Henry A. Wallace (1888–1965): Agricultural Statistician (Econometrician) Extraordinaire

Lawrence Hubert

Department of Psychology
The University of Illinois

Statistical Best Practices Presentation
Leiden, The Netherlands
January, 2014

This talk and the associated handout are at:

`cda.psych.uiuc.edu/wallace_handout.pdf`

`cda.psych.uiuc.edu/wallace_beamer_talk_short_version.pdf`

What follows is a quote from Studs Terkel, the great oral historian:

There are three great Americans of the 20th century; two are household names, Franklin D. Roosevelt and Martin Luther King. The third should be: Henry A. Wallace.

Henry A. Wallace: A Brief Vita

Henry A. Wallace was born in 1888 on a farm in Iowa and into the prominent Wallace family.

His grandfather (Henry Wallace) was the editor and owner of the prominent farm journal, *Wallaces' Farmer*

His father (Henry C. Wallace) worked on the journal; he was also an Associate Professor of Dairy Science at Iowa State, and later the U.S. Secretary of Agriculture from 1921-1924.

Henry A. Wallace:

1910: B.S. in Animal Husbandry, Iowa State College, Ames

1910–1932: Editorial staff of *Wallaces' Farmer* (Chief Editor, 1924–1932) (Motto: Good Farming; Clear Thinking; Right Living)

1933–1940: U.S. Secretary of Agriculture (under Roosevelt and the New Deal)

1941–1945: Vice-President of the United States (under Roosevelt)

1945–1946: U.S. Secretary of Commerce (under Truman)

1948: Progressive Party nominee for the U.S. Presidency (the Party's platform called for the end of segregation; equal rights for women; and a Medicare-type program [that we will call Wallace-care])

The Wallace Legacy to Statistical Best Practices

Henry A. Wallace demonstrated the best practices in (agricultural) statistics and econometrics throughout his whole life, including his time as the New Deal U.S. Secretary of Agriculture (1933–1940)

Wallace demanded an evidence-based set of agricultural policies and practices (through statistics and data gathering) that helped pull the U.S. out of the Great Depression. He was, for example, the prime initiator of the annual *Agricultural Outlook Report* of the USDA

He is arguably the single individual responsible for the first Department of Statistics in the U.S.; this is due to Wallace's connections with George Snedecor.

He was the lead architect of numerical procedures (in the 1920s) for solving the normal equations in large-scale multiple regression.

Hybrid Corn

In 1924, Wallace developed a hybrid corn called Copper Cross that won a gold medal at the Iowa Corn Yield Test.

This was the first time a hybrid had beaten standard open pollinated strains.

To market this, Wallace founded in 1926 the (Pioneer) Hi-Bred Corn Company

Slogan: Developed not discovered; Made to fit — not found by chance

By the early 1940s, 98 percent of all planted corn in Iowa (and elsewhere) was hybrid, with much of it sold by Pioneer.

A small anecdote that shows Wallce's belief in evidence-based actions backed up by statistical analyses (from Mordecai Ezekiel):

During the depths of the depression Wallace built up a market for hybrid seed by offering it to farmers without requiring a cash payment, on condition that they plant half their land in hybrid and half in ordinary corn, and then pay back a fraction of the value of the higher yield on the hybrid acres.

Agricultural Prices (Henry A. Wallace; 1920)

This work has been called the first true econometric study published in the United States.

Wallace commented that economists dealt too much with economic theory and believed that the truth ultimately was to be found in statistics.

Short excepts from the preface and the chapter called the "Technique of the Ratio Method" follow:

Its purpose is to promote a better understanding of the factors which influence prices of farm products and stimulate an intelligent interest in statistical economics.

If this book is used as a college text, it is hoped that it will be by classes which are especially concerned in applying statistical laws to agricultural prices. Such classes should have free access to calculating machines, multiplying tables, etc., and should make a serious effort to work out various ratios and also to work out supply and demand laws for various farm products by means of correlation coefficients, lines of regression, etc.

TECHNIQUE OF THE RATIO METHOD

The fundamental idea of the ratio method is that the price of every product is determined in the long run by the price of some other product or products. The price of hogs is determined in the long run by the price of corn. ...

In its simplest form, the ratio method deals with only two products, as for example, with hogs and corn. As an average of the ten Januarys extending from 1907 to 1916, No. 2 corn on the Chicago market sold for 59.9 cents a bushel, and heavy hogs on the same market sold for \$7 per hundredweight. In other words, as an average of this ten-year period, it has required the value of 11.7 bushels of corn to equal in value one hundred pounds of heavy hog flesh.

Correlation and Machine Calculation (H.A. Wallace and George W. Snedecor; 1925)

A remarkable publication; its story is told in the handout.

It details how to carry-out Gaussian elimination (using a machine) for the solution of the normal equations when you have many independent variables.

This was Snedecor's first publication (in statistics or elsewhere) at the age of 42.

It led directly to the Mathematical Statistical Service at Iowa State (1927).

And in 1933, to the Iowa State Statistical Laboratory, arguably the first "Department of Statistics" in the United States.

In some reminiscences made in 1951 at the Oral History Research Office of Columbia University, Wallace comments:

My work on cycles began in about 1913, when I began to study the relationship of weather to corn yields, of corn supply to corn prices, the relation of corn prices to hog prices, the cycle of hogs, the cycle of cattle, the cycle of horses, and so on. I did that as a preliminary to getting into more serious and careful statistical work. As a result of studying the relationship of corn weather to corn yields, I ran across the work of H.L. Moore, the Columbia University professor. He had put out some very careful statistical analyses involving the relation of independent variables to a dependent variable, expressing them by regression lines and correlation coefficients. Suffice it to say that I became proficient at doing work of that kind, using a key-driven calculating machine to facilitate matters.

I thought that the people at the Iowa State College of Agriculture and Mechanical Arts at Ames were not sufficiently current in that field. I went up and met with several of the professors and sold them on the idea that they should be able to evaluate their experimental work much more accurately if they had more adequate statistical background. As a result, they employed me for ten weeks to conduct a statistical course ... There was no one in the class of some twenty who was not either a professor or a post-graduate student.

Wallace continues:

Then I took another problem which was interesting to them as agricultural people–the relationship between farm land values in the different counties–for which there were census figures–and the yield of corn per acre. We used an average of ten years for which we had crop reporting figures. We used the percentage of the crop land in corn, for which we had census figures; the value of the buildings per acre, for which we had census figures; and so on. We took up various independent variables bearing on the dependent variable of the value of the farm land per acre. That was the problem which I set to them, which later was embodied in a bulletin put out by Iowa State College entitled *Correlation and Machine Calculation*.

Other Notable Occurrences at Iowa State

Noting that H.A. Wallace was head of the USDA from 1933 to 1940, we may have the literal as well as figurative giving of "pork" to Iowa State Statistics.

The Statistics Section of the Iowa Agricultural Experiment Station was organized in 1935, with Snedecor as Section Head.

A hugely lucrative cooperative agreement was made in 1938 between Iowa State and the USDA for joint research in agricultural statistics and related statistical theory:

survey methodology for area sampling and the development of the Master Sample of Agriculture;

methods for forecasting and estimating crop yields.

What's Looks to a Hog (H.A. Wallace, *Wallaces' Farmer*, 1907)

"Corn shows" were very popular at the turn of the century, with major prizes given for the appearance of corn.

As a teenager (16 years old), Wallace test-planted five acres with corn from the "Corn Professor," Perry Holden, that had been judged for "beauty." [Wallace had been experimenting with plant breeding in his own plot since he was ten.]

After meticulous quantitative analysis of the test-planting, Wallace found no correlation between beauty and high yield.

An instance of "what you see is not what you get."

As a result of Wallace's work (and writing about it in *Wallaces' Farmer*), corn shows were gradually replaced by yield contests.

This is a good example where a "surrogate end point" (i.e., corn beauty) is not a good one for the real endpoint of interest (i.e., corn yield).

Corn and Its Early Fathers (Henry A. Wallace and William L. Brown)

William L. Brown was the Chairman and President of Pioneer Hi-Bred International, the company Wallace founded in 1926.

There are two editions of this book; one in 1956 and a second in 1988 after Wallace had died.

Brown revised this text to reflect the contributions of Wallace to the corn story; and to specifically highlight the fact that Wallace's work with corn was part of the worldwide agricultural revolution initiated by Wallace.

There is a remarkable Chapter 14, Small Gardens and Big Ideas, that is given in the handout. A few quotes are given from this chapter:

More and more, we feel that grave danger exists of statistics being used as a substitute for critical observation and thought.

The senior author joins somewhat apologetically in presenting this point of view because he had a lot to do, back in 1923, with starting the present Statistical Laboratory in Iowa State College, one of the better departments of its kind in the nation. Statistics have their place, a very important one, but they can never serve as a substitute for close association with plants. Their real value, it seems to us, is in measuring precisely what we already know in a general way. Statistics tends to be an office art based on machines and figures rather than a field art based on living things.

A great deal of work has gone into the development of mathematical models designed to explain just how hybrid vigor operates. As a mode of attack, we believe, it overlooks completely the more important problem of understanding the specific ways in which hybrid vigor affects the plant itself. We fear that until we return to thinking of corn in terms of what the plant itself is doing, instead of working out neat mathematical formulae to fit what we think its performance should be, no real advance will be made.

What is in the corn judge's mind?

This is a highly cited paper (H.A. Wallace, *Journal of the American Society of Agronomy*, 1923) in the contemporary judgement and decision-making literature; it is reproduced in whole in the handout.

[There is also an article by H.A. Wallace in *Wallaces' Farmer* (March 2, 1923) that tells the story to farmers without actually using the solutions to the normal equations (*What the Corn Judge Can't See*: Factors that Determine Yield Don't Show on the Score Card).]

It may be the first empirical piece to appear anywhere on the distinction between "Actuarial versus Clinical Prediction."

It relies on Sewall Wright's method of path coefficients published in 1921 as "Correlation and Causation"

In the form that Wallace used Sewall's method, he had 500 ears of corn that were assessed using six (independent) variables:

b: length (ear)
c: circumference (ear)
d: weight of kernel
e: filling of kernel at tip
f: absence of blistering
g: absence of starchiness

There were two possible dependent variables:

a: judge's rating

y: yield (field tested)

The intercorrelation matrix follows:

	a	b	c	d	e	f	g	y
a	x	.7	.4	.5	.4	.2	.3	.2
b	.7	x	.3	.3	.2	.2	.2	.2
c	.4	.3	x	.2	.1	.0	.0	.15
d	.5	.3	.2	x	.4	.2	.2	.4
e	.4	.2	.1	.4	x	.5	.6	.4
f	.2	.2	.0	.2	.5	x	.5	.3
g	.3	.2	.0	.2	.6	.5	x	.2
y	.2	.2	.15	.4	.3	.2	.2	x

Matlab corn_script.m


```
rater_vector = [.7;.4;.5;.4;.2;.3]
yield_vector = [.2;.15;.4;.3;.2;.2]
intercorr = [ 1.0 .3 .3 .2 .2 .2; .3 1.0 .2 .1 .0 .0; ...
.3 .2 1.0 .4 .2 .2; .2 .1 .4 1.0 .5 .6; .2 .0 .2 .5 1.0 .5; ...
.2 .0 .2 .6 .5 1.0 ]
beta_rater = inv(intercorr)* rater_vector
beta_rater_abs = abs(beta_rater)
sum_beta_rater_abs = sum(beta_rater_abs)
prop_beta_rater = (beta_rater_abs) ./ (sum_beta_rater_abs)
beta_yield = inv(intercorr)* yield_vector
beta_yield_abs = abs(beta_yield)
sum_beta_yield_abs = sum(beta_yield_abs)
prop_beta_yield = (beta_yield_abs) ./ (sum_beta_yield_abs)
rsquared_rater = (rater_vector')*beta_rater
rsquared_yield = (yield_vector')*beta_yield
```

beta_rater =
0.5437 0.1730 0.2334 0.1726 -0.0830 0.0825
sum_beta_rater_abs =
1.2881
prop_beta_rater =
0.4221 0.1343 0.1812 0.1340 0.0644 0.0641
beta_yield =
0.0480 0.0623 0.3105 0.1123 0.0556 0.0331
sum_beta_yield_abs =
0.6218
prop_beta_yield =
0.0771 0.1002 0.4994 0.1806 0.0894 0.0533
rsquared_rater =
0.6436
rsquared_yield =
0.1946

A Few Notable People at the USDA During Wallace's Tenure

John Kenneth Galbraith worked at the USDA on New Deal initiatives (Berkeley Ph.D. in Agricultural Economics, 1934)

Mordecai Ezekial worked as the Economic Advisor to the Secretary of Agriculture (published *Methods of Correlation Analysis*, 1930; 1941). Speaking of Wallace, he said:

He was one of the great generative forces of the first half of the twentieth century.

Adlai Stevenson worked for the Agricultural Adjustment Administration.

W. Edwards Deming worked at the USDA (1928–1938); he brought Walter Shewhart to lecture (and brought his ideas to the 1940 census)

Deming brought Neyman to the USDA Graduate School to lecture in 1937; this led to *Lectures and conferences on mathematical statistics and probability* (1938; 1952)

Other Notable Facts

Almost fifty percent of the world's eggs are produced by hens that derive from the genetics of Hy-Line poultry stock;

the Hy-Line breed was developed by Wallace in the 1930s; the company was founded in 1936.

Pioneer Hi-bred is now wholly owned by Dupont; it was bought in 1999 for 10 Billion dollars.

During R.A. Fisher's visit to the U.S. and Iowa State in the Summer of 1931, Henry Wallace attended most of Fisher's lectures; word has it however that Fisher was not particularly interested in Wallace's more fanciful ideas, i.e., the placement of the planets and weather in Iowa, and how that all affects corn production.

Items Present in the Handout

1) *Journal of the American Society of Agronomy*, 1923, *15*, 300–304.

WHAT IS IN THE CORN JUDGE'S MIND?

2) Henry Wallace and statistics by Oscar Kempthorne

3) Henry A. Wallace and the Modeling of Expert Judgments (From *A Statistical Guide for the Ethically Perplexed*, Lawrence Hubert and Howard Wainer, 2013, pp. 146–150).

4) corn_script.m and results

5) From *Correlation and Machine Calculation*, Wallace and Snedecor, Part VI: Precautions and Suggestions (1925, pp. 46–47)

6) Excerpt from H.A. David, *Statistics in U.S. Universities in 1933 and the Establishment of the Statistical Laboratory at Iowa State*, *Statistical Science*, 1998, *13*, 66–74.

7) From John C. Culver and John Hyde, *American Dreamer: A Life of Henry A. Wallace*, 2000, pp. 26–29

8) *Corn and Its Early Fathers* (Henry A. Wallace and William L. Brown; 1956 and 1988); Chapter 14: Small Gardens and Big Ideas

9) The Story of a Perplexing and Indomitably Naive Public Servant

AMERICAN DREAMER: The Life and Times of Henry A. Wallace; By John C. Culver and John Hyde; W.W. Norton (2000).

March 12, 2000 — ARTHUR SCHLESINGER JR. —

10) *Early Statistics at Iowa State University*, Jay L. Lush, Chapter 5: The Wallace Lectures