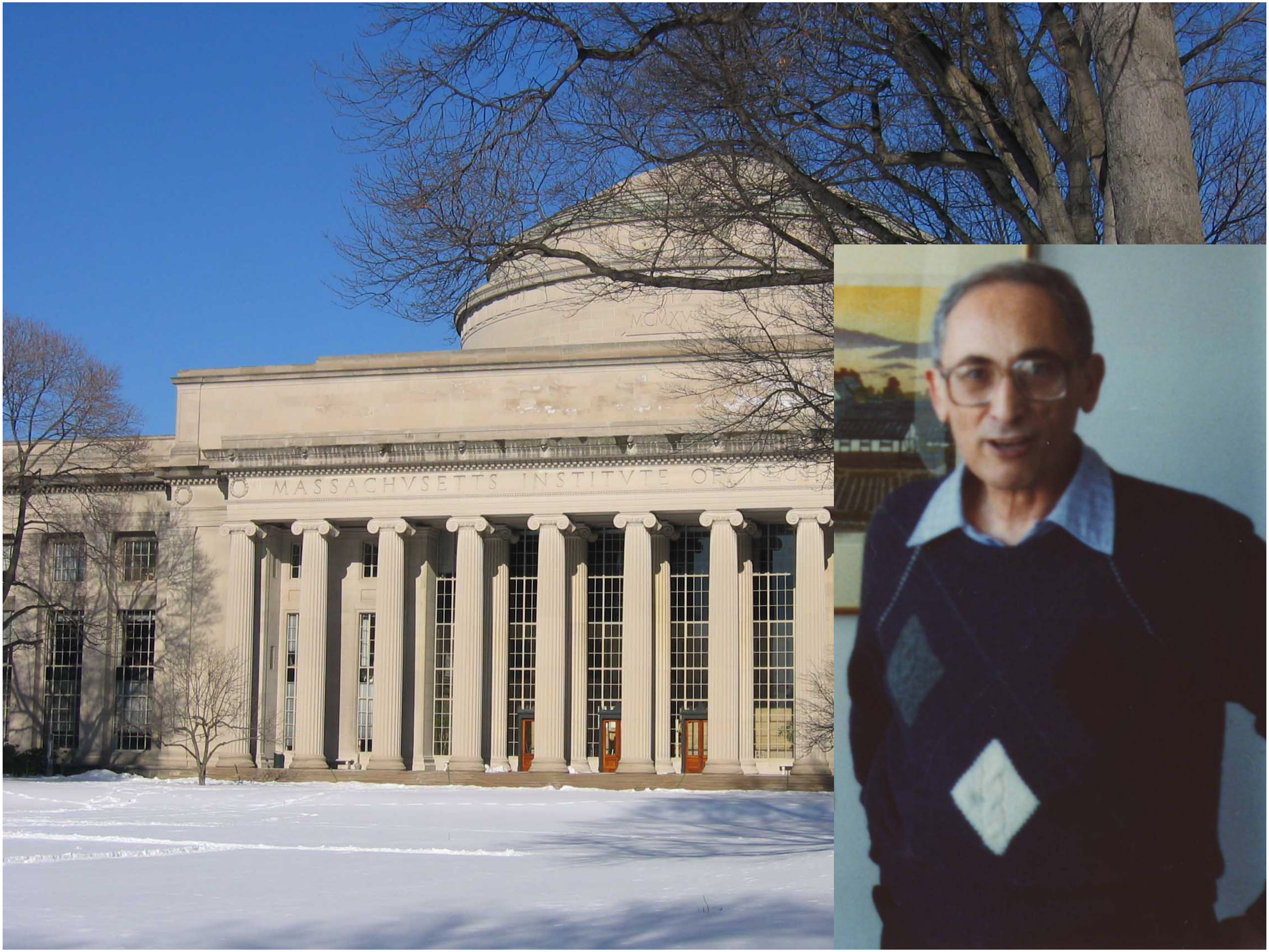
Hierarchical modeling: a unifying framework and some open questions

Andrew Gelman, Department of Statistics and Department of Political Science, Columbia University

Talk for the Harvard Statistics Department 50th anniversary, 26 Oct 2007


MASSACHVSETTS INSTITVTE


PLACE CARDS FOR
JOURNALS ON LOAN
HERE


CAT
HAT
FAT

J. Carlin

DATA ANALYSIS QUESTION

Data were collected in order to develop a means of estimating the total production (biomass) of mesquite leaves using easily measured parameters of the plant, before actual harvesting takes place.

Two separate sets of measurements were taken, one on 26 mesquite bushes (labelled MCD) and the other on 20 bushes (labelled ALS). All the data were obtained in the same geographical location (ranch), but the two groups correspond to different times of the year, and neither constitutes a strictly random sample.

Variables on which data were collected are as follows:

- DIAM1 = canopy diameter (meters) measured along the longer axis of the bush
- DIAM2 = canopy diameter (meters) measured along the shorter axis of the bush
- TOTHT = total height (meters) of the bush
- CANHT = canopy height (meters) of the bush
- DENS = plant unit density (number of primary stems per plant unit)
- LEAFWT = total weight (grams) of photosynthetic material as derived from actual harvesting of the bush

The data are shown on the next page and are available on the Class 3 VAX, in the file [GRADUATE.CARLIN] MESQUITE.DAT.

F. Mosteller

DATA ANALYSIS QUESTION

THE SOLOMON-WYNNE EXPERIMENT

In a study of "traumatic avoidance learning" Solomon and Wynne describe an experiment in which dogs learn to jump a barrier to avoid an intense electric shock [3]. The subjects were 30 "mongrel dogs of medium size" weighing 9 to 13 kilograms. The apparatus was a variation of the Miller-Mowrer shuttle box used for avoidance training of rats. The box consisted of two compartments separated by a barrier and a "guillotine-type gate," which could be raised or lowered. The barrier was adjusted to the height of each dog's back. The floor of the apparatus consisted of steel bars which were wired to the shock circuit.

The conditioned stimulus (CS) consisted of turning out the lights above the compartment the dog was in and simultaneously raising the gate. The other compartment was still illuminated. In ten pretest trials none of the 30 dogs jumped the barrier during a 2-minute exposure of the CS. During training the CS was presented for 10 seconds and was then followed by an intense electric shock applied through the floor to the dogs' feet. The voltage was the "highest possible without producing tetany of the dogs' leg muscles." The current was about 100 to 125 milliamperes for most dogs. The shock was left on until the dog escaped over the barrier into the illuminated compartment, where no shock was administered.*

P. J. Huber

DATA ANALYSIS QUESTION 4

Single Crystal Intensity Measurement Project: (see the attached paper by Abrahams et al., *Acta Cryst.* (1970), A26, 1).

The data deserve to be re-analyzed by robust methods.

1. Keep close to the stated aims of the project.

2. The task is, first of all, descriptive: establish magnitudes and patterns of systematic errors, gross errors, and "ordinary" random errors.

3. Note that the scale of the measurements is arbitrary; the 17 sets were placed on a common scale by a least squares method. About 25% of the data are missing for various reasons.

4. The data are available on VAX3 DISK2:[HUBER2.XRAY].

Table 5(*a*). *Inter-experimental R_{ij} factors*

R_{ij} and $R_{i\mu}$ array $\times 10^3$

i \ j	13	12	11b	11a	14	15	16	1	5	7	2	4	8	3	9	6	10	Diffractometer geometry		λ	$R_{i\mu}\times 10^3$
13	–	497	359	416	430	383	443	401	435	448	482	430	460	434	406	520	508			Mo	478
12	497	–	164	126	138	151	123	128	121	120	135	136	130	130	139	140	162			Mo	141
11b	359	164	–	143	140	093	133	128	135	128	153	171	156	102	100	114	116	Equi-	ω	Mo	092
11a	416	126	143	–	086	118	104	111	100	096	098	093	113	099	109	116	125	inclination	traverse	Mo	093
14	430	138	140	086	–	098	071	085	070	065	106	099	092	059	076	069	088			Mo	062
15	383	151	093	118	098	–	075	071	069	085	126	135	110	066	066	042	034			Mo	058
16	443	123	133	104	071	075	–	044	038	034	070	087	057	040	060	048	071			Mo	048
1	401	128	128	111	085	071	044	–	048	041	046	072	063	062	062	072	083			Mo	053
5	435	121	135	100	070	069	038	048	–	044	048	086	060	038	061	051	057			Mo	046
7	448	120	128	096	065	085	034	041	044	–	060	071	046	043	057	059	080		ω, 2θ	Mo	049
2	482	135	153	098	106	126	070	046	085	060	–	047	061	090	091	120	145	4-circle	traverse	Cu	034
4	480	136	171	093	099	135	087	072	086	071	047	–	068	097	099	128	137			Cu	091
8	460	130	156	113	092	110	051	063	060	046	061	068	–	067	073	074	094			Cu	071
3	434	130	102	099	059	066	040	062	037	043	090	097	067	–	056	042	048	Fixed crystal/fixed counter		Cu	031
9	406	139	100	109	076	066	060	062	061	057	091	099	073	056	–	052	051	Normal beam ω		Cu/Mo	039
6	520	140	114	116	069	042	048	072	051	059	120	128	074	042	052	–	033	Zero-layer only		Mo	040
10	508	162	116	125	088	034	071	083	057	080	145	138	094	048	051	033	–			Cu	049

Table 5(*b*). *Inter-experimental wR_{ij} factors*

wR_{ij} and $wR_{i\mu}$ array $\times 10^3$.

i \ j	13	12	11b	11a	14	15	16	1	5	7	2	4	8	3	9	6	10	Diffractometer geometry		λ	$wR_{i\mu}\times 10^3$
13	–	500	373	416	433	399	435	421	437	435	447	460	443	428	406	496	499			Mo	460
12	500	–	189	176	210	177	208	196	205	218	214	240	244	231	214	124	176			Mo	204
11b	373	189	–	167	221	166	197	171	190	191	190	259	248	152	237	133	213	Equi-	ω	Mo	154
11a	416	176	167	–	134	146	150	133	141	144	134	185	204	144	202	120	177	inclination	traverse	Mo	122
14	433	210	221	134	–	127	110	123	097	108	116	161	180	104	167	079	114			Mo	096
15	399	177	166	146	127	–	090	105	094	099	132	147	178	107	103	047	047			Mo	082
16	435	208	197	150	110	090	–	095	085	092	117	206	153	102	153	046	072			Mo	089
1	421	196	171	133	123	105	095	–	093	094	088	115	171	103	086	116	134			Mo	092
5	437	205	190	141	097	094	085	093	–	077	092	144	143	067	158	070	058			Mo	070
7	435	218	191	144	108	099	092	094	077	–	082	172	155	080	178	061	087		ω, 2θ	Mo	080
2	447	214	190	134	116	132	117	088	092	082	–	156	139	106	171	110	130	4-circle	traverse	Cu	087
4	460	240	259	185	161	147	206	115	144	172	156	–	176	206	238	125	125			Cu	163
8	443	244	248	204	180	178	153	171	143	155	139	176	–	193	167	072	084			Cu	131
3	428	231	152	144	104	107	102	103	067	080	106	206	193	–	197	058	055	Fixed crystal/fixed counter		Cu	080
9	406	214	237	202	167	103	153	086	158	178	171	238	167	197	–	076	078	Normal beam ω		Cu/Mo	138
6	496	124	133	130	079	047	046	116	070	061	110	125	072	058	076	–	044	Zero-layer only		Mo	059
10	499	176	213	177	114	047	072	134	058	087	130	125	084	055	078	044	–			Cu	064

Bayes Estimates for the Linear Model

By D. V. Lindley and A. F. M. Smith

University College, London

[Read before the Royal Statistical Society at a meeting organized by the Research Section on Wednesday, December 8th, 1971, Mr M. J. R. Healy in the Chair]

Summary

The usual linear statistical model is reanalyzed using Bayesian methods and the concept of exchangeability. The general method is illustrated by applications to two-factor experimental designs and multiple regression.

Professor Oscar Kempthorne (Statistical Laboratory, Iowa State University): The authors should be congratulated on their presentation. It will be informative to many. I have no detailed questions or remarks about the formal development, I wish mainly to comment on philosophical issues that underly the whole matter under discussion.

$$(\mathbf{R}+\mathbf{ZDZ}')^{-1} = \mathbf{R}^{-1}-\mathbf{R}^{-1}\mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{D}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}.$$

This result grew out of a Bayesian process which I call legitimate because the model is based on prior knowledge and *not on lack of knowledge*. It is worth noting that Henderson advocates from the viewpoint of computation the maximization with regard to the θ_i of a sort of likelihood equal to $p(\theta_i)\,p(y \mid \theta_i)$, which gives the mode of the posterior distribution.

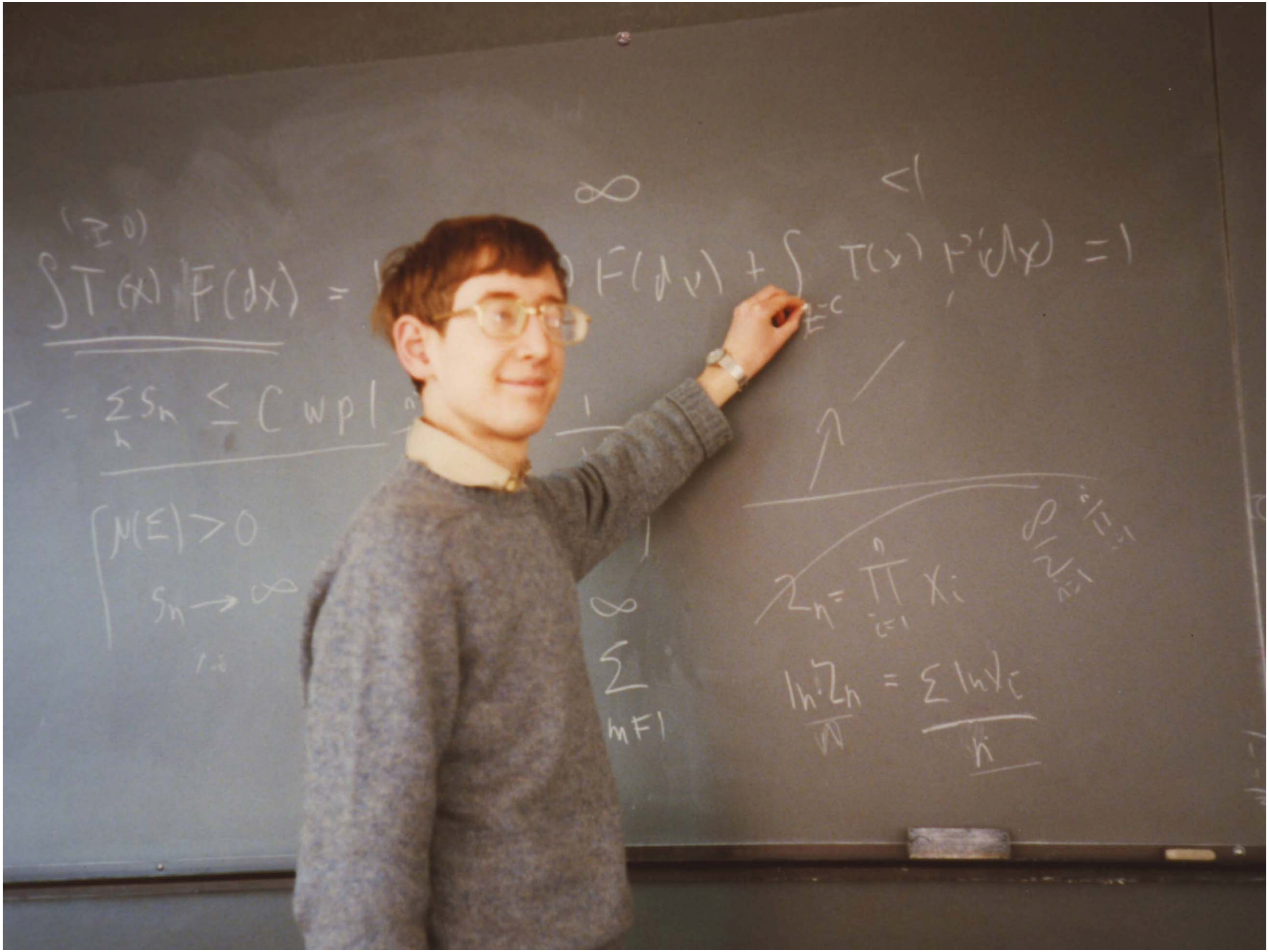
obtains data also on μ_O. I believe this sort of idea will not "sell". I believe the idea will "butcher" the accumulation and condensation of investigational information. I doubt strongly that the authors have talked to workers in *experimental* science and tried to sell them the idea. I surmise, furthermore, that if our informed public were aware of the

I can well surmise the attitude of scientists whose only data input from other workers consists of other workers' Bayesian estimates. Surely the answer will be, "I do not care what Joe thinks about the parameter: I want to know the observational facts or a good condensation of them". *I believe our present authors are not in touch with the processes of science.* They are not aware of the need for the development of interpersonal validity

realistic"? The authors use the phrase. What do they mean? Furthermore, let us turn to the American College Testing Program. Is it "practically realistic" to use an exchangeable prior? Information is available in the records to show that schools differ widely, students of different social and ethnic backgrounds perform differentially on tests, and so on. Information on students is available to show, I think, that exchangeable priors are "practically unrealistic" whatever that means. Are they being used in high school and

Professor LINDLEY and Dr SMITH replied briefly at the meeting and subsequently in writing as follows:

points in detail. We conclude by remarking that his penultimate paragraph reveals that he has not, despite de Finetti's expectation, understood the idea of exchangeability (it does not mean the units are the same), and by expressing the hope that the American College Testing Program can deal with his accusations.



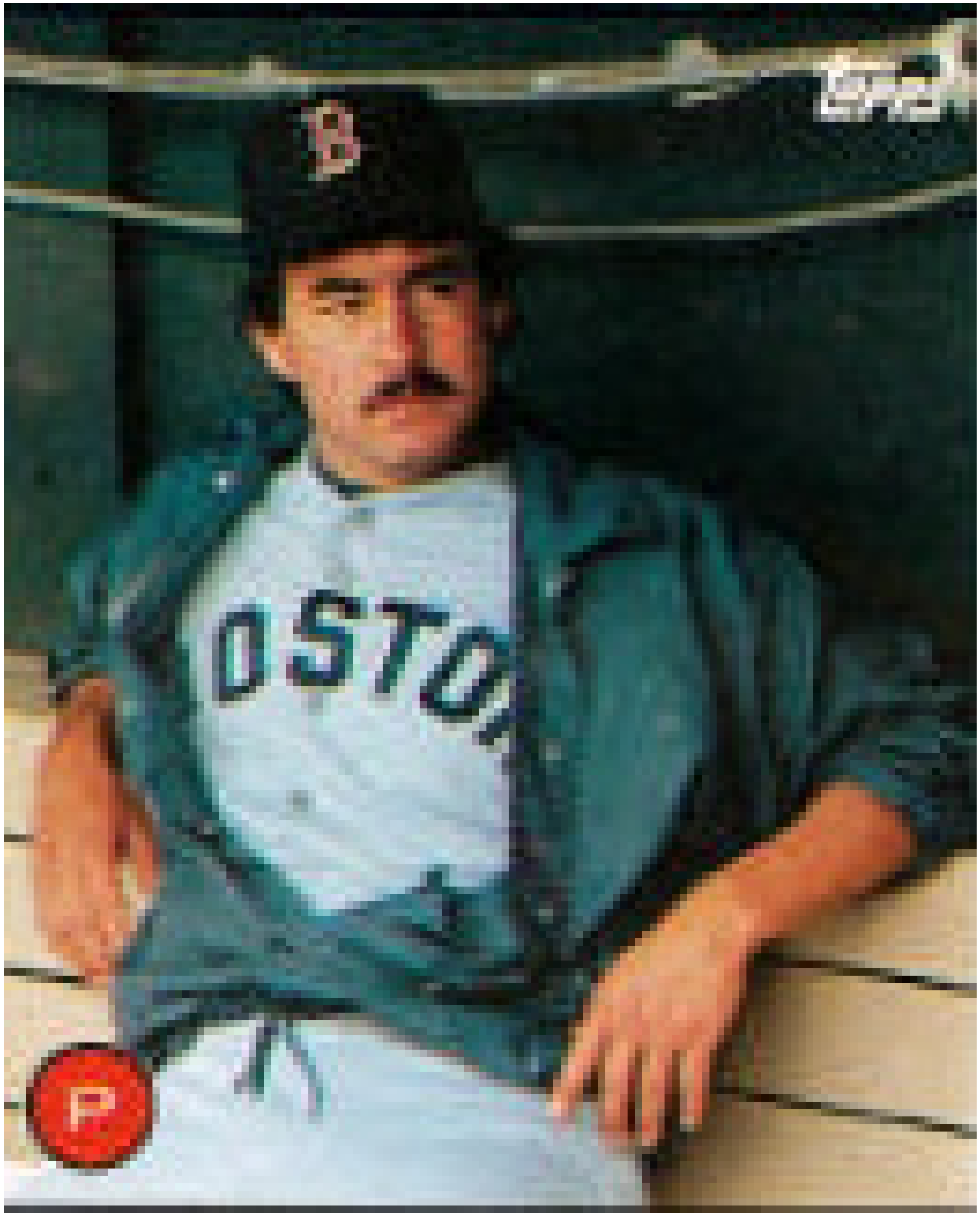
THE ALBANIANS OF THE NORTH AND SOUTH

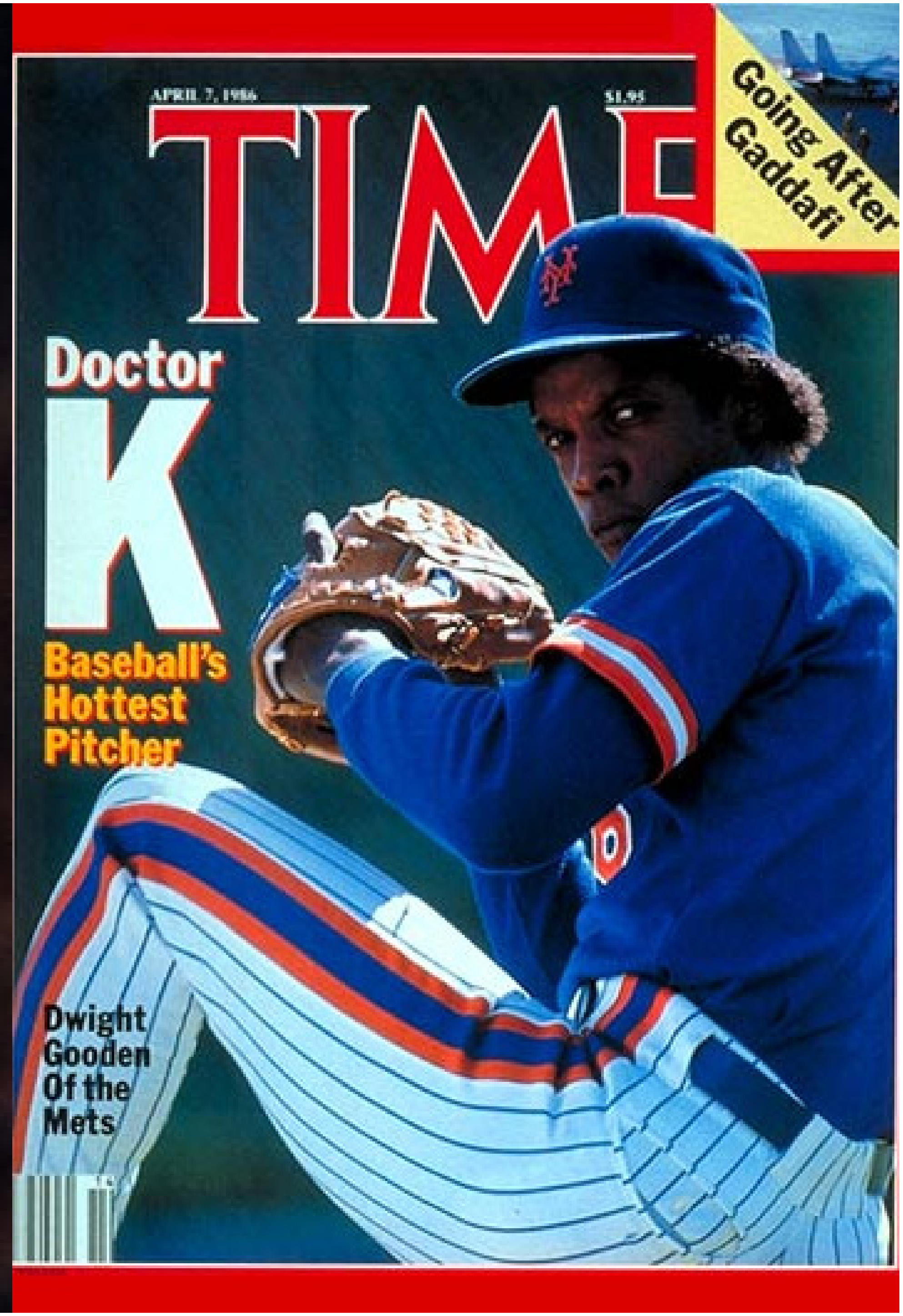
(1) INTRODUCTORY ACCOUNT OF MEASUREMENTS AND PHOTOGRAPHS TAKEN IN 1929.

TABLE I. *Constants for the Characters measured.*

(Measurements in centimetres.)

Character / Absolute Measurements	Albanians of the South: No.	Mean ± Prob. Error	Stand. Deviation ± Prob. Error	Albanians of the North: No.	Mean ± Prob. Error	Stand. Deviation ± Prob. Error
(1) Stature	85	163·72 ± ·449	6·14 ± ·318	77	169·03 ± ·529	6·68 ± ·374
(23) Sitting Height	85	87·94 ± ·275^{-}	3·73 ± ·194	77	89·19 ± ·246	3·20 ± ·174
(17) Span	85	169·32 ± ·495	6·76 ± ·350	77	175·58 ± ·630	8·20 ± ·446
(4) Suprasternal Height	84	133·58 ± ·388	5·28 ± ·275^{-}	77	137·91 ± ·493	6·42 ± ·349
(8) Acromial Height	85	133·29 ± ·400	5·47 ± ·283	77	137·86 ± ·472	6·14 ± ·334
(9) Elbow Height	85	102·80 ± ·318	4·35^{+} ± ·225^{+}	76	106·33 ± ·363	4·69 ± ·257
(10) Wrist Height	85	78·43 ± ·255^{+}	3·49 ± ·180^{6}	77	81·03 ± ·290	3·77 ± ·205^{+}
(11) Finger Height	85	60·93 ± ·235^{-}	3·21 ± ·166	77	62·84 ± ·24[illegible]	3·18 ± ·173
(35) Shoulder Breadth	85	37·06 ± ·121	1·65^{+} ± ·086	75	37·48 ± ·138	1·77 ± ·098
(40) Hip Breadth	85	27·65^{+} ± ·110	1·50^{+} ± ·078	77	28·33 ± ·102	1·33 ± ·072
(36) Chest Breadth	85	25·63 ± ·096	1·31 ± ·068	77	26·05^{+} ± ·097	1·26 ± ·069
(37) Chest Depth	85	19·18 ± ·081	1·11 ± ·058	77	19·70 ± ·095^{-}	1·23 ± ·067
(52) Hand Breadth	85	8·39 ± ·035	·47 ± ·024	77	8·49 ± ·028	·36 ± ·020
(58) Foot Length	85	24·30 ± ·089	1·21 ± ·063	77	25·97 ± ·098	1·27 ± ·069
(59) Foot Breadth	85	10·05^{-} ± ·048	·66 ± ·034	77	10·29 ± ·037	·48 ± ·026




APRIL 7, 1986
$1.95
TIME
Going After Gaddafi
Doctor
K
Baseball's Hottest Pitcher
Dwight Gooden Of the Mets

Hierarchical modeling as a unifying framework

- No longer need to worry about
  - Multiple comparisons
  - Abstract principles for prior distributions
  - Lots of things that used to seem important, once upon a time . . .

- The hierarchical model represents your "reference set"


Datasets:
betcha can't analyze just one.

Doing it again and again

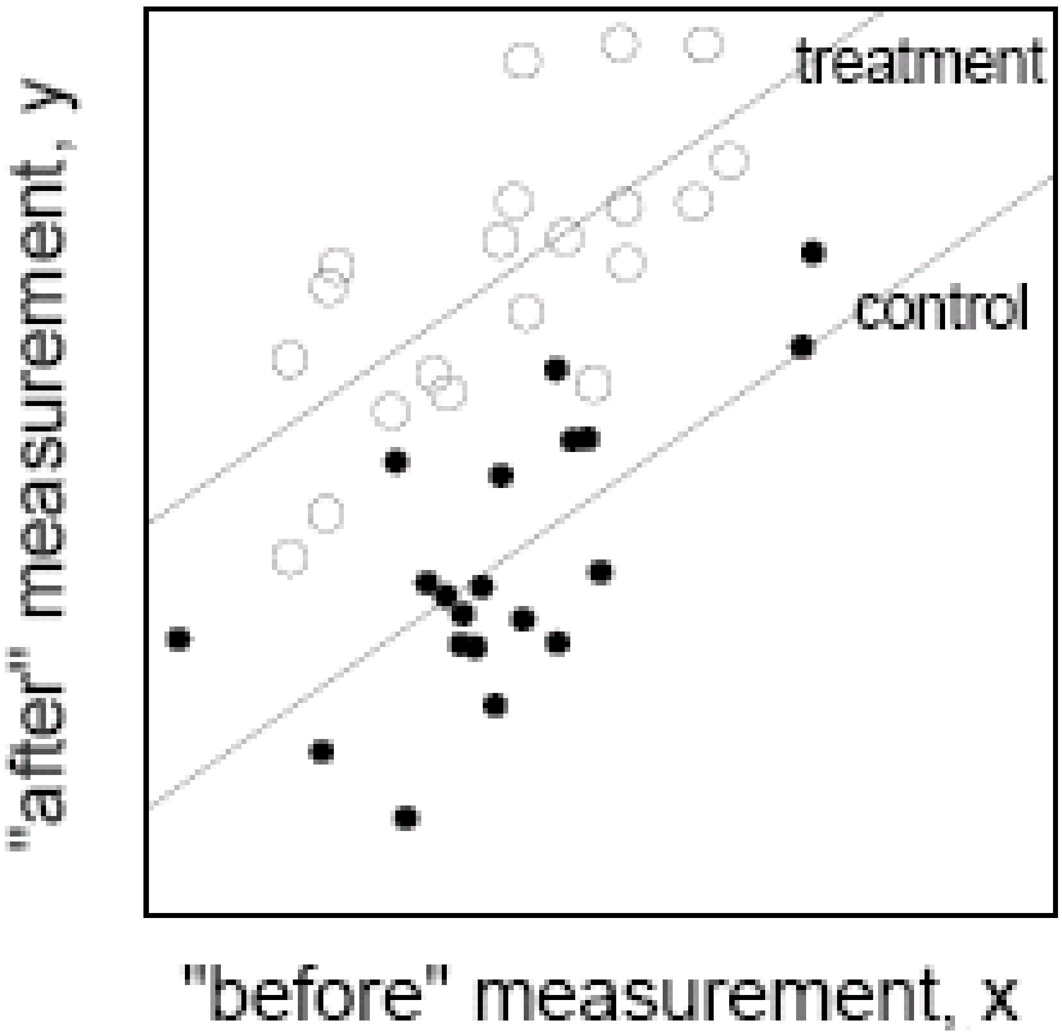
- Division of labor
- Actions and policies
- We’re all frequentist

- Repetition fits into the hierarchical framework

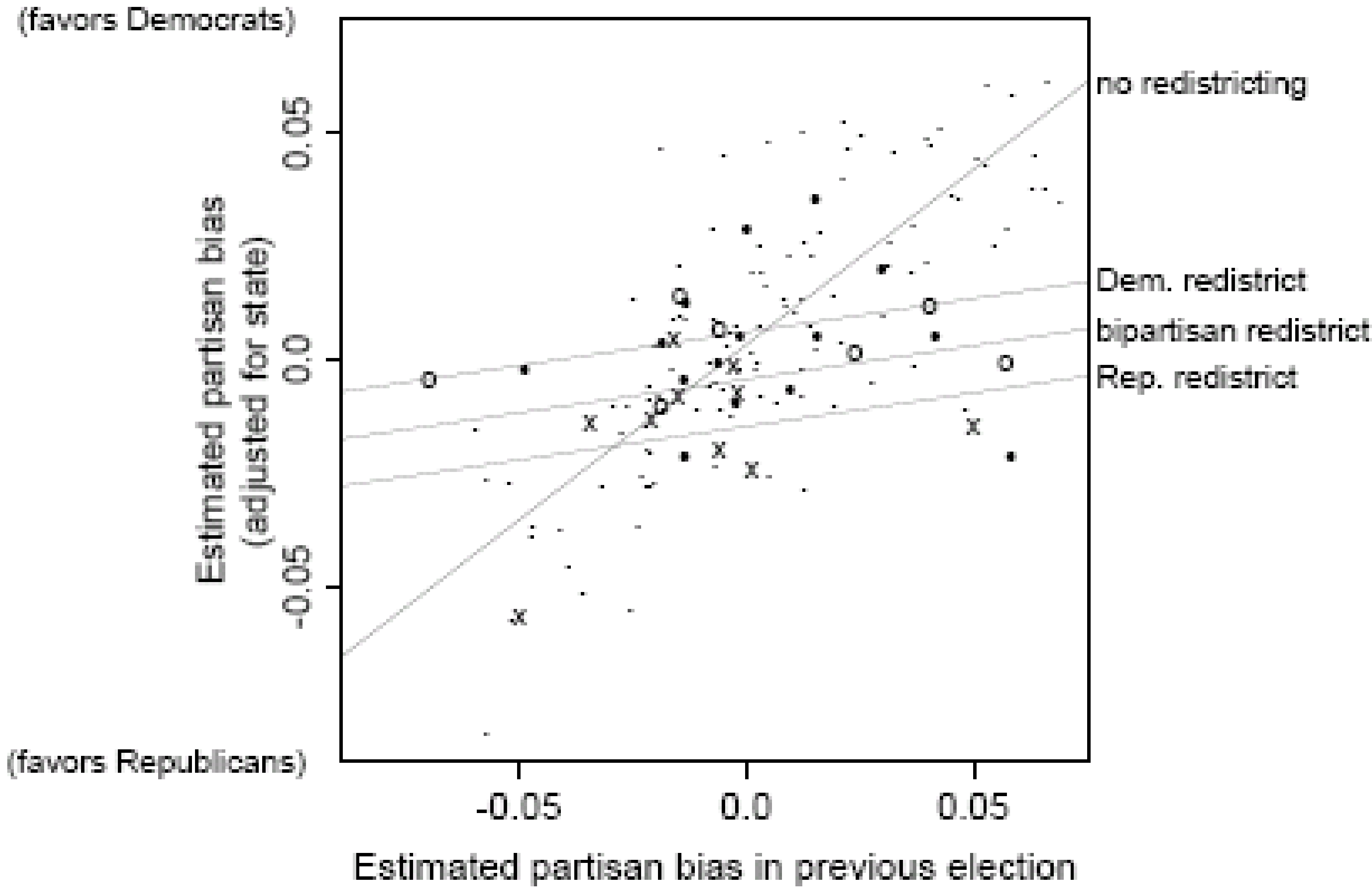
Some open problems

- Computation (of course)
- Priors and invariance to subdivisions
  - (fixed data, increase #groups)
  - taxonomy (tree) models
- Lots of specific models (time series, spatial, networks . . .)
- Deep interactions

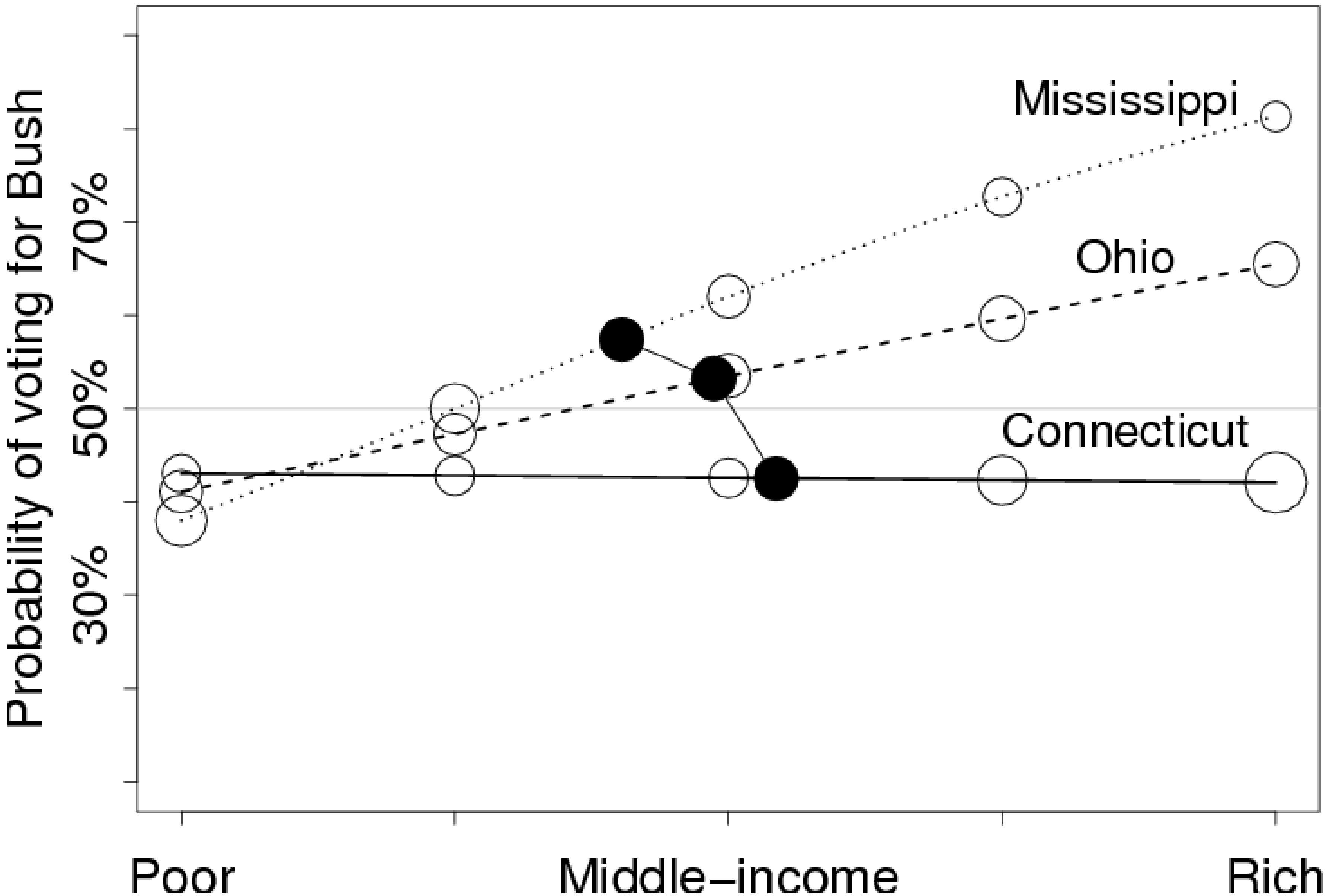
Default no-interaction model



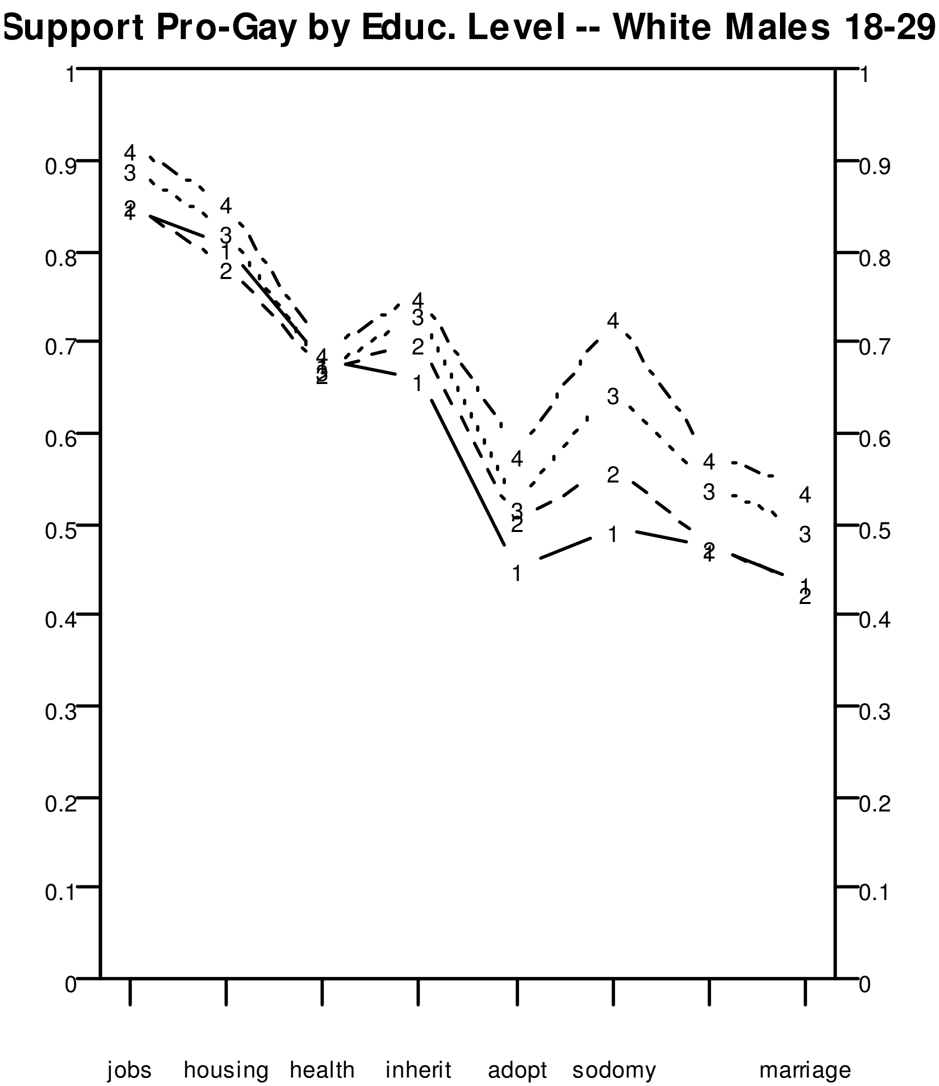
Actual before-after data

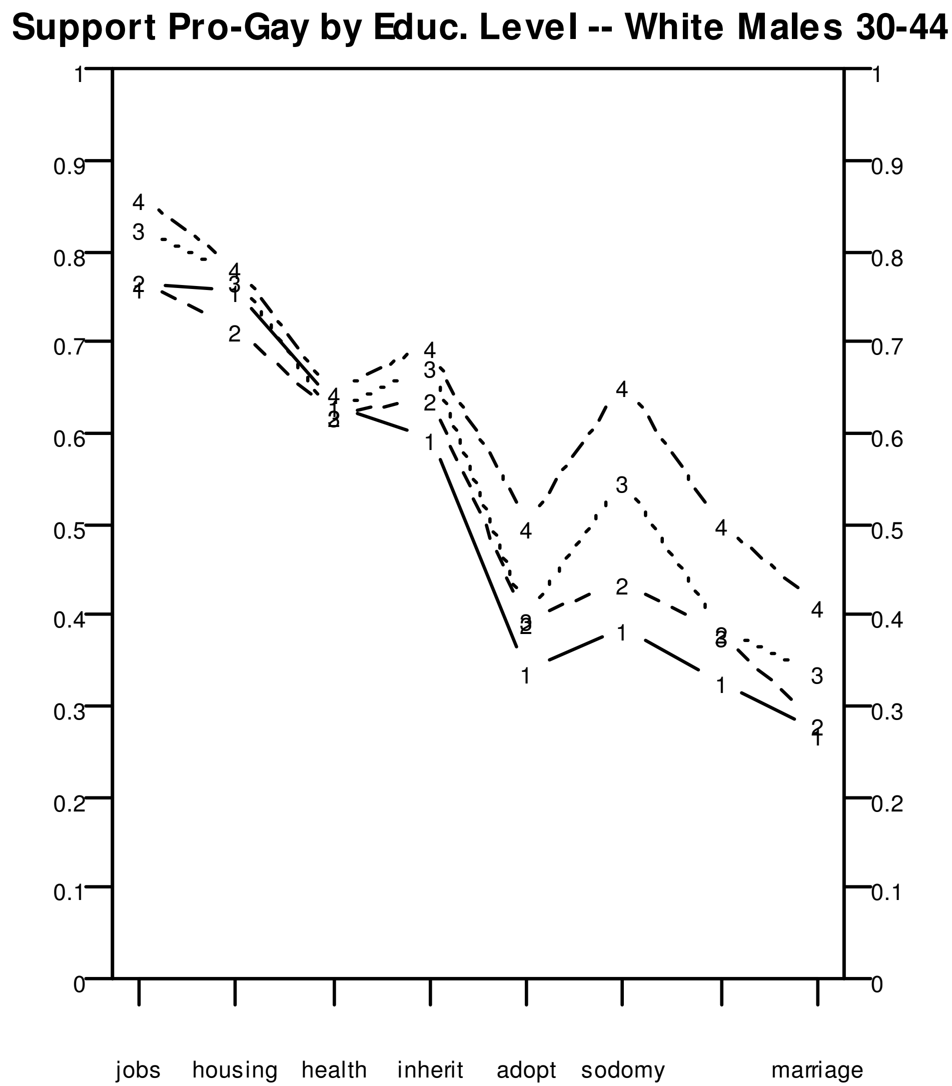


More interactions



Still more . . .


Support Pro-Gay by Educ. Level -- White Males 18-29
jobs
housing
health
inherit
adopt
sodomy
marriage


Support Pro-Gay by Educ. Level -- White Males 30-44
jobs
housing
health
inherit
adopt
sodomy
marriage

Conclusions

- Hierarchical modeling is a way of life
- Transforms subjective choices (whether to combine two datasets) into objective methods
- Challenge: constructing generic models . . .

- Lots more work for Harvard's next 50 years!