

Measuring Poverty: the Question of Standardization

Robert T. Michael
Harris School, University of Chicago

January 2010

“I try not to speak more clearly than I think.”

Niels Bohr (Rhodes 77)

My charge is to discuss the measurement of poverty as an example of the advantages or disadvantages of standardization of a scientific construct. Poverty measurement offers an interesting case-in-point for a conversation about standardization of social science measurement since it is one specific area in which we have nearly a half-century of experience living with a standardized, officially sanctioned measure. We have no better scientists than Sargent Shriver and Richard Nixon to thank for that experience, however. Declaring a “war on ...” something like poverty implied, in the mid-sixties, a need for a body count, so Shriver looked around for a candidate measure of poverty and discovered Mollie Orshansky’s essay in the *Social Security Bulletin* in which she had constructed a measure of poverty for her purposes. He declared it a fitting metric by which to measure success in his war. In 1969 Nixon’s administration made it official – the poverty rate that Orshansky had devised and OEO had used would be “the” official measure of poverty in the US and it would help determine eligibility for several social welfare programs, thus it became both “official” and quite important to the flow of federal dollars to states and to certain of its citizens. Thus, we have lived with a standard measure of the undeniably important social condition of economic impoverishment for quite some time.

Before reflecting on the lessons I think we should learn from that experience, several preliminaries should be addressed. I'll begin with a little primer on how to measure poverty. I'll avoid most of the controversy at this point and lay this out in a way that will be useful as we consider what is and what isn't a role for science in its measurement as well as the role of standardization in the advancement of science. I'll then reflect on the purposes of such a measure and suggest that the answer to the question whether it is wise to have a standardized measure depends on which purpose one has in mind. In that discussion of purposes, I'll discuss a view of how science works since it informs my judgment about the wisdom of standardization of this or most other metrics.

How to measure poverty

Unlike some social science concepts and most of today's exciting physical and life science concepts, the notion of poverty predates the beginnings of the sciences that study it. There has been discussion of poverty since at least the time of the writing of the bible ("the poor will always be with us"). From the very beginning of economic science (assuming the science began with Adam Smith), economists have been discussing how to measure poverty. Smith (1776) framed the issue in the manner almost everyone since his time has done it: in terms of some threshold level of resources the absence of which constitutes impoverishment. He talked about having "necessaries," defined as "not only the commodities which are indispensably necessary for the support of life, but whatever the custom of the country renders it indecent for creditable people, even of the lowest order, to be without." Bringing that same idea forward two hundred years, Peter Townsend (1979) similarly defined economic poverty as "lack of

sufficient income for people to play the roles, participate in the relationships, and follow the customary behavior which is expected of them by virtue of their membership in society.”

These two definitions fall squarely on the side of one of the controversies in measuring poverty, suggesting that it should be defined in “relative” terms not in terms of some “absolute” notion of what a person needs. To emphasize how important that one (controversial) distinction is, consider the work of Seebohm Rowntree, one of the pioneers of poverty research. He attempted to measure poverty in York, England in 1899, and again in 1936 (Rowntree, 1901, 1941). When Rowntree did his study in York in 1936 he wanted to address the question whether poverty in York had increased or decreased over that one-third of a century. One of his comparisons applied the same poverty standard (one he called a “bare subsistence” standard), adjusted only for the movement of prices over that time: York’s poverty rate in 1899 was 9.9%, and by 1936 it had fallen to 3.9%: a big reduction in poverty. But he also used a different definition, a “relative” concept measuring “what human needs are for a minimum necessary for a healthy life,” and by that standard York’s poverty rate rose to an estimated 17.7% of the York population: quite a big increase in poverty. Sadly, but not surprisingly, the definition matters greatly in terms of what we conclude about the level of poverty, its change over time, and about its cause, and thus its alleviation.

Rowntree’s helper, Chapman (1951 referenced in Atkinson, 1983, p.41) explained why the “healthy life” standard is better than the bare subsistence standard for 1936: “the second-hand boots and jacket which were adequate in 1899 would not have been entertained by the worker of 1935; a bath and a garden, absent in almost every working class dwelling in 1899, were by now considered essential.” That is, argued Chapman, poverty is set in reference to the

conditions of the population at the time. (If you wonder about it, the York England poverty threshold was set at £2 per week in 1935 (for a couple with three children) compared to £1.45 per week in 1899.)¹

From Smith, through Rowntree, and on to the official US poverty measure and most all notions of measuring poverty, there are five steps that must be taken in its measurement. *First, choose a concept of poverty.* Economic deprivation is most common but some have argued for other concepts. Amartya Sen, for example, argues “poverty is better seen in terms of capability failure than in terms of the failure to meet the ‘basic needs’ of specified commodities.” (1973) In the 1970s the French promoted the concept of social exclusion that has become widely used in Europe. Bob Haveman argued similarly when he said, “In essence, being incapable of independently securing sufficient income to meet basic needs may reflect a more debilitating and vulnerable situation—and a situation reflecting more social exclusion—than being short of cash income in a particular year, living currently in substandard housing, or temporarily living at a consumption level below a minimally acceptable standard.” So choosing a concept is an important and defining first step. A respectful characterization of the US’s official poverty measure describes it as based on dietary food needs, as determined by the U.S. Department of Agriculture’s “economy food plan” from the 1950s, expanded to cover other items and updated as prices have changed.

Second, select a unit of analysis. There are three alternatives from which to choose. Nearly all scholars agree that one is analytically best but infeasible and so the practical choice rests between the other two. The individual is the best unit since utility and well-being are characteristics of each individual. But there isn’t a workable strategy for determining the

resources available at the *individual* level since so many of us live in families that jointly acquire and consume many goods. Who within the family consumed the refrigerator and the electricity last month? How about the expenditure for that new safety alarm on the front door? Which member of the family got that tax cut or consumed the EITC payment? We cannot tell, even in our own households. There is just too much joint consumption within the household. That leaves us, for implementing and measuring poverty empirically, choosing either the family or the household. (The family is a unit connected by blood or contract; the household is all who live under one roof and pool their expenditures.) Most poverty measures use the family as the unit of analysis with unrelated individuals treated as separate one-person units.

Third, determine the poverty threshold level (T) for a numeraire group and decide how to adjust that level across units, time and location. Essentially, one must make a decision about an appropriate cut-off level below which the unit is “in poverty” and above which it is “not in poverty.” Rowntree thought it was £2 per week for a couple and their three children in York England in 1935; Mollie Orshansky thought it was \$3,100 for a family of four in the U.S. in 1963 (Orshansky, 1965) and her judgment guides the official U.S. poverty figures today. There are, of course, a lot of ideas how one should determine this base threshold, but in the final analysis it is necessary to set some figure of something for someone, hopefully closely linked to the concept of poverty chosen in step #1.

That’s not the only decision required in step #3, however. There are three others. There’s the calibration of what that equivalent level is for all the other types or sizes of units – so what is the equivalent level for a family with one fewer child or with only one adult? This is the issue of an equivalence scale that converts that numeraire threshold to its appropriate level for all

other type and size units. Once decided, this equivalence scale is typically unchanging over time or location.

There's a second decision about how to up-date that threshold over time: prices change, consumption bundles change, products and social norms change, so a decision is needed regarding if and how that key threshold level that defines the margin between poverty and not poverty changes with the passing of time. An "absolute" poverty measure doesn't allow any change in the content of the necessary bundle of goods consumed but typically does allow for changes in the price of that bundle – a "cost of living" or "inflation" adjustment, usually made annually. A "relative" poverty measure attempts to keep the social level fixed but allows the magnitude of the level to vary as incomes rise or fall, as customs change and (as Smith says it) as "whatever the custom of the country renders it indecent for creditable people, even of the lowest order, to be without" changes over time.

And third, there is the issue of location and whether the poverty threshold should be modified for rural or urban living, for hotter or colder climates, for more or less expensive locations generally. Here, the chosen concept from step #1 offers some guidance: if the notion of poverty is strictly one of adequate diet, and if food prices don't vary much across location for example, then there's little reason to have a geographic cost of living adjustment. On the other hand, if the concept includes housing and its price does vary widely from location to location and those differences remain somewhat stable over time, then logically one should adjust for differences in the cost of housing in setting the threshold, say for mid-Manhattan compared to rural Mississippi.

Fourth, determine what resources to include (R). Even if the concept of poverty is economic deprivation, one must yet decide if poverty is to be calibrated in terms of actual “consumption” or in terms of more easily measured “expenditures” or in terms of the “income” that ultimately constrains those expenditures. Of course, the selection of a poverty threshold and the resources available to meet that threshold should be expressed in the same units or terms (even though the official U.S. poverty measure fails to satisfy that obvious condition since its threshold was determined as an expenditure level while the resources are measured as before tax income). From the standpoint of economic analysis, “consumption” is closest to the notion of utility or well-being that motivates the interest in economic deprivation. Consumption, however, is most difficult to measure with any accuracy so the practical choices typically are restricted to expenditures or to income per period of time.

Practically, in the US, expenditures (that have much to recommend it as the measure) are captured by the Consumer Expenditure Surveys while income is collected in the much larger Current Population Surveys, so the choice often hinges on whether one wants to be able to make estimates of poverty at smaller geographic levels than the whole nation. If so, most choose to go with income, measured in the CPS. Again, the threshold and resources should be in the same units: if the income measure is to be before-tax-and-transfer income, then the threshold as well should allow for whatever taxes or savings or other uses of income one deems “necessary.” Alternatively, if the concept and threshold reflect exclusively the necessities of food, clothing and shelter, then after-tax and after-transfer income is the right measure of the relevant resources.

Fifth, and most simply, for each unit compare T and R and if the former is larger in size, that unit is “in poverty.” If R is larger than T, that unit is “not in poverty.” Then, of course, there are the issues of aggregation of units into a proportion of the population “in poverty” or if the ratio of R to T is featured, one can calculate the depth of poverty for an impoverished unit and from this one can determine the amount of R that would be required to move each unit up to its poverty threshold (the sum of which is the “poverty gap”). Here as well science can provide guidance about the methods of aggregation but distributional judgments also inform that choice.

The role of science in the measurement of poverty

Notice that I’ve framed the issue backwards from the standpoint of our workshop: I am asking how science helps us measure poverty not whether a standardized measure of the concept of poverty might help science. The distinction goes to the heart of my central point, but that will come later. Science plays a role; it should guide but it does not dictate the choices made at each step along the way toward measuring poverty. In concept, surely, our scientific understanding of the basis of welfare or well-being guides that choice, but as the discrepant quotations above from perfectly good scientists suggest, there is no right answer and there are several good, scientifically acceptable answers. That is true as well regarding the choice of the unit of analysis, and also regarding the level of the threshold and choice of the type of resources one includes in its measurement. What science does do is rule out some options as illogical, as infeasible, or as less useful than some other choice. Describing many of the requisite decisions along the way as real choices implies that there are scientifically acceptable alternatives.

Science cannot tell us which of several very different measures of poverty is “right.” In fact, the role of science seems to me to be greatest in terms of some of the more arcane elements in its measurement – in informing us about equivalence scales, about how and under what conditions to make regional cost of living adjustments, about how to handle home ownership compared to rental homes, or how to treat lumpy purchases like automobiles if we chose to measure resources by expenditures. We couldn’t measure poverty well without science, but unfortunately science doesn’t give us much guidance in making big decisions like whether or not to work with an absolute or a relative concept of poverty, or at what level to set the initial numeraire threshold which will, after all, determine the level of poverty. Those decisions are essentially political judgments; they reflect our personal views about economic inequality, social justice, equity, or matters of politics.

To come to the question of whether it is useful to standardize the measure of poverty, the question becomes: “For what purpose is that measure to be used?” I will distinguish three. One purpose is of a purely scientific nature where poverty is a dichotomous metric reflecting economic deprivation. That measure is often used in empirical inquiries as a dependent variable in studying the incidence or prevalence of poverty, or as an independent or control variable in the study of some behavior or outcome. There might be some value in having those studies all focus upon a consensus measure of poverty – it would make comparisons across studies easier and would focus on just what it was in one study or another that was correlated with poverty or that caused it to change. But because there is no scientific consensus about the concept of poverty, there is no single scientific measure of poverty. The concept itself encompasses political judgment which explains the limited role of science in its measurement. That description seems to me to apply to every one of the specific topics listed for discussion at

our workshop – high school graduation, race, ethnicity, disability, intergenerational mobility and self-regulation, as well as poverty.

Standardization is helpful in science when it makes cross-study comparisons easier and facilitates communication. An example is when the measure is essentially neutral and arbitrary: standard measures of lengths such as inches or centimetres and miles or kilometres come to mind as an example. But the measure of poverty includes non-scientific judgment about what it is we wish to measure – those choices that underlie the five steps in its measurement. These choices are neither arbitrary nor firmly based in science.

Another purpose for which we often wish to measure poverty is to use that measure as a social indicator of the outcome of economic activity, as a reflection of our social compassion. “A decent provision for the poor is the true test of civilization,” quotes Tony Atkinson of Samuel Johnson (1770) (Atkinson, 1983, p.224) As a social indicator, “poverty” is a measure of a political reality and because it is, it necessarily involves political judgments that are not based in any science. Measuring poverty for this purpose surely has value, and clarity and transparency in measurement is important. But for this purpose, the measure of poverty must be made in a political context and with the political implications unabashedly in mind. Ultimately I think it weakens the public respect for science to claim scientific authority for a decision that is essentially political in nature.

Yet another, third, purpose for which the U.S. official measure of poverty is used is to establish eligibility for certain government social welfare programs. Today, the standard is often in terms of some percentage (often well above 100%) of the poverty threshold as eligibility for

one welfare program or another. Here we have a necessity for standardization in behalf of fairness, but this is of no particular value to science. To be useful in determining eligibility, some specific definition of poverty must be imposed since we can't use one definition for folks in Arizona and another definition for those in Alaska. Both that definition and the level that defines eligibility are properly political judgments. Here, however, there is a separate, generalizable point, one I first heard from Ken Prewitt in regard to the Decennial Census count.

Prewitt suggested that if a scientific measure automatically translates into a public policy outcome, science will lose out to politics. We have folks who don't know much about sample surveys and scientific methods of adjusting the census count weighing in with passionate arguments about the matter. They have interest in the issue precisely because that count translates directly into congressional seats. Serious politics is linked tightly to a scientific methodology. In such a case, science loses and politics will determine the methodology. That, is the principle reason the 1994 NRC recommendations (Citro & Michael, 1995) for improvements in the measure of poverty – based as well as is feasible on solid science – didn't get any traction in Washington: too many eligibility dollars and political bragging rights were automatically linked to the measure of poverty. The NRC recommendations urged uncoupling the two uses – a scientifically sound measure of the level of some reasonable definition of poverty, and a wholly separate system by which eligibility for entitlement programs would be determined. If that had been adopted, the politicians who care greatly about the latter would have had far less interest in the former, just as politicians don't get exercised about how the ratings of academic departments are adjusted by any scientific schemes. They care deeply about how many federal dollars flow to their districts, as they should, and if the determinant of those dollars is some scientific formula that can't be tinkered with post-implementation, then

the formula be damned, the politics will win out. Maybe they should. But a measure that is that political is not likely to be crafted as the best science. It would be well if scientists just focused on getting their metrics right and left to other specialists and politicians how those dollars should be allocated.

What remain is the question that if we stipulate that our measure is for scientific purposes, and that only, would it then be wise to have standardization in the metric by which we measure poverty? No, I don't believe so for several reasons. (1) If there are judgments required about some elements of the definition (as I've shown there does in the case of poverty), it is very unlikely that all scientists will agree on those judgments. There will not be true consensus about that one best or scientifically sanctioned official definition. To attempt to impose harmony where consensus does not exist will result in conflict and needless tensions. I've argued elsewhere (Michael 2001) that making something into a public good – a commonly used, singular way of thinking about the matter – imposes costs that need not be borne if there is no true consensus and no compelling requirement of one standard for all: we have gotten ourselves awfully riled up about some things on which we hold different opinions, in large part because we have chosen to make of them a public good, a common collective item. Other things we disagree about, which we have not made into a common or public good, don't cause much fuss at all: consider brussel sprouts and abortion. We as a nation have quite diverse views of both, but we only expend energy and resources arguing about the latter. (Yes, this is partly due to ethical or moral concerns, but that too is partly only a line of argument to justify making that into a public good or a collective "right.") Where we have different views and no requirement for a common action or for a single common expression of value, we get along fine,

individually expressing our opinions. So it should be with the measure of poverty for scientific purposes.

(2) Standardization has costs of at least two sorts. If a standard is implemented and successfully adopted, that may inhibit alternative ways of thinking, encouraging thinking in lock-step. It may also inadvertently impose faulty or erroneous elements into the measure of a concept. Also, the process of deciding which of several alternatives is sanctioned as the standard, will likely waste much effort. Those who have some sense of ownership of other measures will surely spend effort and energy advocating for their idea instead of getting on with using it in their scientific endeavors.

(3) A free market solution to a problem is preferable if one exists and in the measurement of a scientific concept, I believe choice is best. Consider the experience of the efforts by Simon Kuznets and the National Bureau of Economic Research in defining a set of national income accounting concepts. There was never, so far as I know, a board of scientists who passed judgement on “Net National Income” or “Gross National Product” and awarded it sanctity over all near or competing measures. Yet it was widely adopted, highly successful as a concept and as a metric, and became the standard because it was widely accepted. Why? Because it was a very good metric, reflecting a very useful concept and was adopted because one after another national income accountant thought it was good. That’s true as well for Jacob Mincer’s highly regarded “earnings function.” It too became the standard because it won out in the competition for a good and useful idea, measured with clarity and feasibility. We don’t need a committee to decide the matter. If we were successful in imposing some standard metric on a nice concept, how long should it be applied? Arthur Burns had a highly regarded reference

cycle for the pattern of U.S. business fluctuations and it was widely used until it wasn't useful anymore, but again no one set it up as the standard and no one needed to take it down as the standard when it became outdated.

This gets at the heart of how science works, I think. I have been persuaded by a description of the scientific process attributed to Michael Polanyi and described by Richard Rhodes in The Making of the Atomic Bomb (1986). To do justice to the description, I append it below, quoting Rhodes extensively. In essence, Polanyi argues “The authority of scientific opinion remains essentially mutual; it is established between scientists, not above them.” That allows anyone with a better idea to pursue it and if it works better, it will be adopted and used by others – those insidious metrics of “citations” and “influence” by which we judge one another.

Summary

So let me state as clearly as I can (but mindful of Bohr's concern, no clearer than I think) my position about standardization with regard to the measure of poverty. Poverty as a concept has three distinct and useful purposes: it is a dichotomous measure of economic deprivation, it is a measure of social compassion, and it is a policy tool useful in calibrating the degree of need for social welfare support. In the later purpose, standardization is necessary as the measure should be applied equally to all, and used for this purpose there is a true public goods aspect to the measure of poverty that cannot be avoided. Its measurement is essentially a political judgment, however. While there are elements in which science can offer guidance, science cannot state which of several choices is superior, and scientists have no more, or no less, right or wisdom than others in deciding which choice is best. Determining the level of some resource that should constitute eligibility for social welfare – and calling it poverty – should be in the hands

of political decision makers. As an analytic notion useful in assessing the outcome of economic activity for those at the lower tail of the distribution of economic well-being and as a measure of social compassion there are several scientifically worthy measures that can and do yield quite different results at a point in time and over time. Different competent scientists will find one or another most useful and illuminating. I do not see any grounds for thinking I or you or any collective body has the wisdom to select one of these for use by others. I know which I prefer, and I will use it. You should do the same. If one proves more useful in more circumstances, it will gain the influence it deserves.

References

Atkinson, Anthony B. 1983. The Economics of Inequality. Oxford: Clarendon Press. 2ed Edition.

Atkinson, Anthony. 1989. Poverty and Social Security. Hemel Hempstead, England: Harvester-Wheatsheaf.

Citro, Constance F. & Robert T. Michael, eds. Measuring Poverty: A New Approach. Washington, D.C.: National Academy Press. 1995.

Robert T. Michael. 2001. "Private Sex and Public Policy." in E.O. Laumann and R.T. Michael, eds. Sex, Love, and Health in America: Private Choices and Public Policies. Chicago: Univ. of Chicago Press. pp. 465-492.

Orshansky, Mollie. 1965. "Counting the Poor: Another look at the Poverty Profile". Social Security Bulletin. V.28(1: January): 3-29.

Rhodes, Richard. 1986. The Making of the Atom Bomb. New York: Simon & Schuster.

Rowntree, B. Seebohm. 1901. Poverty: A Study of Town Life. London: MacMillan.

Rowntree, B. Seebohm. 1941. Poverty and Progress. London: Longmans.

Sen, Amartya K. 1973. On Economic Inequality. Oxford: Clarendon Press.

Smith, Adam. 1776. An Inquiry into the Nature and Causes of the Wealth of Nations. London: Routledge. [Modern Library Edition 1993. NY: Random House, Inc.]

Townsend, Peter. 1979. Poverty in the United Kingdom. London: Allen Lane

Appendix: Excerpts from Rhodes

I quote below excerpts from a wonderful book titled The Making of the Atom Bomb by Richard Rhodes [1986, NY: Simon & Schuster Paperback; it was awarded a Pulitzer Prize, the National Book Award and the National Book Critics Circle Award, so I get no credit for recommending it]. At one point (pp.31-34) Rhodes is talking about how science works and discusses the views of chemist Michael Polanyi who argued that physics, indeed, all science “works by sorting among alternatives.” Quoting extensively from Rhodes book:

He [Polanyi] discovered a traditional organization [of science] far different from what most nonscientists suppose. A ‘republic of science,’ he called it, a community of independent men and women freely cooperating, ‘a highly simplified example of a free society.’ Not all philosophers of science, which is what Polanyi became, have agreed. Even Polanyi sometimes called science an ‘orthodoxy.’ But his republican model of science is powerful in the same way successful scientific models are powerful: it explains relationships that have not been clear.

Polanyi asked straightforward questions. How were scientists chosen? What oath of allegiance did they swear? Who guided their research – chose the problems to be studied, approved the experiments, judged the value of the results? In the last analysis, who decided what was scientifically ‘true’? Armed with these questions Polanyi then stepped back and looked at science from outside.

Behind the great structure that in only three centuries had begun to reshape the entire human world lay a basic commitment to a naturalistic view of life. Other views of life dominated at other times and places – the magical, the mythological. Children learned the naturalistic outlook when they learned to speak, when they learned to read, when they went to school. ‘Millions are spent annually on the cultivation and dissemination of science by the public authorities,’ Polanyi wrote once.... “In other words, our civilization is deeply committed to certain beliefs about the nature of things; beliefs which are different, for example, from those to which the early Egyptian or the Aztec civilizations were committed.’

Who then guided that work [of scientists]? The question was really two questions: who decided which problems to study, which experiments to perform? And who judged the value of the results?

Polanyi proposed an analogy. Imagine, he said, a group of workers faced with the problem of assembling a very large, very complex jigsaw puzzle. How could they organize themselves to do the job most efficiently?

Each worker could take some of the pieces from the pile and try to fit them together. That would be an efficient method if assembling a puzzle was like shelling peas. But it wasn’t. The pieces weren’t isolated. They fitted together into a whole. And the chance of any one worker’s collection of pieces fitting together was small. Even if the group make enough copies of the pieces to give every worker the entire puzzle to attack, no one would accomplish as much alone as the group might if it could contrive a way to work together.

The best way to do the job, Polanyi argued, was to allow each worker to keep track of what every other worker was doing. ‘Let them work on the putting the puzzle together in the sight of the others, so that every time a piece of it fitted in by one [worker], all the others will immediately watch out for the next step that becomes possible in consequence.’ That way, even though each worker acts on his own initiative, he acts to further the entire group’s achievement. The group works independently together; the puzzle is assembled in the most efficient way.

Polanyi thought science reached into the unknown along a series of what he called ‘growing point,’ each point the place where the most productive discoveries were being made. Alerted by their network of scientific publications and professional friendships – by the complete openness of their communications, an absolute and vital freedom of speech – scientists rushed to work at just those points where their particular talents would bring them the maximum emotional and intellectual return on their investment of effort and thought.

It was clear, then, who among scientists judged the value of scientific results: every member of the group, as in a Quaker meeting. ‘The authority of scientific opinion remains *essentially mutual*; it is

established *between scientists*, not *above* them.’ There were leading scientists, scientists who worked with unusual fertility at the growing points of their fields; but science had no ultimate leaders. Consensus ruled.

The author thanks John Mickelwright for comments on an early draft.

¹ There is an interesting aside story based on Rowntree’s work, at this present time when we are concerned how best to ascertain the constitutionally mandated fact of the size of the U.S. Population by a complete enumeration or by scientific sampling. The mid-thirties was the time sampling theory was promoted. Bowley had shown that the standard error for a population’s percentage, p , is $\sqrt{p(1-p)/n}$, but Rowntree didn’t believe it would work, so after the York enumeration he took a sample of 1-in-20 from his own data and calculated the proportion in poverty. His complete enumeration yielded 31.1% of his sample, his 1-in-20 random sample yielded 32.6% – within the 1 standard error difference implied by Bowley’s equation. Most researchers shifted to random sampling from then on. George Gallup began his company in the mid-30s, NORC founded in 1941 as a spin-off; Kinsey began collecting sex data in late 1930s, but eschewed using a scientific sample since he asserted (without any evidence) that those selected from a probability sample wouldn’t do his survey, and was subsequently severely criticized for that by a National Academy panel led by Allen Wallis and Fredrick Mosteller. Currently many statisticians believe the U.S. would get as accurate an estimate of the U.S. population by a well-crafted sample, at a fraction of the cost, but again, there’s much scepticism about that; perhaps we need another Rowntree.