

# 인과추론: 실험

서울대학교 정치외교학부  
2018년 겨울 방법론 캠프

---

서강대학교 정치외교학과  
하상응

# 실험?

---

## 경험 연구 설계

- 실험(experiment)
- 관찰(observation; non-experimental methods)

실험 연구 설계에서는 관심있는 독립변수를 실험으로 조작(experimentally manipulate)하여 인과관계를 파악한다

- $Y = \beta_0 + \beta_1 X + \varepsilon$
- $\text{Outcome} = \beta_0 + \beta_1 \text{Treatment} + \varepsilon$

Treatment가 “random assignment” 방식을 통해 조작되었다면  $\text{corr}(\text{Treatment}, \varepsilon) = 0$ 이 달성됨 (in theory; in expectation);  $\beta_1$ 의 추정값인  $b_1$ 에 편향이 없음

# The Logic of Randomized Experiments

---




# 실험의 종류 및 문제점

---

무작위 실험(randomized experiments)은 관찰 연구(observational studies)를 평가할 수 있는 척도!

- 실험실 실험 (lab experiment)
- 현장 실험 (field experiment)
- 서베이 실험 (survey experiment)
- 유사 실험 (quasi-experiment) 혹은 자연 실험 (natural experiment)

실험은 관찰 연구의 내적 타당성(internal validity)과 관련된 문제의 대부분을 해결해 주나 나름의 내적 타당성 문제가 여전히 상존한다 (ABC problems)

- Attrition: 결과값이 누락되는 상황
  - Balance: 처리군과 대조군 간에 차이가 있는 경우
  - (Non-)Compliance: 처리군에 속한 피실험대상 일부가 처리를 받지 못하는 경우
  - ( + Spill-over): 대조군에 속한 피실험대상 일부가 처리를 받는 경우
- 

# Random Assignment

---

피실험대상을 처리군과 대조군으로 나누는 작업

- 동전 던지기
- 제비뽑기 (lottery)
- Random number generator

이 과정을 거치면 처리군과 대조군 간의 차이는 “처리를 받는지의 여부” 밖에는 없을 것이라 기대됨 (피실험대상의 수, 즉  $n$ 이 클수록 이러한 상황이 발생할 가능성은 커짐)

하지만 실제로는 처리군과 대조군 간에 의미있는 차이가 있을 수 있음

주의: random assignment  $\neq$  random sampling (실험대상은 반드시 무작위 추출된 표본으로 구성할 이유는 없음)

# Random Assignment

Subject ID	X	성별	교육수준	세대
1	T	여성	대졸	중년
2	T	남성	고졸	중년
3	C	여성	대졸	중년
4	C	여성	대졸	중년
5	C	남성	대졸	중년
6	T	남성	대졸	중년
7	C	남성	고졸	노년
8	T	남성	대졸	노년
9	T	여성	고졸	노년
10	C	남성	고졸	노년
...	...	...	...	...

# Block Randomization

Subject ID	성별 (block)	X	...	...
1	여성	T	여러 변수를 동시에 blocks으로 활용할 수 있음...	
2	여성	T		
3	여성	T		
4	여성	C		
5	여성	C		
6	여성	C		
7	남성	T		
8	남성	T		
9	남성	C		
10	남성	C		
...	....	...		

# Checking for balance (between T and C)

---

- $Y = \beta_0 + \beta_1 \text{Treatment\_Assigned} + \varepsilon$
- $X = \gamma_0 + \gamma_1 \text{Treatment\_Assigned} + u$ 
  - (X: 자료에 존재하는, 실험으로 조작된 변수와 결과변수를 제외한 나머지 변수들)
- 여기서 얻은  $\gamma_1$ 이 “통계적으로 무의미”하면 처리군과 대조군 간의 balance가 확보된다고 볼 수 있음 (물론 관찰되지 않아서 자료에 포함되지 않은 X에서 imbalance가 있을 가능성을 완전히 배제하지는 못함)



# 실험자료 회귀 분석 시 통제변수의 활용

---

- $Y = \beta_0 + \beta_1 \text{Treatment\_Assigned} + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n + \varepsilon$
- 통제변수들에 딸린 회귀계수는 아무 의미 없음 (해석해 주지 않음)
- 통제변수들을 포함하면  $\beta_1$ 의 추정치( $b_1$ )에 딸린 표준 오차가 작아지는 효과를 얻음 (precision이 높아짐)
  - 예: <https://www.cambridge.org/core/journals/american-political-science-review/article/social-pressure-and-voter-turnout-evidence-from-a-largescale-field-experiment/11E84AF4C0B7FBD1D20C855972C2C3EB>

# Balance Check: 예 (Progresa in Mexico)

---

- $\text{Health} = \beta_0 + \beta_1 \text{Aid} + \varepsilon$
- 만약 어떤 특정 변수에서 처리군과 대조군 간에 의미있는 차이가 발견되면, 그 변수를 통제하는 것이 바람직함
- $\text{Health} = \beta_0 + \beta_1 \text{Aid} + \beta_2 \text{Mother\_Educ} + \varepsilon$
- 하지만 그 변수가 실험을 집행한 이후 측정된 변수라면 주의해야 됨 (post-treatment variable bias)

**TABLE 10.1** Balancing Tests for the Progresa Experiment: Differences of Means Tests Using OLS

Dependent variable	$\hat{\gamma}_0$	$\hat{\gamma}_1$	$t \text{ stat } (\hat{\gamma}_1)$	$p \text{ value } (\hat{\gamma}_1)$
1. Age (in years)	1.61	0.01	0.11	0.91
2. Male	0.49	0.02	1.69	0.09
3. Child was ill in last 4 weeks	0.32	0.01	0.29	0.77
4. Father's years of education	3.84	−0.04	0.03	0.98
5. Mother's years of education	3.83	−0.33	1.87	0.06
6. Father speaks Spanish	0.93	0.01	1.09	0.28
7. Mother speaks Spanish	0.92	0.02	0.77	0.44
8. Own house	0.91	0.01	0.73	0.47
9. House has electricity	0.71	−0.07	1.69	0.09
10. Hectares of land owned	0.79	0.02	0.59	0.55
11. Male daily wage rate (pesos)	31.22	−0.74	0.90	0.37
12. Female daily wage rate (pesos)	27.84	−0.58	0.69	0.49
Sample size:	7,825			

*Results from 12 different OLS regressions in which the dependent variable is as listed at left. The coefficients are from the model  $X_i = \gamma_0 + \gamma_1 \text{Treatment}_i + v_i$  (see Equation 10.2).*

여|

<https://www.aeaweb.org/articles?id=10.1257/0002828042002561>

**TABLE 10.11** Variables for Resume Experiment

Variable	Description
education	0 = not reported; 1 = some high school; 2 = high school graduate; 3 = some college; 4 = college graduate or more
yearsexp	Number of years of work experience
honors	1 = resume mentions some honors
volunteer	1 = resume mentions some volunteering experience
military	1 = Applicant has some military experience
computerskills	1 = resume mentions computer skills
afn_american	1 = African-American-sounding name; 0 = white-sounding name
call	1 = applicant was called back ; 0 applicant not called back
female	1 = female; 0 = male
h_quality	1 = High-quality resume; 0 = low-quality resume

# Balance Check

(1. 하나의 회귀식으로...; 2. OLS 사용)

```
. reg afn_american education yearsexp honors volunteer computerskills h_quality female
```

Source	SS	df	MS	Number of obs	=	4,870
Model	1.44501429	7	.206430612	F(7, 4862)	=	0.83
Residual	1216.05499	4,862	.250114148	Prob > F	=	0.5660
				R-squared	=	0.0012
				Adj R-squared	=	-0.0003
Total	1217.5	4,869	.250051345	Root MSE	=	.50011

afn_american	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
education	-.0013209	.0101267	-0.13	0.896	-.0211738	.018532
yearsexp	.0001329	.0014479	0.09	0.927	-.0027057	.0029715
honors	-.0092397	.032678	-0.28	0.777	-.0733034	.054824
volunteer	.0166645	.0235155	0.71	0.479	-.0294365	.0627655
computerskills	.0411996	.0199083	2.07	0.039	.0021703	.0802288
h_quality	-.0198024	.0236088	-0.84	0.402	-.0660863	.0264814
female	.0075099	.0176768	0.42	0.671	-.0271447	.0421645
_cons	.4677313	.0439395	10.64	0.000	.3815901	.5538726

```
. reg call afn_american
```

Source	SS	df	MS	Number of obs	=	4,870
Model	1.24928131	1	1.24928131	F(1, 4868)	=	16.93
Residual	359.197536	4,868	.073787497	Prob > F	=	0.0000
				R-squared	=	0.0035
				Adj R-squared	=	0.0033
Total	360.446817	4,869	.074028921	Root MSE	=	.27164

call	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
afn_american	-.0320329	.007785	-4.11	0.000	-.0472949	-.0167708
_cons	.0965092	.0055048	17.53	0.000	.0857173	.1073011

```
. reg call afn_american education yearsexp honors volunteer computerskills h_quality female
```

Source	SS	df	MS	Number of obs	=	4,870
Model	4.59499233	8	.574374041	F(8, 4861)	=	7.85
Residual	355.851825	4,861	.073205477	Prob > F	=	0.0000
				R-squared	=	0.0127
				Adj R-squared	=	0.0111
Total	360.446817	4,869	.074028921	Root MSE	=	.27057

call	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
afn_american	-.0312952	.0077588	-4.03	0.000	-.046506	-.0160844
education	-.0036209	.0054786	-0.66	0.509	-.0143614	.0071197
yearsexp	.0025694	.0007833	3.28	0.001	.0010337	.0041051
honors	.0746091	.0176792	4.22	0.000	.0399499	.1092683
volunteer	-.0094764	.0127227	-0.74	0.456	-.0344187	.0154658
computerskills	-.022233	.0107753	-2.06	0.039	-.0433574	-.0011086
h_quality	.0204678	.0127734	1.60	0.109	-.0045739	.0455094
female	.0104551	.0095635	1.09	0.274	-.0082936	.0292038
_cons	.088974	.024047	3.70	0.000	.041831	.1361169

# Manipulation Check

---

연구자가 의도한 대로 실험 조작(experimental manipulation of the treatment variable)이 되었는지를 확인하는 작업

예:

- 기부 요청자의 외모 → 기부 여부/기부금 액수
  - <https://academic.oup.com/qje/article-abstract/121/2/747/1884044>
- 기부 요청자의 “외모” 판단은 각 개인마다 다름
- 피실험대상자가 아닌 일반들을 대상으로 실험에 사용될(?) 기부 요청자의 외모를 평가하게 함

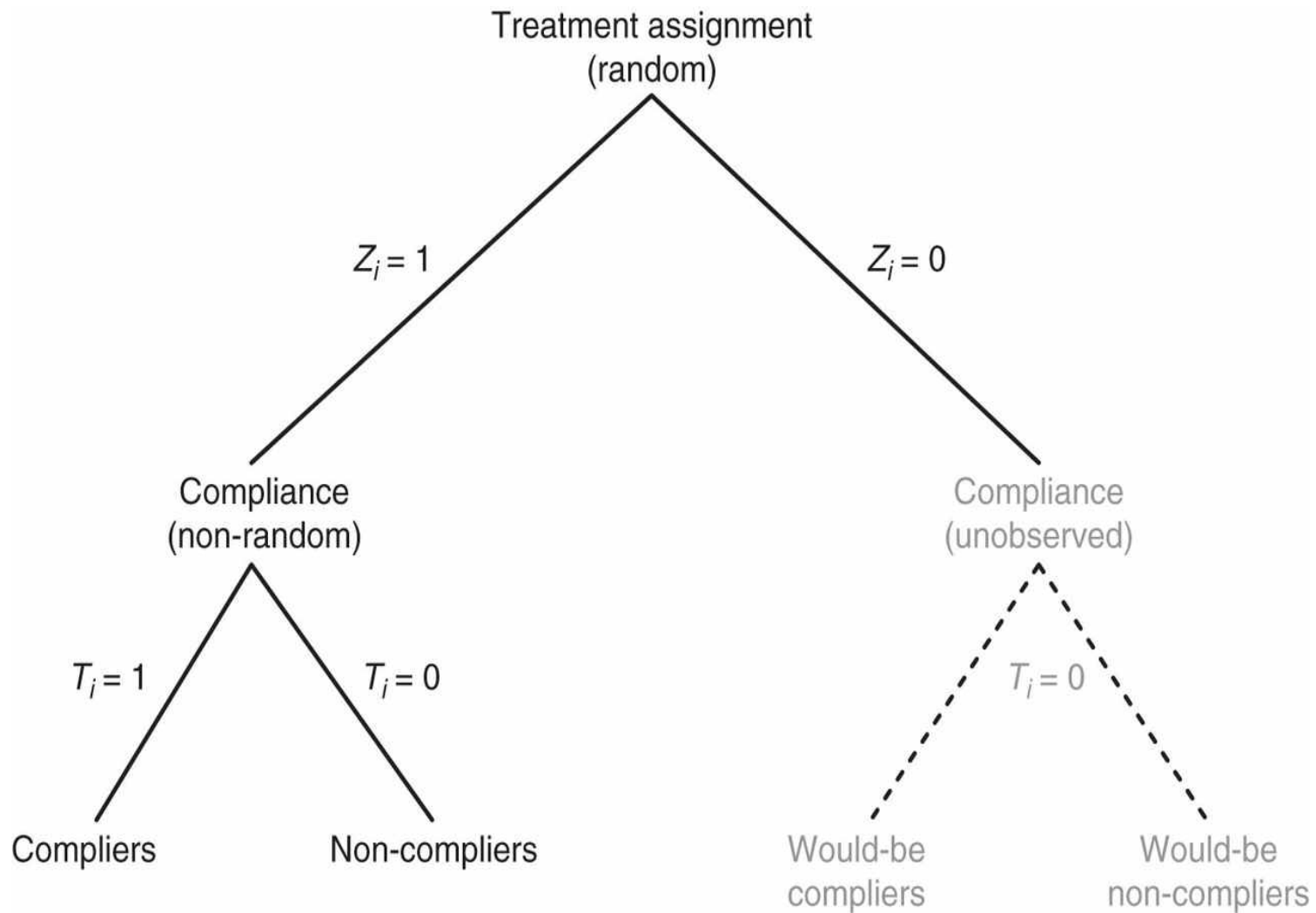
# Non-Compliance 문제

- 
- $Y = \beta_0 + \beta_1 \text{Treatment\_Assigned} + \varepsilon$ 
    - 처리를 받아야 되는 피실험자가 처리를 못 받거나 혹은 처리 받기를 거부할 때 생기는 문제
    - 이 경우 complier와 non-complier는  $\varepsilon$ 의 구성 요인에 따라 나뉘는 것이 보통 (실험 설계로부터 얻고자 하는 내적 타당성에 위협)

## 1990년대 뉴욕시의 school voucher 프로그램

- 낙후된 지역에 사는 학생들에게 인근 사립학교 입학이 가능한 school voucher를 추첨을 통해 나누어 줌
- $\text{Academic\_Performance} = \beta_0 + \beta_1 \text{Voucher\_Assigned} + \varepsilon$ 
  - **Complier**: school voucher를 이용해 인근 사립학교로 전학간 학생
  - **Non-complier**: school voucher를 받고도 인근 사립학교로 전학가지 않은 학생





# Non-compliance 해결책

---

1.  $T = 1$ 과  $T = 0$ 만 비교 (실험이 아님...)

2.  $T = 1$ 과  $((T = 0) + (Z = 0))$ 를 비교 (두 개의 서로 다른 통제군)

3.  $Z = 1$ 과  $Z = 0$ 을 비교 (non-compliance 무시)

- Intention-to-Treat (ITT) Model
- 실제 처리의 효과보다 작게 나옴

4. 도구변수 회귀를 이용

- $\text{Treatment\_Completed} = \gamma_0 + \gamma_1 \text{Treatment\_Assigned} + u$  (first-stage)
- $Y = \beta_0 + \beta_1 \text{Treatment\_Completed}(\text{hat}) + \varepsilon$  (second-stage)

## 예: GOTV Experiment

**TABLE 10.10** Variables for Get-out-the-Vote Experiment  
from Gerber and Green (2005)

Variable	Description
Voted	Voted in the 1998 election (voted = 1)
ContactAssigned	Assigned to in-person contact (assigned = 1)
ContactObserved	Actually contacted via in-person visit (treated = 1)
Ward	Ward number
PeopleHH	Household size

# ITT

```
. reg Voted ContactAssigned
```

Source	SS	df	MS	Number of obs	=	29,380
				F(1, 29378)	=	11.14
Model	2.75999013	1	2.75999013	Prob > F	=	0.0008
Residual	7275.90863	29,378	.247665213	R-squared	=	0.0004
				Adj R-squared	=	0.0003
Total	7278.66862	29,379	.247750727	Root MSE	=	.49766

Voted	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ContactAssigned	.0243592	.007297	3.34	0.001	.0100568	.0386616
_cons	.4476808	.0032405	138.15	0.000	.4413294	.4540323

```
. ivregress 2sls Voted (ContactObserved = ContactAssigned), first
```

#### First-stage regressions

```
Number of obs      =    29,380
F(   1,  29378)     =    9114.33
Prob > F            =    0.0000
R-squared           =    0.2368
Adj R-squared       =    0.2368
Root MSE           =    0.1991
```

ContactObserved	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ContactAssigned	.2787366	.0029197	95.47	0.000	.273014	.2844593
_cons	2.91e-15	.0012966	0.00	1.000	-.0025413	.0025413

#### Instrumental variables (2SLS) regression

```
Number of obs      =    29,380
Wald chi2(1)       =    11.18
Prob > chi2        =    0.0008
R-squared           =    0.0037
Root MSE           =    .49681
```

Voted	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
ContactObserved	.0873915	.0261338	3.34	0.001	.0361702	.1386129
_cons	.4476808	.0032349	138.39	0.000	.4413405	.4540211

Instrumented: ContactObserved

Instruments: ContactAssigned

# 자연실험 (natural experiment)

---

연구자가 처리군-대조군을 나누지 않고 “자연”이 무작위 배분을 한 상황을 이용한 실험 (“as-if random”)

예

- Houston, TX의 학군이 New Orleans, LA의 학군보다 좋다고 알려져 있음.
- 그런데 학군의 질 차이가 학생들 성적에 직접적인 인과 효과를 주는 요인인지는 알 수 없었음. (지역 소득 수준, 지역 교육 수준 등의 confounding factors...)
- 2005년 Hurricane Katrina가 New Orleans의 일부에 큰 피해를 줌 (일부 학생들이 인근 학군으로 임시로 옮겨감)
- 비교 (1) Houston으로 옮겨간 New Orleans 학생; (2) New Orleans에 남은 학생

# 자연실험의 예: 1969 Vietnam war draft Lottery

---

Joshua Angrist

- <http://www3.nccu.edu.tw/~hmlie/pfinance/pf1/readings/draft.pdf>
- Are veterans adequately compensated for their service? (the effect of veteran status on civilian earnings)
- The earnings of white veterans were approximately 15% less than the earnings of comparable nonveterans

Robert Erikson

- <https://www.cambridge.org/core/journals/american-political-science-review/article/caught-in-the-draft-the-effects-of-vietnam-draft-lottery-status-on-political-attitudes/37B0E3788769BF032C516E6F93794F97>
- Males holding low lottery numbers became more antiwar, more liberal, and more Democratic in their voting compared to those whose high numbers protected them from the draft.

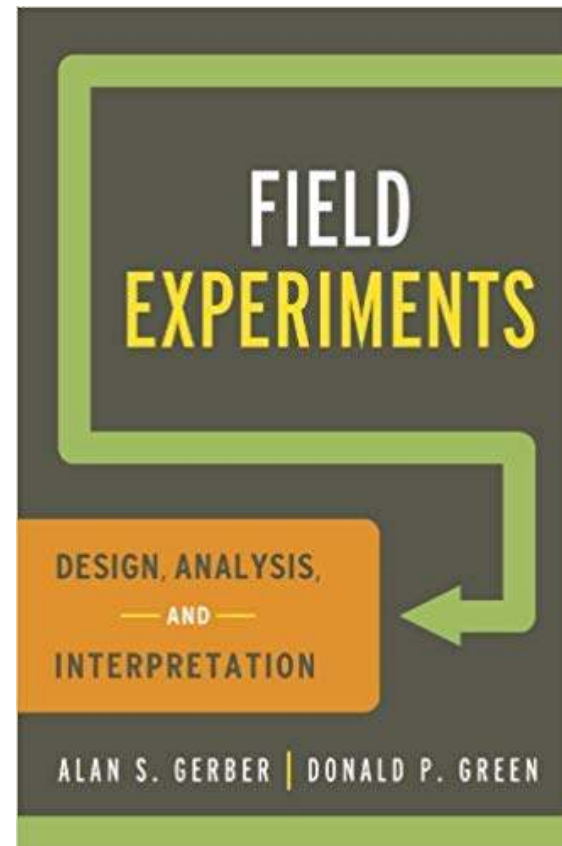
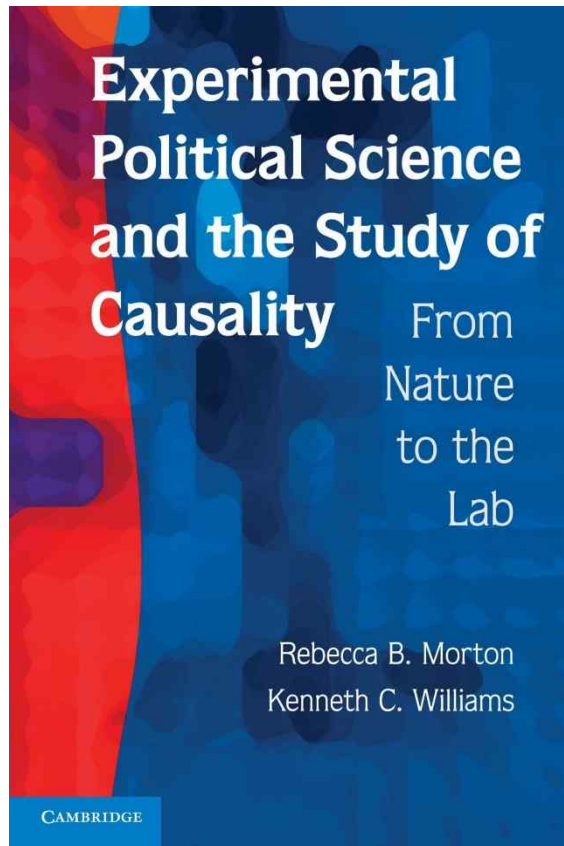
**1970 RANDOM SELECTION SEQUENCE, BY MONTH AND DAY**

	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1	305	086	108	032	330	249	093	111	225	359	019	129
2	159	144	029	271	298	228	350	045	161	125	034	328
3	251	297	267	083	040	301	115	261	049	244	348	157
4	215	210	275	081	276	020	279	145	232	202	266	165
5	101	214	293	269	364	028	188	054	082	024	310	056
6	224	347	139	253	155	110	327	114	006	087	076	010
7	306	091	122	147	035	085	050	168	008	234	051	012
8	199	181	213	312	321	366	013	048	184	283	097	105
9	194	338	317	219	197	335	277	106	263	342	080	043
10	325	216	323	218	065	206	284	021	071	220	282	041
11	329	150	136	014	037	134	248	324	158	237	046	039
12	221	068	300	346	133	272	015	142	242	072	066	314
13	318	152	259	124	295	069	042	307	175	138	126	163
14	238	004	354	231	178	356	331	198	001	294	127	026
15	017	089	169	273	130	180	322	102	113	171	131	320
16	121	212	166	148	055	274	120	044	207	254	107	096
17	235	189	033	260	112	073	098	154	255	288	143	304
18	140	292	332	090	278	341	190	141	246	005	146	128
19	058	025	200	336	075	104	227	311	177	241	203	240
20	280	302	239	345	183	360	187	344	063	192	185	135
21	186	363	334	062	250	060	027	291	204	243	156	070
22	337	290	265	316	326	247	153	339	160	117	009	053
23	118	057	256	252	319	109	172	116	119	201	182	162
24	059	236	258	002	031	358	023	036	195	196	230	095
25	052	179	343	351	361	137	067	286	149	176	132	084
26	092	365	170	340	357	022	303	245	018	007	309	173
27	355	205	268	074	296	064	289	352	233	264	047	078
28	077	299	223	262	308	222	088	167	257	094	281	123
29	349	285	362	191	226	353	270	061	151	229	099	016
30	164	---	217	208	103	209	287	333	315	038	174	003
31	211	---	030	---	313	---	193	011	---	079	---	100

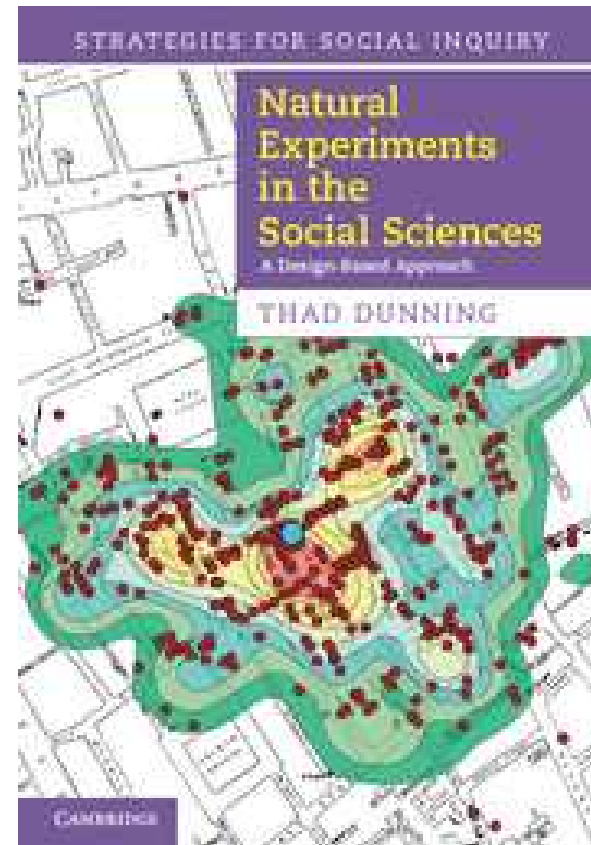


## 참고문헌

---

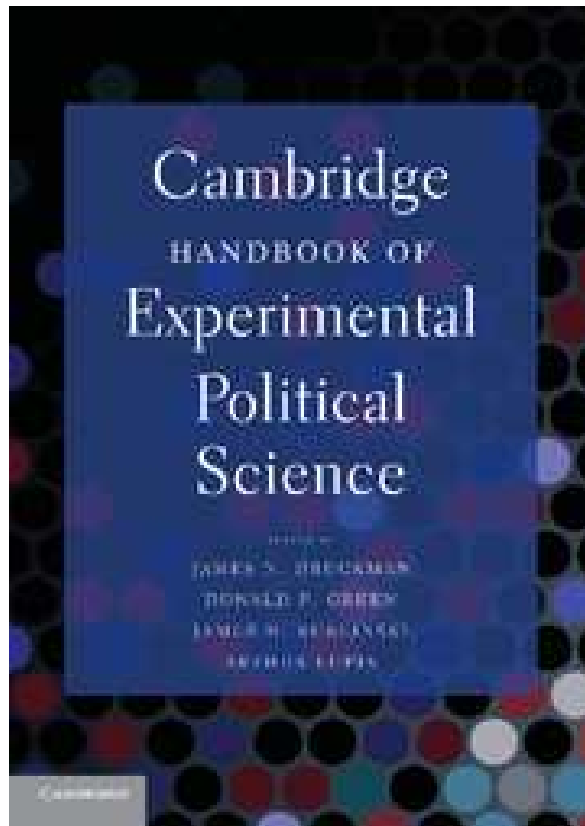


# 참고문헌



## 참고문헌

---



# Replication

---

Chen, J. , Pan, J. and Xu, Y. (2016), Sources of Authoritarian Responsiveness: A Field Experiment in China. *American Journal of Political Science*, 60: 383-400.

Article (Main Text):

<https://onlinelibrary.wiley.com/doi/abs/10.1111/ajps.12207>

Replication Material:

<https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/UMIBSL>