不可能でも
 - 傾向スコア分析が使える。
- そのために必要なこと
 - RCTや傾向スコア分析を意識した政策デザイン。
 - 綿密な交絡変数の洗い出し。
 - 緻密なデータ収集。



6. 政策デザイン、データの問題点

米国との比較から見えること

航空局で政策レビューに関わり感じた問題点

• 政策デザイン

- 評価対象の政策（施策）自体が、RCTや傾向スコア分析を意識した政策デザインになっていない（もしくは、できない）。
 - 例：施策が全航空会社を対象としている。
- そのため、対照群が存在しない。

• データ

- 活用を意識したデータ収集・整理が不十分。
 - 例：航空輸送統計年報は会社別データを使えない。
データベース形式でデータが整理されていない。

米国の状況（航空関係政策に限定）

- 政策デザイン

- 日本のそれと同様の課題が認められる。
- RCTや傾向スコア分析を意識した政策デザインになっているとは言い難いものが多いと感じられる。

- データ

- 分析への活用を意識したデータ収集・整理が、かなり徹底して実施されている。

米国運輸省運輸統計局のポータルサイト

- 運輸関係データのポータルサイト。
- 航空に限っても、膨大なデータにアクセスできる。
 - このポータルからアクセスできない運輸関連データももちろんある。
- データベース形式で整理されているため、PCですぐに分析に取り掛かることができる。

Bureau of Transportation Statistics

Ask-A-Librarian | A-Z Index

Search this site:

Advanced Search

Resources

- Database Directory
- Glossary
- Upcoming Releases
- Data Release History

Data Finder

By Mode

- Aviation
- Maritime
- Highway
- Transit
- Rail
- Pipeline
- Bike/Pedestrian
- Other

By Subject

- Safety
- Freight Transport
- Passenger Travel
- Infrastructure
- Economic/Financial
- Social/Demographic
- Energy
- Environment
- National Security

Quick Answers

- Carrier Snapshots
- Airline Fuel Cost and Consumption
- Air Freight Summary
- Employment
- Airport Snapshots
- Holiday Flight Delays
- Inter-Airport Distances
- Tarmac Times

Airline Activity : National Summary (U.S. Flights)

	2017 *	2018 *	Change
Enplaned Passengers (million)	732	764	4.4%
Departures (000)	8,654	8,744	1.0%
Freight/Mail (million lbs)	21,320	21,917	2.8%
Load Factor (%)	84.6	84.7	0.1 points
Airlines with scheduled service	99	98	-1.0%

* 12 months ending July of each year

Airline Domestic Market Share August 2017 - July 2018

Domestic Revenue Passenger Miles (billions)

Airlines	Share
American	18.1%
Southwest	17.9%
Delta	16.8%
United	14.9%
JetBlue	5.5%
Alaska	5.4%
Spirit	3.7%
Frontier	2.7%
SkyWest	2.7%
Hawaiian	1.6%
Other	10.7%

Mouseover each bar for the total.

At a Glance

Flight Delays more...

Percent of U.S. Flights On Time (2017-2018)

Click a bar for details. Mouseover it for percentage.

Average Air Fares more...

Average Domestic Airline Fares

Click a bar for details. Mouseover it for total.

TranStats

<https://www.transtats.bts.gov/>

航空関係データの例

- Air Carrier Financial Reports (Form 41 Financial Data)
 - 航空各社財務データ(四半期: 1990年以降)
- Air Carrier Statistics (Form 41 Traffic)
 - 航空各社国内・国際運航データ(月次: 1990年以降)
- Airline On-Time Performance Data
 - 航空各社国内定時運航データ(月次: 1987年以降)
- Airline Origin and Destination Survey (DB1B)
 - 航空各社の国内運賃データ(10%サンプル)
(四半期: 1990年以降)

Air Carrier Statisticsと航空輸送統計年報との違い

Air Carrier Statistics

1	YEAR	QUARTER	MONTH	ORIGIN	ORIGIN_C	ORIGIN_C	ORIGIN_S
2	2000	1	1	ABE	Allentown	2200	PA
3	2000	1	1	ABE	Allentown	2200	PA
4	2000	1	1	ABE	Allentown	2200	PA
5	2000	1	1	ABE	Allentown	2200	PA
6	2000	1	1	ABE	Allentown	2200	PA
7	2000	1	1	ABE	Allentown	2200	PA
8	2000	1	1	ABE	Allentown	2200	PA
9	2000	1	1	ABE	Allentown	2200	PA
10	2000	1	1	ABE	Allentown	2200	PA
11	2000	1	1	ABE	Allentown	2200	PA
12	2000	1	1	ABE	Allentown	2200	PA
13	2000	1	1	ABE	Allentown	2200	PA
14	2000	1	1	ABE	Allentown	2200	PA
15	2000	1	1	ABE	Allentown	2200	PA
16	2000	1	1	ABE	Allentown	2200	PA
17	2000	1	1	ABE	Allentown	2200	PA
18	2000	1	1	ABE	Allentown	2200	PA
19	2000	1	1	ABE	Allentown	2200	PA
20	2000	1	1	ABE	Allentown	2200	PA
21	2000	1	1	ABE	Allentown	2200	PA
22	2000	1	1	ABE	Allentown	2200	PA

航空輸送統計年報

1	第4表 国内定			
2				
3				
4				
5	区分	区間 (km)	運航回数	運航キロメートル
6	月別			
7		日本航空	東京(羽田)	一札幌(千歳)
8	1	東京-千歳	786	657984
9	2	894Km	754	674076
10	3		762	631228
11	4		782	654408
12	5		788	704472
13	6		791	707154
14	7		889	750066
15	8		821	733974
16	9		844	754586
17	10		827	739888
18	11		788	700002
19	12		828	735762
20	年	東京-千歳	4740	4237560
21	年	千歳-東京	4760	4255440
22	計	計	9500	8493000
23	1	東京-千歳	798	708942
24	2		776	698744
25	3		811	725084
26	年	東京-千歳	4806	4296564
27	年	千歳-東京	4822	4310868
28	計	計	9628	8607432
29		全日本空輸	東京(羽田)	一札幌(千歳)
30	1	東京-千歳	820	733080
31	2	894Km	790	706260

- Air Carrier Statisticsは各社月次集計データ。データベース形式。
- 航空輸送統計年報も各社月次集計データ(但し2010年まで)。データ整形が必要。

On-time performance dataと定時運航率等データとの違い

On-time performance data

1	Year	Quarter	Month	DayofMon	DayOfWee	FlightDate	UniqueCa	AirlineID	Carrier	TailNum	FlightNum	OriginAirp	OriginAirp	OriginCity	Origin	OriginCity
2	2014	1	1	1	3	2014/1/1	AA	19805	AA	N338AA	1	12478	1247802	31703	JFK	New York,
3	2014	1	1	2	4	2014/1/2	AA	19805	AA	N338AA	1	12478	1247802	31703	JFK	New York,
4	2014	1	1	3	5	2014/1/3	AA	19805	AA	N323AA	1	12478	1247802	31703	JFK	New York,
5	2014	1	1	4	6	2014/1/4	AA	19805	AA	N327AA	1	12478	1247802	31703	JFK	New York,
6	2014	1	1	5	7	2014/1/5	AA	19805	AA	N323AA	1	12478	1247802	31703	JFK	New York,
7	2014	1	1	6	1	2014/1/6	AA	19805	AA	N319AA	1	12478	1247802	31703	JFK	New York,
8	2014	1	1	7	2	2014/1/7	AA	19805	AA	N328AA	1	12478	1247802	31703	JFK	New York,
9	2014	1	1	8	3	2014/1/8	AA	19805	AA	N323AA	1	12478	1247802	31703	JFK	New York,
10	2014	1	1	9	4	2014/1/9	AA	19805	AA	N339AA	1	12478	1247802	31703	JFK	New York,

定時運航率等データ

9	(1) 定時運航率(全体の便数に占める出発予定時刻以降15分以内に出発した便数の割合)												(単位:%)		
10	日本航空(株) (法人番号 7010701007666)		全日本空輸(株) (法人番号 1010401099027)		日本トランスオーシャン航空 (法人番号 3360001001727)		スカイマーク (法人番号 7010801019529)		AIRDO (法人番号 6430001021797)		ソラシドエア (法人番号 2350001002669)				
11															
12	H29.4~6	94.38 (94.65)	92.18 (89.77)	92.39 (89.67)	93.03 (92.13)	95.31 (91.15)	90.95 (91.90)								
13	H29.7~9	90.66 (91.92)	87.43 (88.70)	84.60 (88.55)	92.18 (88.55)	87.93 (86.11)	86.30 (87.18)								
14	H29.10~12	87.66 (89.85)	88.47 (88.53)	84.96 (86.44)	93.90 (88.79)	90.14 (81.19)	86.75 (88.41)								
15	H30.1~3	87.40 (90.42)	85.93 (89.94)	83.72 (86.68)	93.19 (89.62)	86.58 (85.43)	87.00 (87.55)								
16	H29年度上期	92.49 (93.26)	89.77 (89.23)	83.47 (88.51)	92.59 (90.25)	91.57 (88.61)	88.62 (89.50)								
17	H29年度	90.01 (91.68)	88.52 (89.23)	83.90 (87.55)	93.06 (89.72)	90.06 (86.00)	87.75 (88.74)								
18															
19															
20	スターフライヤー (法人番号 6290801006558)		Peach Aviation (法人番号 7120101047384)		ジェットスター・ジャパン (法人番号 3040001076850)		バニラエア (法人番号 6010401095509)		春秋航空日本 (法人番号 7010601043349)		エアアジア・ジャパン (法人番号 6180001113372)		計		
21	H29.4~6	92.89 (94.58)	83.00 (80.31)	87.11 (80.67)	80.26 (80.59)	89.82 (92.40)	-	-	92.09 (91.06)						
22	H29.7~9	89.69 (90.74)	77.87 (78.81)	86.53 (76.52)	84.63 (83.52)	79.65 (85.02)	-	-	88.36 (88.67)						
23	H29.10~12	88.62 (92.37)	79.87 (77.17)	86.90 (81.75)	82.47 (81.01)	83.55 (74.84)	90.00 (88.01)		88.06 (88.01)						
24	H30.1~3	88.27 (91.01)	76.49 (70.14)	79.99 (80.47)	76.02 (77.80)	74.49 (74.70)	90.71 (86.27)		86.27 (88.59)						
25	H29年度上期	91.25 (92.61)	80.35 (79.55)	86.80 (78.47)	82.55 (86.74)	84.74 (88.45)	-	-	90.19 (89.93)						
26	H29年度	89.86 (92.15)	79.17 (76.75)	85.17 (79.75)	80.94 (83.02)	83.22 (80.17)	90.42 (89.07)		88.70 (89.07)						
27		(注)1. 割合については、運航便数に対する割合。													
28		2. ()内は、前年同期の数字。													
29		3. 「日本航空」は、日本航空、ジェイエア及び平成28年10月30日以降の北海道エアシステムの合計。「全日本空輸」は、全日本空輸、ANAウイングスの合計。													
30		4. エアアジア・ジャパンは、平成29年10月29日より運航を開始。													

- 米国のOn-time performance dataは各社1便ごとのデータ。データベース形式。
- 日本の定時運航率等データは四半期集計データ。データ整形が必要。

豊富なデータの公開は何を可能にするのか

- 世界中の誰でもデータを自由に利用・分析。

例

- Airline On-Time Performance Dataを利用
 - Fukui, H., & Nagata, K. (2014). Flight cancellation as a reaction to the tarmac delay rule: An unintended consequence of enhanced passenger protection. *Economics of Transportation*, 3(1), 29–44.
- Air Carrier Financial Reportsを利用
 - Fukui, H., & Miyoshi, C. (2017). The impact of aviation fuel tax on fuel consumption and carbon emissions: The case of the US airline industry. *Transportation Research Part D: Transport and Environment*, 50, 234–253.
- Airline Origin and Destination Survey (DB1B)を利用
 - Fukui, H. (2018). How do slot restrictions affect airfares? New evidence from the US airline industry. *Economics of Transportation*, Under Review (Revised Manuscript).

豊富なデータの公開は何を可能にするのか（続）

- データに基づくより客観的な論争が可能に。
 - Fukui & Nagata (2014)
 - ターマック上の遅延に関する規則の効果を分析。
 - 同規則の目的：旅客サービス水準の向上。
 - 分析結果：意図せざる効果が発生、すなわち、目的に反する副作用（欠航増大）が生じている。
- ↑
- 米国運輸省からの挑戦状
 - 2016年3月2日にOffice of Inspector GeneralのエコノミストBetty Krier氏からメール。
 - 議会の命令により、規則が欠航に及ぼした影響を分析した。我々の分析結果についてコメントを頂きたい。

福井のコメントに対してOIGはどう反応したか

- Office of Inspector Generalの分析結果
 - 先行研究と方向性は同じ。
 - しかし、欠航増大の規模について、より控えめな推定結果。
- 福井は、分析方法やデータセット作成方法に関する問題点を指摘。
- しかし、Office of Inspector Generalは問題点の指摘に対応せず分析結果を公表。

出典：US DOT. (2016). Effects of the Tarmac Delay Rule on Flight Cancellations and Delays (Report Number: ST-2017-003). Washington, DC.

<https://www.oig.dot.gov/library-item/33850>

とはいえ、データに対する米国の姿勢は尊敬に値する

- 米国

- 豊富なデータを使いやすい形で公開。

- 世界中の人々が自発的に分析。
- 自由な政策分析を促進。
- 結果的に、政策評価の改善も促進。

- 日本

- 限定的なデータを使いにくい形で公開。

- 海外の人だけでなく日本人も分析できない。
- 政策分析の停滞。
- 結果的に、政策評価の改善にもつながらない。



7. 私見

「遅い思考 = 熟慮」 としての
政策評価の重要性

私見（1）政策評価が直面する課題

- 政策レビュー（総合評価方式）について
 - 評価枠組の検討と系統的なデータ収集が必要。
 - テーマの決定と同時に評価枠組（RCT、傾向スコア分析等）も検討し、系統的にデータを収集することが望ましい。
 - 評価枠組・手法の知識・技能の蓄積・継承が必要。
 - 人事異動によりそれが難しいのであれば、外部人材の力を借りてもよいのではないか。

私見（2）政策評価が直面する課題（続）


- 政策レビューと政策との関係について
 - 評価結果を評価対象である政策に反映させることはどこまで可能か？
 - 変化の速い政策領域では、政策形成・実施が政策評価の先を歩み続ける可能性がある。
 - 評価結果を、評価対象である政策に反映させることは重要だが、かなり大きな時間的ずれが生じる。

私見 (3) それでは政策レビューの位置づけは？

- ダニエル・カーネマン『ファスト&スロー：あなたの意思はどのように決まるか?』（原著：2012年）を参考に
 - 人間には2つの思考モードがある。
 - 速い思考(ファストな「直感」)： 日常の多くの判断
 - 遅い思考(スローな「論理」)： 熟慮
 - 速い思考はおおむね適切。
 - しかし、衝動的に発動しバイアスの影響を受けがち。
 - 遅い思考はより正確な判断を下せる。
 - しかし、遅くて効率が悪い。

私見（4）政策レビューに期待する役割

- 「遅い思考 = 熟慮」としての政策評価の役割
 - 政策形成・実施の現場では、速さ（「速い思考 = 瞬発力」）が求められることが多いと感じられた。
 - しかし、「速い思考」で形成・実施される政策はバイアスの影響を受けている部分もあるだろう。
 - 「遅い思考」としての政策評価は、「速い思考」のバイアスを検出し、より長期の広い視点から政策の改善を促す役割を与えられるべきではないか。



8. 「遅い思考 = 熟慮」 としての
政策評価の実現に必要なものは


「遅い思考」としての政策評価を官庁は実施できるか？

- 難しいのではないか。
 - 「速い思考」のバイアスの検出とは、具体的には、政策がもたらした(しばしば意図しない)ネガティブな効果を明らかにする作業。
 - 身内の批判は難しいだろう。


「遅い思考」としての政策評価を実現するには

- 外部（出向者）の力を借りる。
 - ただし、自己評価の限界を免れるわけではない。
- 豊富なデータを使いやすい形で公開する。
 - 世界中の人々が自発的に分析に取り組み、政策評価の改善も促進される。
 - もっとも、必要性の高い政策評価が行われるとは限らない。

まとめ：いずれにせよ、政策評価の改善に重要なのは



良質なデータ の 収集・整理・公開



9. 主要参考文献

主要参考文献

1. ダニエル・カーネマン『ファスト&スロー：あなたの意思はどのように決まるか?』(上)(下) (早川書房、2014年)
2. 国土交通省航空局『LCCの事業展開の促進』 (国土交通省、2017年)
3. D. J. トーガーソン & C. J. トーガーソン『ランダム化比較試験(RCT)の設計：ヒューマンサービス、社会科学領域における活用のために』 (日本評論社、2010年)
4. Fukui, H. (2018). How do slot restrictions affect airfares? New evidence from the US airline industry. *Economics of Transportation*, Under Review (Revised Manuscript).
5. Fukui, H., & Miyoshi, C. (2017). The impact of aviation fuel tax on fuel consumption and carbon emissions: The case of the US airline industry. *Transportation Research Part D: Transport and Environment*, 50, 234–253.
6. Fukui, H., & Nagata, K. (2014). Flight cancellation as a reaction to the tarmac delay rule: An unintended consequence of enhanced passenger protection. *Economics of Transportation*, 3(1), 29–44.
7. Guo, S., and Fraser, M. W. (2014). *Propensity score analysis: Statistical methods and applications*, 2nd Edition. Sage.
8. US Department of Transportation (US DOT). (2016). *Effects of the Tarmac Delay Rule on Flight Cancellations and Delays* (Report Number: ST-2017-003). Washington, DC.