Projecting California's population: The road ahead

Simulation of demographic outcomes under uncertainty

Ethan Sharygin
Demographic Research Unit,
California Department of Finance
ethan.sharygin@dof.ca.gov

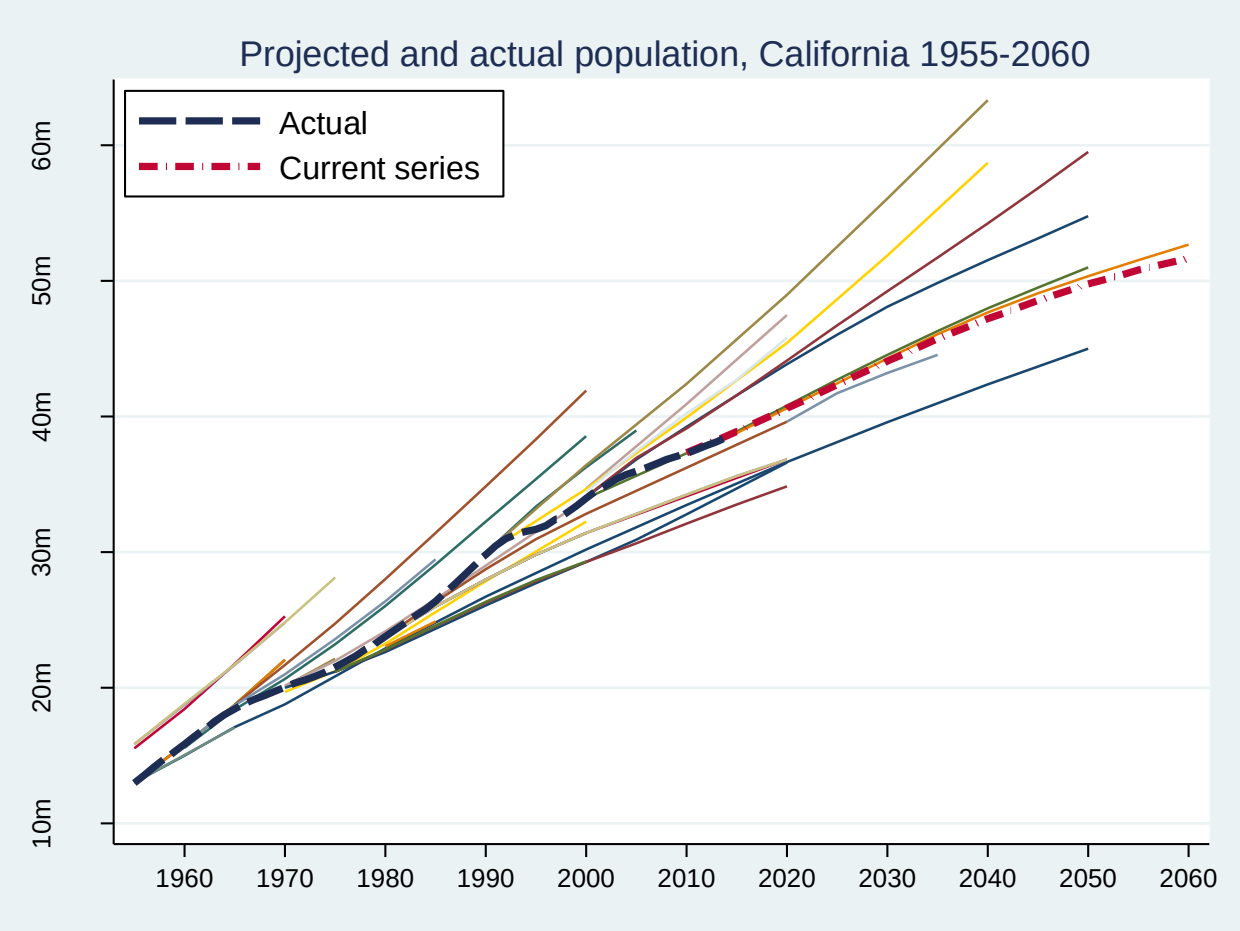
CA SDC Annual Meeting / Oct 8, 2015

Summary

Talk overview:

(1) Background
(2) Content review
(3) Methodology
(4) Assumptions

Background: Long-term forecasts, 1970--2060



2013 Baseline Projections (1/2015)

40 million by 2020, 50 million by 2052

2013 Basis = 2010 Census + 2011-2013 Estimates

Available files (http://www.dof.ca.gov/research/demographic/projections/):

- P-1 State and County Totals (overall; race/eth; age groups)
- P-2 State and County Totals by race/eth + age
- P-3 Complete file (race/eth + age + sex)
- P-3 Median age by race/eth + sex
- P-4 State and County Households, GQ, PPH
- State and County Births
- State and County Public K-12 Enrollment and HS Graduates

Next projections: road ahead

Review the following aspects of our projections process:

(1) Content
(2) Methods
(3) Assumptions

Content review

"Get your **facts** first, and then you can distort them as much as you please."

Rudyard Kipling, *An Interview with Mark Twain* (1899)

Assessing population change

Population balancing identity:

$$P_1 = P_0 + (B - D) + (IM - OM)$$

Elements:

(1) Starting population

Census

(2) Net natural increase

Births, deaths

(3) Net migration

In-, out-migration

Specificity

Uncertainty

Data inputs

Vital statistics database and reports from CDPH

Problems:

(1) Increasing share missing

(2) Linking parents' and child race

(3) Comparability of traits at birth and death

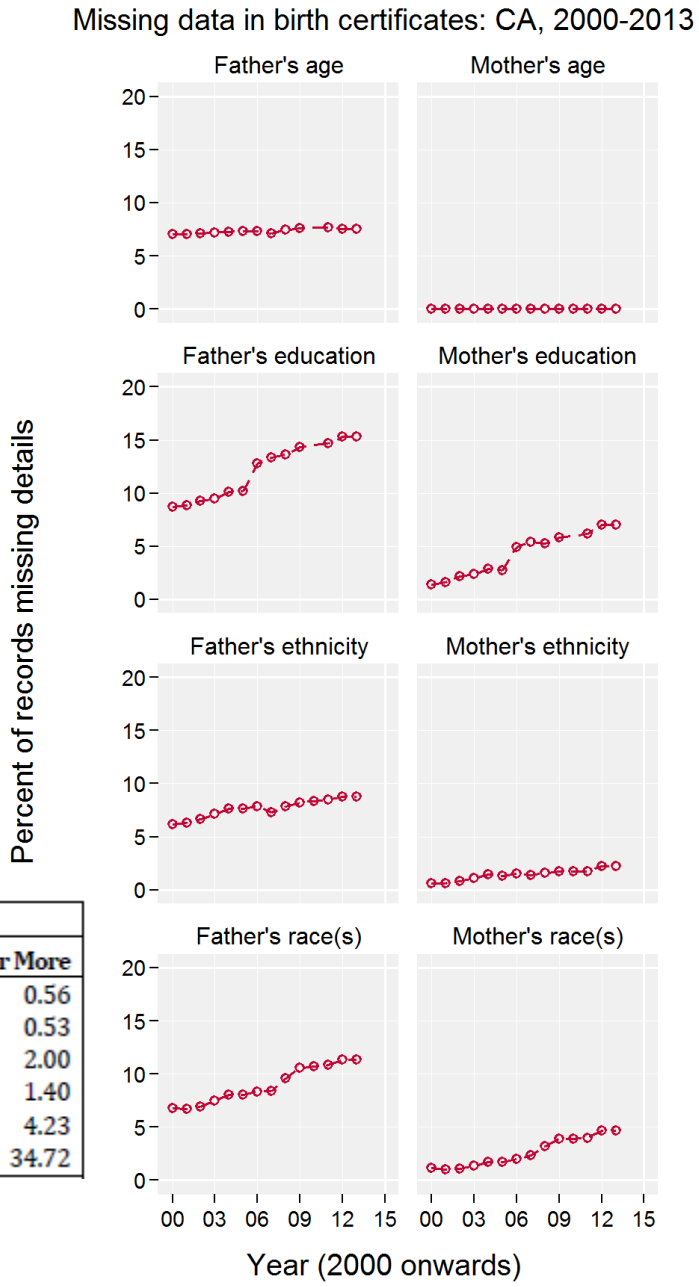
Data processing:

(1) Multiple imputation

(2) Race bridging (probabilistic)

(3) Multiple races (instead of multi-race)

Race of mother	Race of Father					
	White alone	Black alone	AIAN alone	Asian alone	NHPI alone	Two or More
White alone	97.58	0.98	0.42	0.40	0.05	0.56
Black alone	2.70	96.45	0.14	0.14	0.03	0.53
AIAN alone	34.65	2.99	59.56	0.58	0.22	2.00
Asian alone	14.19	1.29	0.15	82.79	0.19	1.40
NHPI alone	19.14	4.80	0.65	3.77	67.41	4.23
Two or more	47.34	10.11	2.07	4.62	1.14	34.72



Data inputs

Migration data (ACS, CPS, IRS, DMV)

Problems:

(1) Lack of specificity

(2) Conflicting definitions/totals

Solution:

(1) Traits from ACS

(2) Totals from DRU

(3) Iterative proportional fitting to reconcile

Migration to/from California (State—state)			
Years	ACS	DMV	IRS
2005-09	-184,592	-109,636	-118,991

Migration within California (County—county)			
Years	ACS	DMV	IRS
2005-09	1,125,961	819,352	835,764

Note: figures are annual averages

Data outputs

Central projection	+	Variants	+	Scenarios

VARIANTS are designed to test sensitivity of the projection

Examples of projection variants:

(1) Fertility +/- 0.5 child

(2) Net migration rate +/- 5 persons per 1,000

SCENARIOS are intended to model comprehensive changes

Examples of alternate scenarios:

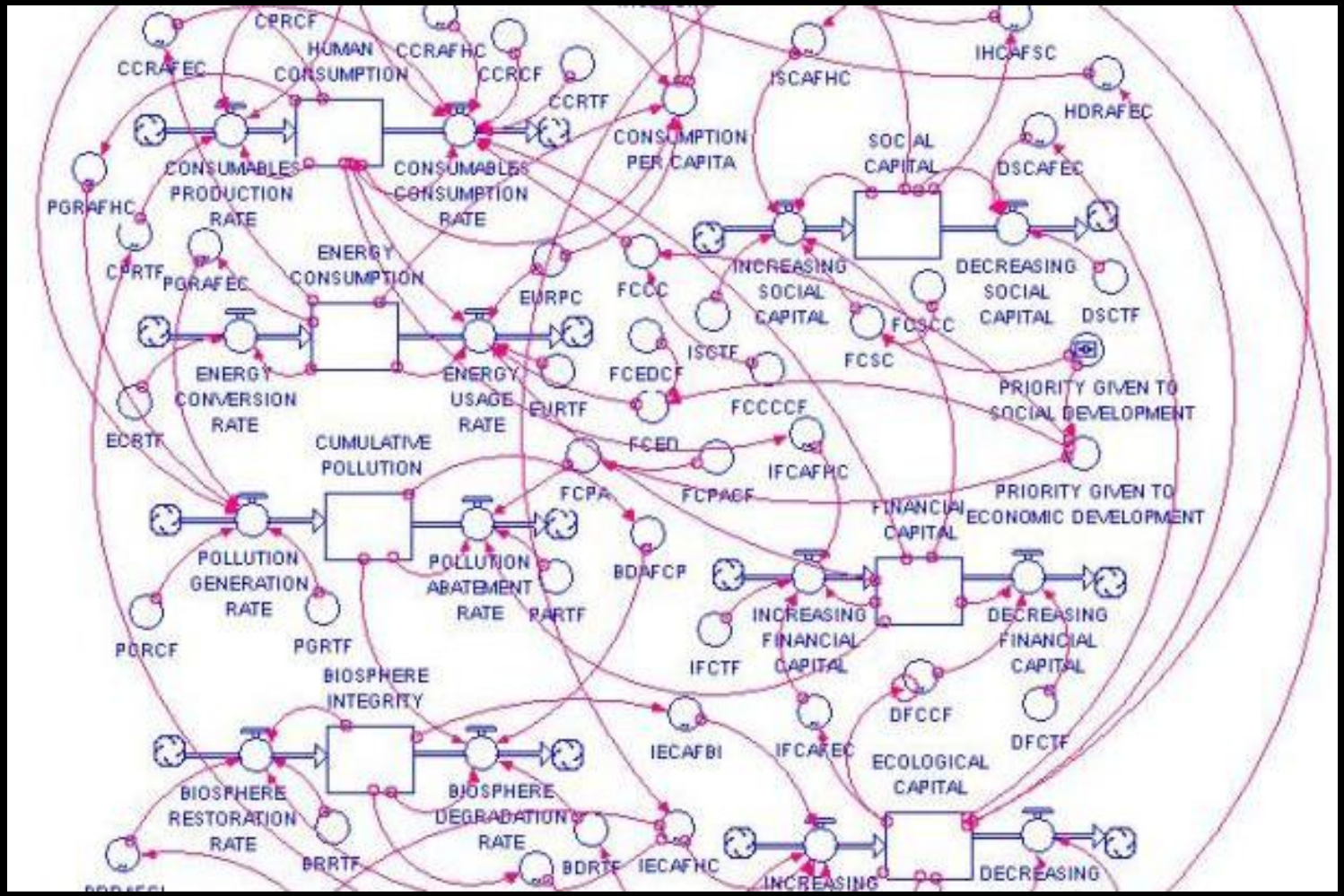
(1) Constant fertility

(2) Balanced net long term migration

(3) Stationary population

(4) Economic, housing, policy conditions

Methodology Review



Modeling population change (revisited)

Population balancing identity:

$$P_1 = P_0 + (B - D) + (IM - OM)$$

Elements:

(1) Starting population
    Census
(2) Net natural increase
    Births, deaths
(3) Net migration
    In-, out-migration

Specificity

Uncertainty

Specificity

STOCK refers to the cross-sectional snapshot of the population. FLOW refers to the hazard rates for demographic events.

Stock should be specified by following state spaces:

- Age (~0—100+)
- Sex (~male, female)
- Fertility status (~parity)
- Location (~county, out-migrant, emigrant)
- Living arrangement (~household, prison, dormitory, hospital, etc.)
- Nativity (~place of birth)
- Race/ethnicity (~15 combinations of bridged race, Hispanic)
- Education (~primary, lower secondary, higher secondary, tertiary)
- Marital status (~never married, married, divorced, widowed)
- Mortality status (~alive, dead)

Flow should be specified by parsimonious state spaces

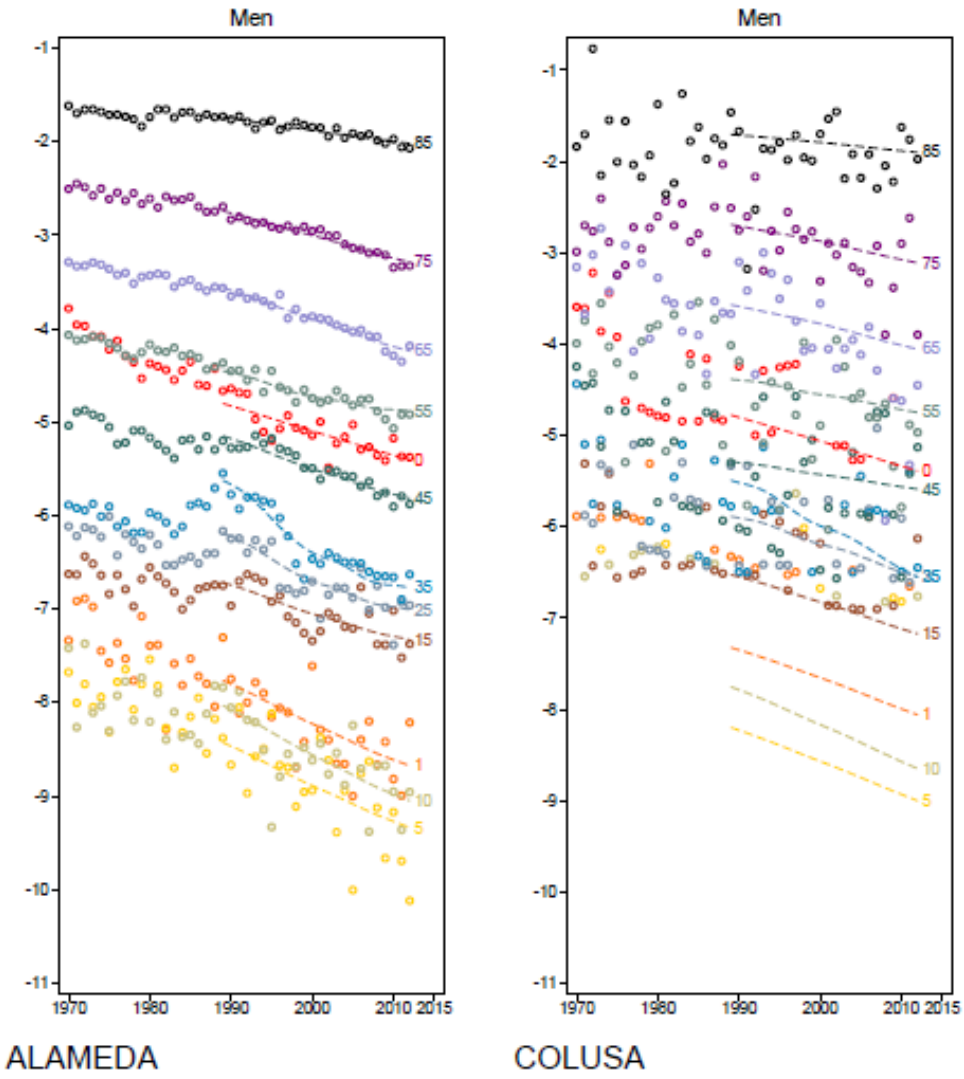
Specificity

Flow should be specified
by parsimonious state spaces:

- Age
- Sex
- County
- Race/ethnicity
- (Education)

E.g., mortality model:

$$\begin{aligned}\ln(\mu_{ijt}) = \ln(N_{ijt}) \\ +(\beta_1 + \eta_j)YR_t \\ +\beta_2 EDU_{it} \\ +\beta_3 INC_{it} \\ +\rho_i^{geo} \\ +\eta_j \\ +\eta_i \\ +\varepsilon_{ijt}\end{aligned}$$

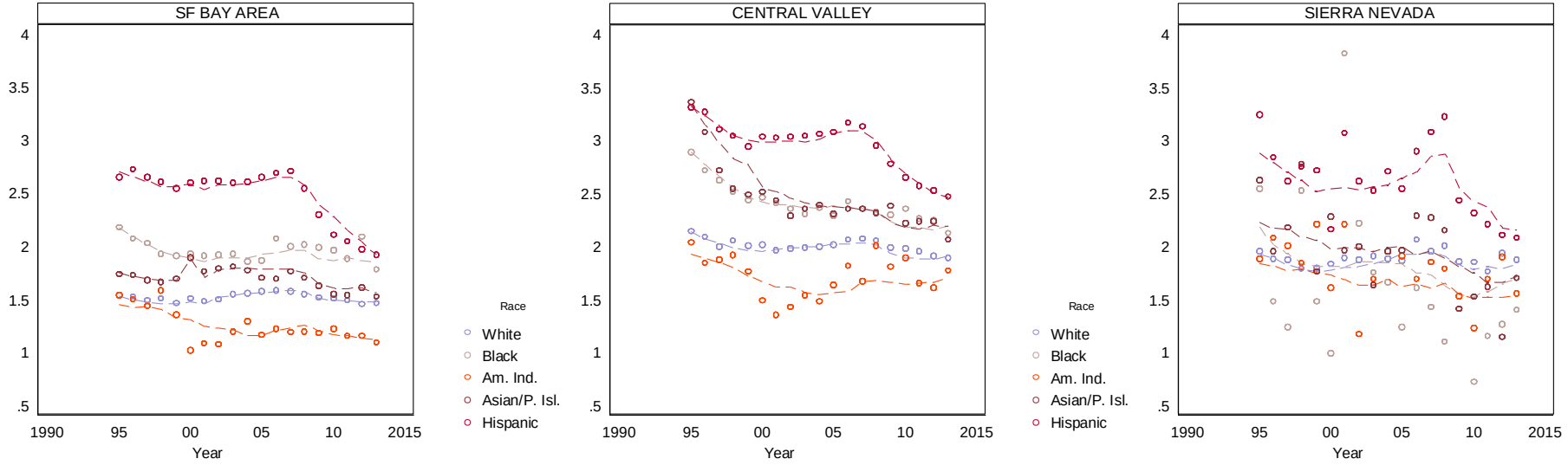


Specificity

Flow should be specified
by parsimonious state spaces:

- Age
- Sex
- County
- Race/ethnicity
- (Education)

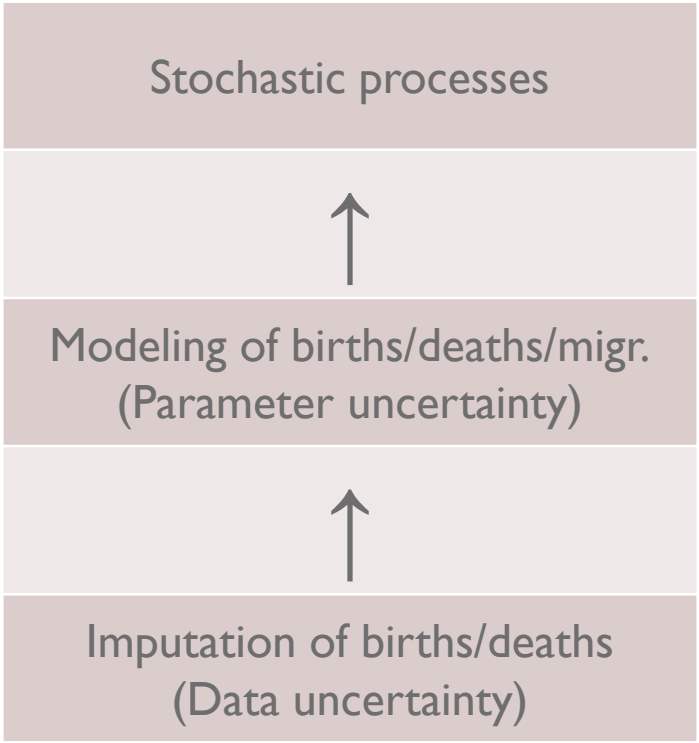
Fertility example: Total Fertility Rate by Region + Race/Ethnicity

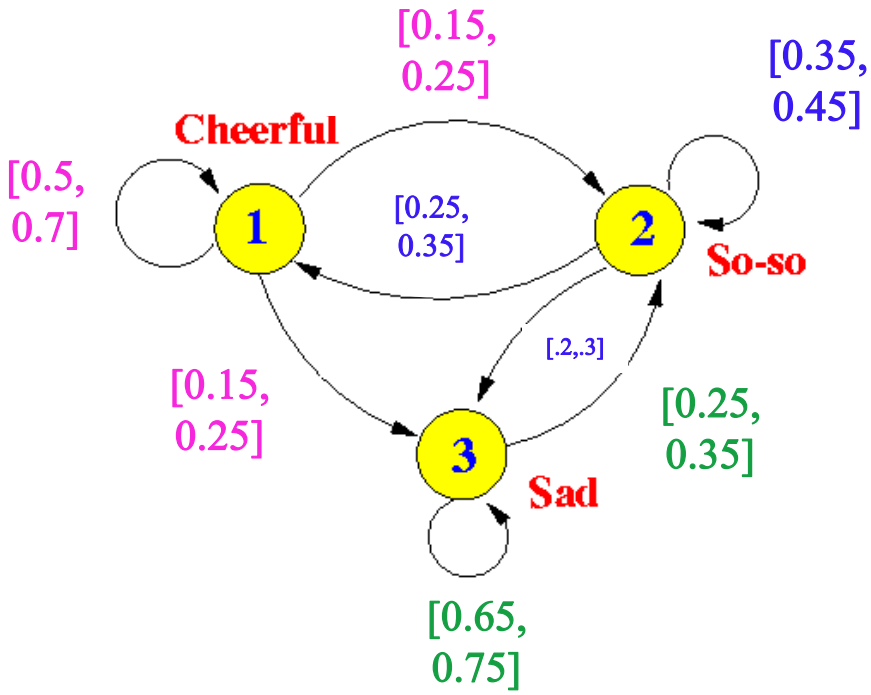


Uncertainty

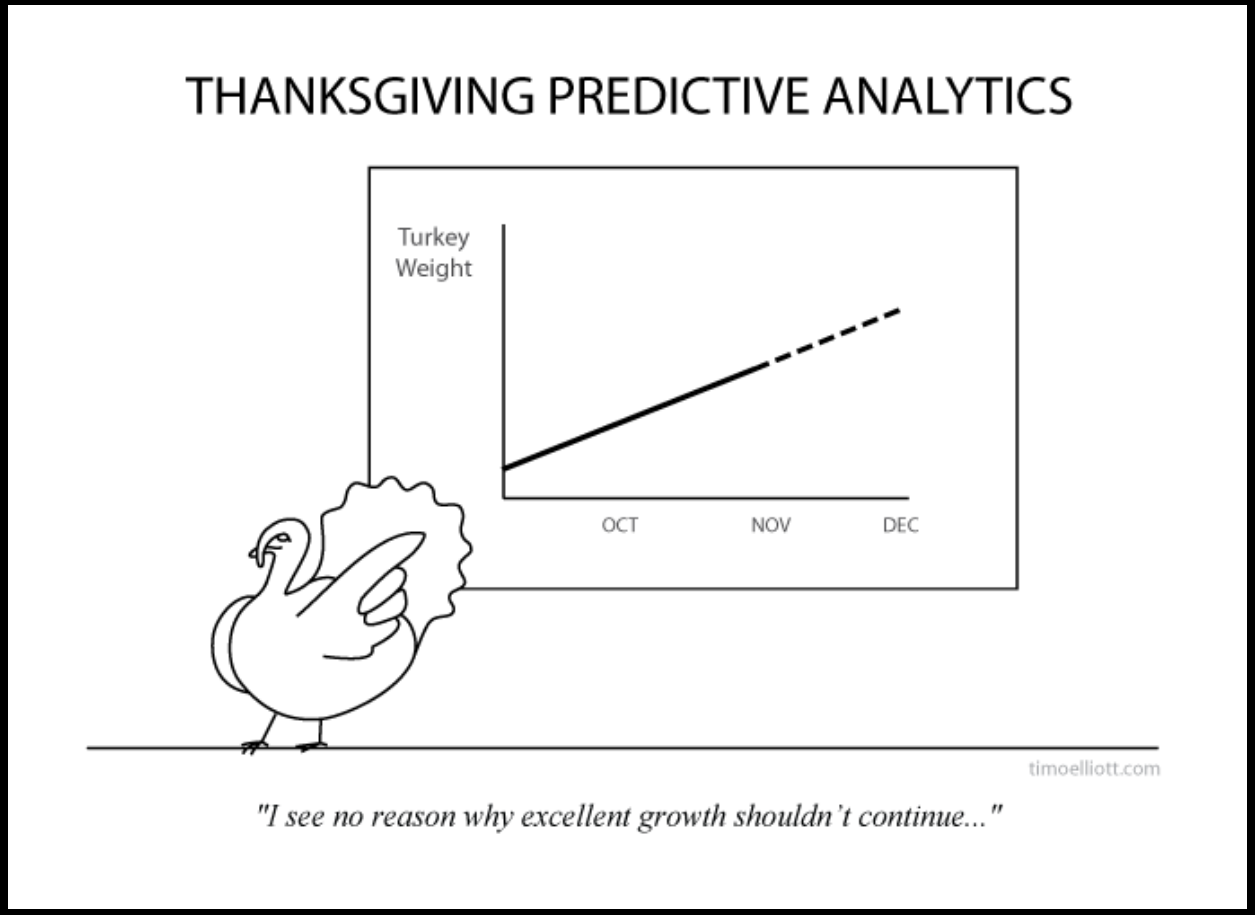
UNCERTAINTY refers to the quantifiable likelihood of outcomes other than the most likely one.

Uncertainty should be propagated from the following sources:





Revisiting assumptions



Forecasting assumptions: relating the past and future

Models allow better assessment of past rates

- Accurate; complete; internally consistent; dynamic
- … but does not tell the future.

Review past sources of error

- Too influenced by contemporary trends
- Lack of feedback mechanisms (i.e., agent behavior)

In first stage, incorporate three elements in review:

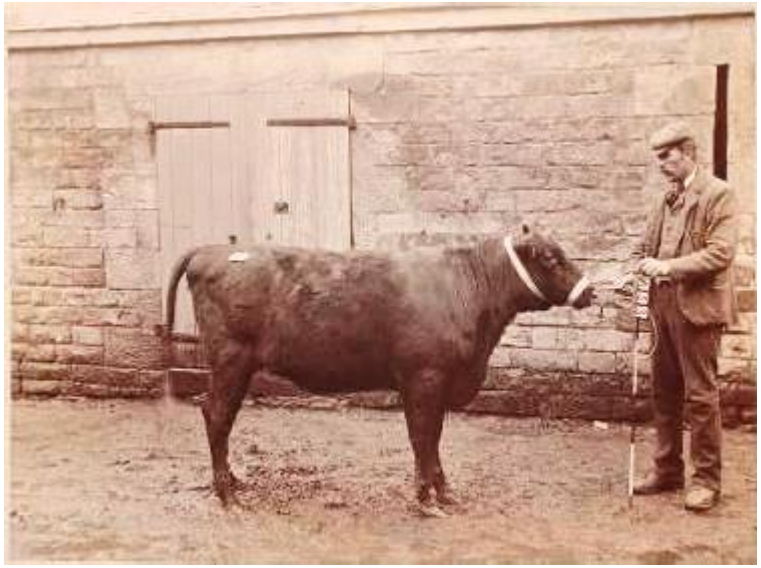
(1) Historical trends

(2) Demographic/sociological theory

(3) Expert judgment

Consultations and incorporating expert judgment

An approach to uncertainty: wisdom of crowds



he mean
ıd of the
ıture for
st month

cal year-
5). Both
s years.
—Bulletin
contains
al station
esults of
ations in
1851 and
ely in a
station is
, and is
ral appli-
ico with
tempera-
naximum
absolute
absolute
ual rain-
ount was
05. Most
eptember.
unt being
by J. D.

e results
Deutsche
he second

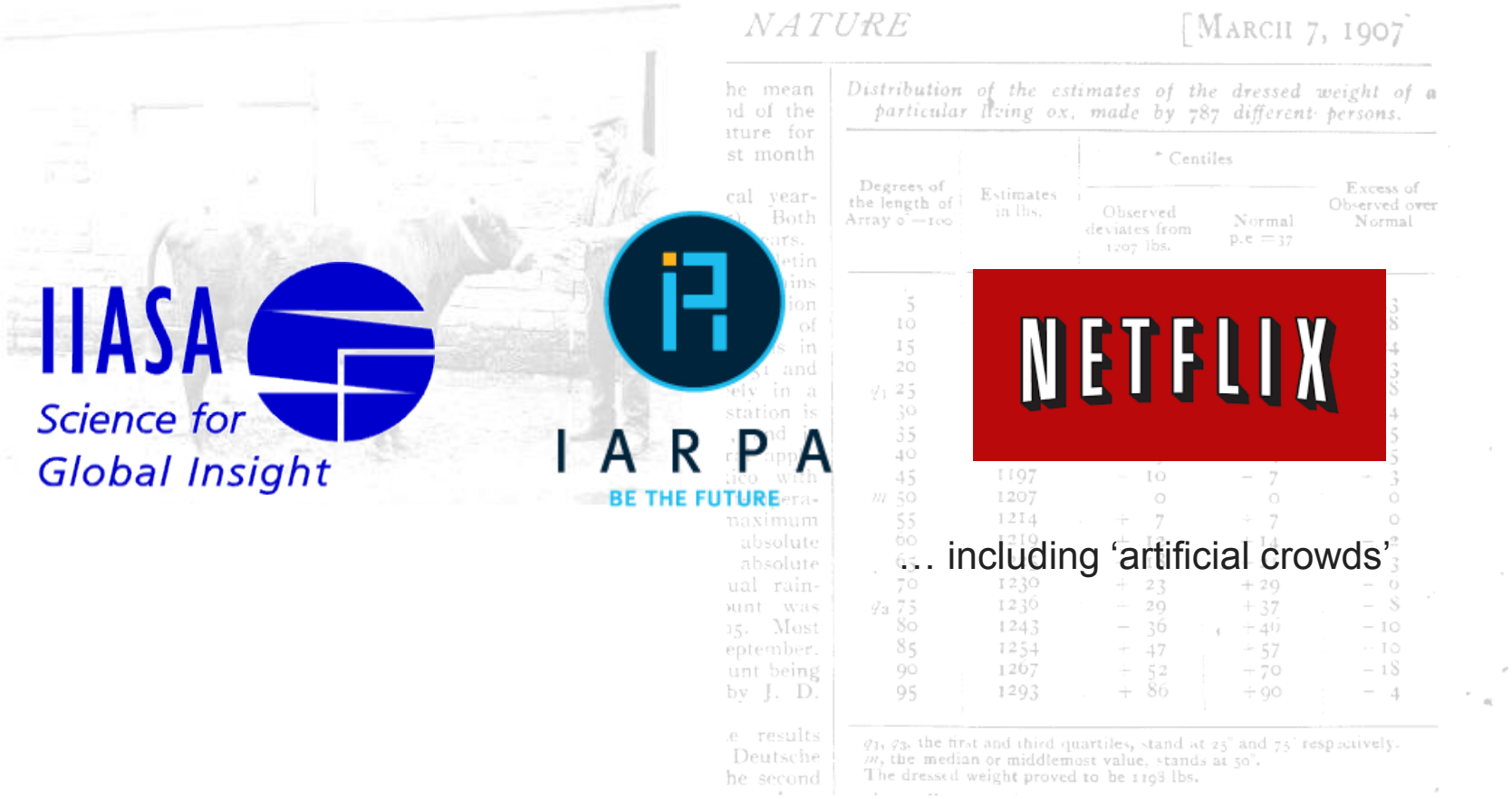
Distribution of the estimates of the dressed weight of a particular living ox, made by 787 different persons.

Degrees of the length of Array 0°—100°	Estimates in lbs.	* Centiles: Observed deviates from 1207 lbs.	* Centiles: Normal p.e = 37	Excess of Observed over Normal
5	1074	−133	−90	+43
10	1109	− 98	−70	+28
15	1126	− 81	−57	+24
20	1148	− 59	−46	+13
q_1 25	1162	− 45	−37	+ 8
30	1174	− 33	−29	+ 4
35	1181	− 26	−21	+ 5
40	1188	− 19	−14	+ 5
45	1197	− 10	− 7	+ 3
m 50	1207	0	0	0
55	1214	+ 7	+ 7	0
60	1219	+ 12	+14	− 2
65	1225	+ 18	+21	− 3
70	1230	+ 23	+29	− 6
q_3 75	1236	+ 29	+37	− 8
80	1243	+ 36	+46	−10
85	1254	+ 47	+57	−10
90	1267	+ 52	+70	−18
95	1293	+ 86	+90	− 4

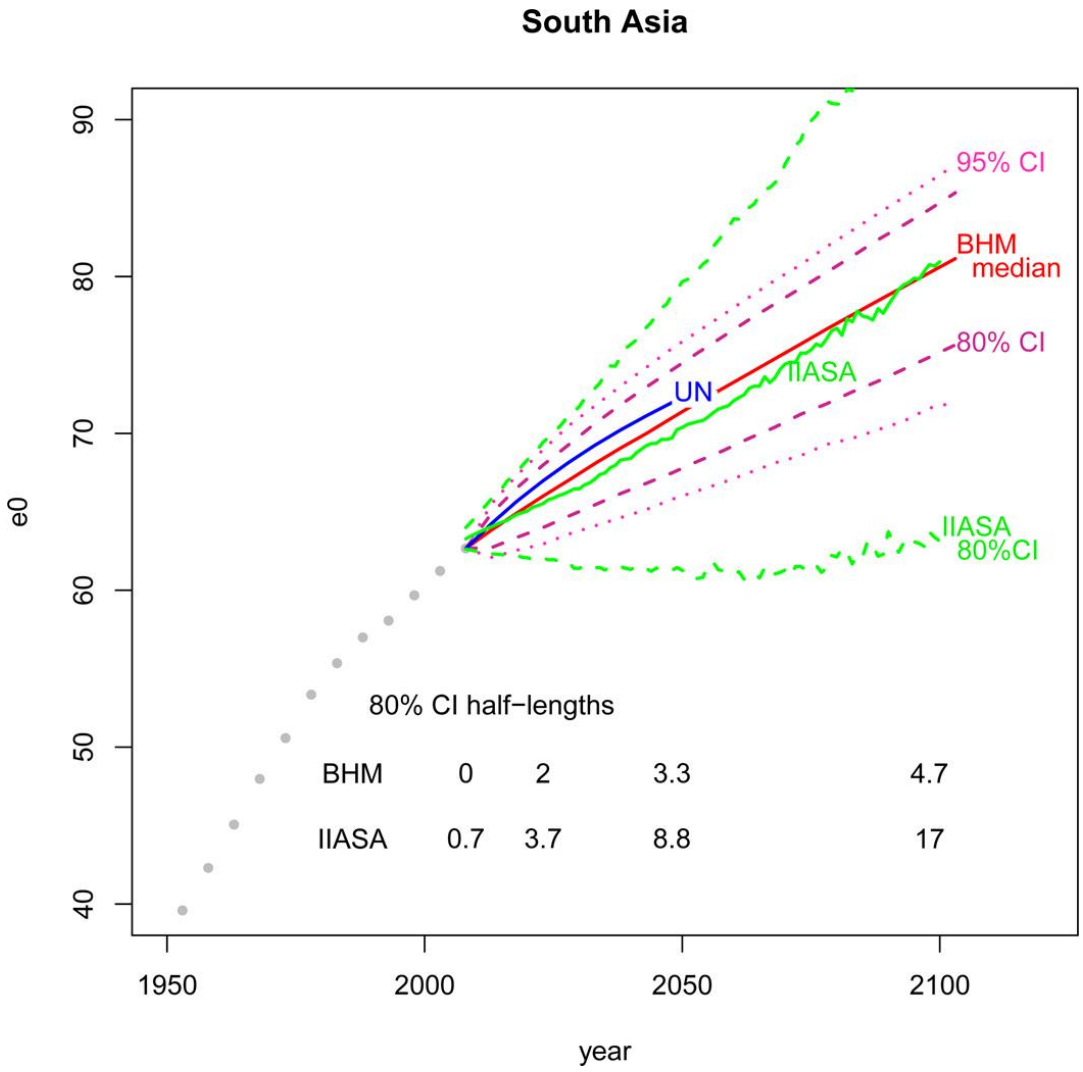
q_1, q_3, the first and third quartiles, stand at 25° and 75° respectively.
m, the median or middlemost value, stands at 50°.
The dressed weight proved to be 1198 lbs.

Consultations and incorporating expert judgment

An approach to uncertainty: wisdom of crowds



Consultations and incorporating expert judgment



Consultations and incorporating expert judgment

Collective wisdom of California experts in attendance