Guinnessometrics:
The Economic Foundation of "Student's" *t*

by Stephen T. Ziliak[1]

> *The History of Science has suffered greatly from the use by teachers of second-hand material, and the consequent obliteration of the circumstances and the intellectual atmosphere in which the great discoveries of the past were made. A first-hand study is always instructive, and often . . . full of surprises.*
> Ronald A. Fisher 1955b, p. 6

In "The Probable Error of a Mean" (1908a) "Student" made one of the great contributions in the history of science. His contribution is the test and table of "Student's" *t* or, as Ronald A. Fisher called it, *"Student's" test of significance* (Student 1908a). "I have purposely laid emphasis on one scientific achievement with which ['Student's'] name will always be associated," Fisher confessed in a 1939 obituary notice for "Student." "[I]t is the 'Student' of *'Student's' test of significance* who has won, and deserved to win, a unique place in the history of scientific method" (Fisher 1939, p. 9; italics in original).

Unlike some other great achievements, such as continental drift and flying bicycles, "Student's" test of significance did not divide the scientific community. "'Student's' treatment of the problem of uncertainty," said the great Bayesian physicist, Sir Harold Jeffreys, writing in full agreement with the classical Fisher, "has the further merit of being accepted by all schools" (Jeffreys 1961, p. 378).[2]

But Fisher and Jeffreys did not mention that Fisher himself was commercially speaking a big *cause* of "Student's" place in history. Though published in Francis Galton's and Karl Pearson's new and highly acclaimed journal, *Biometrika* (1901-), "Student's" test of significance was shelved until Fisher picked it up and made it central to *Statistical Methods for Research Workers* (1925).[3]

Other works by Fisher got noticed and built upon by top-flight theorists, including top economic theorists.[4] But in general, the influence of the "Student"-enriched *Statistical Methods for Research Workers* on how we argue is evidently challenged only by *The Design of Experiments* (1935a), another Fisher product. And in *Design*, as much or more than in *Methods*, "Student's" test of significance was *again* made central to the forms of Fisher's arguments. The "lady tasting tea" is a famous 1935 illustration of "Student's" test (Fisher 1935a, pp. 11-26). But Fisher's second classic book was like his first, literally steeped in significance.[5] "The feeling induced by ["Student's"] test of significance has an objective basis," the author later explained. "The level of significance fulfils," he said in a difficult platitudinous philosophical language, "the conditions of a measure of the rational grounds for the disbelief it engenders" (Fisher 1956, p. 43).

Precisely what Fisher means by that is still highly uncertain.[6] One thing for sure: blue skies, nothing but blue skies, for "Student's" test of significance. Sir Ronald Fisher himself, a Copley Medal winner and super-meme of science, swept up "Student's" test and made it his own (Dawkins 2006, p. 163, 434). He *volunteered* to be its biggest commercial and philosophical communicator.

It would be surprising to discover a big difference between "Student" and Fisher.

Philosophers and Porters

"The very different genius which appears to distinguish men of different professions," wrote Adam Smith in *The Wealth of Nations*, "is not upon many occasions so much the cause, as the effect of the division of labour." "The difference between the most dissimilar characters," Smith continued, "between" he said "a philosopher and a common street porter, for example, seems to arise not so much from nature, as from habit, custom, and education" (Smith 1776, Book I, 2.4).

Smith's insight may be applied to any difference of genius or profession. In the history and philosophy of econometrics, an intriguing if unexplored difference can be found between William Sealy Gosset (1876-1937) aka "Student" of "Student's" *t,* and one of "Student's" own adopted students, his friend the great mathematical geneticist and statistician, Ronald A. Fisher (1890-1962).[7]

"Student" was not a common porter, true. But "Student" was an anonymous *brewer* of porter who, we've commonly heard since the 1920s, somehow "discovered" the *t*-table and test.[8] How or why he discovered *t* we do not know.[9]

Ronald Fisher was not a professional philosopher, either, but he is widely credited as such.[10] Sir Ronald is regarded as a mature, if not our most mature, philosopher of scientific inference. Indeed, the great champion of "Student's" test is considered by eminent statisticians from Maurice Kendall (1978) to Bradley Efron (1998) to be the chief visionary and architect of modern statistics. And Fisher's *Statistical Methods*—"Student's" biggest commercial, they agree—is the "serious man's" introduction to it (Savage 1971, pp. 441-2).[11]

Inter-scholastic disputes exist, as have existed under the radar long before Savage's last anguish.[12] Still, in the normal distribution of scientific credit, all schools accept that Fisher was the mathematical David who defeated a great and powerful Karl Pearson, by remaking biology *and* statistics—including "Student's" test—in his own image.[13]

We are all Fisherians now, and especially as regards our covenant on the "significance" of a difference from a null, guided "objectively" and "exactly" by "Student's" *t.*[14]

The social significance of "Student's" test is illustrated by the epistemological status it is accorded by editors and referees.[15] In economics and other sciences after Fisher, "statistical" significance is by custom, habit, and education a necessary and sufficient condition for proving a result. The case is not exaggerated: "significance" rules.[16] Fisher's "rule of 2" or, what is similar, his "5 percent rule," applied to a difference—$t \geq 2.0$ or $p \leq .05$—has in reality become rule of law, deciding cases as it does in Supreme Courts of the United States.[17] The rule is, in brief: "If the variable doesn't fit/you have to acquit."

Significantly, it was Fisher the philosopher, not "Student" the porter brewer, who invented the ubiquitous rule and with it, a rote procedure for science and jurisprudence. Like

Smith said, "The difference of talents comes then to be taken notice of, and widens by degrees, till at last the vanity of the philosopher is willing to acknowledge scarce any resemblance."

And at last, the difference between "Student" and Fisher was scarcely acknowledged by the talented Fisher.[18] At the centenary mark of "Student's" path-breaking article, a first-hand study of "Student," his circumstances, and his actual approach to *t* might be instructive.[19]

"It will be seen then that the difference between Prof. Fisher and myself," "Student" prophesied, "is not a matter of mathematics—heaven forbid—but of opinion" (Student 1938, p. 367).

"Student"—A Porter Brewer's Tale

"Student" is the pseudonym used in 19 of 21 published articles by William Sealy Gosset, who was by pre- and real-occupation a chemist, brewer, inventor, and self-trained statistician, agronomer, and designer of experiments (Student 1942).

Gosset was born in 1876 in Canterbury, England. Gosset and Vidal are recognized families of Kent and Devon, old of French Huguenot descent. The eldest of five children born to Colonel Frederic Gosset, R.E. (Royal Engineer), and Mrs. Agnes Sealy (Vidal) Gosset, William was cradled in the gentry and educated at Winchester School and Oxford University.[20] By all accounts it was a happy and comfortable childhood and life.

The great unknown of statistical science worked his entire adult life—1899 to 1937—as an experimental brewer. He had one employer, Arthur Guinness, Son & Company, Ltd., Dublin, St. James's Gate.

It sounds like fantasy. But Gosset—the statistical "Student" of "Student's" test of significance—was a master brewer. He was a businessman, too, and rose in fact to the top of the top of the brewing industry—Guinness, St. James's Gate—that global maximum, that fermentation kingdom, that looming dark brick factory of good jobs and great beers, perched high above the Dublin average—yet friend of rich and poor—where he was hailed by more than golden harps and smiling toucans.

Gosset was in 1899 a energetic—if slightly loony—23 year old gentleman scientist.[21] An obsessive observer, counter, cyclist, and cricket nut, the self-styled brewer had a sizzle for invention, experiment, and the great outdoors. "All who have known him agree," recorded a master of statistics, "that [Gosset] possessed more of the characteristics of the perfect statistician than any man of his time" (E.S. Pearson 1939, pp. 248-9).[22] A friend from childhood remembered that Gosset stood stoutly in the heart, too, possessing, he said, "an immovable foundation of niceness."[23]

Twenty years before Fordism would really hit the gas, Guinness was going great guns giving its brewery a radical make-over, and the nice chemist was one of a few good men it brought in to help.[24]

To wit, economies of scale had not been fully tapped. The look-touch-and-sniff approach of 18th century craft brewers had satisfied a Guinness bottom line for 150 years. No longer. (The firm was established in 1759; it went public in 1886.) The extent of Guinness's market was in a sense already large—larger than Bass, their chief competitor, largest in Ireland, and largest in the world. By 1914 annual production at Gosset's brewery would surpass 2 million hogsheads or 4.35 billion pints.[25] But until the 1890s, the extent of the market was limited by the guildsmen's division of labor (Dennison and MacDonagh, p. 23, 38; O'Grada, pp. 304-5).

The Guinness future was in "scientific brewing"—large-scale, industrially controlled brewing—wherein all factors of production, from barley breeding to taste testing, are controlled, improved, and confirmed by experimental science.[26] Danish and German breweries were transforming similarly—perhaps a little ahead of Guinness's imposing pace.[27] And worldwide, it's true, a scientific enthusiasm was in the general air, for industry and poor laws as much as for agriculture and beer laws.[28]

But Guinness was unique in the industry in a number of ways, and especially—for the future development of statistics—by vesting its exclusively chosen brewers with economic authority. With scientific brewers in managerial position, the theory was, experiments could shine a light on the bottom line and the bottom line a light on the experiments.[29]

On August 4, 1899 C. D. La Touche, then managing director, recorded flatly in a note to his Chief Executive Officer, Sir Reginald Guinness, that "Mr Gosset" graduated Winchester as "Scholar of New College [Oxford University], [earning a] 1st Class in Mathematical Moderations, Trinity Term 1897, and 1st Class in Chemistry, July 1899. He is short-sighted and wears spectacles," added La Touche. "Seems generally speaking suitable."[30] Reginald Guinness traveled to London to meet Mr. Gosset in person and agreed: Gosset's apprenticeship in experimental brewing began in October, 1899.

The great experimentalist was short-sighted when mocking his own "maths."[31] With it, you can see, he mocked neither mathematics nor Oxford (the number theorist G. H. Hardy was a classmate and pal) but his own abilities. In a letter to Karl Pearson he confessed, "I don't feel at home in more than three dimensions."[32] He told Fisher, "My mathematics stopped at Maths. Mods. at Oxford, consequently I have no facility therein."[33] True, Gosset did not have the facility of Émile Borel. But that is not the same as "*no* facility": open a page of Gosset, 1907 to 1937, and one finds his jokes beside the point (Student 1942).[34]

Regardless, the inventor of "Student's" *t* possessed a wickedly fertile imagination and more energy and focus than a St. Bernard in a snowstorm. He worked in relative obscurity. He won no prizes and held no academic posts; his bust ornaments no column or portico. But he went on to invent or inspire solution concepts that have evolved into six or seven different research fields in the modern disciplines of statistics, agronomy, brewing, and industrial quality control.[35]

Gosset made significant scientific and commercial contributions, too, embodied in Guinness beer. Other contributions are less opaque and yet lamentably unknown.[36]

Strangely nowhere in our statistical education do we learn that "Student" was a major scientist and muse of 20th century statistics who once lorded over the great Irish brewery. The correlation between "Student's" test and Guinness is positive and large. Log-logistically Gosset rose in rank to become both Head Brewer and Head Statistician. Pounding out up to 100 million gallons of Guinness annually, Gosset brought to perfection the quantitative side of scientific brewing and with it a storehouse of theory and tools. "Student's" test made the "Student" name world-famous but the test was nothing like Gosset's last or favorite word on the subject of uncertainty.[37]

From 1899 to 1906 Gosset was Apprentice Brewer, mostly in the "Experimental Brewery," a miniature brewery within the Main.[38] In 1904 he began to tackle the problem of making an inference from small samples of malt and hops, two of the crucial inputs. With Board endorsement he spent in 1906-1907 a sabbatical year at Karl Pearson's Biometrics Laboratory, University College London, where his general logic of small samples—and "The Probable Error of a Mean"—fermented.[39] (He married in 1906 a famous hockey player and coach, Marjory Surtees Phillpotts, with whom he sired three children.)

In 1907 Gosset returned to Dublin as Head Experimental Brewer, a position he held through calm and turbulent times until 1935. (He volunteered to fight in the War but, like Fisher, his application was denied by reason of short-sightedness.) In the early 1920s Gosset became Head of the Statistics Department he established. (His first task: to estimate the effect of advertising on sales.)[40] Gosset was promoted to Head Brewer in 1935. As Head Brewer *and* Head Statistician, he controlled product at two plants—St. James's Gate and the new plant at London, Park Royal (now closed). Gosset died of a heart attack in Beaconsfield, England, October 16, 1937, his actual identity in science being eclipsed for more reasons than the pseudonym itself can explain.[41]

The Economic Origins of "Student's" *t*

Small samples came microeconomically to Gosset—an organic case of competition and choice arising from scarcity. His job was to show a profit in experimental porter without raising the price of the porter. That, in short, is how he discovered *t.*

Porter is a name given to a dark and bitter beer with a good head on it. The origin of its name is equally mysterious. The economic historian Oliver MacDonagh (1964, p. 530) traces porter to early 18th century London. "Black" beers, MacDonagh finds, were revealed preferred by common street porters, and publicans and brewers seized an opportunity. In the 19th century and early 20th century the eponymous porter was alternatively called "stout" or "stout porter." Normally speaking the different names did not signify a real deviation of product.[42]

"Stout" is the name used now to describe a beer such as Guinness that is bitter on the up-take (bitterness being a function of both the quantity and quality of hops added per barrel of malt) and yet smooth, mellow, and slightly smoky on the finish. Black-ruby tint arises like the smoky finish from roasted barley or "malt"—the distinguishing ingredient. As Arthur Guinness himself stated before a Parliamentary Committee of 1783, "A porter brewer buys none but the best, as none else will answer."[43]

But "best"-practice brewing was, scientifically speaking, anyone's guess. Until the 1890s, the industry ran on faith and folk theorems. William Sealy Gosset had to be invented.

Brewing "experimentally" introduced challenges and trade-offs. For example, in Gosset's time, Guinness stout was a completely natural and unpasteurized beer. No artificial ingredients or preservatives. But what is good for health or taste is bad for the life of the beer: in keg, cask, or bottle, the life of a natural beer is numbered in days. Yet Guinness's beer was shipped worldwide, on an increasingly large scale, a problem. Hops acts as a natural and effective preservative, true. But it is bitter, and brings bacteria, pests, and other costs to the beer. Other things equal, each pound of hops added to the mash tuns induced life *and bitterness* in some (as yet) unknown functional form. A heavily hopped beer, such as "Foreign Extra Stout," would continue to condition on the ocean voyage. By Port Royal, Jamaica, or Rio de Janeiro, a hopped beer may be safe but *too* bitter—not Irish.[44] A trade-off to be estimated.

Another challenge was that Guinness, a wholesale dealer only, pursued an unusual pricing strategy: constant nominal price (measured in Dublin prices).[45] Product price was fixed. But between 1887 and 1914, output more than *doubled*. Plant size expanded, too, and with it the capital/output ratio. So the profit-searching question was: how can experimental science advance economies of scale in brewing? How can inferential statistics help to control and improve product, while at the same time reduce average total costs for the firm?

Take hops, for example. The method of choosing the ancient yellowing bells based on looks or fragrance wasn't treated to a pauper's burial. But neither was it efficient or reliable on the large scale. When in 1898 results of Guinness's first hops experiments arrived the brewers were elated. "Until quite recently, no attempt has been made to analyse hops commercially," began a report, "Hops, Season 1897."[46]

The author of the report was Thomas B. Case, an older graduate of Winchester and Oxford.[47] Case and his team felt they had found a method "proving that those hops which contained the most soft resins gave the most stable beer"—the most shelf life.

"The earliest Brewery experiments were directed towards getting the same results with the same sample of hops." So Case analyzed the average percentage of soft and hard resins found in samples of 50 grams (dry weight) taken from several seasons of American and Kent hops (Case 1898, p. 47). He compared his samples with those of Mr. L. Briant, a cooperating scientist. (Extracting resins from hops is a chemical procedure too involved to explain here.) Table 1 summarizes the results:

Table 1. Results of Guinness's First Hops Experiment, 1897-98
(Average % Resins Content)

Analyzed by T.B.C.	Sample size	Soft resins	Hard resins	Total
Kents, 1896	6	7.3	5.4	12.7
Kents, 1897	11	8.1	4.2	12.3
American, 1895	2	8.0	9.2	17.2
American, 1896	10	10.2	6.8	17.0
American, 1897	5	11.5	5.4	16.9
Analyzed by L.B.				
Kents, 1896	-	-	-	-
Kents, 1897	14	8.4	7.1	15.5
American, 1895	2	8.7	10.2	18.9
American, 1896	3	10.4	8.9	19.3
American, 1897	7	11.1	8.6	19.7

Case couldn't celebrate. "We could not . . . support the conclusion that there are no differences between pockets of the same lot."

The Table shows why. Case examined 11 samples of Kents, 1897, finding 8.1% soft resins; Mr. Briant examined 14 samples of the same lot, finding 8.4% soft resins—a difference of .3%. The mean difference between their two samples of "American, 1895" was even higher, at .7% (soft) and 1.0% (hard); and for some reason, Case consistently found *less* resins. Case worried about "defects" in his procedure, especially the "difficulty of sampling."

The main portion of the work concerned "brewing value." The goal, after all, was to employ soft resins content as a predictor of optimal hops for purchase and thus brewing value.

Case offered a "provisional" figure, suggesting that each "1 per cent of soft resins is worth 7s. 4d"—seven shillings, four pence—per hundredweight. Given the 1898 input of 42,108 hundredweight (4.72 million imperial pounds), knowing the true resin-to-shilling ratio was commercially important. Yet "the great difficulty" of sampling kept Case agnostic (p. 50).

It seems the real difficulty, however, was a lack of knowledge about inferential statistics, period. Lacking a theory of error, what difference could Case infer, and with what accuracy, beyond the unmeasured judgment of big, small, or "I can't say"?

Other experiments proceeded apace: on barley yield and quality ("variety trials"), on the chemistry of malt extract, on the temperature of kiln drying.

One of the great difficulties with all these experiments was the small size of samples, and for two economically related reasons. First, in field experiments, the number of acres available was limited by the opportunity costs of farmers and brewers and by the economic division of scientific labor. The scientists experimented on a two acre plot at St. James's Gate (Case 1908, p. 2).[48] But commissioned farmers and maltsters themselves, such as Gosset's long-time collaborator, the eminent E.S. Beaven, allocated no more than four acres (Beaven 1947, p. 164; Gosset 1937, Student 1938). Second, in mixing new beers and malts, time plus the call of profit placed a limit on the samples. As Gosset told Fisher in 1915:

> Experiments naturally required a solution of the mean/S.D. problem and the Experimental Brewery which concerns such things as the connection between the analysis of malt or hops, and the behavior of the beer, *and which takes a day to each unit of the experiment, thus limiting the numbers*[49]

The specter of probable error haunted all these tests of porter but in 1898 Case and his team did not know how.

All this difficulty (save a dependence on small samples) changed rapidly in 1904 when Gosset circulated an internal report entitled "The Application of the 'Law of Error' to the Work of the Brewery" (Nov. 3, 1904):

> Results are only valuable when the amount by which they probably differ from the truth is so small [he said] as to be insignificant for the purposes of the experiment. What the odds should be depends—
>
> 1. On the degree of accuracy which the nature of the experiment allows, and
>
> 2. On the importance of the issues at stake.[50]

Gosset's report was remarkable for a number of reasons.[51] Two features can be highlighted. First, the self-trained statistician perceived an *economic* relationship between "importance" and odds ratios (what level of statistical significance a brewer will accept in any given case). Second—and fully four years before his 1908 article—Gosset underscored a positive correlation in the normal distribution curve between "the square root of the number of observations" and the level of statistical significance. Other things equal, he wrote, "the greater the number of observations of which means are taken, the smaller the [probable] error" (p. 5). "And the curve which represents their frequency of error," he illustrated, "becomes taller and narrower" (p. 7).

Since its discovery in the early 19th century, the normal probability curve had been tabled for comparatively large samples only.[52] Helmert's pioneering work sat for years unread by English statisticians, and seems anyway beside the point (Fisher 1939, p. 4). Gosset had a problem to solve, and a solution for it could not be found among books or friends. Karl Pearson himself dealt in up to thousands of biometric observations, and took "3σ [three standard deviations] as 'definitely significant.'"[53] But with such large samples, Gosset found, nearly everything is *statistically* "significant" (though not, in Gosset's terms, "important"). Regardless, Gosset did not have the luxury of large samples. Gosset's first big drink at the Gaussian table began in 1904 at *n*=2 (Gosset 1904, p. 7). He wasn't out on a statistical limb; he was out for profit.

For example, he said, "it might be maintained" that malt extract "should be [estimated] within .5 of the true result with a probability of 10 to 1" (p. 7). Malt extract = ([Specific gravity of the wort] – 1000) x 4.67, measured in degrees saccharine per barrel of 168 lbs. malt.[54] In the decade before the First World War, an extract in the neighborhood of 133º gave the targeted level of alcohol content. Controlling extract was important for maintaining it and the nominal price rigidity: the higher the extract (and thus alcohol content), the higher the excise tax; ± .5 was a difference or error in malt extract that Guinness and its customers could swallow so Gosset (in the style of statistical astronomy) made the economically and aesthetically harmless range of error a part of the significance test.[55]) With mean differences of extract values Gosset generated with samples from the Main and Experimental breweries, he then calculated the odds of observing the stipulated accuracy for small and large numbers of observations. At:

Odds in favour of
smaller error than .5 [are]
2 observations 4:1
3 " 7:1
4 " 12:1
5 " 19:1
82 " practically infinite[56]

And he correctly concluded, "In order to get the accuracy we require [that is, 10 to 1 odds with .5 accuracy], we must, therefore, take the mean of [at least] four determinations" (*n*= 4; the odds at *n*=3 being insufficiently low).[57]

Circulated by Gosset when Fisher was a "practically blind" 14 year old school boy (Box, p. 14), the report was instantly hailed by the Board. Gosset himself wasn't convinced. "We have been met with the difficulty," he cautioned, "that none of our books [on the theory of errors] mentions the odds, which are conveniently accepted as being sufficient to establish any conclusion." He said, "It might be of assistance to us to consult some mathematical physicist on the matter" (p. 12). Board Endorsement No. 62, signed March 9, 1905, explains that "Mr. Case will make arrangements for Mr. Gosset to have an interview with Prof. Karl Pearson."

Professor Pearson was willing to meet Mr. Gosset at Pearson's summer home, July 1905. The great Pearson, who was the celebrated author of *The Grammar of Science* (1892), Galton protégé, founding editor of *Biometrika*, designer of the first statistics department, and recent inventor of chi-squared (1901), requested a letter detailing his question.[58] Gosset's letter of 1905 is not known. The neglect is unfortunate for the first century of "Student's" *t* and the economic approach to the logic of uncertainty:

> *My original question and its modified form*. When I first reported on the subject [of "The Application of the 'Law of Error' to the Work of the Brewery" (1904)], I thought [said Gosset] that perhaps there might be some degree of probability which is conventionally treated as sufficient in

such work as ours and I advised that some outside authority in mathematics [such as Karl Pearson] should be consulted as to what certainty is required to aim at in large scale work. However it would appear that in such work as ours *the degree of certainty to be aimed at* must depend on the *pecuniary advantage to be gained by following the result of the experiment, compared with the increased cost of the new method, if any, and the cost of each experiment.* This is one of the points on which I should like advice.[59]

Setting the level of significance is not to be made "conventionally" or by "some outside authority in mathematics." What is considered "significant" or not, Gosset told Pearson, what is worth retaining for analysis and decision, depends on the opportunity cost of following a result as if true, added to the opportunity cost of conducting the experiment itself. The limit of significance, in other words, is not to be set at .05 or some other arbitrary convention; how the limit should be set in any given case is a matter of quality-adjusted net pecuniary gain. Gosset never deviated from this central position.

The July meeting took place. But Pearson did not answer the "modified" question. (He never did, though printing in *Biometrika* two-thirds of Gosset's published output.) It seems the talented porter brewer was the "outside authority" giving advice.

The next year Gosset nearly perfected his logic of small samples while on sabbatical at Pearson's Laboratory. "Nearly," one says, because in "The Probable Error of a Mean" (1908a), Gosset effectively tabled "Student's" *z* (as he originally called it) as:

$$z = (X' - \mu)/S,$$

where $X_i, \ldots, X_n$ are i.i.d. with normal distribution $N(\mu, \sigma^2)$,

$$X' = \Sigma (X_i)/n, \text{ and } S^2 = \Sigma (X_i - X')^2/n,$$

and where *n* is total sample size.[60] (Years later, in 1922, Gosset and Fisher agreed to call Gosset's revised and expanded table "*t*", which is the old "Student's" *z* reduced by (*n*-1), "degrees of freedom."[61]) Still the son of a combat engineer could stand alone. Said his editor Pearson in a letter, it made little difference whether the standard error was divided by *n* or the rigorously correct (*n*-1) "because only naughty brewers take *n* so small that the difference is not of the order of the probable error!" But the naughty brewer did not retreat from his opportunity cost approach to small samples. As a profit center, he couldn't.

Gosset began to use regression analysis in the small sample context as early as 1904. In 1908 he used it to revisit the 1898 'hops input' versus 'life of beer' question. In a fantastic analytical leap beyond Case, Gosset—assisted only by slide rule, mechanical calculator, and a few good books—estimated parabolas of the form "$L = A + BH^2$,

where L = life [of beer] in days

A = life in days of no-hopped beer depending on conditions

H = lbs. of hops

B = a constant [slope parameter] depending on the hops and on the conditions"

(Gosset 1908, p. 145).[62]

"No-hop brewings" (A) could survive between "12.2 and 16.7 days," he found after numerous repetitions of the experiment under same and different conditions, whereas "hopped" (B) could live for a month or beyond. Good so far as it goes. Yet here spills forth additional evidence of

Gosset's persistently economic approach. Despite the analytical leap, he a few months later admitted the "advantage to replace a vague character like increase in life . . . . by a single definite value which can be directly converted into £ *s. d.*"—pounds, shillings, and pence (Gosset 1909 *Lab. R.*, p. 211).

So Gosset did take advantage, as was his usual, and replaced the vague character. From partial correlations he calculated on the percentage of soft resins and "[brewing] value," he argued, "We can find an equation giving the probable value for any given percentage of soft resins. The equation is V = 2.82 + 10.78S, where V is the per cent Value compared with the 'standard' hop, and S is the soft resins measure from 9 per cent [$n = 40$]. It will be seen," he said in a language that was practically taboo to a mind like Pearson's, "that each one per cent. of soft resins makes a difference of 10.78 per cent. in the value of the hops. This," he said, "at the average [1909] price of hops, represents about 8*s.* per [hundredweight]." At a 1909 input of 6.79 million pounds of hops, Gosset discovered a big economic difference (Case's guess, being close to it, was better than a large sample regression bias would predict).

"The probable error of the prediction is large," Gosset however cautioned, "being about 6.6 per cent" (compare Pearson's 1% rule and Fisher's 5% rule). But the noisy resins variable did not stop Gosset from making a judgment about resins' *economic* importance to brewing value. "Of this [probable error] some 3.2 per cent. is due to errors of analysis and sampling," he said, "leaving [a residual experimental error] due to brewing errors and other factors not included in the analysis" (p. 212). But with the new if still highly imperfect and noisy method of making inferences from small samples—Gosset called his method "life regressions"—Gosset was able to reject about one-third of the 'standard' hops that unscientific methods had previously commended (pp. 212-13). One third. Again the Board cheered.[63]

What Gosset did later for the bottom line in barley is hopped up by several orders of magnitude (Ziliak 2007). His "extract" regressions and variety trials helped to establish, for example, the quantitative and chemical relationships between ready formed sugars and malt quality and between nitrogenous material in barley and brewing value. To repeat, three barleys he selected and proved in experiments with E. S. Beaven were by 1947 grown on "well over five million acres."

Precision matters, as for example in soft resins and malt extract. But a low probable error—high statistical significance—ranked low in "Student's" ordering. His had no rules. "Student" did not close the brewery on grounds that a result reached or failed to reach an arbitrary level of statistical significance.[64] Minimizing *real error* –the loss function—was the porter-brewer's rule and test. Samples heap up aplenty. Two more shall suffice. On May 18, 1929, Gosset wrote a letter to a very different adopted student and friend, Egon S. Pearson, repeating to the son what had been lost upon the father:[65]

> I fancy you give me credit for being a more systematic cove than I really am in the matter of limits of significance. What would actually happen would be that I should make out P_t (normal) and say to myself "that would be about 50:1; pretty good but as it may not be normal we'd best not be too certain" . . . and whether one would be content with that or would require further work would *depend on the importance of the conclusion and the difficulty of obtaining suitable experience.*
>
> Gosset to E. S. Pearson, reprinted in Pearson 1939, p. 244; italics added

In 1937 Gosset again wrote to Egon (then editor of *Biometrika*), this time using no uncertain terms:

> Obviously the important thing . . . is to have a low real error, not to have a [statistically] "significant" result at a particular station. The latter ["Student" said] seems to me to be nearly valueless in itself. . . . Experiments at a single station [that is, tests of statistical significance on a single set of data] *are* almost valueless. . . . You want to be able to say not only "We have significant evidence that if farmers [or brewers or whomever] in general do this they will make money by it", but also "we have found it so in nineteen cases out of twenty and we are finding out why it doesn't work in the twentieth." To do that you have to be as sure as possible which is the 20th—your real error must be small
>
> Gosset to E. S. Pearson 1937, in Pearson 1939, p. 244. Emphasis in original.

Statistical "significance"—high *t* values—meant that little to "Student." In fact, few realize that Neyman-Pearson "power"—which to Fisher's consternation estimates a significance-level-adjusted trade-off with Type II error as the experimental result deviates from the null—can also be traced to "Student."[66] "Student" intuited the idea of power in two letters of May, 1926 to Egon Pearson (Pearson 1966: 4-11)[67] Despite *that* analytical leap, the power function, too, was to "Student" just one of dozens of checks to put on the size or opportunity cost of "real error." Statistical significance was in any case "nearly valueless" in itself.

Fisher's 5% Rule

Not so with Fisher. If Karl Pearson was resistant to "Student's" cost considerations, Ronald Fisher was ideologically opposed to them. Sir Ronald's is a non-economic and even—by the 1950s—a disturbingly *anti*-economic approach. Though claiming to teach "Student's" small sample method, Fisher rejected "Student's" natural ingredients, added some unnatural ones, and re-presented Guinnessometrics in a foreign, philosophical tongue. Out of his mash popped a rote method and ton of science. Said Fisher in successive years, 1925 to 1955:

> The value for which P=.05, or 1 in 20, is 1.96 or nearly 2; *it is convenient to take this point as a limit in judging* whether a deviation is to be considered significant or not. Deviations exceeding twice the standard deviation [said Fisher] *are thus formally regarded as significant*
>
> [Fisher was a rhetorical magician and Mendelian: notice how his 5 percent rule evolved in one stroke of his pen from a simple "convenience" to a "formal regard." Look again at the .05 rule and recall Gosset's hop-life regression, at P = .066. Fisher's rule is irrational: it says to disregard Gosset's soft resins, though doing so loses profits and life in the beer]
>
> Fisher 1925a [1941], *Statistical Methods for Research Workers*, p. 42, italics supplied.

It is convenient to draw the line at about the level at which we can say: "*Either there is something in the treatment, or a coincidence has occurred* such as does not occur more than once in twenty trials." . . . Personally, the writer prefers to set a low standard of significance at the 5 per cent point, and *ignore entirely all results which fail to reach this level* [including, then, Gosset's valuable hops and extract regressions].

Fisher 1926b, p. 504, italics supplied.

It is usual and convenient for experimenters to take 5 per cent. as a standard level of significance, in the sense that they are prepared to *ignore all results which fail to reach this standard*, and, by this means, *to eliminate from further discussion* the greater part of the fluctuations which chance causes have introduced into their experimental results

[But an arbitrary rule about sampling error wasn't "usual" to anyone save a budding philosopher. Again, ignoring all results which fail to reach this standard causes untold economic and social waste. Passing a significance test is by itself "almost valueless" is how the brewer expressed it to Egon Pearson in 1937.]

Fisher 1935a [1960], *The Design of Experiments*, p. 13, italics supplied.

A null hypothesis may, indeed, contain arbitrary elements, and in more complicated cases often does so: as, for example, if it should assert that the death-rates of two groups of animals are equal, without specifying what these death-rates actually are. *In such cases it is evidently the equality rather than any particular values of the death-rates that the experiment is designed to test, and possibly to disprove*

["Student" did not employ Fisher's null procedure. He wanted to know what the death-rates actually are (Student 1931a). He focused on estimation, prior probability, alternative hypotheses, questions of "how much" (e.g., Student 1926 [1942], p. 126). As a Bayesian it was natural for him to see the illogic of Fisher's procedure: the probability of the hypothesis, given the data, is not the same as the probability of the data, given the hypothesis. Yet in Fisher's procedure the prior probability is neglected and the identity forced: contrast Jeffreys's discussion of his own Gosset-inspired "porter" test, the posterior odds ratio test of economic relevance: 1961, pp. 378-9; see also Leamer 1978, and Ziliak and McCloskey 2007, references to Jeffreys's *d,* prior probability, and "real error."]

Fisher 1935a, in Savage 1971a, pp. 471-2, italics supplied.

> *Finally, in inductive inference we introduce no cost functions for faulty judgments* [Sir Ronald intoned in the 1950s, despite the cost-conscious "Student"] . . . *In fact, scientific research is not geared to maximize the profits* of any particular organization [such as Guinness?], but is rather an attempt to improve public knowledge *undertaken as an act of faith . . . We make no attempt to evaluate these consequences, and do not assume that they are capable of evaluation in any currency*
>
> [But the philosopher leaves us with a "vague" foundation for econometrics and no real conclusion. Contrast finally and again "Student," at the largest for-profit brewery in the world, in 1905: "the degree of certainty to be aimed at must depend on the *pecuniary advantage to be gained by following the result of the experiment, compared with the increased cost of the new method, if any, and the cost of each experiment."* An aristocratic faith could not understand.]
>
> Fisher 1955a, p. 75, italics supplied.

Reasons for the widespread acceptance of Fisher's 5 percent philosophy are too complex to disentangle in a brief article. (See Ziliak and McCloskey 2007, chps. 20-23.)[68] The fact of acceptance in economics and other areas is obvious.

Conclusion

The difference between "Student" and Fisher seemed to "Student" to be one of opinion. The facts support the view that the difference between them was real and big and canonical. But as Adam Smith long ago suggested, the difference probably did not arise naturally but rather as the product of habit, custom, and education. The mathematical geneticist did not habitually respond to the call of profit but rather to a brand of rule-based philosophy. As Bruno de Finetti once observed in a symposium about the esteemed Fisher corpus, "The economic approach seems (if not rejected owing to aristocratic or puritanic taboos) the only device apt to distinguish neatly what is or is not contradictory in the logic of uncertainty" (de Finetti 1971, pp. 486-7). Here we seem to have a case in point.

It is not obvious what to do with this difference; it depends on one's prior. For example, it is important to remember that not every student you meet, beginning with Ronald Fisher himself, deigns to pass the "porter" test. But if you hand him a test of significance and a pint of Guinness, and ask him to explain the logic and history of each, he will, if he is careful not to stumble, come upon William Sealy Gosset and the economic foundation of "Student's" *t*.[69]

References

Primary Sources
(in-text citations)

Beaven Collection, The Museum of English Rural Life, University of Reading
Galton Papers, University College London (UCL), Special Collections.
Gosset Collection, Guinness Archives, Dublin (Diageo).
Gosset 1962: *Letters of William Sealy Gosset to R. A. Fisher. Vols. 1-5,* Eckhart Library, University of Chicago. Private circulation.
Guinness Archives, Dublin (Diageo)
Especially the fifteen volumes of *Laboratory Reports,* 1898-1912 (abbreviated as *Lab. R.*)
Pearson Papers (Karl Pearson, Egon Pearson, Gosset, Fisher, Neyman files), University College London (UCL), Special Collections.

Secondary Sources

Barbacki, S., and R. A. Fisher. 1936. "A Test of the Supposed Precision of Systematic Arrangements," *Annals of Eugenics* 7: pp. 189-93.
Bennett, J. H., ed. 1990. *Statistical Inference and Analysis: Selected Correspondence of R. A. Fisher.* Oxford: Clarendon Press.
Berger, James O. 2003. "Could Fisher, Jeffreys and Neyman Have Agreed on Testing?" *Statistical Science* 18(1): 1-32.
Boland, Philip J. 1984. "A Biographical Glimpse of William Sealy Gosset." *The American Statistician* 38(3, Aug.): 179-83.
Box, Joan Fisher. 1987. "Guinness, Gosset, Fisher, and Small Samples." *Statistical Science* 2(1, Feb.): 45-52.
Box, Joan Fisher. 1997. "Fisher, Ronald A." Pp. 99-108 in N. L. Johnson and S. Kotz, eds., *Leading Personalities in Statistical Sciences.*
Camerer, C., G. Loewenstein, and M. Rabin, eds. 2003. *Advances in Behavioral Economics*. Princeton: Princeton University Press.
Darnell, Adrian C., ed. 1990. *The Collected Economics Articles of Harold Hotelling.* New York: Springer-Verlag.
Dawkins, Richard. 2006. *The Blind Watchmaker.* New York: W. W. Norton.
De Finetti, Bruno. 1971 [1976]. "Comments on Savage's "On Rereading R. A. Fisher." *Annals of Statistics* 4(3): 486-7.
DeGroot, Morris H. 1989. *Probability and Statistics*. Reading, MA: Addison-Wesley
Dennison, S. R., and Oliver MacDonagh. 1998. *Guinness 1886-1939*. Cork: Cork University Press.
Efron, Bradley. 1998. "R. A. Fisher in the 21st Century." *Statistical Science* 13(2): 95-122.
Elliott, Graham, and Clive W. J. Granger. "Evaluating Significance: Comments on 'Size Matters.'" *Journal of Socio-Economics* 33(5): 547-550
Fienberg, Stephen E., and Arnold Zellner, eds. 1975. *Studies in Bayesian Econometrics: In Honor of Leonard J. Savage*. Amsterdam: North Holland Publishing.

Fisher, R. A. 1915. “Frequency Distribution of the Values of the Correlation Coefficients in Samples from an Indefinitely Large Population.” *Biometrika* 10: 507-21.
Fisher, R. A. 1918. “The correlation between relatives on the supposition of Mendelian inheritance.” *Trans. Roy. Soc. Edinb*. 52: 399-433.
Fisher, R. A. 1921. “On the "probable error" of a coefficient of correlation deduced from a small sample.” *Metron* 1: 3-32.
Fisher, R. A. 1922a. “On the Mathematical Foundations of Theoretical Statistics.” *Philos. Trans. Roy. Soc. London Ser. A* 222: 309-368.
Fisher, R. A. 1922b. “On the Interpretation of χ^2 from Contingency Tables, and the Calculation of P,” *Journal of the Royal Statistical Society* 85 (Jan.): 87-94.
Fisher, R. A. 1923. “Statistical tests of agreement between observation and hypothesis.” *Economica* 3: 139-147.
Fisher, R. A. 1925a [1941]. *Statistical Methods for Research Workers*. New York: G. E. Stechart and Co. Originally published Edinburgh: Oliver and Boyd. Eighth of thirteen editions, in at least seven languages.
Fisher, R. A. 1925b. “Applications of ‘Student’s’ distribution.” *Metron* V(3, Dec.): 90-104.
Fisher, R. A. 1925c. “Expansion of ‘Student’s’ Integral in Powers of n^{-1}.” *Metron* V(3, Dec.): 110-120.
Fisher, R. A. 1926a. “Arrangement of Field Experiments.” *Journal of Ministry of Agriculture* XXXIII: 503-13.
Fisher, R. A. 1926b. “Bayes’ Theorem.” *Eugenics Review* XVIII: 32-33.
Fisher, R. A. 1931. “Letter to Arne Fisher.” Page 313, in J. H. Bennett, ed., 1990.
Fisher, R. A. 1932. “Family Allowances in the Contemporary Economic Situation.” *Eugenics Review* 24: 87-95.
Fisher, R. A. 1935a. *The Design of Experiments.* Edinburgh: Oliver & Boyd. Reprinted in eight editions and at least four languages.
Fisher, R. A. 1935b. “The Mathematical Distributions Used in the Common Tests of Significance,” *Econometrica* 3 (4): 353-65.
Fisher, R. A. 1936. “The significance of regression coefficients” [*abstract*], *Colorado College Publ. Gen. Ser. 208*, pp. 63-67. Cowles Commission Research Conference on Economics and Statistics.
Fisher, R. A. 1939. “’Student’.” *Annals of Eugenics* 9: 1-9.
Fisher, R. A. 1950. *Contributions to Mathematical Statistics*. New York: Wiley. Ed. Walter A. Shewhart.
Fisher, R. A. 1955a. “Statistical Methods and Scientific Induction.” *Journal of the Royal Statistical Society, Series B (Methodological)*, Vol. 17, No. 1, pp. 69-78.
Fisher, R. A. 1955b. Page 6 in G. Mendel’s *Experiments in Plant Hybridisation.* Edinburgh: Oliver & Boyd.
Fisher, R. A. 1956 [1959]. *Statistical Methods and Scientific Inference.* New York: Hafner. Second edition.
Fisher, R. A., and Frank Yates. 1938 [1963]. *Statistical Tables for Biological, Agricultural and Medical Research.* Edinburgh: Oliver and Boyd. Sixth edition.
Friedman, Milton. 1996. “Milton Friedman [interview],” p. 132, in B. Snowdon and H. Vane, *Conversations with Leading Economists* (Edward Elgar, 1999).
Geweke, John. 2005. *Contemporary Bayesian Econometrics and Statistics.* New York: Wiley.

Giere, Ronald N., and Richard S. Westfall, eds. 1973. *Foundations of Scientific Method: The Nineteenth Century.* Bloomington: Indiana University Press.
Gigerenzer, Gerd. 2004. "Mindless Statistics." *Journal of Socio-Economics* 33(5): 587-606.
Gigerenzer, Gerd, Zeno Swijtink, Theodore Porter, Lorraine Daston, John Beatty, and Lorenz Kruger. 1989. *The Empire of Chance*. Cambridge: Cambridge University Press.
Gosset, William S. [see Student]. 1936. "Co-Operation in Large-Scale Experiments." *Supplement to the Journal of the Royal Statistical Society* 3(2): 115-36.
Hacking, Ian. 1990. *The Taming of Chance.* Cambridge: Cambridge University Press.
Heckman, James. 2005. "The Scientific Model of Causality," *Sociological Methodology* 35 (1): 1-98.
Hendry, David, and Mary S. Morgan, eds., 1995. *The Foundations of Econometric Analysis.* Cambridge: Cambridge University Press.
Horowitz, Joel L. 2004. "Comments on 'Size Matters'." *Journal of Socio-Economics* 33(5): 551-554.
Hotelling, Harold. 1927 to 1939. Reviews of Fisher's books in *Journal of American Statistical Association:*
1927 22, pp. 411-12—*SMRW*
1928 23, p. 346 " "
1930 25, pp. 381-2 " "
1933 28, pp. 374-5 " "
1935 30, pp. 771-2— *Design*
1937 32, pp. 218-9— *SMRW*
1937 32, pp. 580-2— *Design*
1939 34, pp. 423-4— *SMRW*
Hotelling, Harold. 1930. "British Statistics and Statisticians Today," *Journal of the American Statistical Association* 25 (170, June): 186-90.
Hotelling, Harold. 1931. "The Generalization of Student's Ratio." *Annals of Mathematical Statistics* 2: 360-78.
Hotelling, Harold. 1938. "Review of Fisher and Yates, *Statistical Tables for Biological, Agricultural and Medical Science*." *Science* 88: 596-7
Hotelling, Harold. 1940 [1960]. "The Teaching of Statistics." Pp. 11-24 in I. Olkin, et al., eds., *Contributions.* First published in *Annals of Mathematical Statistics,* 1940, 11, pp. 457-70.
Hotelling, Harold. 1951. "The Impact of R. A. Fisher on Statistics," *Journal of the American Statistical Association* 46 (243, Mar.): 35-46.
Hotelling, Harold. 1958. "The Statistical Method and the Philosophy of Science," *The American Statistician* 12 (5, Dec.): 9-14.
Howie, David. 2002. *Interpreting Probability: Controversies and Developments in the Early Twentieth Century.* Cambridge: Cambridge University Press.
Hunter, Herbert. 1938. "William Sealy Gosset, 1876-1937." *Journal of the Royal Statistical Society* 101(1): 248-49.
Jeffreys, Harold. 1919. "On the Crucial Test of Einstein's Theory of Gravitation," *Monthly Notices of the Royal Astronomical Society* 80 (December): 138-54.
Jeffreys, Harold. 1931 [1973]. *Scientific Inference.* Cambridge: Cambridge University Press.
Jeffreys, Harold. 1934. "Letter to Ronald Fisher." Page 150, in J. H. Bennett, ed., 1990.

Jeffreys, Harold. 1961. *Theory of Probability*. Third revised edition. London: Oxford University Press.

Jeffreys, Harold. 1939. “Random and Systematic Arrangements.” *Biometrika* 31(1/2, July): 1-8.

Jeffreys, Harold. 1963. Review of L. J. Savage, et al., *The Foundations of Statistical Inference* (Methuen, London and Wiley, New York, 1962), in *Technometrics* 5(3), August 1963, pp. 407-410.

Joyce, James. 1914 [1961]. *Ulysses.* Modern Library Edition.

Kendall, Maurice G. 1978. “The History of Statistical Method,” pp. 1093-1101, in W. Kruskal and J. Tanur, eds., *International Encyclopedia of Statistics, Vol. 2.*

Kruskal, William H. 1968a. “Tests of Statistical Significance." Pp. 238-250, in David Sills, ed., *International Encyclopedia of the Social Sciences* 14. New York: MacMillan.

Kruskal, William H. 1968b. “Statistics: The Field." Pp. 206-224, in David Sills, ed., *International Encyclopedia of the Social Sciences* 15. New York: MacMillan.

Kruskal, William H. 1978. “Formulas, Numbers, Words: Statistics in Prose.” *The American Scholar* 47(2): 223-29.

Kruskal, William H. 1980. “The Significance of Fisher: A Review of R. A. Fisher*: The Life of a Scientist*, *Journal of the American Statistical Association*, Vol. 75, No. 372 (Dec.):1019-30.

Kruskal, William H. 2002. Personal interview. August 16, 2002, University of Chicago.

Kruskal, William H., and Judith M. Tanur. 1978. *International Encyclopedia of Statistics*, *Vols. 1-2.* New York: The Free Press.

Leamer, Edward E. 1978. *Specification Searches: Ad Hoc Inference with Non-Experimental Data*. New York: Wiley.

Leamer, Edward E. 2004. “Are the Roads Red? Comments on ‘Size Matters,’” *Journal of Socio-Economics* 33(5): 555-558.

Lehmann, Erich. 1999. “Student” and Small-Sample Theory,” *Statistical Science* 14 (4): 418-26.

Levy, Daniel, and Andrew T. Young. 2004. “The Real Thing: Nominal Price Rigidity of the Nickel Coke, 1886-1959,” *Journal of Money, Banking, and Credit* 36: 765-99.

Lynch, Patrick, and John Vaizey. 1960. *Guinness's Brewery in the Irish Economy, 1758-1876* (London: Cambridge University Press).

MacDonnell, William R. 1901. “On Criminal Anthropometry and the Identification of Criminals.” *Biometrika*, Vol. 1, No. 2. (Jan., 1902): 177-227.

Mahalanobis, P. C. 1964. “In Memorium: Ronald Aylmer Fisher, 1890-1962.” *Biometrics* 20 (2, June): 368-71.

Mahalanobis, P.C. 1938. “Professor Ronald Aylmer Fisher.” Pp. 365-72, in R. A. Fisher, *Contributions to Mathematical Statistics* (New York: John Wiley & Sons, 1950). Walter A. Shewhart, editor.

Manski, Charles. 1995. *Identification Problems in the Social Sciences.* Cambridge, MA: Harvard University Press.

Matthias, Peter. 1959. *The Brewing Industry in England, 1700-1830*. London: Cambridge University Press.

Mayr, Ernst. 1982. *The Growth of Biological Thought: Diversity, Evolution, and Inheritance.* Cambridge, MA: Belknap and Harvard University Press

Mayo, Deborah G. 1999. *Error and the Growth of Experimental Knowledge.* Chicago: University of Chicago Press.
McCloskey, Deirdre N. 1985a. "The Loss Function Has Been Mislaid: The Rhetoric of Significance Tests." *American Economic Review*, Supplement 75 (2, May): 201-205.
McCloskey, Deirdre N. 1985b [1998]. *The Rhetoric of Economics.* Especially chapters 7 and 8 of the 1985 edition and chapters 8 and 9 of the 1998 edition. Madison: University of Wisconsin Press.
McCloskey, Deirdre N., and Stephen T. Ziliak. 1996. "The Standard Error of Regressions." *Journal of Economic Literature* 34 (March 1996): pp. 97-114.
McCloskey, Deirdre N., and Stephen T. Ziliak. 2008. "Signifying Nothing: Reply to Hoover and Siegler," *Journal of Economic Methodology*, forthcoming.
McMullen, Launce. 1939. "'Student' as a Man." *Biometrika* 30(3/4, Jan.): 205-10.
Moore, David S., and George P. McCabe. 1993. *Introduction to the Practice of Statistics.* New York: Freeman.
Morgan, Mary S. 1990. *The History of Econometric Ideas.* Cambridge: Cambridge University Press.
Neyman, Jerzy, and E. S. Pearson. 1933. "On the problem of the most efficient tests of statistical hypotheses." *Phil. Trans. Roy. Soc., A,* 231, pp. 289-337.
Neyman, Jerzy, and E. S. Pearson. 1938. "Note on Some Points on 'Student's' Paper on 'Comparison Between Balanced and Random Arrangements of Field Plots.'" *Biometrika* 29(3/4, Feb.): 379-88.
Neyman, Jerzy. 1938. "L'estimation statistique traite comme un probleme classique de probabilité." *Actualites Scientifiques et Industrielles,* No. 739, pp. 25-57.
Neyman, Jerzy. 1956. "Note on an Article by Sir Ronald Fisher," *Journal of the Royal Statistical Society. Series B (Methodological)* 18 (2):288-94.
Neyman, Jerzy. 1957. "'Inductive Behavior' as a Basic Concept of Philosophy of Science." *Review of the Mathematical Statistics Institute* 25: 7-22.
Neyman, Jerzy. 1960. "Harold Hotelling—A Leader in Mathematical Statistics." Pp. 6-10 in I. Olkin, et al., eds., *Contributions.*
Neyman, Jerzy. 1961. "Silver Jubilee of my Dispute with Fisher," *Journal of Operations Research* (Society of Japan) 3: 145-54.
Neyman, Jerzy. 1981. "Egon S. Pearson (August 11, 1895-June 12, 1980): An Appreciation," *Annals of Statistics* 9 (1, Jan.): 1-2.
O'Grada, Cormac. 1995. *Ireland: A New Economic History* (Oxford: Oxford University Press).
Pearson, Egon S. 1936. "Karl Pearson: An Appreciation of His Life and Work, Part I: 1857-1906." *Biometrika 28* (3/4, Feb.): 193-257.
Pearson, Egon S. 1938. "Karl Pearson: An Appreciation of His Life and Work, Part II: 1906-1936." *Biometrika 29* (3/4, Feb.): 161-248.
Pearson, Egon S. 1939. "'Student' as Statistician." *Biometrika* 30 (3/4, Jan.): 210-50.
Pearson, Egon S. 1966. "The Neyman-Pearson Story: 1926-34." Pp. 1-24 in F. N. David, ed., *Research Papers in Statistics: Festschrift for J. Neyman* (1966).
Pearson, Egon S. 1968. "Studies in the History of Probability and Statistics. XX: Some Early Correspondence Between W. S. Gosset, R. A. Fisher, and Karl Pearson, with Notes and Comments." *Biometrika 55* (3, Nov.): 445-57.

Pearson, Egon S. 1990 [posthumous]. *'Student': A Statistical Biography of William Sealy Gosset.* Oxford: Clarendon Press. Edited and augmented by R. L. Plackett, with the assistance of G. A. Barnard.

Pearson, Egon S., and N. K. Adyanthaya. 1929. "The Distribution of Frequency Constants in Small Samples from Non-Normal Symmetrical and Skew Populations." *Biometrika 21*(1/4, Dec.): 259-86.

Pearson, Egon S., and M. C. Kendall, eds. 1970. *Studies in the History of Statistics and Probability.* London: Charles Griffin.

Pearson, Karl, Francis Galton, William Weldon, and Charles Davenport. 1901. "Editorial: The Scope of Biometrika." *Biometrika* 1(1): 1-2.

Pearson, Karl. 1892 (1951). *The Grammar of Science.* London: J. W. Arrowsmith.

Pearson, Karl. 1914. *Tables for Statisticians and Biometricians, Part I* (London: Biometrics Laboratory, University College London).

Pearson, Karl. 1931. "Further Remarks on the '*z'* Test." *Biometrika* 23(3/4, Dec.): 408-15.

Pesaran, M. Hashem. 1990. "Econometrics," pp. 1-34, in J. Eatwell, M. Milgate, and P. Newman, eds., *The New Palgrave Encyclopedia of Econometrics* (New York: W.W. Norton).

Popper, Karl R. 1934 [1959]. *The Logic of Scientific Discovery.* London and New York: Routledge.

Porter, Theodore M. 2004. *Karl Pearson: The Scientific Life in a Statistical Age.* Princeton: Princeton University Press.

Press, S. James. 2003. *Subjective and Objective Bayesian Statistics.* New York: Wiley.

Reid, Constance. 1982 [1998]. *Neyman: A Life.* New York: Springer.

Savage, Leonard J. 1954 [1972]. *The Foundations of Statistics.* New York: John Wiley & Sons [Dover edition].

Savage, Leonard J. 1971a [1976 posthumous]. "On Re-Reading R. A. Fisher." *Annals of Statistics* 4(3): 441-500.

Savage, Leonard J. 1971b [1974 posthumous]. "Elicitation of Personal Probabilities." Pp. 111-56 in Fienberg and Zellner, eds., *Studies in Bayesian Econometrics.*

Senn, Stephen, and William Richardson. 1994. "The First *t*-Test." *Statistics in Medicine* 13: 785-803.

Smith, Adam. 1776. *The Wealth of Nations.* 1904 Edward Canaan edition. www.econlib.org.

Spanos, Aris. 1999. *Probability Theory and Statistical Inference*. Cambridge: Cambridge University Press.

Steward, Dwight, and S. W. O'Donnell. 2005. "Evaluating the Statistical and Economic Significance of Statistical Evidence in Employment Discrimination Cases." *Expert Evidence Report* 5(5, March): 117-. (http://www.bna.com/)

Stigler, Stephen. 1986. *The History of Statistics. Cam*bridge: Harvard University Press.

Stigler, Stephen. 1996. "The History of Statistics in 1933." *Statistical Science* 11: 244-52.

Stigler, Stephen. 1999. *Statistics on the Table.* London: Harvard University Press.

Stigler, Stephen. 2000. "The Problematic Unity of Biometrics." *Biometrics* 56 (3): 272-7.

Stigler, Stephen. 2005a. "Fisher in 1921." *Statistical Science* 20 (1): 32-49.

Stigler, Stephen. 2006. "Milton Friedman (1912-2006)," http://www.chicagoasa.org/

Student [W. S. Gosset]. 1907. "On the Error of Counting with a Haemacytometer." *Biometrika* 5(3, Feb.): 351-60.

Student. 1908a. "The Probable Error of a Mean." *Biometrika* VI (1, March): pp. 1-24.
Student. 1908b. "The Probable Error of a Correlation Coefficient." *Biometrika* (2/3, Oct.): pp. 300-310.
Student. 1911. "Appendix to Mercer and Hall's paper on 'The Experimental Error of Field Trials,'" *Journal of Agricultural Science* IV: 128-131.
Student 1914. (See Pearson 1914.)
Student. 1917. "Tables for Estimating the Probability that the Mean of a Unique Sample of Observations Lies Between - ∞ and Any Given Distance of the Mean of the Population from which the Sample is Drawn." *Biometrika* 11(4, May): 414-17.
Student. 1923. "On Testing Varieties of Cereals." *Biometrika* 15 (3/4, Dec.): 271-293.
Student. 1925. "New Tables for Testing the Significance of Observations." *Metron* V(3, Dec.): 105-108.
Student. 1926 [1942]. "Mathematics and Agronomy," *Journal of the American Society of Agronomy* 18; reprinted: Pp. 121-34 in E. S. Pearson and J. Wishart, eds., *Student's Collected Papers* (1942).
Student. 1927. "Errors of Routine Analysis." *Biometrika* 19(1/2, July): 151-64.
Student. 1931a. "The Lanarkshire Milk Experiment." *Biometrika* 23(3/4, Dec.): 398-406.
Student. 1931b. "On the '*z'* Test." *Biometrika* 23(3/4, Dec.): 407-8
Student. 1938 [posthumous]. "Comparison between Balanced and Random Arrangements of Field Plots." *Biometrika* 29 (3/4, Feb.): 363-78.
Student. 1942 [posthumous]. Student's *Collected Papers* (London: Biometrika Office). Eds. E. S. Pearson and John Wishart.
Wooldridge, Jeffrey M. 2004. "Statistical Significance is Okay, Too: Comment on 'Size Matters,'" *Journal of Socio-Economics* 33(5): 577-80.
Yates, Frank. 1951. "The Influence of *Statistical Methods for Research Workers* on the Development of the Science of Statistics," *Journal of the American Statistical Association* 46**:** 19-34.
Yates, Frank, and Kenneth Mather. 1963. "Ronald Aylmer Fisher," *Biograph. Mem. Fell. R. Soc. Lond.* 9: 91–129.
Yeats, W. B. "Gratitude to the Unknown Instructors," c. 1931, page 249 in *The Collected Poems of Yeats* (MacMillan, 1933 [1979])
Zabell, Sandy. 1989. "R. A. Fisher on the History of Inverse Probability." *Statistical Science* 4(3): 247-63.
Zellner, Arnold, ed. 1980. *Bayesian Analysis in Econometrics and Statistics: Essays in Honor of Harold Jeffreys.* North-Holland Publishing.
Zellner, Arnold. 1984. *Basic Issues in Econometrics*. Chicago: University of Chicago Press.
Zellner, Arnold. 1997. *Bayesian Analysis in Econometrics and Statistics.* Cheltenham, UK: Edward Elgar.
Zellner, Arnold. 2004. "To Test or Not to Test and If So, How?: Comments on 'Size Matters,'" *Journal of Socio-Economics* 33(5): 581-86.
Zellner, Arnold. 2005a. "Some Thoughts about S. James Press and Bayesian Analysis." Prepared for presentation at a Conference in Honor of Professor S. James Press, University of California at Riverside, May 14, 2005.
Zellner, Arnold. 2005b. Personal communication, University of Chicago, Oct. 20, 2005.

Ziliak, Stephen T. 2004. "Self-Reliance before the Welfare State: Evidence from the Charity Organization Movement," *Journal of Economic History* 64 (2): 433-61.
Ziliak, Stephen T. 2007. *"Student's" Methods: Science Before Fisher.* Book in progress. Roosevelt University, Department of Economics.
Ziliak, Stephen T., and Deirdre N. McCloskey. 2004a. "Size Matters: The Standard Error of Regressions in the *American Economic Review*." *Journal of Socio-Economics* 33(5): 527-46.
Ziliak, Stephen T., and Deirdre N. McCloskey. 2004b. "Significance Redux." *Journal of Socio-Economics* 33(5): 665-75.
Ziliak, Stephen T., and Deirdre N. McCloskey. 2007. *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives* (Ann Arbor: University of Michigan Press, 2007).

1 The author is Professor of Economics at Roosevelt University. *Acknowledgments*: In the archives I've been ably assisted by Eibhlin Roche and Clare Hackett, at Guinness (Diageo), Dublin, St. James's Gate; by Gillian Furlong and Steven Wright, at University College London, Special Collections Library; and by staff at the Eckhart Library, University of Chicago, and the Museum of English Rural Life, University of Reading. For stimulating conversations I've many colleagues and venues to thank but especially John Aldrich, Deirdre McCloskey, Arnold Zellner, Jim Ziliak, and faculty at the 2006 holiday party for the University College Dublin, Department of Economics. A seminar at the National University of Ireland, Galway, Department of Economics (August 2007), was particularly helpful and inspiriting. For travel assistance I thank Roosevelt University. Any errors are my own.

2 A major figure of science, Harold Jeffreys (1891-1989) was Plumian Professor of Astronomy and Experimental Philosophy, and Reader in Geophysics, at Cambridge University. Jeffreys participated in the biggest controversies of his day, including Einstein's theory of relativity, quantum theory, continental drift, cosmogony, and the philosophy of inference. His own "Studentized" tables of earthquake travel times (the 70 year old "JB tables") are standard in seismology. Arnold Zellner reminds us in his wonderful way that Jeffreys (*Theory of Probability* 1961) whipped up a Bayesian economic alternative to the Bayesian "Student's" test, the *p*osterior *o*dds *r*atio *t*est of *e*conomic *r*elevance, or "porter" test (Zellner 1980, ed., 1984, 1997, 2004, 2005a, 2005b; Ziliak and McCloskey 2007, chps. 1, 20). It was a conscious improvement over *"Student's"* test, said Jeffreys, who also greatly admired what he knew of the porter brewer's designs of experiments (1961: 29-31, 269-77; Jeffreys 1939; e.g., Student 1923, Student 1938). Like "Student" before him, Jeffreys defended his improvement against comers, including Fisher and Karl Popper (e.g, Letters between Jeffreys and Fisher, Feb. 24, 1934, p. 150, and elsewhere, in J. H. Bennett, ed., 1990; Lindley 1990, 1991, Savage 1971). Jeffreys, then, not Fisher, would appear to be the natural successor to "Student." The Plumian Professor himself all but said so in 1961: especially p. 379, 393, 369-400, *passim*. See also Zellner 1984, p. 277n, and Howie 2002, chps. 3-5.

3 Samuelson's *Foundations of Economic Analysis* (1947) looks like a flash in the pan when compared to *Statistical Methods for Research Workers,* a blockbuster book in every science

(e.g., Yates 1951, Yates and Mather 1963, Ziliak and McCloskey 2007). The themes were anticipated by Fisher 1915, 1918, 1921, and especially Fisher 1922a. See, e.g., Stigler 2005a.

4 Ronald A. Fisher (1890-1962) was employed as a mathematical statistician and geneticist at Rothamsted Agricultural Station when he wrote his influential methods books (Fisher references). In 1933 he replaced the retiring Karl Pearson as Galton Professor of National Eugenics, University College London, before acceding in the 1940s to the Balfour Chair of Genetics, Cambridge, and, finally, in the decade of the 1950s, to President of Gonville and Caius College, Cambridge, where Milton Friedman—a Bayesian critic of statistical significance—"got to know Ronald Fisher quite well" (Stigler 2006, p. 2; Friedman 1996, p. 132). It wasn't by chance. Friedman's mentor from Minnesota, the great economist and teetotaler, Harold Hotelling (1895-1973), was Fisher's Saint Paul (Hotelling references; A. C. Darnell 1990, ed.). Hotelling—who had worked with Fisher at Rothamsted (Box 1978, p. 139)—was later joined by Koopmans, Haavelmo, and Company to engineer at the Cowles Commission a "probabilistic consensus," as it's been called, distributed point-wise around Fisher's statistical values (D. Hendry and M. Morgan, eds., 1995, §IV, VII, VIII; Ziliak and McCloskey 2007, chp. 10; Hotelling 1930, 1951). Fisher's values are now being seriously challenged: e.g., Pesaran (1990, pp. 24-6); Manski (1995), Zellner (1997), Spanos (1999, pp. 681-728), Mayo (1999), Berger (2003), Camerer, Loewenstein, and Rabin, eds.(2003), Leamer (2004), Geweke (2005), Heckman (2005).

5 For example, in *Design,* p. 45, Fisher says: "It is, indeed, demonstrable that, as a test of this hypothesis, the exactitude of 'Student's' *t* test is absolute." Contrast Jeffreys, speaking of Fisher's reengineering of the Bayesian "Student's" test: "the test is not independent of previous knowledge, as Fisher is liable to say in other places" (Jeffreys 1961 [1939], p. 382). As Fisher says again, in *Design:* "In the field of pure research [his uniformly least powerful straw man] *no assessment of the cost of wrong conclusions can conceivably be more than a pretence*, and in any case such an assessment would be *inadmissible* and *irrelevant* in judging the state of scientific [read: economic] evidence" (pp. 25-6; italics added). Fisher is very clear: "minimizing losses" (p. 25)—in for example, the production of beer or inflation-control—is *not* admissible or relevant to science. "Prior information" (p. 25, 26), moreover, he adds, is "known to be lacking," by which he means "ignore entirely Bayes's rule." Fisher did not learn these prior attitudes from "Student," a Bayesian brewer and profit-seeker. By claiming in no uncertain terms that "Student's" test is "appropriate" (p. 44) for Fisher's own null and 5% significance rule (and none others), Fisher greatly confused matters. He falsely equated the methods of the two very different men, despite the friendship and free tuition (p. 34n).

6 Fisher's philosophy of *t* was thinly veiled attack. Fisher was in 1956 trying to thwart the intellectual growth and careers of decision theorists and Neo-Bayesians—that is, Fisher campaigned against the actual students of Gosset. Contrast Neyman 1956, 1961, Pearson 1966, and Savage 1971a, 1971b.

7 A young Fisher introduced himself to Gosset in a letter in 1912, when Fisher, a student of astronomy and biology as well as of mathematics at Cambridge, inquired about a "discrepancy" (Box 1978, p. 72) Fisher found in Gosset's 1908 formula (Box 1978, pp. 70-3; cf. E. S. Pearson,

ed., 1990, pp. 44-7). “Degrees of freedom,” as John Aldrich reminds me, was not yet clear to Fisher (cf. Fisher 1922a, 1922b). But as Box observed, “Fisher showed his recognition, even at this early date, of the concept of degrees of freedom” (p. 83). Gosset was grateful for a second letter and a mathematical proof from Fisher, and asked Karl Pearson to publish Fisher’s paper in *Biometrika* (Pearson did—it wasn’t 1922 yet), and a friendship between Gosset and Fisher developed (Fisher 1915). Over 150 letters survive, mostly from Gosset to Fisher (Gosset 1962). Richard Dawkins has called Fisher “the greatest of Darwin’s successors.” Regardless of that plausible assessment (cf. Mayr 1982), Fisher is not necessarily the greatest of *“Student’s”* successors. Joan Fisher Box (1978), Fisher’s historian daughter, is helpful on the private side of Fisher. But on the scientific sides of originality and authenticity, especially as concerns his relations with “Student,” see Kruskal’s 1980 review of Box (1978). Also see: Ziliak and McCloskey 2007, chps. 20-23, and Ziliak 2007.

8 For example, in D. Moore and G. McCabe (1998, p. 505), in S. J. Press (2003, p. 218), and in virtually every book and article on statistics after R. A. Fisher (1925c, 1925a, p. 19; Ziliak and McCloskey 2007, chps. 21-22). Most methods books give no space at all to “Student” (*ibid,* chp. 22). A century on, the brewer’s actual achievement is dimly understood.

9 Hotelling (1930, p. 189) had to laugh more than two decades after “A Probable Error of a Mean” was published: “I have heard guesses in this country,” he said, “identifying ‘Student’ with Egon S. Pearson and the Prince of Wales.” But Hotelling, who was Fisher’s chief American disciple, and personally acquainted with “Student” himself (Hotelling 1930, 1951, 1958), did not inquire into significant differences between Fisher and “Student.”

10 See, for instance, Hotelling (1930, 1931), and again, Hotelling’s effusive reviews of Fisher’s books (Hotelling 1927-1939). Also see K. R. Popper 1959, p. 330, 388, and Jeffreys 1961, p. 388, 393-5. Jimmie Savage reflected late in life: “My statistical mentors, Milton Friedman and W. Allen Wallis, held that Fisher's *Statistical Methods for Research Workers* (1925) was the serious man's introduction to statistics. They shared that idea with their own admired teacher, Harold Hotelling. They and some others, though of course not all, gave the same advice: "To become a statistician, practice statistics and mull Fisher over with patience, respect, and skepticism” (Savage 1971a, pp. 441-42; Savage 1954). Zellner (1980, p. 4) gives the alternative hypothesis. Quoting I. J. Good, he holds: “Jeffreys’s book *Theory of Probability* is . . . of greater practical importance . . . than nearly all the books on probability written by professional philosophers *lumped together*.”

11 Against the comparative obscurity of “Student” stands a towering mythology about Fisher, originating largely, however strange, in the hundreds of articles and books by Fisher himself. Maurice Kendall’s (1978) entry in the *International Encyclopedia of Statistics* is typical. But compare Yates and Mather (1963, p. 93, 113), and especially E.S. Pearson (1939), J. Neyman (1956, 1961), L. J. Savage (1971), W. Kruskal (1980), S. J. Press (2005, pp. 218-20) and, again, H. Jeffreys (1961, p. 388-94). All suggest a significant Oedipal bias against “Student.”

12 Savage's 1971 "On Re-Reading R. A. Fisher" was the subject of an old symposium in *Annals of Science* (1976 posthumous). His sympathy for Fisher was equaled by his anguish and outrage, and echoed by other participants, such as the great Italian statistician and economist, Bruno de Finetti (1906-1985). Surprisingly, neither Savage nor Kendall knew about the "Student"-Fisher relationship (Kendall 1978, pp. 1096-8).

13 Fisher (1918, 1922a, 1922b) were the big blows. The bruised Goliath returned the favors by offering Fisher a job while refusing to publish his articles (Porter 2005, p. 8, 251, 256-7, 275-78).

14 The words are Fisher's: 1925a, 1925b, 1935a, 1939, *passim*. The factual claim is established in Zellner 1984, pp. 277-80, McCloskey and Ziliak 1996, and Ziliak and McCloskey 2004a, 2004b, 2007, 2008. Also see the comments on Ziliak and McCloskey (2004a) by Elliott, Gigerenzer, Granger, Horowitz, Leamer, Thorbecke, Wooldridge, Zellner, and others in a *Journal of Socio-Economics* symposium (vol. 5 (3), 2004). Conclusion: we're all Bayesians now, too, a contradiction that cannot be sustained.

15 Ziliak and McCloskey 2007, chp. 10; McCloskey and Ziliak 1996.

16 As Zellner said, "The rationale for the 5% 'accept-reject syndrome' which afflicts econometrics and other areas requires immediate attention" (1984, p. 277). Compare James O. Berger, in his 2001 Fisher Lecture (2003, p. 4): "The harm from the common misinterpretation of $p = 0.05$ as an error probability is apparent." Apparently the harm is not apparent enough. "Student's" successor, Harold Jeffreys, was again on point: 1961, pp. 388-9.

17 Steward and O'Donnell (2005). For instance, Castenda v. Partida, 430 U.S. 482 (1977), concerning jury discrimination. The Court held, "as a general rule for such large samples [note], if the difference between the expected value and the observed number is greater than two or three standard deviations, then the [null] hypothesis would be suspect to a social scientist, 430 U.S. at 496 n.17."

18 An exception is Barbacki and Fisher (1939). But the "difference" Fisher acknowledged was not about "Student's"*t,* on which he built a career. It was about "Student's" *balanced* or *economic* design of experiments versus his own error-prone random blocks, another career. "Student" was not impressed by the one exception (Gosset 1936, Student 1938). Seeking to minimize economic loss, "Student" rejected randomization in design back in 1908.

19 Note: whether the difference is statistically significant is not in question. To "Student," who was to repeat a life-long Bayesian, the *existence* of a difference is already announced by the *prior* probability (see, e.g., Gosset 1909 *Lab. R.*, pp. 205-6; Student 1908b [1942], pp. 35-6; E. S. Pearson 1990, *passim*; cf. Jeffreys 1961, p. 389: "It is already known that varieties habitually differ and that treatments have different effects, and the problem is to decide which is the best"). The independent "Student" told the great Karl Pearson: "if I didn't fear to waste your time I'd fight you on the *a priori* probability and give you choice of weapons! But I don't think the move is with me; I put my case on paper last time I wrote and doubt I've much to add to it" (Student

1915, quoted in Pearson 1990, pp. 26-7; cf. Letter no. 5, Gosset to Fisher, April 3, 1922, in Gosset 1962). Few statisticians have noticed "Student's" preference for the Bayesian approach, another by-product, historians agree, of the Fisher monopoly (e.g., S. Zabell 1989). By contrast, Harold Jeffreys, Karl Pearson, Egon Pearson, Arnold Zellner, Dennis Lindley, and James Press—not to mention R. A. Fisher himself—definitely did notice the Bayesian start to small sample estimation. The point is that Fisher would not say so in his allegedly "Studentized" methods books.

20 Guinness Archives GDB/C004.06/0001.04 (File: "William Sealy Gosset, Memoranda and correspondence regarding Gosset's recruitment to Junior Brewer").

21 "Loony" is a trait he never outgrew, or had a need to. Gosset loved animals. And in a piece on "routine errors," published in Karl Pearson's journal, he offered readers "unfamiliar with the term 'kurtosis'" a "*memoria technica*" he himself uses to "bear [kurtosis] in mind" (Student 1927 [1942], p. 145n). "Bear," he said. His *memoria technica* is a caricature drawing of a representative "platypus" (body squat, center wide, tails stubby). Gosset's nearly uniformly distributed platypus hovers in space above two *more* creatures, identical "kangaroos," standing up kissing (bodies tall, centers svelt, tails long and thin). You get, distribution of moments-wise, what the loon tries to convey: "The important property," he said, "is that platykurtic curves have shorter 'tails' than the normal curve of error and leptokurtic longer 'tails' (kangaroos being "noted for 'lepping'" [he was forced to add for metaphoric continuity], "though, perhaps, with equal reason [he added to play] they should be hares!" (p. 145n). Gosset's loony sense of humor had its moments you "toucan" see.

22 Contrast the turbulent life and character of Ronald Fisher. The Harrow School standout and freakishly talented geometer was left "motherless" by the age of 14 (Joan Fisher Box 1978, p. 14). His father George did not fill the void. George Fisher's skills at business gambles were not equal to his taste, bankrupting him and Ronald's youth. Ronald as a young adult continued to languish in an emotional and economic chokehold. He was helped in the late 1910s by the charity of Leonard Darwin (eugenical leader, son of Charles) and finally by a research job at Rothamsted, which Gosset had recommended him for. His persistently difficult personality and mathematical cum theological standards of proof in science brought turbulent relations with colleagues and these three traits are known to be highly correlated in Fisher: e.g., Yates and Mather 1963; Neyman 1956, 1961; Savage 1971; Box 1978, Chp. 1, *passim*; Reid 1982, *passim*; Kruskal 1980; McCloskey 1985 [1998]; Ziliak and McCloskey 2007, chps. 21-22. "I would rather give my beard a Messianic cut," Fisher told a Danish actuary, "if this would encourage [Walter] Shewhart and [Harold] Hotelling to go ahead and do their job in a workmanlike manner" (Letter of R. A. Fisher to Arne Fisher, May 1931, in Bennett, ed., p. 313.) On alternative philosophies of inference—including "Student's"—Fisher was "obtuse" (Kruskal 1980 p. 1021), "obsessed" as he was "with ambition and personal bitternesses" (p. 1029). Unlike a humble and tolerant "Student."

23