The History of the First American Census and the Constitutional Language on Censustaking: Report of a Workshop

Margo Anderson
Department of History
University of Wisconsin–Milwaukee
Milwaukee, WI 53201

Stephen E. Fienberg
Department of Statistics
Carnegie Mellon University
Pittsburgh, PA 15213-3890

February 23, 1999

Contents

1 Background and Overview

On Friday November 13, and Saturday, November 14, 1998, with support from the Donner Foundation, the Woodrow Wilson International Center for Scholars hosted a workshop on the history of the First American Census in 1790. The workshop was attended by scholars with expertise in the history of statistics, the history of censustaking and early American history. See Appendix A for brief biographical statements of the principal participants and a list of other attendee/participants. The group considered the format, capacities, and impact of the First American (1790) Census. Individual presentations by small groups of scholars on particular topics were made and were supported by background documents. The group then attempted to come to consensus on several substantive questions about the nature of the 1790 census experience, and to define what was still open to interpretation, dispute, or would require further research. A brief report on the workshop appeared in *The Wilson Quarterly*, Winter (1999) and is included here as Appendix B.

The questions that prompted the workshop on the 1790 census experience derive largely from contemporary issues of censustaking in America, and the conflict between congressional Republicans and the Clinton administration over the proper format for taking the 2000 census. The controversy emerged in 1996 and intensified in early 1998 when the U.S. House of Representatives sued the Clinton administration over its plan for the 2000 census. The historical arguments about what transpired in 1790 and the intents of the framers are embedded in the legal filings in the case, *United States House of Representatives et al., vs. United States Department of Commerce, et al.*, [24]. In those filings, attorneys for the U.S. House of Representatives challenged the Clinton administration's plans for employing methods of statistical estimation in the 2000 Census, and alleged that sampling would be inconsistent with both the 1976 Census statute and the Constitution's census clause. A three judge district court upheld the House claims. An additional suit, *Glavin, Barr et al., vs. Clinton et al.*, raised similar claims and was also decided in favor of plaintiffs. Both suits were appealed to the Supreme Court which heard oral arguments on November 30, 1998, and decided the cases on January 25, 1999.[1] While the Supreme Court resolved the cases at hand, on the basis of its interpretation of the 1976 Census statute, it sidestepped the historical and constitutional issues.

It is worthwhile looking at the way the arguments were framed in this litigation. The House brief of April 1998 claimed, for example, that the framers ac-

[1] Department of Commerce v. House of Representatives (98-404) No. 98-404, 11 F. Supp. 2d 76, appeal dismissed; No. 98-564, 19 F. Supp. 2d 543, affirmed.

cepted errors in the census because they had bound themselves to take an actual enumeration: “Even though [Thomas] Jefferson [Secretary of State and the official in charge of the 1790 census] was familiar with methods of statistical estimation, having used them effectively in his 1782 survey of Virginia’s population, he did not adjust the 1790 census numbers despite his knowledge that the ‘omissions . . . had been very great.”’ The Clinton administration on its part claimed that the framers intended an accurate count, and that a census method deemed to be inaccurate would also fail a test of constitutionality. The Supreme Court in its opinions in these cases chose not to address this constitutional issue directly although several justices in their written opinions did discuss the matter. Instead, it concluded that the 1976 statute authorizing the taking of the census by the U.S. Bureau of the Census precluded the use of sampling for the purposes of apportionment. The major use of sampling in the Census 2000 plan would be via a postenumeration survey, the results of which would be used to adjust those from the “traditional” enumeration. Since there remain major differences of opinion regarding this use of sampling in the census context, between the congressional Republicans and the Clinton administration and some state and local governments, it is reasonable to expect the constitutional questions to be raised again in litigation in the future.

The workshop addressed three primary issues:

- The first involved the intent of the framers of the Constitution in drafting the census clause, and in particular what was meant by an actual enumeration. Two related questions relate to the state of censustaking and the state of scientific sampling and statistical estimation in the 18th century.
- A second question involved exactly how the 1790 census was conducted, from the drafting of legislation in early 1790 to the first apportionment of Congress in April 1792.
- A third question involved how the Washington administration, Congress, and later observers evaluated the 1790 experience, particularly as that experience affected censuses in the early 19th century. We consider these issues in turn, and then attempt to make several points about possible further research.

2 Constitutional Grounding: Why the Framers Wrote the Census Clause

The 1787 Constitutional Convention confronted a number of issues in constituting the new national government which the census clause was designed to solve. First, the framers debated the proper political grounding for legislative authority, and concluded that the most acceptable grounding of governmental authority and power derived from the people. They also debated how to accommodate the disparate and changing sizes of the various states, that is, the problem of distributing power among the disparately sized states, and how to revalue that distribution of power in order to accommodate new states and population growth. The framers also sought the proper basis on which to levy taxes, cognizant that the efforts of the federal government under the Articles had not been able to raise revenue effectively.

The framers also knew that population counting was a venerable tradition in late 18th-century America, since the Crown had required population counts from many colonial governors since the founding of the colonies (Cassedy [7], Cohen [8],Wells [25]). The Crown had ordered population counts as a mechanism for evaluating the success of these far off plantations and more particularly to determine if the colonies were defensible if attacked. The fledgling colonies were by no means guaranteed to be demographic successes, and were expensive to found and maintain in their early years. During the American Revolution, the Continental Congress had proposed using a population census, including a three-fifths clause to count slaves, as a substitute for the land requisitions that had been the basis on which to levy funds from the states for the support of the union (Rakove, [17] [18]). The census clause was the solution to many of these issues. Rakove [16] has provided the background of the debate. The term actual enumeration was added to the final version of the Constitution by the Committee on Style, to place emphasis on the framers' intent that one of the first tasks of the first Congress and President would be to write legislation and take the enumeration which would lead to reallocation or reapportionment of Congress from the 65 seats allocated in the constitution itself (see Table 1 for the list). There is little evidence that the framers considered or even anticipated the logistical and technical problems that might plague an enumeration. The Constitution did not specify an apportionment formula or size of the House of Representatives beyond the requirement that there be a minimum of 30,000 people per representative.

3 The State of Population Estimation in the 18th Century

Nevertheless, the fact that many colonial governors had counted the population before the American Revolution resulted in population data that could provide the raw material for further population counting and estimation. And in fact these data were widely used. At the same time, Europeans also began to explore the nature of population distribution, growth and change among the several nations of

Table 1: Congressional Apportionment, 1790s

	Apportionment		State Share of Apportionment	
State	Constitutional	1792	Constitutional	1792
Virginia	10	19	15.38	18.09
Maine	8	14	12.30	13.33
Pennsylvania	8	13	12.30	12.38
New York	6	10	9.23	9.52
North Carolina	5	10	7.69	9.52
Maryland	6	8	9.23	7.61
Connecticut	5	7	7.69	6.66
South Carolina	5	6	7.6	9 5.71
New Jersey	4	5	6.15	4.76
New Hampshire	3	4	4.61	3.81
Rhode Island	1	2	1.53	1.90
Georgia	3	2	4.61	1.90
Vermont		2	.	1.90
Kentucky		2	.	1.90
Delaware	1	1	1.53	0.95

Europe. Thus there was an emerging demography in early America (see Cassedy [7]) which could have influenced the shape of the 1790 count. As we quoted above, one of the claims that sampling cannot be used as a census method derives from a comparison of Jefferson's methods of statistical estimation in his *Notes on the State of Virginia*, [13], and his practice as the Secretary of State and in charge of administering the 1790 count. In his *Notes*, written in 1781-82 and published in 1785, Jefferson extrapolated a population count for Virginia from partial data. The claim is thus that existing knowledge would have permitted the framers to use statistical estimation in the constitution, or the 1790 census legislation, but that they consciously chose not to do so. The consensus of the participants at the workshop was that such a claim is historically inaccurate. We review the claim below.

4 The State of Statistical Sampling and Estimation in the 18th Century

The workshop addressed at length some of the issues relating to the history of sampling and censustaking, so that the efforts of the founding fathers to establish a periodic census as part of the constitution could be set in methodological context. Appendix C, prepared by Professor Eugene Seneta, of the University of Sydney, Sydney, Australia, provides considerable statistical background and commentary relevant to the issues. Appendix D presents related materials on the history of sampling by Dan Cork of the Department of Statistics, Carnegie Mellon University. We provide a brief summary here.

The questions addressed in Appendix C are:

- What might Jefferson have known of the methodology of statistical estimations, and of enumeration, of populations at the time of ratification of the U.S. Constitution and of the first American Census?
- What was the understanding and application of sampling ideas to populations in France and the Russian Empire?

The succinct answer is that the framers of the Constitution, and Jefferson in particular, knew nothing of statistical sampling as such, since the scientific methodology of probability sampling would not emerge for more than another century (e.g., see Bellhouse [6] and Desrosières, [9] p.25 and p. 210, as well as Appendix D to this report on the History of Sampling authored by Dan Cork). Further, while the Frenchman Pierre Simon de Laplace was writing about some specific uses of sampling-like ideas in a very technical fashion in the 1780s, there is no evidence to indicate that these ideas were either accessible to or discussed by any of the framers. In particular, there is no evidence that Jefferson had any knowledge of them.

As Desrosières [9] notes:

> The technique of conducting surveys by means of probability samples did not appear until the late nineteenth century with the work of the Norwegian Kiaer, and even then only in a rudimentary form, more intuitive than formalized. . . . Surveys based on a small number of individuals had been conducted for much longer, throughout the nineteenth century in particular, often by persons with a high degree of scientific knowledge . . . for whom the elements of probability calculus necessary for the "intuition" of the sampling method ought not to have constituted an insurmountable problem. Laplace, moreover, had use of them as early as the late eighteenth century to estimate the French population, but this experiment had not been followed up for more than a hundred years.

Laplace did not actually implement his ideas until after the French Revolution in the early 19th century, and even then, as Desrosières notes, they seem to have had little influence on censustaking methodology.

Thus we conclude that the notion that the framers considered sampling as an alternative to actual enumeration and rejected its use, has no historical or methodological basis.

5 Legislative History of the 1790 Census Law

Further clues to the framers' intentions and practice may be gleaned from an examination of the conduct of the 1790 count itself. Workshop participants explored what is known about the legislation for, administration of, and impact of the first census. Although there are a number of brief standard historical treatments of the 1790 count (e.g., William Rossiter, [20]; Carroll Wright and William C. Hunt, [27]; and Hyman Alterman, [1]) as well as the writings of several of the participants (Margo Anderson, [2]; Cohen, [8]; Cassedy, [7]), none directly addressed the specific questions posed at the workshop. Thus it is useful to review the history of the 1790 count in some detail with those questions in mind.

The first federal Congress convened on March 4, 1789 and faced the enormous tasks involved in establishing the machinery of the new government. These included developing such basic government institutions as the tax and budget system, the federal court system, and the post office. The census bill was on the agenda as well, and was referred to committee. (Rossiter [20]) No bill was reported before the first session adjourned on September 29, 1789. When Congress reconvened at the second session on January 4, 1790, the old committee was dismissed, a new committee created, and the press reported that a bill was drafted by January 18, 1790. In the discussion of the census on the floor of the House on January 25, James Madison offered to revise the bill. Madison had not been a member of either of the original committees which received the charge to write the bill. He reported a substantially revised bill on January 26, which was debated, amended and passed by the House as HR34 on February 8, 1790.

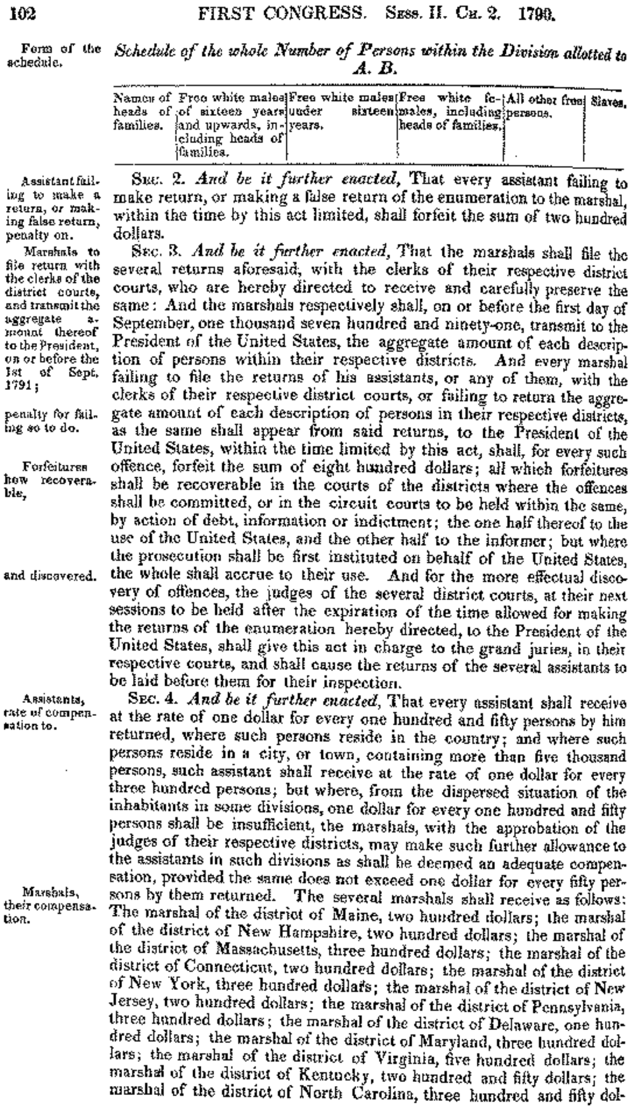
From there the bill went to the Senate, where it was amended and simplified, particularly by deleting the ambitious census of occupations that James Madison had proposed. The Senate also made a number of stylistic changes to the bill, and voted on February 8 to delete the word "actual" from the title and text of the bill. Senate debates in the first Congress were not public, so we do not have a clear explanation of why the word "actual" was deleted, but presumably it was considered redundant. Hence the 1790 census act as passed by the Senate and then agreed to by the House and signed by the President on March 1, 1790 was titled "An Act Providing for the Enumeration of the Inhabitants of the United States." The act required the United States marshals to begin the enumeration on August 1, 1790 and return the results to the President in nine months. (See Figure 1 for a copy of the schedule and Table 2 for a list of the United States marshals.) The census asked six basic questions, dividing the population by age, race, sex and civil status. It required only the name of the household head to be listed on the form, and the household head to respond.

Table 2: United States Marshals at the Time of the 1790 Census (Appointed September 24th-25th, 1789, unless otherwise noted).

Connecticut	Philip Bradley
Delaware	Allan McLean
Georgia	Robert Forsyth
Kentucky	Samuel McDowell junr.
Maine	Henry Dearbourn
Maryland	Nathaniel Ramsay
Massachusetts	Jonathan Jackson
New Hampshire	John Parker
New Jersey	Thomas Lowry
New York	William S. Smith
North Carolina	John Skinner (June 7th, 1790)
Pennsylvania	Clement Biddle
Rhode Island and Providence Plantations	William Peck (July 2nd, 1790)
South Carolina	Isaac Huger
Vermont	Lewis R. Morris (March 4th, 1791)
Virginia	Edward. Carrington

*Alternative spelling for some names.

George Washington designated the Secretary of State to administer the census in his name. In September 1789, Washington had appointed Thomas Jefferson Secretary of State, and Jefferson had left Paris where he had been serving as Ambassador for the trip back to the United States. In late March 1790, Jefferson arrived in New York and took up his duties as Secretary. Jefferson thus played no role in the legislative debates on the census bill, nor did the debates in the House address any questions which could be seen as statistical estimation. In contrast, though, Jefferson and the Congress did face other administrative issues in starting



102 FIRST CONGRESS. Sess. II. Ch. 2. 1790.

Form of the schedule.

Schedule of the whole Number of Persons within the Division allotted to A. B.

Names of heads of families.	Free white males of sixteen years and upwards, including heads of families.	Free white males under sixteen years.	Free white females, including heads of families.	All other free persons.	Slaves.

Assistant failing to make a return, or making false return, penalty on.

Sec. 2. *And be it further enacted,* That every assistant failing to make return, or making a false return of the enumeration to the marshal, within the time by this act limited, shall forfeit the sum of two hundred dollars.

Marshals to file return with the clerks of the district courts, and transmit the aggregate amount thereof to the President, on or before the 1st of Sept. 1791; penalty for failing so to do. Forfeitures how recoverable, and discovered.

Sec. 3. *And be it further enacted,* That the marshals shall file the several returns aforesaid, with the clerks of their respective district courts, who are hereby directed to receive and carefully preserve the same: And the marshals respectively shall, on or before the first day of September, one thousand seven hundred and ninety-one, transmit to the President of the United States, the aggregate amount of each description of persons within their respective districts. And every marshal failing to file the returns of his assistants, or any of them, with the clerks of their respective district courts, or failing to return the aggregate amount of each description of persons in their respective districts, as the same shall appear from said returns, to the President of the United States, within the time limited by this act, shall, for every such offence, forfeit the sum of eight hundred dollars; all which forfeitures shall be recoverable in the courts of the districts where the offences shall be committed, or in the circuit courts to be held within the same, by action of debt, information or indictment; the one half thereof to the use of the United States, and the other half to the informer; but where the prosecution shall be first instituted on behalf of the United States, the whole shall accrue to their use. And for the more effectual discovery of offences, the judges of the several district courts, at their next sessions to be held after the expiration of the time allowed for making the returns of the enumeration hereby directed, to the President of the United States, shall give this act in charge to the grand juries, in their respective courts, and shall cause the returns of the several assistants to be laid before them for their inspection.

Assistants, rate of compensation to. Marshals, their compensation.

Sec. 4. *And be it further enacted,* That every assistant shall receive at the rate of one dollar for every one hundred and fifty persons by him returned, where such persons reside in the country; and where such persons reside in a city, or town, containing more than five thousand persons, such assistant shall receive at the rate of one dollar for every three hundred persons; but where, from the dispersed situation of the inhabitants in some divisions, one dollar for every one hundred and fifty persons shall be insufficient, the marshals, with the approbation of the judges of their respective districts, may make such further allowance to the assistants in such divisions as shall be deemed an adequate compensation, provided the same does not exceed one dollar for every fifty persons by them returned. The several marshals shall receive as follows: The marshal of the district of Maine, two hundred dollars; the marshal of the district of New Hampshire, two hundred dollars; the marshal of the district of Massachusetts, three hundred dollars; the marshal of the district of Connecticut, two hundred dollars; the marshal of the district of New York, three hundred dollars; the marshal of the district of New Jersey, two hundred dollars; the marshal of the district of Pennsylvania, three hundred dollars; the marshal of the district of Delaware, one hundred dollars; the marshal of the district of Maryland, three hundred dollars; the marshal of the district of Virginia, five hundred dollars; the marshal of the district of Kentucky, two hundred and fifty dollars; the marshal of the district of North Carolina, three hundred and fifty dol-

Figure 1: Reproduction of the Schedule for the 1790 Census

the census. One involved passing a second bill in June 1790 making provision for the United States marshal in Rhode Island to take the enumeration since Rhode Island had ratified the constitution and joined the union.

From a modern perspective, the 1790 census seems a primitive administrative operation. It is generally believed that Thomas Jefferson forwarded the statute to the marshals, and the marshals in turn used the statute as their only guidance in appointing their assistants to canvass the population. Jefferson did not print or distribute forms or instructions for taking the census. Jefferson was also responsible for many other activities at the same time, including foreign policy, and patents and copyrights. The Secretary's staff was extraordinarily small: a half dozen people at most compared with the roughly 70 in Alexander Hamilton's Treasury Department of the time. (Leonard D. White, [26]) Hence little systematic information exists on the problems and procedures in the 1790 field enumeration. What is currently known derives from very slim correspondence between Jefferson and the marshals, from the general correspondence of Jefferson, Washington, and other high government officials, and from scattered press reports of returns. The available information is thus generally in the form of side comments in longer letters and reports, rather than as systematic reportage or analysis of the ongoing count. Nevertheless some inferences can be made on procedures and on how the principals involved viewed the census process.

It is possible to get a general sense of the timing of the count from the dates on the oaths signed by the assistant marshals, or from the publication of the results in the press. In Boston, for example, the assistant marshal began on August 2, and finished counting the 18,000 inhabitants on August 21. The population of the towns in Rhode Island was published in October 1790, and the population for the Charleston, SC on November 12, 1790. (Rossiter, [20], pp. 45-47.) In late January, 1791, Thomas Jefferson wrote to William Short noting that "the census has made considerable progress" and predicted "that our numbers will be between 4 and 5 millions." By the midsummer 1791, however, Jefferson had moderated his expectations somewhat, both for the completion of the count and for the final population result. In correspondence to Lewis Littlepage in July 1791, he predicted, "our numbers to be about three millions and a half." In correspondence with Thomas Mann Randolph in early August 1791, he estimated the census results at "upwards of 3,800,000 inhabitants." He still did not have final returns from Vermont (a new state which started its enumeration in April 1791); Virginia (the largest state), and New Jersey and South Carolina. Less populous areas were harder to count than eastern seaboard cities.

The tendency toward population boosterism was evident in the correspondence of other federal officials as the leaders of the new nation tried to bolster their place in the world of nations. Alexander Hamilton used the likely results of the census to impress British agent George Beckwith in October 1790. In December 1790 George Washington exuberantly wrote Gouverneur Morris that "The numbers of our people as far as they can be ascertained from the present stage of the Census, will not fall short, it is said, of five millions, some think more." Washington advised Gouverneur Morris to consider the size of the population as evidence of the viability of the new nation.

But in later correspondence in July 1791 Washington also warned Morris, who was about to become ambassador to France, that he had overestimated the census results when the first returns came in and had thus misled him earlier in his December letter. The quote is embedded in a much longer letter to Morris. The section on the census is worth repeating in its entirety, since it has been widely quoted to support the claim that the framers opposed statistical estimation.

> "In one of my letters to you the account which I gave of the number of inhabitants which would probably be found in the United States on enumeration, was too large. The estimate was then founded on the ideas held out by the Gentlemen in Congress of the population of their several States, each of whom (as was very natural) looking thro' a magnifying glass would speak of the greatest extent, to which there was any probability of their numbers reaching. Returns of the Census have already been made from several of the States and a tolerably just estimate has been formed now in others, by which it appears that we shall hardly reach four millions; but one thing is certain our real numbers will exceed, greatly, the official returns of them; because the religious scruples of some, would not allow them to give their lists; the fears of others that it was intended as the foundation of a tax induced them to conceal or diminished theirs, and thro' the indolence of the people, and the negligence of many of the Officers numbers are omitted. The authenticated number however is far greater, I believe, than has ever been allowed in Europe, and will have no small influence in enabling them to form a more just opinion of our present and growing importance than has yet been entertained there." (Washington to Gouverneur Morris, July 28, 1791)

Washington, in other words, was providing advice to his future ambassador on the proper "spin" to be put on the census results: that is, even with an "undercount," the results "will have no small influence in enabling" Europeans "to form a more just opinion of our present and growing importance." He was not discussing the possibility or advisability of estimating error or adjusting the count for the various states.

6 Reporting the Results and Reapportioning Congress

Washington took a similar tack when he presented the census results in his annual address to Congress in October 1791. "The completion of the Census of the Inhabitants," he wrote, "will give you the pleasing assurance that the present population of the United States borders on four Millions of persons."(Washington's Third Annual Address to Congress, October 25, 1791). Even then, he cautioned that the returns from South Carolina had not been received. Further Congressional action was required extending the deadline for taking the count into the spring of 1792; Congress also exempted the marshal from the penalties for failing to submit timely returns.

Congress did not wait for the South Carolina returns to begin debating an apportionment bill. Rather, Congress began debating apportionment bills in November, and continued until a bill was passed in April. The concerns about the census faded into the background as a new issue emerged from the census clause: namely what the appropriate house size and apportionment method should be. As noted above, the Constitution was silent on these issues, and reasonably quickly a partisan dispute emerged with Thomas Jefferson and his supporters favoring one method and Alexander Hamilton and his supporters favoring another. The bill finally sent to George Washington called for a House of 120 members and was supported by the Hamiltonians. On the advice of Thomas Jefferson, George Washington vetoed the bill, using the presidential veto for the first time. Congress failed to override, and a compromise bill providing for 105 members was passed in late April. As Balinski and Young [5] have shown, there were strong partisan elements to the debates about apportionment; what did not occur, though, was a challenge to the accuracy of the census itself, or proposals to change the census numbers because they might be too low. In other words, Congress saw very quickly the distributional effects of various apportionment methods and proposals for sizes of the House. They did not challenge the basic adequacy of the census in the debate.

7 The Census in Later Years

Little is currently known about how the framers understood their experience with censustaking and apportionment beyond the obvious fact that the 1790 methods were repeated fairly consistently for the next 50 years. The U.S. marshals and their assistants continued to serve as the field enumeration staff of the census, collecting the data, and tallying the results for their local areas and reporting the results to national officials. The Secretary of State continued to oversee the count in Washington. And officials continued the practice of merely listing the name of the household head and the numbers of household members in various demographic groups. By 1820, instructions on census procedures were sent from Washington; by 1830, the Secretary of State printed the forms and distributed them. Congress required increasingly fine grained breakdowns of age, sex, and race cohorts, but added relatively few other questions to the count.

Only in 1840 did a methodological challenge to the accuracy of the census emerge which had political ramifications, and it did not affect the basic population count. In 1840, as Cohen [8] has described, the addition of a tabulation on insanity by race and local area resulted in the publication of data that seemed to show that free blacks in the North displayed much higher rates of insanity than did slaves. Southerners accepted the results as demonstrating that the institution of slavery embodied the appropriate civil status for African Americans; northerners, and anti-slavery activists in particular, accused John C. Calhoun, the South Carolinian then serving as Secretary of State, as fabricating the data. The controversy raged for years, and the data were not changed in the official publications, but they did make Congress aware of the possibilities of a political controversy over slavery could be affected by census data. And that realization in turn led to efforts to reform the census machinery in 1850.

In other words, the relative absence of such a challenge in 1790 or in the censuses up to 1840 strongly suggests that officials at the first census simply did not conceive of methods of statistical estimation as methods of improving the count. Nor did complaints that people were missed lead to charges that any one state's population was counted better than another's. Rather, officials recognized that the census was imperfect, but they had no conceptual alternative adequate to the task.

Two hundred years later, of course, the history of innovations in census taking tells a different story. It has been possible to do a better job of counting and the methods have been repeatedly refined. It is also clear that even a dispute over the disposition of one House seat can result in major political controversy, "fierce" political competition (Balinski and Young, [5], p. 21), and hence calls for changes in census procedures.

We conclude with a brief list of issues that could usefully be further investigated:

1. Examination of the practices of the states in the 1780s as they drafted constitutions, some of which called for types of population counts.

2. More systematic exploration of newspaper accounts of the receipt of returns.
3. Systematic review of the marshals' correspondence for mention of the census.
4. Systematic examination of the problems with the South Carolina count which prompted the delay in returns.
5. Examination of the oaths and payment schedules for the assistant marshals as a mechanism for tracking the progress of the count.
6. Review of the debates on the 1800 census with an eye to historical references to the experience in 1790 as lessons for the 1800 count.

A Biographical Sketches of the Workshop Participants

Margo Anderson is currently Professor of History & Urban Studies at the University of Wisconsin-Milwaukee. She is an American social historian specializing in the history of statistical accounting systems, including censuses and surveys, and the agencies of the federal statistical system which produce data. This year she is a fellow at the Woodrow Wilson International Center for Scholars exploring the history of federal survey research on families and households. She teaches American social and urban history, historical quantitative analysis, and women's and family history. She received a B.A. from Bucknell University in 1967, served in the Peace Corps in Maharashtra, India, and received a Ph.D. in History from Rutgers University in 1978. She has also taught at Vassar College, Case Western Reserve University, and in the Summer Program in Quantitative Methods of the Inter-university Consortium for Political and Social Research (ICPSR) at the University of Michigan.

Anderson's major publications include *The American Census: A Social History* (Yale, 1988); *The United States Census and Labor Force Change* (UMI Research Press, 1980); and, with Maurine Greenwald, eds., *Pittsburgh Surveyed: Social Science and Social Reform in the Early Twentieth Century* (University of Pittsburgh Press, 1996); and, with Stephen Fienberg, *Who Counts? The Politics of Census-Taking in Contemporary America* (Russell Sage, forthcoming). She has received major grant support from the National Endowment for the Humanities, the American Council for Learned Societies, and the Russell Sage Foundation. Anderson's other activities include a service as a member of the National Academy of Sciences Panel on Census Requirements for the Year 2000 and Beyond (1992-94); Chair of the Social Statistics Section of the American Statistical Association (1998); Program Chair for the Annual Conference of the Social Science History Association (1992), and Chair of the History Department (1992-95), Chair of the Executive Committee of the faculty (1996-98), and Director of the Center for Women's Studies (1983-89) at the University of Wisconsin–Milwaukee.

David Bellhouse is Professor of Statistics and Chair of the Department of Statistical and Actuarial Sciences at the University of Western Ontario, London, Canada. He is currently serving on the Advisory Committee on Statistical Methods for Statistics Canada and is an Associate Editor of Survey Methodology. He is the current President of the Statistical Society of Canada. Previously, he served as Chair of the Statistics Grant Selection Committee of the Natural and Engineering Sciences Research Council of Canada. In 1986-88 he served on a panel of Committee on National Statistics which dealt with the redesign of surveys of scientists and engineers. Bellhouse's research interests are in the theory and methods of sample surveys and in the history of statistics. He has published extensively on survey sampling. He has also written several papers on the development of probability and statistics prior to 1800. Recently, with Christian Genest he has completed a paper on the formative years of the Statistical Society of Canada. He is currently researching the statistical movement in pre-Confederation Canada.

Charlene Bickford is coeditor of the *Documentary History of the First Federal Congress, 1789-1791 (DHFFC)*, and has been the Director of the First Federal Congress Project at the George Washington University since 1984. An editor on 13 volumes of the *DHFFC*, she was the coeditor with Helen E. Veit of volume IV, *Legislative Histories*, which contains the official history of the Enumeration (Census) Act of 1790. Bickford is the coauthor with Kenneth R. Bowling of *Birth of the Nation: The First Federal Congress, 1789-1791* and was a curator on the exhibit by the same name. She has written articles and lectured widely on the subjects relating to the First Federal Congress and the presidency of George Washington. Her particular interests are the precedents set by the First Federal Congress and the relationship between the legislative and executive branches. Bickford currently teaches a graduate course in historical documentary editing for George Mason University, has taught at the George Washington University, and has been a visiting lecturer for numerous teacher training institutes. She is a former president of the Society for History in the Federal Government and the Association for Documentary Editing.

Kenneth R. Bowling is coeditor of the twenty volume documentary history of the *First Federal Congress* (which passed the Enumeration Act in 1790) at George Washington University. He received his Ph.D. from the University of Wisconsin in history in 1968. He is the author of *The Creation of Washington DC: The Idea and Location of the American Capital* and has written extensively on the politics of the American Revolution following the Revolutionary War.

James Cassedy is a historian with the National Library of Medicine in Bethesda, Maryland. In addition to personal research and writing, he is Editor of the Library's *Bibliography of the History of Medicine* as well as of the related database, HISTLINE. Prior to coming to NLM in 1968 Cassedy was a science administrator for the National Institutes of Health, spending several years as Executive Secretary of the History of the Life Sciences Study Section, followed by a term as Deputy Chief of NIH's European Office in Paris. In a still earlier existence, during the 50s, in a contract relationship with the U.S. Information Agency, he served as Director of binational cultural centers in several countries, respectively, in Haiti, Burma, and

Pakistan. With graduate degrees in American Civilization, he also gave courses in American history and literature at universities in those countries as well as in the U.S. for short periods.

In pursuit of his scholarly career Cassedy's professional activities have been principally in the fields of the history of medicine, science, demography, and public health. He has been active in appropriate professional groups and served as president of several. He has also done editorial work on several large historical publications, most recently for the forthcoming *American National Biography*. Early in his career Cassedy became interested in the extensive historical relationships of medicine, health, and science with statistics, particularly applied statistics. As a result, much (most)of his research and writing since 1959 has focussed on these connections. His book-length studies dealing significantly with the history of statistics include: *Charles V. Chapin and the Public Health Movement* (1962); *Demography in Early America, Beginnings of the Statistical Mind 1600-1800* (1969); *American Medicine and Statistical Thinking, 1800-1860* (1984); and *Medicine and American Growth 1800-1860* (1986). He is presently writing a volume of essays on the late 19th century polymath John Shaw Billings. This will include a substantial section (probably two chapters) on Billings as a statistician, with a principal focus upon his extensive contributions to the censuses of 1880 and 1890.

Patricia Cline Cohen is Professor of History and Women's Studies at the University of California at Santa Barbara. She served as Chair of Women's Studies for five years followed by two years (1996-1998) as the acting Dean of Humanities and Fine Arts on her campus. She is an early American historian, specializing in the cultural study of numeracy in 18th and early 19th-century America as well as the history of women in America, from 1600-1890. Cohen's first book, *A Calculating People: The Spread of Numeracy in Early America* (University of Chicago Press, 1983), traced the rise of a propensity to use numbers, quantities, and statistics in both political and daily life in America. The book includes a chapter on American political arithmetic and the history of the early censuses and a chapter history of arithmetic instruction in the period, with attention to gendered cultural notions of math talent and math applications. Cohen has published articles on math education and women and on the origins of the American Statistical Association in 1839. She is also the co-author of a recent US history survey textbook (*The American Promise*, Bedford Books, 1998) for which she wrote six chapters covering the years 1754-1840. Her most recent book is in women's history: *The Murder of Helen Jewett: The Life and Death of a Nineteenth-Century Prostitute in New York City* (Alfred A. Knopf, 1998).

William C. diGiacomantonio is an associate editor of *The Documentary History of the First Federal Congress*, 1789-91. A graduate of Georgetown University in International Affairs (B.S., 1985) and the University of Chicago in History (M.S., 1988), he has been employed at the First Congress Project since the great Bicentennial Year, 1989. (The First Congress Project is co-sponsored by the National Historical Publications and Records Commission of the National Archives, and by The George Washington University, in Washington, D.C.) As an editor, he has helped with the document searching, transcription, annotation, proof reading, and indexing of the five volume House Debates series, and the two volume petition series. The entire staff is now preparing the five volume correspondence series of letters to and from members of the First Congress. The House Debate series includes all known floor debate over the Enumeration Act of 1 March 1790. Volume 4 of the series (published in 1986) includes the complete legislative history of the act, including a calendar, the motions, the bill, and the final enrolled act.

diGiacomantonio has published articles and lectured on Quaker antislavery lobbying in the First Congress, congressional biography, early petitioning to the federal government, Washington, D.C. history, Quaker history, and–more to the point of the conference–the formation of national character by public support for the arts, sciences, and morality in the First Federal Congress. The latter covers the Enumeration Act, Madison's proposed ambitious first census schedule, and links it with his earlier support for a Library of Congress as a way of increasing the knowledge about the new nation by the new nation, for political as well as sociological purposes.

Stephen E. Fienberg is Maurice Falk University Professor of Statistics and Social Science, Carnegie Mellon University in Pittsburgh. He has served as Dean of the College of Humanities and Social Sciences at Carnegie Mellon and as Vice President for Academic Affairs at York University in Toronto, as well as on the faculties of the University of Chicago and the University of Minnesota. He has served as Applications and Co-ordinating Editor of the *Journal of the American Statistical Association* and was a founding co-editor of the popular-oriented statistical magazine, *Chance*. He has served as Vice President of the American Statistical Association and as President of the International Society for Bayesian Analysis, and is currently President of the Institute of Mathematical Statistics. He served as chair of the Committee on National Statistics at the National Academy of Sciences from 1980-1986, and on two of that Committee's panels on methodology for census taking.

Fienberg has published extensively on statistical methods for the analysis of categorical data, and on aspects of sample surveys and randomized experiments.

His research interests include the use of statistics in public policy and the law, the role of statistical methods in census taking, Bayesian inference, and the history of statistics. His most recent book, *Intelligence, Genes, and Success: Scientists Respond to THE BELL CURVE*, (Copernicus/Springer-Verlag, 1997), which he edited with colleagues at Carnegie Mellon, blends a number of these interests. Fienberg is currently working on a number of different aspects of data disclosure limitation and has married these to his longstanding interest in categorical data problems. He is also in the process of completing a book with Margo Anderson entitled *Who Counts? The Politics of Census-Taking in Contemporary America*, (Russell Sage, forthcoming).

Joseph B. Kadane is Leonard J. Savage Professor of Statistics and Social Sciences at Carnegie Mellon University in Pittsburgh. He earned a B.A. in Mathematics at Harvard (1962) and Ph.D. in Statistics at Stanford (1966). He has been a faculty member of the Department of Statistics at Carnegie Mellon since 1971. He served as Applications and Coordinating Editor of the *Journal of the American Statistical Association* from 1983 to 1985, and was an Associate editor for the Special Section on Census Undercount Research in 1991-1992. He has served as a member and chair of the American Statistical Association's Advisory Committee to the U.S. Census Bureau, and as a member of the Technical Advisory Committee to the U.S. Census Bureau on the Undercount. He was a participant on the National Research Council/National Academy of Science Panel on Decennial Census Methodology (1984-1989).

Kadane's publications include "Using Administrative Lists to Estimate Census Omissions" (*Journal of Official Statistics*, 1986), "Estimating the Population in a Census Year: 1980 and Beyond" (*Journal of the American Statistical Association*, 1985), "Sensitivity Analysis of Local Estimates of Undercounts in the 1980 U.S. Census" in *Small Area Statistics* (*Wiley*, 1987), all joint with E. Ericksen; "Adjusting the 1980 Census of Housing and Population" (*Journal of the American Statistical Association*, 1989) with E. Ericksen and J. Tukey; "Reconstructing the Adjusted Census for Florida: A Case Study in Data Examination" (*Journal of Statistical Computation and Graphics*, 1992), and "Evaluation of a Reconstruction of the Adjusted 1990 Census for Florida" (*Journal of Official Statistics*, 1995), both with M. Meyer; and "A Bayesian Approach to Designing U.S. Census Sampling for Reapportionment" (*Journal of Official Statistics, 1996*). Kadane has testified before the Subcommittee on Census and Population, Committee on Post Office and Civil Service, House of Representatives, on several occasions.

Eugene Seneta is Professor of Mathematical Statistics at the University of

Sydney, Australia, to which he came as Head of Department in 1979 after a long period at the Australian National University, Canberra. He has been Visiting Professor at the Virginia Polytechnic Institute and State University, Colorado State University, the University of Virginia, the University of Chicago and the Ecole des Hautes Etudes en Sciences Sociales, Paris. He has served on the Editorial Board of the Encyclopedia of Statistical Sciences, and is currently on the Editorial Boards of *Historia Mathematica* and *Theory of Probability and Mathematical Statistics (Kiev).* Born in Western Ukraine in 1941, he came to Australia with his parents in 1949, and studied in Adelaide until 1965.

Seneta has published extensively in the theory of Markov chains and of discrete branching processes, and the associated theory of finite and infinite non-negative matrices on which he has published a book. He also published a book on the theory of regularly varying functions in 1976, which was subsequently translated into Russian. Other areas of research have included population genetics, path analysis, sequentially rejective tests of hypotheses, and Bonferroni-type multivariate probability bounds. One of his many contributions to the history of probability and statistics was the book, *I.J. Bienayme: Statistical Theory Anticipated* (1977), that he coauthored with C.C. Heyde. This was in effect a history of probability and statistics in the 19th and early 20th centuries. Seneta's area of special interest is the work in France and the Russian Empire in this period, and is facilitated by his linguistic backgrounds. He is an elected Fellow of the Australian Academy of Science since 1985, and in 1998 was awarded the Pitman Medal of the Statistical Society of Australia.

Herbert Sloan is Associate Professor of History and Director of the First-Year Seminar Program at Barnard College, Columbia University. Educated at Stanford (B.A.) and Columbia (Ph.D.), he is interested in the political and financial history of the Early Republic. Among his publications are "The Earth Belongs in Usufruct to the Living," in Peter S. Onuf, ed., *Jeffersonian Legacies* (Charlottesville, Va., and London: University Press of Virginia, 1993), and *Principle and Interest: Thomas Jefferson and the Problem of Debt* (New York and Oxford: Oxford University Press, 1995), both of which pay particular attention to the ways in which Jefferson drew striking political consequences from the notion of generational rights, a notion whose practical effects depended on Jefferson's ability to calculate the "life" of a generation. Recently, like Patricia Cline Cohen, he has found crime the perfect relief from numbers and is now working on the life of Nancy Randolph, a late-eighteenth-century Virginian who may or may not have been guilty of incest, infanticide, and murder.

Daniel Scott Smith is professor of history at the University of Illinois at Chicago where he teaches early American history and various topics in social history. He received his Ph.D. in history from Berkeley in 1973 and a Certificate in Demography from the Office of Population Research, Princeton University, in 1974. He has received fellowship support from the Population Council, the American Council for Learned Societies, the Newberry Library, and the Center for Advanced Study in the Behavioral Sciences in Stanford, California. During the 1980's he edited the journal, *Historical Methods*, and in 1986-87 served as president of the Social Science History Association. Currently he is on the editorial boards of *Historical Methods, Social Science History*, and *The History of the Family*, and previously served on the boards of *Continuity and Change* and *The American Historical Review*.

He has published extensively in American family, social, and demographic history on such topics as old age and the family, the distinctive features of the preindustrial American family and demographic system, and the role of cultural factors in the decline of fertility in the United States during the 19th century. Among current projects are studies on demography of participation in the Continental Army during the American Revolution and for the North during the Civil War. Next year the *William and Mary Quarterly* will publish his article on the generational demography advanced by Thomas Jefferson in support of his proposition that "the earth belongs in usufruct to the living."

Robert V. Wells is the Chauncey H. Winters Professor of History and Social Sciences at Union College in Schenectady, New York, where he has taught since 1969. He received his Ph.D. from Princeton University in 1969, with a study of Quaker demography in the eighteenth century. He is a specialist in both early American and American demographic history. He teaches courses in both American demographic history and on the era of the American Revolution, including on the writing and ratification of the Constitution. He has been a visiting professor at Stanford University and the State University of New York at Albany. He has just returned from a year at the Center for American Studies at Odense University in Denmark, where he was a Distinguished Professor under the Fulbright program. While in Europe, he traveled to Cambridge, England, Paris, France, and Lisbon, Portugal to present his current work. He has also been a fellow at the Charles Warren Center at Harvard University and of the John Simon Guggenheim Memorial Foundation. Other grants have come from the National Institutes of Health and the American Philosophical Society.

His publications include: *The Population of the British Colonies in America before 1776: a Survey of Census Data* (Princeton, 1975); *Revolutions in Americans' Lives* (Greenwood, 1982); *Uncle Sam's Family* (SUNY, 1985) and numerous articles on various aspects of American demographic history. His most recent work is on American experiences with and attitudes toward death, and will be published by Cambridge University Press in 1999 under the title, *Facing the King of Terrors: Living with Death in an American Community*, Schenectady, New York, 1750-1990.

Other Workshop Participants:

Michelle Ash – Minority Counsel, Committee on Government Reform, U.S. House of Representatives, Washington, D.C.
Stuart Banner – Fellow, Woodrow Wilson International Center for Scholars, Washington, D.C.; and Associate Professor of Law, Washington University, St. Louis, MO.
Tom Brunell – APSA Congressional Fellow, Subcommittee on the Census, Committee on Government Reform, U.S. House of Representatives, Washington, D.C.
Carolee Bush – Bureau of Transportation Statistics, Washington, D.C.
Dan Cork – Graduate Student, Department of Statistics and the Heinz School of Public Policy & Management, Carnegie Mellon University, Pittsburgh, PA.
Tom Hofeller – Staff Director, Subcommittee on the Census, Committee on Government Reform, U.S. House of Representatives/Washington, D.C.
Richard John – Fellow, Woodrow Wilson International Center for Scholars, and Associate Professor of History, University of Illinois, Chicago, IL.
Marvin Kosters – American Enterprise Institute, Washington, D.C.
Michael Lacey – Director, United States Studies Program, Woodrow Wilson International Center for Scholars, Washington, D.C.
Robert K. Landers – Senior Editor, *The Wilson Quarterly*, Woodrw Wilson International Center for Scholars, Washington, DC.
TerriAnn Lowenthal – Project Consultant, Census 2000 Initiative, Washington, D.C.
David McMillen – Subcommittee on the Census, Committee on Government Reform, U.S. House of Representatives, Washington, D.C.
Janet Norwood – Urban Institute, Washington, D.C.; and former Commissioner of Labor Statistics, U.S. Bureau of Labor statistics, Washington, D.C.
Kenneth Prewitt – Director, U.S. Bureau of the Census, Washington, DC
Representative Tom Sawyer-U.S. House of Representatives, Washington, D.C.
Charles Schultze – The Brookings Institution, Washington, D.C.
Ed Spar-Executive Director, Council of Professional Associations on Federal Statis-

tics, Alexandria, VA.
Miron Straf – Director, Committee on National Statistics, Washington, D.C.
Katherine K. Wallman – Chief Statistician, Office of Management & Budget, Washington, D.C.
Andrew A. White – Deputy Director, Committee on National Statistics, Washington, D.C.
Don Woffensberger – Fellow, Woodrow Wilson International Center for Scholars, Washington, D.C.; and former Chief of Staff, Committee on Rules, U.S. House of Representatives, Washington, D.C.
Andrea Worden – Attorney, O'Melveny & Myers, LLP, Los Angeles, CA.

B Report on Workshop from *The Wilson Quarterly*, Winter 1999, p.141

"The First American Census in Methodological Perspective"
A workshop, November 12-14, 1998, at the Woodrow Wilson International Center for Scholars, Washington, D.C. The workshop was conducted with support from the Donner Foundation.

The litigation and controversy over the prospective use of statistical sampling in the 2000 census have sent scholars back to the Constitution and the first U.S. census, in 1790, in search of guidance. The main conclusion to emerge from this workshop–cochaired by Margo J. Anderson, a Wilson Center Fellow and historian at the University of Wisconsin-Milwaukee, and Stephen E. Fienberg, a professor of statistics and social science at Carnegie Mellon University–seemed to be that, contrary to some claims, the Founding Fathers have precious little guidance to offer. The GOP-controlled House of Representatives has challenged in court the Clinton administration's plans to use statistical sampling to correct for the large number of black Americans–an estimated 5.7 percent in 1990, compared with 1.3 percent of whites–and other minorities who will not be counted in the census. (The uncounted are mostly in poor urban neighborhoods.) In November, the Supreme Court heard arguments in that case, as well as in a related lawsuit brought by private plaintiffs. Lower courts had ruled against the administration in both cases.

The Constitution originally provided (Article 1, Section 2) that members of the House of Representatives were to be apportioned among the states "according to their respective Numbers," and that, "The actual Enumeration" would take place "within three Years after the first Meeting of the Congress," and every 10 years thereafter. Opponents of statistical sampling have put much weight on the phrase "actual Enumeration," contending that the Framers wanted a head count, not an estimation. Thomas Jefferson, who as secretary of state was in charge of the first census, "was familiar with methods of statistical estimation, having used them effectively in his 1782 survey of Virginia's population," the House brief in the current lawsuit claims, but did not use them to adjust the 1790 census results.

However, while Jefferson had "demonstrated considerable practical ingenuity in producing estimates in the absence of a census," says Daniel Scott Smith, a historian at the University of Illinois at Chicago, he did not draw inferences from a sample. The French mathematician Pierre Simon de Laplace was at work on probability theory in France, but Eugene Seneta, a professor of mathematical statistics at the University of Sydney, Australia, says, after an investigation of the matter, that there is no evidence that Jefferson had any knowledge of it.

The federal government has never attempted to make a physical headcount of everyone in the country, Anderson and Fienberg note. Rather, heads of households have been asked, in person or by mail, to report on their households. Nor does the phrase “actual Enumeration” seem laden with any great significance. The Framers, observes Seneta, “knew nothing of sampling as such, and could not have rejected its use.” Reviewing the legislative history of the 1790 law authorizing the nation’s first census, Charlene Bickford, director of the First Federal Congress Project, points out that the Senate struck out the word actual from both the title and the text of the law. Apparently, the Senate did not consider the adjective as adding anything vital to the noun.

The Framers of the Constitution seem to have paid little attention to how the census was to be carried out. Indeed, censuses, conducted at England’s request and in various ways, were common occurrences in the colonies during the 18th century, notes Robert V. Wells, a historian at Union College.

Robert Landers is a senior editor at the Wilson Quarterly.

C Statistical Notes On The First American Census in Methodological Perspective

by
Eugene Seneta[2]
School of Mathematics and Statistics
University of Sydney, NSW 2006, Australia.
E.Seneta@maths.usyd.edu.au

C.1 Questions Addressed

The U.S. Constitution, was drafted in 1787 in Philadelphia and ratified in 1788, with input from **Thomas JEFFERSON (1743-1826)**, who was U.S. Minister to France, 6th August 1784 - 18 September 1789, and a polymath in his interests, with some statistical work already present in his *Notes on the State of Virginia* (1782). He took up duties in New York as Secretary of State on March 21, 1790 and was in this capacity the official in charge of the first American Census of 1790. The U.S. Constitution determines that Congress is responsible for Census enumerations. In the review article "Statistical work of the Federal Government of the United States" by John Cummings, Statistician, United States Bureau of Census [3] Jefferson is not mentioned, although the first census is discussed (pp. 670-671).

The questions addressed in this appendix after Section /refsec:appcsec2 below are:

- What might Jefferson have known of the methodology of statistical estimations, and of enumeration, of populations at the time of ratification of the U.S. Constitution and of the first American Census?
- What was the understanding and application of sampling ideas to populations in France and the Russian Empire?

C.2 Use of Sampling in the U.S. Census in the Year 2000

The U.S. Constitution determines that Congress is responsible for Census enumeration, hence the legal challenges on the basis of U.S. Federal Law to the introduction

[2]Acknowledgements: I thank for advice at short notice: Steve Stigler, Susan Linacre, Bernard Bru, Pierre Crépel, Ian Castles, Jane Barton, Fred Hoppe and Ludmilla Seneta; and Viola Chao for word-processing.

[3]In J. Koren (ed.)(1918) *History of Statistics*, American Statistical Association, pp. 573-689.

of statistical estimation procedures, and the appeal hearing in the U.S. Supreme Court on November 30th, 1998. In Australia regular PES (post-enumeration surveys) using sampling have long been held to determine for each State its share of the total population and hence its share of House of Representatives seats at Federal level. (Elections occur at greater frequency than Censuses.) In contrast to the U.S., the Australian Constitution (of 1901, the time of Federation of separate Colonies) largely influenced by that of the U.S., vests responsibility for this in the Commonwealth Statistician, who heads the Australian Bureau of Statistics. The sampling procedures have not been controversial, or challenged by the States.

C.3 The Statistical Background for the U.S. Constitution and First Census

When writing his *Notes on the State of Virginia* before coming to Paris, Jefferson used the *Histoire naturelle, générale et particuliére* of **George Louis Leclerc BUFFON (1707-1788)**, who asserted that the certainty of a physical truth is to be measured by the probability of the corresponding facts, and in whose *Supplément à l'Histoire naturelle, Vol.IV* the "Essai d'arithmétique morale" of 1777 is concerned with "the measurement of uncertain things". Buffon had already fallen back on statistical methods in his paper "De la vieillesse et de la mort" in Vol. II, 1749, of the *Histoire naturelle* in which he studied human mortality in general, using Dupré de St. Maur's mortality tables. His statistical interests were in life tables rather than population enumeration. Jefferson used Buffon's life table in his letter from Paris of Sept. 6, 1789, to Madison.

Jefferson in his *Notes* took issue with Buffon's conclusions as a naturalist, particularly in reference to the United States, and their later scientific discussions in Paris were on natural history. I have found no evidence that Buffon's probabilistic thinking influenced Jefferson on population enumeration issues. In his own catalogue of his library, the Buffon works are under Chapter 12: Natural-History-Animals.

Chapter VIII ("The number of its inhabitants") of Jefferson's *Notes* of 1782 has two distinct statistical themes. One is projection of population size on the basis of the estimate that the population of Virginia is doubling every 27.25 years, an estimate, as he says, "with a considerable degree of precision". This idea has the spirit of **Thomas Robert MALTHUS (1766-1834)** about it; and indeed Jefferson eventually had Malthus' book (in 2 volumes) in his library, but that was first published in 1798, and a 2-volume edition not till 1806. The second theme is how *to estimate* the total number of inhabitants of Virginia in 1782 (he arrives at a final figure of 567,614), from an enumeration of some kind giving 53,289 free males above the age of 21, 211,698 slaves of all ages and sexes, and 23,766 "said to be tytheable slaves" - that is, above the age of 16; with "no returns from eight counties". He makes rather crude assumptions on the basis, for example, of "a former experiment" to estimate his final figure, from what one might consider to be a (highly non-random) *sample* of size 288,753, which is just 51% of his final result. [4]

Jefferson's contacts from 1785 with **Marie-Jean-Antoine-Nicholas Caritat de CONDORCET (1743-1794)** were frequent as the two men had similar intellectual interests. Condorcet's (1785) *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix* appeared in Paris during Jefferson's appointment, and a copy may have been presented to him there. He possessed a copy on his return to the U.S. Parts of the book are on voting systems and public choice, are now recognized as being pioneering in these areas, but appear to have been beyond the interest or understanding of Jefferson or Madison. There was fruitful exchange of ideas of political and technical kind, but Condorcet's writings on probability and public choice were not among them.

It is worth mentioning that **John ADAMS (1735-1826)** was from 1785 first US Minister to Great Britain and was also writing back to the Constitutional Convention, was on splendid terms with **Richard PRICE (1723-1791)** who was the leading British authority on mortality tables and pensions, as well as a civil libertarian. Adams however was not inclined to quantitative matters; and regarded Condorcet as a "mathematical charlatan".

I have found no evidence that Jefferson during his sojourn in Paris, met with, or knew of the work on the probabilistic estimation of population size by **Pierre-Simon de LAPLACE (1749-1827)**. Condorcet and Laplace detested each other. But Jefferson may have visited **Jean-Baptiste-François de la MICHODIÈRE (1720-1797)** who was, prior to the French Revolution, a leading authority (writing under the pseudonym Messance) on the estimation of population size (though not by probabilistic methods) and knew Laplace.

The work of la Michodière and Laplace is most relevant in addressing the Questions above; it is described by Bernard Bru. [5]

A number of other potential, but not probabilistic, statistical influences on Jefferson in regard to populations include **Pierre-Samuel DUPONT DE NEMOURS (1739-1817)**. A *Comité de l'Agriculture* created in 1782 ordered Dupont (as was then his name) to present to it a summary of the value of the crops of the kingdom. Another member of the then *Comité*, the celebrated chemist **Antoine Lau-**

[4] The number of free males under sixteen should read as 71,052. In at least some printings it reads as 17,052.

[5] Bru, B. Estimations Laplaciennes. Un Exemple: La Recherche de la Population d'un Grand Empire. 1785-1812. *Statistique et Analyse des Données*. 1988. Vol. 13, No.2, pp.3-42.

rent LAVOISIER (1743-1794) prepared a large work by 1784 building on this, which after the Revolution the Constituent Assembly ordered printed in 1791 [6] which alluded to methods of determining the population size. On the basis of this the Constituent Assembly resolved to proceed to a complete enumeration of the French population. (However no census was held till 1801, after the establishment of a Central Statistical Office under another régime.) Dupont was on a *Comité du Commerce* during Jefferson's time in Paris as now *inspecteur général du commerce* and Councillor of State. The task was in part was to study problems of Franco-American trade. Dupont de Nemours emigrated to the U.S. in 1799, returned to France in 1802 helping Jefferson with the Louisiana purchase, fled to the U.S. in 1815, and died there. His economic theories had some influence on U.S. policies. The now multinational chemical company bears his name. His close colleagues Condorcet and Lavoisier died in the Terror.

According to several authorities, Jefferson's expression in mathematics went little beyond arithmetic calculations. His own catalogue [7] accompanying the sale of his books to the Library of Congress reveals an enormous breadth, even for a polymath; we focus on books relevant to the preceding discussion, as he described them: Under "Mathematics. Pure. Arithmetic": De Moivre's Doctrine of Chances; L'analyse de la Probabilité des décisions, par Condorcet; Mémoires Mathématiques de Diderot; Price on Annuities. Under "Political Economy - General - Statistics - Commerce - - - Finance": Petty's Political Arithmetic; Blodget, Statistical Manual for the U.S. of America; Statistique élémentaire de la France, par Peuchet, 1805 [Peuchet was an authoritative and overt critic of the first French Census of 1801 - E.S.]; Malthus on the principles of population, 2v. The only work of Laplace in Jefferson's catalogue is under "Astronomy": Exposition du Système du Monde, par la Place, 2v. The other works of Condorcet, and of Dupont, tend to be linked and there are many, possibly all: for example, under "Modern History. Foreign" and under "Special Governments, Modern". Under "Polygraphical" there is: Encyclopédie de Diderot et D'Alembert, 39v. Lausanne.

It can be said with a high degree of certainty that the Founding Fathers of the U.S. Constitution knew nothing of sampling as such, and could not have rejected its use. There is a once much-quoted opinion [8] of **Moreau de Jonnes** " ... that American founders looked only to practical ends ... a careful search through the 'Madison papers' has failed to reveal that any member of the Convention considered the census in its scientific bearings".

[6] *Résultats extraits d'un ouvrage intitulé: de la Richesse territoriale du royaume de France.*

[7] Gilreath, J. and Wilson, D.L. (Eds.) (1989) *Thomas Jefferson's Library*: L.C., Washington.

[8] Cummings, p.670.

Further, from the forgoing discussion, one could presume that Jefferson, with his scientific and inventive turn of mind, had he known of and understood la Michodière's and Laplace's work in the U.S. context, would not have been averse to the cautious use of probabilistic sampling and estimation. His *Notes* of 1782 already support this view.

C.4 France: Sampling

Lavoisier (1791) refers to the books of la Michodière (Messance) and de Montyon (Moheau) regarding the number of inhabitants, and estimates without details the population of France for about 1784 at about 25 million. These two were the most remarkable of the 32 *intendants* of the provinces (*généralités*) towards the end of the *ancien régime*. The *intendant* was the supreme administrative head of each province. Their archives form principal sources for statistical studies of pre-revolutionary France. Only a few provinces attempted head-by-head enumeration, twelve used a census of households, some borrowed figures from tax registers. Enumeration was regarded so difficult and expensive as being unreasonable; and an approximation to population size of France was obtained by synthesizing figures from the provinces. By *mid 18th century* it was agreed to adopt as an aid "the least uncertain index", the number of births, since registers of these were carefully kept.

The procedure was as follows for a province. A number of parishes was "carefully selected", and in these a complete enumeration of persons made. This number was divided by the mean number of births over the six preceding years in the parishes. Because of the variability in this ratio (it fluctuated wildly between provinces), an "average figure" was obtained to be used in every province, which was then multiplied by the total number of births in the province to give an estimate of total population size in the province. Thus in modern terms the essence of this was to calculate, on the basis of a *sample* (hardly random) of parishes, a *ratio estimator* of the population size. For the whole of France, these province totals could be added; but, it seems, de Montyon, Necker, and des Pommelles multiplied the total number of births in France (an average over a number of years) by a different ratio to get the ratio estimate (one thinks in modern terms of the "combined" and "separate" *ratio estimators from a stratified sample*). The estimates obtained were 24,000,000 (1774); 24,802,500 (1781); 25,065,883 (1784).

Laplace's involvement in the question of determination of population size appears to begin with a paper of 1786 in which he uses la Michodière's ratio estimate of 26 for the population size to the number of births in France and multiplies it by the average yearly number of births for 1781-1782, 973,054.5 to obtain 25,299,417. The innovation is that the ratio estimator is complemented by *proba-*

bility limits on its accuracy, and his primary purpose is the theoretical derivation of these. He determines the sample size of total number of persons required (771,469) to give an absolute deviation of at most a given size (500,000) with high probability (1000/1001). The prediction methodology is, as usual for the times, Bayes' Theorem with uniform prior; and the model used presupposes randomness in the sampling.

It is unlikely that even Laplace's mathematical colleagues understood at the time the significance of what he was trying to do; nor Jefferson, already in Paris, who might just have heard of it, as of the earlier ratio-estimation-sampling work, from la Michodière if he had met him.

The Bastille fell on July 14, 1789 while Jefferson was still in Paris. Laplace adjusted to the several changes of régime, and for a short period was Minister of Interior, thus, one supposes, with an intimate knowledge of demographic statistics of his time, before being replaced as Minister by Lucien Bonaparte who ordered the first census of the French population in 1801. This met with difficulties and the results were not in for some two years. In the meantime, Laplace persuaded the Government to carry out a *sample survey* to estimate the population size. This was carried out on September 22, 1802; about 2,000,000 persons were involved in the sample. The number of births used was for the 3 years preceding, September 22, 1799 to September 22, 1802. The results were reported, rather casually, in the editions 1812-1820 of Laplace's *Théorie analytique des probabilités*. Another unreliable Census had been held in 1806; then not repeated till 1821. If anything, Laplace's work had cast *doubts on the accuracy of a complete enumeration*. Laplace's own sampling procedure had the practical form of *a two-stage cluster sample*, but the random (representative sample) character was lacking: thirty *départements* were chosen and in each of these a number of places where the mayor was intelligent and zealous.

By this time the first American Census of 1790 and its successors were well and truly over. Jefferson resigned as Secretary of State on December 31, 1793, was then Vice-President (1797-1801), and in retirement in the interim.

In fact the use of sampling to provide information of demographic and social kind went into severe decline as the method of total enumeration gained ground. **Georg von MAYR (1841-1925)** and his school objected to *indirect methods* such as sampling, and it began to gain ground only in the early years of the 20th century.

C.5 The Russian Empire: Sampling

Relevant Census and sampling activity here began considerably later than in France or the U.S.[9]

The first and only general enumeration, prior to Soviet times, of population, according to the authoritative **A.A. KAUFMANN** [10] **(1864-1919)**, took place in 1897. Before this time Russian official statistics had to rely on means such family registers kept for the purpose of recording persons eligible for military service. Administrative and police estimates "were made through the aid of local statistical offices, ... , and the district community administration". The data was "compared with the data in the Central Office [Central Statistical Office, est.1858, within the Ministry of Interior] and published after having been carefully tested". Plans for a complete enumeration of population date to the 1860's. There was no use of sampling methodology in this context.

There were, however, studies dating from 1877 of the agricultural and the rural economy coordinated by the Central Statistical Office. A Russian Agricultural Census took place in 1916. Some 18,000,000 landholding cards were obtained. As part of the investigation of variability within the rural economy it was decided to *sample* these cards. Only the first (methodological) part of the investigation was published (in 1917) due to the Russian Revolution. Its author **S.S. KOHN (1888-1933)**, Assistant Director, Russian Agricultural Census, ascribes the formula for *variance of sample mean for a simple random sample without replacement* to his former teacher at the St. Petersburg Polytechnic Institute, **A.A. CHUPROV** (or **TSCHUPROW) (1874-1926)**.

Sampling had been used earlier in statistical investigations of the rural economy in Russia within institutions of local government called *zemstvos*, established from 1864. (There as an analogy with the provinces of the *ancien régime* in France of a century before.) However the measure-of-precision approach akin to Laplace's in France, but now of a truly random sample without replacement, is credited to S.S. Kohn and A.A. Chuprov.

Within the zemstvo investigations, the idea developed that it was not necessary to put "certain of the questions contained in the community schedule in every instance... it would be ordered that every tenth or twentieth person in alphabetical or

[9]We asked Professor Seneta to describe this particular component of the development of scientific sampling primarily to illustrate how remote the developments actually were from the events surrounding the writing of the U.S. Constitution and the taking of the the first U.S. Census. Other notable contributtors to the development of sampling at the end of the 19th and beginning of the 20th centuries include: Kiaer (Norway), Bowley (England), Neyman (England and Poland). *The Editors.*

[10]In Koren (ed.)(1918), pp. 469-534; *Novyi Entsiklopedicheskii Slovar* **18**, [1911-1916], cols. 617-632.

other mechanical order should be questioned". In 1896 **A.V. PESHEKHONOV (PJESCHECHONOW) (1867-1933)** took a mechanical (that is, as we now say, *systematic*) sample of every 10th landholding, and **V.G. GROMAN** (or **GROHMANN)** used systematic sampling extensively in 1911-1913. It should be kept in mind that this method amounts to the *random selection of a cluster*.

The above account alludes to the use of the systematic sample and purely random sample as independent methods of statistical investigation, dictated by necessity in the presence of vast data sets. The more general tendency in Russia as elsewhere was "monographic" sampling. This pervaded early attempts at representative sampling. The "monographic method" amounted to selecting a cluster of elements "typical" of a population in relevant respects, and their extensive investigation. The method was generally used as an adjunct to a foregoing (local) complete enumeration, which might be used to determine the "typicality" of the elements to be sampled. Complete enumeration would be the only method used to check the precision of estimates obtained from such monographic samples. **A.I. CHUPROV (1842-1908)** proposed the idea in 1894 of *monographic sampling as a useful independent* means of investigation (not merely as an adjunct to complete enumeration) but the idea of a *random sample* did not begin to take root till a paper of his son A.A. Chuprov in 1910, and was extinguished by the aftermath of the Russian Revolution.

Kaufmann's article (of 1918) makes comparison with the American system (p.528: "The prevailing method ... in collecting zemstvo statistics is modelled on the American correspondent system") whereby a village representative fills in a schedule. (p.531: A very special technique of enumeration has been devised which at all points departs from the west European forms and in many respects approach the American). There was therefore an awareness of what was happening in the U.S., and, pre-Revolution, good progress in sampling methodology in the Russian Empire. It had no influence outside of the Russian Empire, however, and was implemented in just a few zemstvos. Although A.I. Chuprov tried to popularize among zemstvo statisticians the work on sample surveys of Kiaer, the influence of this work in Russia was minimal.

The idea of using some kind of sampling to supplement complete enumeration, however, was significant in the late 1800's.

D A Review of Statistical Sampling, From Laplace to Neyman

by
Daniel Cork
Department of Statistics and Heinz School of Public Policy & Management
Carnegie Mellon University, Pittsburgh, PA
cork@stat.cmu.edu

D.1 Overview

The publication of Jerzy Neyman's 1934 paper "On the two aspects of the representative method" [15] stands as a landmark event in the development of modern statistical sampling theory. Among the statistical community, the Neyman paper did much to resolve decades, even centuries, of contentious debate over the appropriateness of basing inference on samples rather than complete enumerations. Random sampling was given a strong theoretical foundation in the paper, and its advantages over other attempts at "representative" sampling were clearly demonstrated.

In the notes that follow, we briefly review the state of statistical sampling before Neyman, from early work in ratio estimation to the pioneering work in sample surveys carried out in Norway and Britain. In particular, we discuss the results of two influential meetings of the International Statistical Institute (ISI) in 1895 and 1925, the first of which led to a savage rebuke of the notion of sampling and the second of which represented an official embrace of the idea. We then describe the "representative method" as it stood in 1925, concluding with an overview of Neyman's findings and contributions in the 1934 paper.

D.2 The Measure of Empires

The notion of drawing inference from sample information dates back to at least 1662, with the development of ratio estimates of population. The basic idea of the estimate was to combine official (and seemingly reliable) vital statistics data with samples from a local area; by dividing the official number of births in an area by a sample estimate of the local birth rate, one could approximate the area's total population. The French statistician Pierre-Simon de Laplace derived the distribution of the ratio estimator by assuming the sample and the population to follow independent binomial distributions. Given his prestige and prominence in French society, Laplace was able to convince the newly-formed Central Statistical Bureau of France in 1800 to subsidize a sample survey to estimate France's population. Specifically, Laplace advocated constructing a sample of 30 departments (administrative divisions) that was representative in terms of geography; a complete enumeration of the residents in the sampled department and the mean number of births culled from local birth registers could then be used to obtain the population estimate [11].

However, censuses became the preferred and trusted mode of obtaining demographic and economic data. In post-Napoleonic 19th Century Europe, the various governments became fascinated with measuring their own growth and quantifying the characteristics of their people; as Laplace himself noted, the prevailing idea was that "the population is one of the surest means for judging the prosperity of an empire" [11]. Prussia's King Frederick William III created the continent's first true statistical bureau after being impressed by amateur statistician Leopold Krug's thirteen-volume *Topographical-Statistical-Geographical Dictionary of the Entire Prussian State*, with the French Central Statistical Office forming shortly thereafter in 1800. In general, the bureaucrats taking the helms at emerging statistical agencies saw their goal as "getting, by extensive use of census-taking and other enumerations, as complete a picture as could be obtained. This, as a rule, only called for elementary methods and they felt no need for complicated mathematical formulae" [11].

Dissatisfaction with both the time required and the final result of France's first census led the French government to approve Laplace's proposed sample-based estimate. But the French experiment was the exception rather than the rule, as the idea of complete enumeration came to dominate statistics collection. A notable clash between the competing ideas of enumeration and sampling occurred in Belgium in 1827, where the statistician Adolphe Quetelet grew influenced by Laplace's methods during a long visit in Paris three years earlier [23, 11]. Quetelet, perhaps best known today for originating the concept of *l'homme moyen* or "the average man," used his post as an adviser to Belgium's statistical office to promote a Laplace-style study to estimate the Belgian population. But another adviser, the Baron de Keverberg, urged the abandonment of this idea. He argued that the complex mix of factors that could plausibly affect local birth rates are inhomogeneous. The only way to account for this diversity would be to select units from a much finer partition of the country than large departments—such a fine partition that the time and cost would be the same as a complete census. Quetelet ultimately deferred to de Keverberg and advocated a complete census, performed in 1829.

D.3 Kiaer and Bowley

Toward the end of the 19th Century, Norway's Central Bureau of Statistics began the move toward generating official statistics through sample surveys. In particular, this initiative was pushed forward by Anders Nicolai Kiaer, who directed the agency from its inception in 1876 through 1913. The impetus for using samples versus complete enumerations was simply the increased volume of duties and questions put to the agency; in addition to conducting a decennial population and agricultural censuses, the agency was frequently charged with surveying the Norwegian population for other investigations. It was a request to study the population's views on a proposed pension and insurance system that became the first of many studies directed by Kiaer that drew from a sample rather than a complete census. This initial survey included 120,000 respondents; as Kiaer's confidence in sampling grew, the sample size was reduced to 10,000 by his second major survey.

Kiaer's mode of sampling the Norwegian population has, in retrospect, been described as a "well-worked-out method of stratification, the stratification factors being geographic, social and economic" [21]. It was not, however, a random sample; rather, Kiaer promoted the idea of drawing samples as if to create the world in miniature; he would purposively select areas of the country (districts, towns, etc.) in a seemingly "representative" way, and then systematically sample units within those strata (e.g., choosing particular ages and first letters of surnames). His samples were constructed so that they "agreed in important characteristics with the population at large"—these characteristics being findings from earlier censuses [10]. Kruskal and Mosteller note that, "for example, if a sample had a deficiency of cattle farmers, he would add more of them" [14].

In later work, Kiaer anticipated other features of modern sampling, such as replication and subsampling. At the ISI's 1901 meeting in Budapest, Kiaer advocated checking the representativeness of a sample by "[dividing] the investigation into two or three separate parts, each itself a representative study. If the parts give similar results, one cannot deny the validity thus gained" (quoted in [14]).

After the turn of the century, the "representative method" pioneered by Kiaer drew gradual acceptance, most forcefully in the work of Sir Arthur Lyon Bowley. Bowley linked survey sampling and statistical inference and played a key role in establishing basic ground rules of representative sampling. In 1906, he used a version of the Central Limit Theorem to show that, for large samples from a finite population, the population proportion is approximately normally distributed. In a chapter of his 1915 book *Livelihood and Poverty*, Bowley broke new ground by discussing measurement and sampling error, and in proposing estimation by intervals of plus or minus three times the sampling error came very close to developing the theory of confidence intervals two decades ahead of its time [14].

Perhaps most importantly, though, Bowley's work showed an eye for detail in conducting surveys set a high standard for future samplers to follow. As Sir Maurice Kendall noted, sampling "still rests on the principles on which [Bowley] insisted, the greater care which could be taken to secure accurate responses with trained interviewers confining their efforts to a moderately sized sample, the importance of avoiding bias, and the ascription of precision in quantified terms to the results" [14].

D.4 Formal Acceptance of Sampling: The ISI meetings of 1895, 1903, and 1925

The early reluctance to accept sampling and its ultimate approval by the statistical community is best seen in the records of three meetings of the International Statistical Institute (ISI). In 1895, Kiaer traveled to Berne to present his paper "Observations and experiences with representative surveys"—an advocation of sampling—at the ISI meeting. His audience, the ISI, was then largely composed of leaders of government statistics bureaus—staunch believers in complete enumeration. Predictably, Kiaer's presentation touched off a heated response. In the discussion, Georg von Mayr of the University of Munich was particularly vocal in denouncing the concept of sampling, regarding the idea as "most dangerous." The value of random samples is "restricted to terrain already illuminated by full coverage," von Mayr continued; "one cannot replace by calculation the real observation of facts. A sample provides statistics for the units actually observed, but not true statistics for the entire nation.. . . Perhaps for legislative or administrative goals sampling may have uses [but] we must remain firm and say: no calculations when observations can be made" (quoted in [14]).

The following decade would cool tempers somewhat, though the von Mayr school of statisticians continued to opposed the "representative method." By 1903, though, the ISI took an early step toward legitimizing sampling in 1903. At its meetings that year, a subcommittee passed a resolution stating that "the Committee, considering that the correct application of the representative method, in a certain number of cases, can furnish exact and detailed observations from which the results can be generalized, within certain limits, recommends its use." That use, though, was conditional on a clear explanation in the publication of results of "the conditions under which the selection of the observation units is made" [21].

By the time of the 1925 ISI meeting in Rome, and following the continued success of Kiaer and Bowley in using sample surveys, "the discussion is not whether to do sampling, but how to do it" [14]. In 1924, the ISI named a Commission to study the representative method and report at the Rome meeting. Bowley served on the commission, as did several government statisticians. This commission generated three papers, including Bowley's "Measurement of the precision attained in sampling". Bowley's paper itemizes a set of ground rules that are still in use today: the construction of a sampling frame that defines exactly the population being surveyed, the selection of a method for sampling, the observation of all sampled units, and attention to possible biases and errors arising from violations in the sampling plan [11]. In his remarks at the 1925 meeting, Bowley stressed the fact that researchers engaged in sampling must have "a complete absence of prejudice and a perfect willingness to accept the results, however unpalatable.... The universe from which the selection is made must be defined with the utmost rigour, [and] the rules of selection... must be followed with perfect strictness" [14].

The end result of the 1925 ISI meeting was the passage of a formal resolution, repeating the 1903 resolution's wish that investigators detail their sampling method and urging that "the investigation should be so arranged, wherever possible, as to allow of a mathematical statement of the precision of the results" (in [11]). Crucially, though, the ISI report did not advocate random sampling but rather considered that the "representative method" could be employed in one of two ways: either random selection of units from the population or the "purposive selection of groups of units which it is presumed will give the sample the same characteristics as the whole" (quoted in [15]). Essentially, the ISI "recommended random selection of elements and very small clusters such as households and purposive selection of larger clusters" [12]. The real advantages of random sampling, and the real flaws of purposive selection, remained to be demonstrated.

D.5 The Neyman Contribution

One of the drawbacks of complete enumerations, particularly in the pre-computer era, is that the sheer volume of returns is unwieldy and takes up a great deal of space. The attempts of two governments, Italy and Poland, to deal with the sheer bulk of census returns, led to the landmark 1934 paper by Jerzy Neyman.

In November 1926, the Italian statisticians Corrado Gini and Luigi Galvani reported their approach to deal with a problem put to them by the Italian government: discarding most of the completed returns of the 1921 Italian general census while retaining a small sample (about 15%) to accommodate future, unforeseen research questions. Naturally, the hope was to retain a sample that was representative of the entire country. Given that the data were already sorted along a geographic hierarchy—provinces, districts (circondarî), and communes—Gini and Galvani elected to retain the census records from 29 out of 214 circondarî. The

29 circondarî "were found such that their average values for seven important characteristics (birthdate, death rate, altitude, etc.) were all close to the corresponding averages for the entire country" [14]. The problem, Gini and Galvani concluded, was that this sample did not seem particularly representative: when the sample was compared to population data on other demographic characteristics, there were large differences. Further, they found that the sample agreed on the seven characteristics *only* in mean, not in variability or association. Armed with these findings, Gini in a 1928 article "concludes... with a sharp criticism of sampling as such, almost a return to the older von Mayr approach" [14].

In 1932, confronted with dire economic crisis, the government of Poland faced a similar problem. Though they had the results of their first general census after World War I, the Polish government now needed to cull information quickly and efficiently; specifically, they wanted to obtain data on types and numbers of workers eligible for social insurance. To address this question, the quasi-governmental Institute for Social Problems employed Jerzy Neyman, a professor at the University of Warsaw. The Russian-born Neyman eagerly accepted the contract, one of many such assignments he accepted to provide continued economic backing for his students. Neyman wrote up the details of his proposed sampling system in a small booklet in 1933; in describing the work to his friend and statistical collaborator Egon Pearson in England, Neyman said that he was asked to investigate the representative method and that, in doing so, he "pushed a little the theory". He adds, "nevertheless there are many things, which seem to be new. Such is the point of view on the accuricy [*sic*] of the results.... I suppose that some paragraphs of the booklet are worth publishing somewhere in an international language" [19].

Shortly after moving to England to teach at University College, London, Neyman revised his booklet, "On the two different aspects of the representative method," for presentation before the Royal Statistical Society on June 19, 1934. In the 1934 paper, Neyman pits random sampling (carefully describing simple random sampling, stratified sampling, and cluster sampling) versus purposive selection (as carried out by Gini and Galvani) against each other. Neyman carefully outlined the conditions under which purposive sampling estimates could be expected to be unbiased and then demonstrates that those conditions generally fail to hold. Under stratified random sampling, though, Neyman found that the consistency of the estimates depends only on sampling from every stratum. He further derives the optimal number of items to draw from each stratum in order to minimize the variance of the estimates; this result is now known as "Neyman allocation."

From the statistical standpoint, Neyman's 1934 paper is memorable for one of the tools developed by Neyman to assess his results—the first complete theory of confidence intervals. Indeed, the Royal Statistical Society's discussion immediately following Neyman's presentation dealt almost exclusively by comments and criticisms on the new confidence intervals. But the importance of the paper to statistical sampling is enormous. As Smith [22] notes, "the only major features of current survey design that [Neyman] failed to introduce were multi-stage sampling and variable probability (p.p.s.) sampling), but these followed logically from his work." Aided by economic pressures and increased sentiment for sampling in general, Neyman's ideas rapidly took root; "within a period of 10 years the Neyman approach to inference had become almost universally accepted by practising survey statisticians" [22].

References

[1] Hyman Alterman. *Counting People: The Census in History*. Harcourt, New York, 1969.

[2] Margo Anderson. *The American Census: A Social History*. Yale, 1988.

[3] Margo Anderson and Stephen E. Fienberg. To sample to or not to sample: The 2000 Census controversy. *Journal of Interdiscipinary History*, 30(1):1–35, 1999.

[4] Margo Anderson and Stephen E. Fienberg. *Who Counts? The Politics of Census-Taking in Contemporary America*. Russell Sage Foundation, 1999.

[5] Michel Balinski and H. Peyton Young. *Fair Representation: Meeting the Ideal of One Man, One Vote*. Yale University Press, New Haven, 1982.

[6] D.R. Bellhouse. A brief history of random sampling methods. In P.R. Krishnaiah and C.R. Rao, editors, *Handbook of Statistics*, volume 6, pages 1–14. Elsevier Science Publishers, 1988.

[7] James H. Cassedy. *Demography in Early America, Beginnings of the Statistical Mind 1600-1800*. Harvard University Press, Cambridge, 1969.

[8] Patricia C. Cohen. *A Calculating People: The Spread of Numeracy in Early America*. University of Chicago Press, 1983.

[9] Alain Desrosières. *The Politics of Large Numbers*. Harvard University Press, Cambridge, MA, 1998. Translated by Camille Nash.

[10] Gerd Gigerenzer, Zeno Swijtink, Theodore Porter, Lorraine Daston, John Beatty, and Lorenz Krüger. *The Empire of Chance: How Probability Changed Science and Everyday Life*. Cambridge University Press, Cambridge, England, and New York, 1989.

[11] Anders Hald. *A History of Mathematical Statistics from 1750 to 1930*. John Wiley and Sons, New York, 1998.

[12] Morris H. Hansen and William G. Madow. Some important events in the historical development of sample surveys. In D.B. Owen, editor, *On the History of Statistics and Probability*, pages 75–102. Dekker, New York, 1976.

[13] Thomas Jefferson. *Notes on the State of Virginia. With an appendix relative to the murder of Logan's family*. Trenton, Printed by Wilson & Blackwell, 1785.

[14] William Kruskal and Frederick Mosteller. Representative sampling iv: The history of the concept in statistics, 1895–1939. *International Statistical Review*, pages 169–195, 1980.

[15] Jerzy Neyman. On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection. *Journal of the Royal Statistical Society*, pages 558–625, 1934.

[16] Jack Rakove. *The Beginning of National Politics: An Interpretive History of the Continental Congress*. Knopf, New York, 1979.

[17] Jack Rakove. *Original Meanings: Politics and Ideas in the Making of the Constitution*. Knopf, New York, 1996.

[18] Jack Rakove. Declaration submitted in glavin vs. clinton and city of los angeles, ca, proposed intervenor-defendants, 98-207-a (eastern district of va, alexandria division), April 28 1998.

[19] Constance Reid. *Neyman*. Springer-Verlag, New York, 1998.

[20] Rossiter. *A Century of Population Growth*. Government Printing Office, Washington, DC, 1909.

[21] You Poh Seng. Historical survey of the development of sampling theories and practice. In Maurice Kendall and R.L. Plackett, editors, *Studies in the History of Statistica and Probability*, volume 2, pages 440–457. Macmillan Publishing Company, New York, 1977.

[22] T.M.F. Smith. The foundations of survey sampling: A review. *Journal of the Royal Statistical Society, Series A*, 139:183–204, 1976.

[23] Stephen M. Stigler. *The History of Statistics: The Measurement of Uncertaity before 1900*. Belknap Press of Harvard University Press, Cambridge, MA, 1986.

[24] United States House of Representatives et al., vs. United States Department of Commerce, et al.; Glavin, Barr et al., vs. Clinton et al. and as argued before the Supreme Court: Department of Commerce v. House of Representatives, 11 F. Supp. 2d 76, appeal dismissed; No. 98-564, 19 F. Supp. 2d 543, affirmed.

[25] Robert Wells. *The Population of the British Colonies in America before 1776: A Survey of Census Data*. Princeton, 1975.

[26] Leonard White. *The Federalists: A Study in Administrative History*. MacMillan, New York, 1948.

[27] Carroll Wright and William C. Hunt. *History and Growth of the United States Census*. GPO, Washington, DC, 1900.