MINNESOTA POPULATION CENTER

UNIVERSITY OF MINNESOTA

Capturing the American People: Census Technology and Institutional Change, 1790-2020

Steve Ruggles†
University of Minnesota

Diana L. Magnuson
Bethel University

October 26, 2018

Working Paper No. 2018-2
https://doi.org/10.18128/MPC2018-2

†Address correspondence to Steve Ruggles, University of Minnesota, ISRDI, 50 Willey Hall, 225 19th Ave S., Minneapolis, MN 55455 (email: ruggles@umn.edu). We appreciate the helpful comments and suggestions of Margo Anderson, Miriam King, and David McMillen. Support for this work was provided by the Minnesota Population Center at the University of Minnesota (P2C HD041023).

Abstract

This article traces the history of the methods and technology used by the Census Bureau to convert individual census responses into published tabulations. We argue that political considerations not only shaped the content and applications of the census, but also the mechanics of census taking. By focusing on federal responses to a specific technical challenge over a very long span, our narrative illuminates the long-run effects of shifting societal preoccupations on bureaucratic decision-making. Technical progress was contested and uneven, with numerous setbacks; nevertheless, through the interplay of politics and necessity, for more than a century U. S. census operations defined the leading edge of innovation in data processing technology. From 1790 to 1990, census employees developed and refined a series of novel approaches to data capture that transformed data processing worldwide. In recent decades, the Census Bureau has relinquished this leadership position. We attribute the declining success of census data capture mainly to ideological shifts of the late 20th century that redefined the role of government. Beginning in the mid-1990s, the Census Bureau increasingly turned to outside vendors from the private sector for data capture. The privatization of data capture led to rapidly escalating costs, reduced productivity, near catastrophic failures of the 2000 and 2010 censuses, and high risks for the 2020 census.

James Madison, the chief architect of the first U.S. Census, argued that the enumeration of the population was "an opportunity of marking the progress of society, and distinguishing the growth of every interest." Accordingly, in January 1790 he proposed two census schedules. The first schedule covered the demographic basics, counting the number of "Free white males, under 16; free white males, above 16; white females; free blacks, and slaves." Madison's second schedule covered occupations, "specifying the number of persons employed in different professions and arts," including merchants, mechanics, and manufacturers.[1]

Madison's demographic schedule became the template for the 1790 Census, but the proposed occupational schedule was defeated amid concerns that the questions "would excite the jealousy of the people" and that the extra expense was simply "gratifying an idle curiosity." Madison wrote to Secretary of State Thomas Jefferson, who was responsible for directing the census, "It was thrown out of the Senate as a waste of trouble and supplying materials for idle people to make a book."[2]

A census is a political construct that reflects the ideological orientation of its creators. Legislators, intellectuals, and the public have contested the content and purposes of the United States census for 230 years. In each period, the meaning and uses of the census reflected the politics and priorities of the moment. In the 1850s, census planners suppressed information about slavery at the behest of Southern legislators; in the 1880s, the census director promoted nativist theories of race suicide; and in the 1940s census officials helped plan Japanese internment. The census is inherently political: Its original purpose was reapportionment of political representation, and in virtually every decade, winners and losers of the demographic contest have debated the legitimacy of the results. In one case—the Census of 1920—the results were ignored altogether and no reapportionment took place, as rural legislators feared losing power to the cities.[3]

---

[1] "Census," January 25 and 26, 1790, Charles F. Hobson, et al. eds., *The Papers of James Madison*, vol. 13 (Charlottesville, 1981); "Census of the Union," *Congressional Register*, January 25, 1870, pp. 1077-78.

[2] "Census of the Union," *Congressional Register*, February 2, 1790, pp. 1106-09. "Census," February 2, 1790, Hobson, *The Papers of James Madison*, vol. 13; Letter to Thomas Jefferson, February 14, 1790, Hobson, *The Papers of James Madison*, vol. 13.

[3] Margo Anderson, *The American Census: A Social History* (New Haven, 2015 2e). Miriam L. King, "'The quantum of happiness': The politics of American population debates, 1850-1930" (Ph.D. diss., University of Pennsylvania, 1990). Margo Conk, "The 1980 Census in Historical Perspective" in *The Politics of*

Political considerations not only shaped the content and applications of the census, but also the mechanics of census taking. This essay traces the history of U. S. census data capture, which we define *as the methods and technologies used to transform raw census responses into statistical tables.* By focusing on federal responses to a specific technical challenge over a very long span, our narrative illuminates the long-run effects of shifting societal preoccupations on bureaucratic decision-making.

Technical progress was contested and uneven, with numerous setbacks; nevertheless, through the interplay of politics and necessity, for more than a century U. S. census operations defined the leading edge of innovation in data processing technology. From 1790 to 1990, census employees developed and refined a series of novel approaches to data capture that transformed data processing worldwide. In recent decades, the Census Bureau has relinquished this leadership position. We attribute the declining success of census data capture mainly to ideological shifts of the late 20th century that redefined the role of government. Beginning in the mid-1990s, the Census Bureau increasingly turned to outside vendors from the private sector for data capture. The privatization of data capture led to rapidly escalating costs, reduced productivity, near catastrophic failures of the 2000 and 2010 censuses, and high risks for the 2020 census.

Household Enumeration, 1790-1840

The U. S. Census was the earliest regularly scheduled national-level population enumeration, and Madison's design included a highly efficient system for data capture. There were only a few earlier censuses to serve as models, and it is unlikely that Madison knew much about them. European powers carried out several enumerations beginning in the late 18th century. In general, these previous censuses listed all individuals in each household, beginning with the household head and followed by his wife, children, other relatives, servants, and slaves, if any.[4] The censuses usually reported the age of each person along with the

Numbers, ed. William Alonso and Paul Starr (New York, 1987), 155-186. Miriam King and Steven Ruggles, "American Immigration, Fertility, and Race Suicide at the Turn of the Century," *Journal of Interdisciplinary History*, XX:3 (Winter, 1990), 347-369. Paul Starr, "The Sociology of Official Statistics," in *The Politics of Numbers*, ed. William Alonso and Paul Starr (New York, 1987), 7-57.

[4] Gunnar Thorvaldsen, *Censuses and Census Takers—A History* (New York, 2018).

occupation of the head and the family relationship of other household members.

Madison's design represented a distinct break with this format of enumeration. Instead of listing the name and characteristics of each individual, the 1790 census listed the name of the household head and the number of household members with each characteristic (white men 16+, white men under 16, white females, other free persons, and slaves). This layout allowed for highly efficient data capture.

There was no standard printed form in 1790. Six hundred and fifty assistant marshals gathered the information using whatever paper was available to draw forms with the column headings specified by Madison's legislation. Although there was some variation from place to place, in general assistant marshals followed a consistent plan, such as the page from the enumeration of Greenville, South Carolina shown in Figure 1. The first household head listed is William Woody, and in his household, there was one white male aged 16 or older and one white female. Another William Woody—probably the first William's son—headed the second household. The second Woody household had the same configuration as the first, with the addition of five white boys under the age of 16.

The marshals kept a continuous count of the number of persons with each set of characteristics, updating the totals with the completion of each page. The first row of each column, labeled "Brought Over," or "Brought Forward," represents the number of persons of each type recorded on prior pages. The final row of each page gives the updated total, including all the families enumerated on the page. Those numbers were then copied to the top of the subsequent page. When an assistant marshal completed his division, he was instructed to "cause a correct copy, signed by himself, of the schedule, containing the number of inhabitants of his division, to be set up at two of the most public places within the same" so that the public would have an opportunity to make corrections. Once those copies had been publicly accessible for a reasonable period, the assistants reported the totals for each category of persons to the U.S. Marshal for their district, and claimed their payment of one dollar for every 150 persons enumerated.[5]

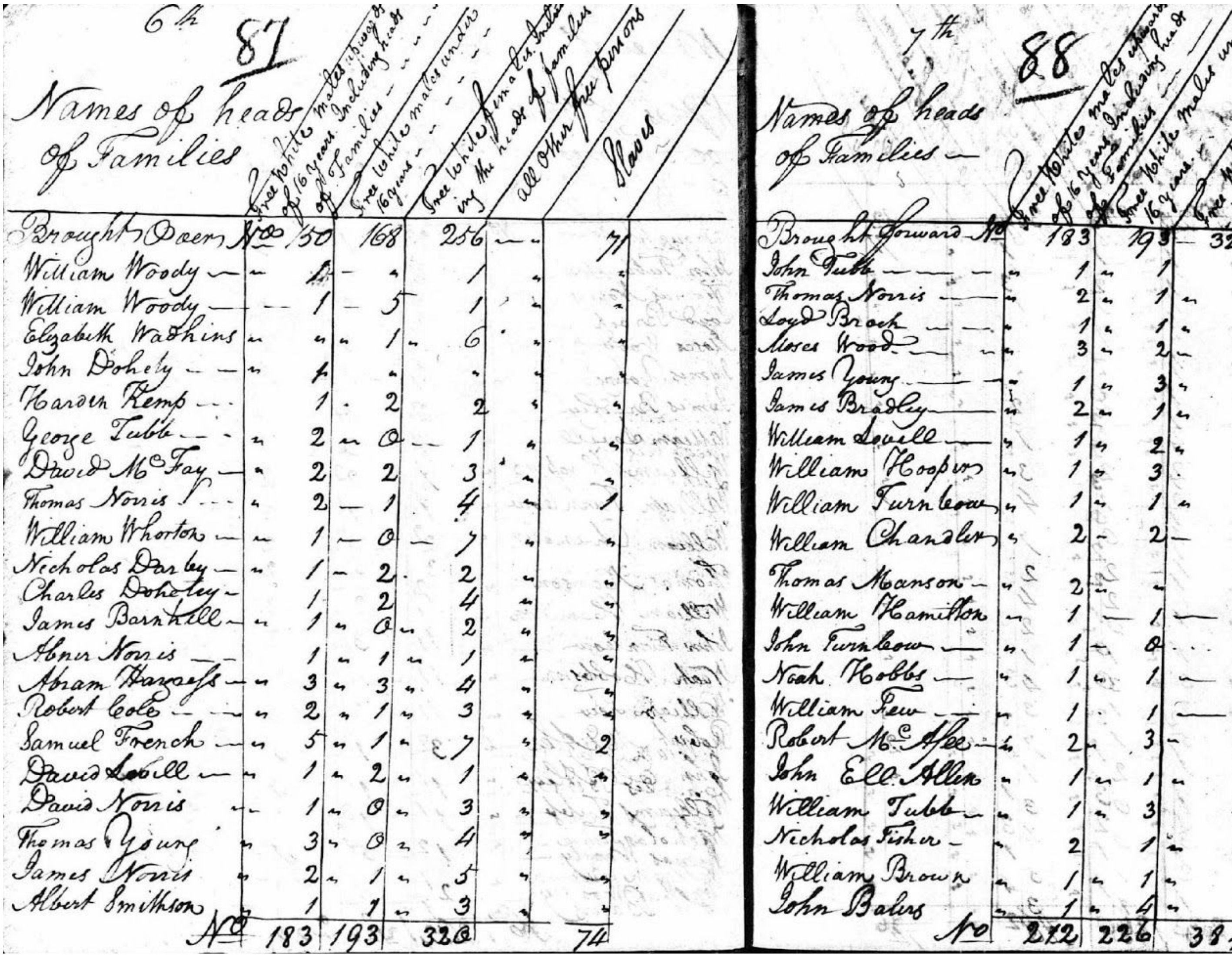
[5] *The Public Statutes at Large of the United States of America, from the Organization of the Government in 1789, to March 3, 1845,* Edited by. Richard Peters. Boston: Little and Brown, 1845, 101-103.

Under this system, data processing was decentralized. The assistant marshals carried out the tabulations for their divisions simply by summing each column as the pages were completed, and reporting the grand totals to the marshals. The marshals then had responsibility to "transmit to the President of the United States, the aggregate amount of each description of persons within their respective districts," along with the returns from each assistant. Secretary of State Jefferson published the results in a fifty-six page report in 1791, virtually just as he received them from the marshals.[6]

As the 1800 census approached, Congress came under pressure to collect more detail about the population. In January 1800, Yale College President Timothy Dwight submitted a recommendation on behalf of the Connecticut Academy of Arts and Sciences that the census should include seven age categories for males and females, occupational information, marital status, and detailed geographic information. Vice President Jefferson then submitted a memorial on behalf of the American Philosophical Society calling for

---

[6] *Return of the whole number of persons within the several districts of the United States, according to "An Act Providing for the Enumeration of the Inhabitants of the United States," Passed March the First, One Thousand Seven Hundred and Ninety-One* (Philadelphia: Childs and Swaine, 1791).



6th 87

Names of heads of Families	Free white males of 16 years & upwards including heads of Families	Free white males under 16 years	Free white females including the heads of families	all Other free persons	Slaves
Brought Over No	150	168	256	–	71
William Woody	1	–	1	–	–
William Woody	1	5	1	–	–
Elizabeth Watkins	–	1	6	–	–
John Dohety	1	–	–	–	–
Harden Kemp	1	2	2	–	–
George Tubb	2	0	1	–	–
David McFay	2	2	3	–	–
Thomas Norris	2	1	4	–	1
William Whorton	1	0	7	–	–
Nicholas Darley	1	2	2	–	–
Charles Doholey	1	2	4	–	–
James Barnhill	1	0	2	–	–
Abner Norris	1	1	1	–	–
Abram Harvess	3	3	4	–	–
Robert Cole	2	1	3	–	–
Samuel French	5	1	7	–	2
David Lovell	1	2	1	–	–
David Norris	1	0	3	–	–
Thomas Young	3	0	4	–	–
James Norris	2	1	5	–	–
Albert Smithson	1	1	3	–	–
No	183	193	320		74

7th 88

Names of heads of Families	Free white males of 16 years & upwards including heads of Families	Free white males under 16 years	Free w[…]
Brought Forward No	183	193	32[…]
John Tubb	1	1	
Thomas Norris	2	1	
Loyd Brock	1	1	
Moses Wood	3	2	
James Young	1	3	
James Bradley	2	1	
William Lovell	1	2	
William Hooper	1	3	
William Turnbow	1	1	
William Chandler	2	2	
Thomas Manson	2	–	
William Hamilton	1	1	
John Turnbow	1	0	
Noah Hobbs	1	1	
William Few	1	1	
Robert McAfee	2	3	
John Ell Allen	1	1	
William Tubb	1	3	
Nicholas Fisher	2	1	
William Brown	1	1	
John Balus	1	4	
No	212	226	33[…]

Figure 1. 1790 Census Enumeration Page, Greenville South Carolina
Source: First Census of the United States. National Archives and Records Administration microfilm publication M637-11, p. 87. Record Group 29. Washington, 1965.

Congress to provide information for twenty-two categories of age and sex. Such detailed age distributions, Jefferson argued, could provide insight into both mortality and population growth. Jefferson also called for information about foreign birth, citizenship, and nine occupational categories.[7]

Over the next several decades, Dwight and Jefferson got most of what they wanted. In 1800, Congress approved collection of some information on age distribution, with ten age categories, five for white males and five for white females. Following Dwight's recommendation, the 1800 census also recorded information on county, parish, township, town, or city. As a result, the number of columns of data to be collected rose from six to fourteen. The 1810 census added a separate enumeration of manufacturing establishments. In 1820 the census included three occupational categories and an enumeration of foreigners not naturalized, and for the first time collected age and sex information for both free colored people and slaves.[8] With these additions, there were thirty-three columns of data in 1820.

For the 1830 census, President John Quincy Adams requested still more age detail.[9] Congress approved this expansion, providing thirteen age groups ranging from 0-4 years to 100 and older for white males and white females and six age groups for slaves and colored persons of each sex. In addition, the 1830 census added the first questions on "defective classes": deaf and dumb (in three age groups) and the total number of blind persons. After these changes, the 1830 census had fifty-seven columns of data.[10]

To manage the greatly increased complexity of the 1830 Census, the Census Office introduced a major innovation: the printed census form. Before 1830, assistants to the marshals used whatever paper they happened to have on hand, hand ruling their paper and writing in the headings. With the expansion of the

---

[7] Garfield Report to Congress from the Committee on the Ninth Census. House Reports, Forty-first Congress, second session, Vol I, No. 3, p. 36. Letter from Connecticut Academy of Arts and Sciences, December 24, 1799, New Haven, CT. Barbara B. Oberg, ed., *The Papers of Thomas Jefferson*, vol. 31, 1 February 1799 – 31 May 1800 (Princeton, 2004), 293–295.

[8] Carroll D. Wright and William C. Hunt, *The History and Growth of the United States Census* (Washington, 1900), 17-27.

[9] James D. Richardson, ed., *A Compilation of the Messages and Papers of the Presidents, 1789-1908* (Bureau of National Literature and Art, 1908), 421.

[10] Wright and Hunt, *The History and Growth of the United States Census*, 28-32.

1830 census to fifty-seven columns, this approach was no longer feasible: large sheets of paper were necessary, and even then considerable precision was required to ensure sufficient space for all the columns.[11] The printed form did not make Adams' additions "trifling," but it made them possible.

The 1840 census expanded again, with new questions on Revolutionary War pensioners, literacy, occupations, schools, colleges, and the "insane and idiotic." Altogether, the printed form squeezed in eighty columns of summary information about thirty-one families onto two sides of a census form that was 18. 5 inches wide by sixteen inches long (Figure 2). Most columns were ⅜ of an inch (9. 5 mm) wide, and the rows were ⅜ inch tall.[12] As soon became clear, enumerators frequently made errors on the cramped form by recording information in the wrong row or column. The results were catastrophic: A post-mortem report to Congress by a committee of the American Statistical Association concluded that "it would have been far better to have had no census at all, than such a one as has been published."[13]

The most notorious error of the 1840 census was a finding that Northern blacks had much higher rates of idiocy and insanity than those in the South. Indeed, the rate of idiocy and insanity among Northern blacks increased directly with distance from the Mason-Dixon line: in Maine, the census indicated that one in fourteen colored persons were insane or idiotic; in Massachusetts, it was one in forty-three, and in New Jersey, one in 297. The rates were far lower in the Southern states, ranging from one in 1,229 in Virginia down to one in 4,310 in Louisiana.[14]

---

[11] "Instructions to Marshals—Census of 1830," March 24, 1830, reprinted in Wright and Hunt, *The History and Growth of the United States Census*, 139; U.S. Bureau of the Census, *The Story of the Census, 1790-1916* (Washington, 1916), 11; Wright and Hunt, *The History and Growth of the United States Census*, 29.

[12] U.S. Bureau of the Census, *200 Years of Census Taking: Population and Housing Questions, 1790-1990* (Washington, 1989), 20-21. A printed census form was in use in Denmark as early as 1801. Thorvaldsen, *Censuses and Census Takers*, 42.

[13] Edward Jarvis, William Brigham, and J. Wingate Thornton (1844). *Memorial of the American Statistical Association Praying the Adoption of Measures for the Correction of Errors in the Returns of the Sixth Census*. Public Documents Printed by Order of the Senate of the United States, Second Session of the Twenty-Eighth Congress. (Washington, 1844).

[14] Edward Jarvis, "Statistics of Insanity in the United States." *Boston Medical and Surgical Journal* XXVII (1842), 116-121. Albert Deutsch, "The first US census of the insane (1840) and its use as pro-slavery

This finding was widely discussed.[15] One observer attributed the results to climate, concluding that cold weather affected the "cerebral organs of the African race." Most Southern commentators pointed to "moral causes," arguing that freedom itself was the cause of lunacy among blacks. Secretary of State John C. Calhoun wrote that the census provided "unquestionable sources" demonstrating that "in all instances in which the States have changed the former relation between the two races, the condition of the African, instead of being improved, has become worse. They have invariably sunk into vice and pauperism, accompanied by the bodily and mental inflictions incident thereto—deafness, blindness, insanity, and idiocy—to a degree without example."[16]

The real source of the finding about race and insanity was uncovered by physician and statistician Edward Jarvis. Examining the returns closely, Jarvis found that many Northern towns reporting idiotic or insane blacks also reported a black population of zero. He concluded that assistant marshals who intended to record idiotic or insane whites sometimes inadvertently entered the information in the column intended for the colored population, an easy error to make on a form with eighty cramped columns. Where extremely few blacks resided, a small number of such errors had enormous consequences for the *rate* of idiocy and insanity among blacks. The opposite error doubtless occurred as well, but where blacks were rare and whites predominated, random mix-ups of the columns had no discernable impact on the insanity rate for whites.

---

propaganda," *Bulletin of the History of Medicine*, 15 (1944), 472. Leon F. Litwak, "The Federal Government and the Free Negro, 1790-1860," Journal of Negro History 43 (1958), 261-278. Patricia Cline Cohen, *A Calculating People: The Spread of Numeracy in Early America* (Chicago, 1982), 175-204.

[15] Wright and Hunt, *The History and Growth of the United States Census*, 37-38.

[16] Leon F. Litwack, *North of Slavery: The Negro in the Free States* (Chicago, 1961), 41-43. Deutsch, "The first US census of the insane (1840) and its use as pro-slavery propaganda," 471-474.

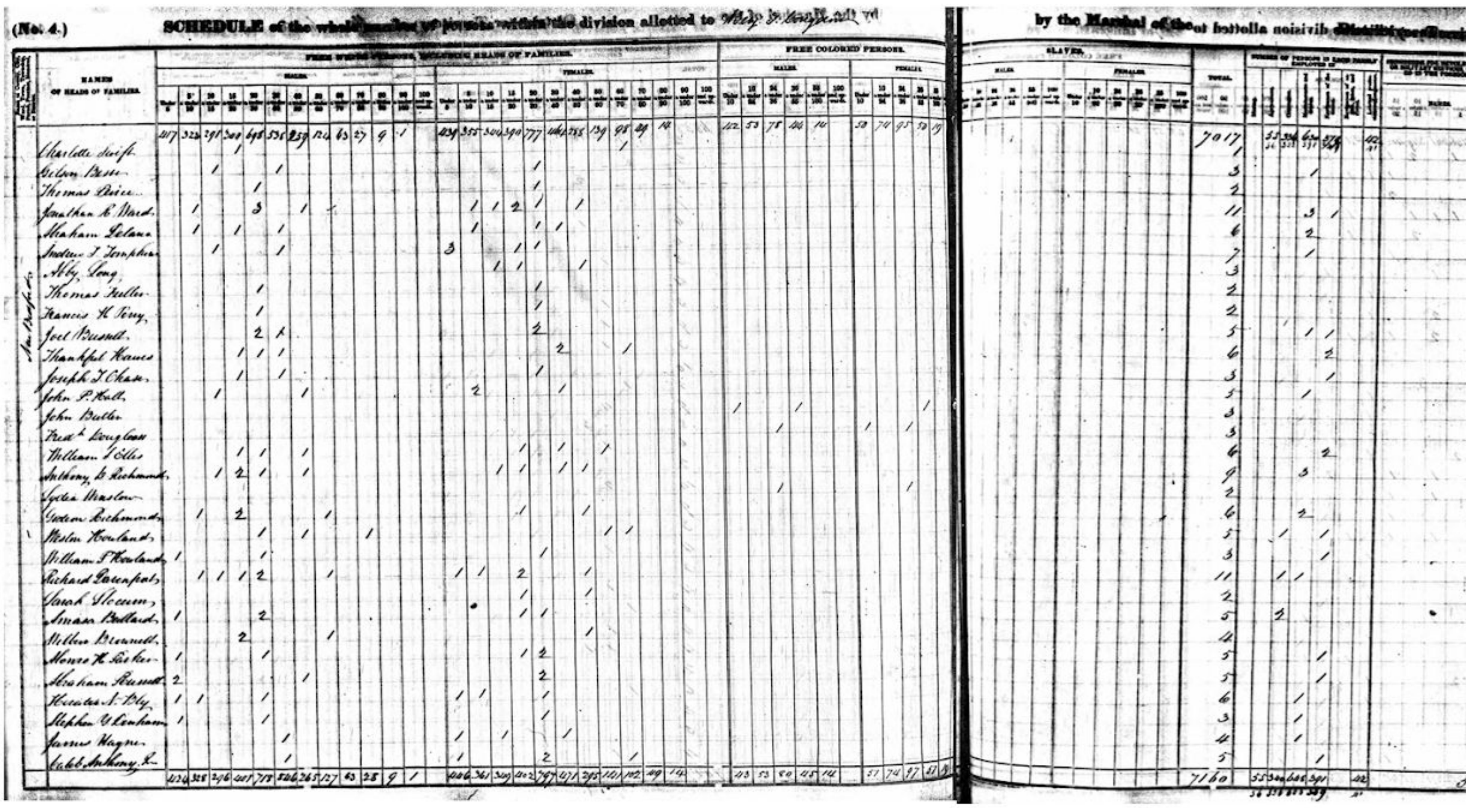


Figure 2. 1840 Census Enumeration Page, New Bedford, Massachusetts
Source: Sixth Census of the United States. National Archives and Records Administration microfilm publication M704-177, p. 410. Record Group 29. Washington, 1965.

Such errors also had no discernable impact on blacks in the South, where both the black and the white populations were substantial. Jarvis demanded, in the name of the nation's honor, medical science, and truth that the census must be fixed to avoid such errors in the future.[17]

The revelation of Jarvis' findings fueled a storm of protest. Representative and former President John Quincy Adams, citing the "atrocious misrepresentations" of the census, demanded that Calhoun, who as Secretary of State was responsible for the census, make the needed corrections. Adams wrote in his diary that Calhoun "writhed like a trodden rattlesnake on the exposure of his false report to the House that no material errors have been discovered in the census of 1840."[18] Nevertheless Calhoun refused to change the reported census results, calling them "unimpeachable."[19] Despite "the manifest and palpable, not to say gross errors" of the 1840 census, the results were allowed to stand.[20]

Census reformers were determined that future enumerations would avoid the egregious errors of the 1840 Census.[21] For the 1850 Census, leading statisticians supported the adoption of changes that addressed the problems caused by the cramped 1840 census form. The redesign, however, completely dismantled the straightforward data capture system that Madison had developed a half-century before. As a result, the redesign created a crisis of data capture, a crisis ultimately resolved by a series of Census Office innovations that transformed data processing in the United States and around the world.

---

[17] Litwack, *North of Slavery: The Negro in the Free States*, 43-44. Cohen, *A Calculating People: The Spread of Numeracy in Early America*, 191-204. Deutsch, "The first US census of the insane (1840) and its use as pro-slavery propaganda," 474-476.

[18] Charles Francis Adams, ed. *Memoirs of John Quincy Adams: comprising portions of his diary from 1795 to 1848*. Vol. XII. (Philadelphia, 1874-1877), 29.

[19] Deutsch, "The first US census of the insane (1840) and its use as pro-slavery propaganda," 476-478.

[20] Wright and Hunt, *The History and Growth of the United States Census*, 40.

[21] *Report of the Superintendent of the Census* (December 1, 1852), 23. Diana L. Magnuson, "The Making of a Modern Census: The United States Census of Population, 1790-1940," (Ph.D. diss., University of Minnesota, 1995), 32-33.

The Tabulation Bottleneck, 1850-1880

When Congress began to debate the 1850 Census, some legislators initially advocated a repeat of the 1840 census legislation, with merely the elimination of the contentious questions relating to disabilities. The nascent statistical community pushed back, and in March of 1849, Congress established a Census Board to determine the content of the 1850 Census. The Board—consisting of the Secretary of State, the Attorney-General, and the Postmaster-General—was charged with developing "forms and schedules for collecting in statistical tables, under proper heads, such information as to mines, agriculture, commerce, manufactures, education, and other topics as will exhibit a full view of the pursuits, industry, education, and resources of the country."[22] After consulting with members of the American Statistical Association, including Edward Jarvis, the Census Board recommended a system of six separate schedules to enumerate the free population, the slave population, mortality, agriculture, industry, and social statistics. This development represented a vast expansion in the scope of the census.

A reorganization of the 1850 form at once addressed the problems of the 1840 form and facilitated the collection of more data.[23] Instead of summarizing statistics for an entire family on each line, the 1850 form listed each individual on a separate line (Figure 3). Thus, a one-person family required just one line, but a ten-person family took up ten lines. Although the number of columns on the form was reduced from eighty to just thirteen, the quantity of information collected dramatically increased.

On Schedule One—describing the free population—names were recorded for each family member, along with exact age in years (and in months for infants), race, sex, whether married in the past year, school

---

[22] Magnuson, "The Making of a Modern Census: The United States Census of Population, 1790-1940," 34-35.

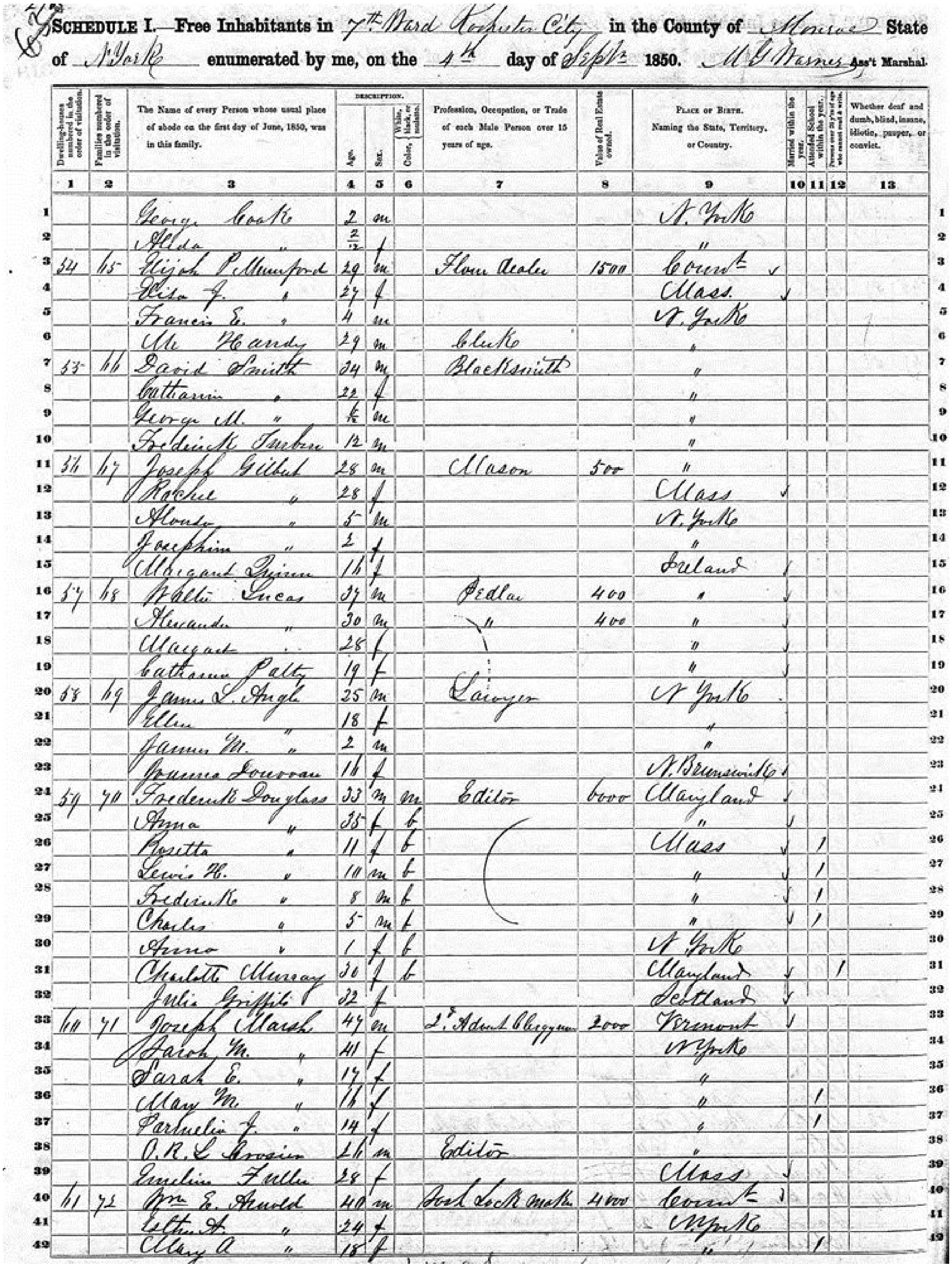
[23] The redesigned form borrowed from census designs used in in Massachusetts and Europe in the 1830s and 1840s. Walter F. Wilcox, "Lemuel Shattuck, Statist, Founder of the American Statistical Association," *Journal of the American Statistical Association* 35: 209 (1940), 224-235; Thorvalsden, *Censuses and Census Takers*. In addition to the obvious desire to avoid the debacle of 1840, a growing statistical community also advocated for a reworking of the census schedules for scholarly purposes. Magnuson, "The Making of a Modern Census: The United States Census of Population, 1790-1940," 36-39.

attendance, literacy, value of real property, and whether deaf and dumb, blind, idiotic, insane, pauper, or convict. Occupations and birthplaces were recorded as open-ended responses, providing vastly greater detail than had been available in the broadly categorized responses to previous censuses.

The Census Board proposed that Schedule Two—for enumerating slave inhabitants—should include many of the same questions as Schedule One: name of each slave, age, sex, color (black or mulatto), birthplace, and whether deaf, dumb, blind, insane, or idiotic. In addition to these overlapping questions, the slave schedules were also to include the number of children ever born to each female slave, and of those, the number known to be alive or dead; the degree of removal from "pure blood" (e. g., quadroon, octoroon, etc.); whether the slave was a fugitive; the name of each owner, and the number of slaves the owner had manumitted.

While dodging the controversies that engulfed the previous decade, the proposed slave schedule triggered a new and more explosive controversy. A protracted and contentious debate ensued in the detailed questions on the slave schedule were designed to provide ammunition for abolitionists.

Southern congressmen objected to the collection of names, arguing that on large plantations, no plantation owner could be expected recall the names of all his slaves. They further demanded removal of the questions on birthplace, fertility, and child survival. Senator William King of Alabama asserted that the question on



SCHEDULE I.—Free Inhabitants in 7th Ward Rochester City in the County of Monroe State of N York enumerated by me, on the 4th day of Septr 1850. M G Warner Ass't Marshal.

	Dwelling-houses numbered in the order of visitation.	Families numbered in the order of visitation.	The Name of every Person whose usual place of abode on the first day of June, 1850, was in this family.	Age.	Sex.	Color, (White, black, or mulatto.)	Profession, Occupation, or Trade of each Male Person over 15 years of age.	Value of Real Estate owned.	Place of Birth. Naming the State, Territory, or Country.	Married within the year.	Attended School within the year.	Persons over 20 y'rs of age who cannot read & write.	Whether deaf and dumb, blind, insane, idiotic, pauper, or convict.	
	1	2	3	4	5	6	7	8	9	10	11	12	13	
1			George Cook	2	m				N. York					1
2			Alida "	2/12	f				"					2
3	54	65	Elijah P. Mumford	29	m		Flour dealer	1500	Connt					3
4			Eliza J. "	27	f				Mass.					4
5			Francis E. "	4	m				N. York					5
6			Mc Candy	29	m		Clerk		"					6
7	55	66	David Smith	34	m		Blacksmith		"					7
8			Catharine "	22	f				"					8
9			George M. "	1/2	m				"					9
10			Frederick Tarbell	12	m				"					10
11	56	67	Joseph Gilbert	28	m		Mason	500	"					11
12			Rachel "	28	f				Mass					12
13			Alonzo "	5	m				N. York					13
14			Josephine "	2	f				"					14
15			Margaret Quinn	16	f				Ireland					15
16	57	68	Walter Lucas	37	m		Pedlar	400	"					16
17			Alexander "	30	m		"	400	"					17
18			Margaret "	28	f				"					18
19			Catharine Patty	19	f				"					19
20	58	69	James L. Angle	25	m		Lawyer		N York					20
21			Ellen "	18	f				"					21
22			James M. "	2	m				"					22
23			Joanna Donovan	16	f				N. Brunswick					23
24	59	70	Frederick Douglass	33	m	m	Editor	6000	Maryland					24
25			Anna "	35	f	b			"					25
26			Rosetta "	11	f	b			Mass		/			26
27			Lewis H. "	10	m	b			"		/			27
28			Frederick "	8	m	b			"		/			28
29			Charles "	5	m	b			"		/			29
30			Anna "	1	f	b			N York					30
31			Charlotte Murray	30	f	b			Maryland			/		31
32			Julia Griffith	32	f				Scotland					32
33	60	71	Joseph Marsh	47	m		2d Advent Clergyman	2000	Vermont					33
34			Sarah M. "	41	f				N York					34
35			Sarah E. "	17	f				"					35
36			Mary M. "	16	f				"		/			36
37			Parmelia J. "	14	f				"		/			37
38			O. R. L. Crosier	26	m		Editor		"					38
39			Emeline Fuller	28	f				Mass					39
40	61	72	Wm E. Arnold	40	m		Bank Lock maker	4000	Connt					40
41			Eliza A. "	24	f				N York					41
42			Mary A. "	18	f				"		/			42

Figure 3. 1850 Census Page, Rochester New York

Source: Seventh Census of the United States. National Archives and Records Administration microfilm publication M432-531, p, 318b. Record Group 29. Washington, 1965.

Congress, and again the census was entangled in the sectional conflict.[24] Southerners were convinced that children born was infeasible because "The woman herself, in nine out of ten cases, when she has had ten or fifteen children, does not know how many she has actually had."[25]

The South won the argument. The only questions that appeared on the slave schedule were name of the slave owner, the number of manumitted and fugitive slaves of that owner, and for each slave, age, sex, color, and disabilities. The debate about slavery in the 1840 census exposed the need for redesign of the census form, but the politics of slavery continued to dictate the form and content of the 1850 Census. The other five schedules survived debate virtually intact.

The reorganization and massive expansion of the other 1850 census schedules created an unanticipated data capture crisis. Unlike the previous censuses where assistant marshals completed the tabulation in the field, the marshals were now directed to forward their completed raw schedules to the Census Office in Washington. The number of rows to be tabulated on the free population schedule rose from 3.1 million in 1840 to twenty-three million in 1850. Moreover, the new layout of the 1850 population schedule meant there were no simple columns of counts to tally; rather, each individual response had to be classified into multiple categories and then tallied. Thus, responsibility for tabulating the results shifted from the assistant marshals to Census Office clerks in Washington. Over one hundred tons of census forms had to be transformed into statistical tables for publication.[26]

By previous standards of post-enumeration processing conducted by the Census Office, the tabulation was a massively complex operation. To generate statistical tables, clerks made tally-marks in groups of five on pre-printed "condensing forms" for each county. There were seven different condensing forms for the population schedules, printed on sheets of usually fifteen by twenty inches, and in some cases

---

[24] *The Congressional Globe*, April 9, 1850, pp. 671-680. Anderson, *The American Census: A Social History*, 45-49.

[25] *The Congressional Globe*, April 9, 1850, p. 674.

[26] Wright and Hunt, *The History and Growth of the United States Census*, 47-48 and 148-153. *An Act providing for the taking of the seventh and subsequent Census of the United States*, May 23, 1850.

on larger sheets.[27] The first form, for example, included the number of free persons in five-year age groups by sex and color; number of free persons born in each state and country; number of free persons married in the past year; and the number of free persons attending school, illiterate, pauper, or convict by race and nativity. Each of the 700,000 paper enumeration forms had to be handled at least seven times, one for each condensing form. To get the job done, the Census Office hired dozens of clerks, creating a pop-up tabulating operation of unprecedented scale. By the end of 1851, the Washington office had a total of 170 staff, about 10% of the entire federal workforce in Washington and nine times the number needed to process the 1840 Census.

The operation did not go smoothly. The Census Office was quickly overwhelmed by the massive scale of the work. For each of the previous censuses, the tabulated results were presented to Congress almost immediately after the returns came in, but in 1850 the new census schedules created a procedural bottleneck. In the spring and summer of 1852, oblivious to the enormity of the work before the temporary Census Office clerks, the Democratic-controlled Congress conducted a partisan investigation into the presumed inefficiency of the Whig-appointed Census Office. Predictably, Congress's investigation did not look favorably on the Superintending Clerk of the Census, Joseph Kennedy, and his staff.[28]

Given the magnitude of the task, the limited and temporary nature of the clerical staff, and the partisan charged atmosphere of the early 1850s, it was no small feat that the Census Office managed to produce: a preliminary count of the population of each state in December 1851; a 160-page "Abstract of the Seventh Census" the following year; a massive volume with over 1,000 pages of tables in 1853; and the compendium of the Seventh Census in 1854. Despite a sharp reduction in staffing by 1853, the Office soldiered on, completing a final volume on the manufacturing schedule in December 1859.[29]

---

[27] J.D.B. DeBow, "Blank Forms Used in the Census Office For Condensing Information—1850," *1850 Census: The Seventh Census of the United States* (Washington, 1853), xiii-xiv. Wright and Hunt, *The History and Growth of the United States Census*, 48.

[28] Anderson, *The American Census: A Social History*, 53 and 56-57.

[29] Wright and Hunt, *The History and Growth of the United States Census*, 48 and 50. *Report of the Superintendent of the Census for December 1, 1852* (Washington, 1853). DeBow, *The Seventh Census of*

Between the Censuses of 1850 and 1870, the tabulation bottleneck worsened. The procedure for data capture remained the same: clerks in the Census Office recorded tally marks on condensing sheets, or "spread-sheets" as they became known, and summed the tally marks to construct tables recording the number of persons in each place with each combination of characteristics. The scale of the problem, however, increased dramatically. The free population almost doubled between 1850 and 1870, partly because of the abolition of slavery in 1865. The number of questions asked on each census schedule also grew. For the population schedule, the questions asked of each individual rose from eleven in 1850 to eighteen in 1870. To meet the demands of post-enumeration data processing, the Census Office raised a "clerical force" through a system of examinations; 719 applicants took the test, and 465 of them passed. The maximum number of clerks rose from 170 in 1850 to 438 in 1870.[30]

The first technological innovation designed to alleviate the tabulation bottleneck was the Seaton Device, patented in 1872 by the Chief Clerk of the Census, Charles W. Seaton. The device was a simple wooden box with rollers, designed to expose eight columns of a large spreadsheet (see Figure 4). The eight columns were adjacent, reducing the time a clerk needed to locate the correct spot on the spreadsheet for a particular tally. When the largest column of tally marks was full, the clerk advanced the roller to start a new

the United States: 1850, An Appendix (Washington, 1853). J.D.B. DeBow, *Compendium of the Seventh Census* (Washington, 1854).

[30] Francis A. Walker quoted in Leon E. Truesdell, *The Development of Punch Card Tabulation In the Bureau of the Census, 1890-1940* (Washington, 1965), 22. W.R. Merriam, "The Evolution of American Census Taking," *Century Magazine*, LXV (1903), 831-842.

column.[31]

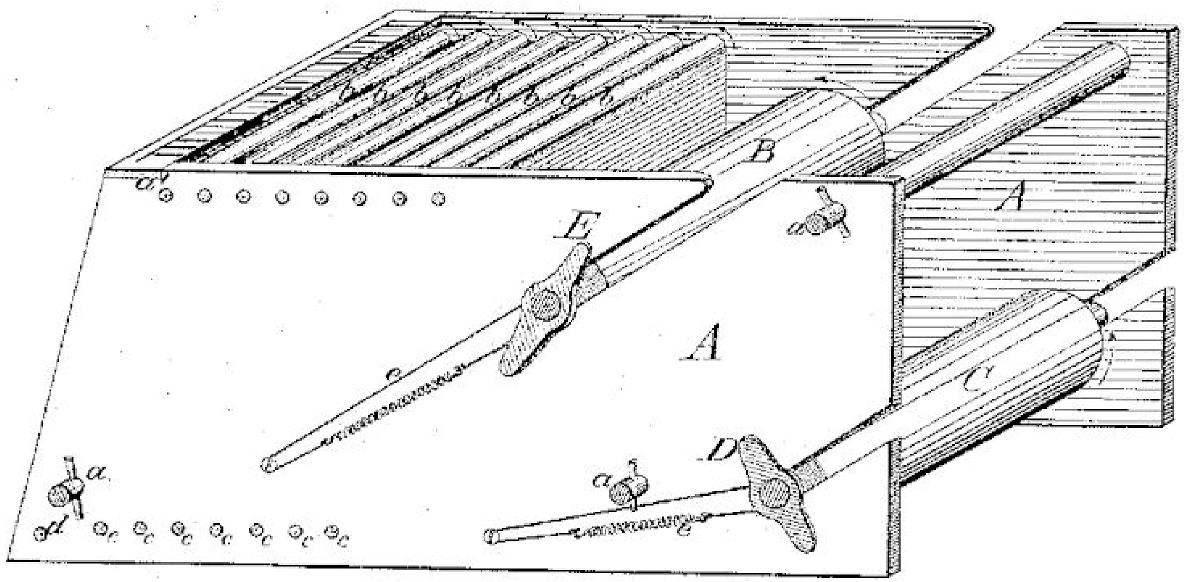


Figure 4. Seaton Tabulating Device.
Source: Charles W Seaton, US Pat. No. 127,435, "Improvement in Tabulating Devices," June 4, 1872.

Superintendent of the Census Francis Walker tested the Seaton Device by comparing results from its use with the "normal spread-sheet" method. Walker found that the machine increased productivity for each clerk from twenty-nine pages to 124 pages of tabulation per day, an improvement of 428%. Although a later analysis suggested this claim was greatly exaggerated, Seaton received a bonus of $15,000 for his invention by a special act of Congress. The amount was justified based on the presumed savings the device would produce: $15,000 was equivalent to the annual salary of twenty-nine census clerks, and it was estimated that the device would save twenty-nine person-years of work.[32]

The Seaton Device did not resolve the data capture bottleneck. In 1880, the number of questions on the population schedule rose again, from eighteen to twenty-four. More importantly, the demand for detailed statistics expanded, Congress became more willing to fund processing such statistics, and Walker was

[31] Truesdell, *The Development of Punch Card Tabulation In the Bureau of the Census, 1890-1940*, 19.

[32] Truesdell, *The Development of Punch Card Tabulation In the Bureau of the Census, 1890-1940*, 22. Wright and Hunt, *The History and Growth of the United States Census*, 68.

enthusiastic about producing them. Consequently, the number of census volumes published increased from five volumes of various sizes in 1870 to twenty-two large quarto volumes in 1880, plus a compendium. This ambitious program required a massive increase in clerical staff, which grew from a peak of 438 for the 1870 census to 1,495 for the 1880 Census. As Census Director William Rush Merriam later remarked, it had become clear "that a point must be reached, before many more decades had passed, where complete tabulation within the census period [before the next enumeration began] would be actually impossible."[33]

Unit Record Machines, 1890-1950

In 1889, Superintendent of the Census Robert P. Porter decided that a new tabulation system was needed for the count of the 1890 census, and he organized a competition to solicit the best ideas. Three inventors responded. Each had developed competing tabulation systems based on a common idea. Charles F. Pidgin, chief clerk of the Massachusetts Bureau of Labor Statistics, developed a system using cardboard chips printed in different colors. Census information was transcribed onto the chips, using symbols to represent different characteristics. Then the chips could be sorted into piles and counted. The Pidgin system was used successfully for the 1885 Massachusetts State Census.[34] William C. Hunt, who had worked on the 1885 Massachusetts Census, offered a simplified version of the Pidgin system, using paper slips with colored inks rather than chips. Finally, Herman Hollerith, who had worked for the Census Office in 1880, invented a machine for electric tabulation of cards using holes punched in the cards, a system he had used for tabulating mortality records in Baltimore, New York, and New Jersey in the late 1880s.

The three systems were all based on the concept of "unit records," which are separate records for each case being processed. For all three systems, information on each enumerated individual was transferred from the enumeration schedule to a piece of paper or cardboard. Instead of using tally-marks on a large

---

[33] Wright and Hunt, *The History and Growth of the United States Census*, 57 and 68; Merriam, "The Evolution of American Census Taking," 838-839.

[34] Truesdell, *The Development of Punch Card Tabulation In the Bureau of the Census, 1890-1940*, 24-25. Lars Heide, *Punched-Card Systems and the Early Information Explosion, 1880-1945* (Maryland, 2009), 23.

spreadsheet, the unit records were then counted. The initial transcription added an extra step to the process, but yielded new efficiencies whenever there was more than one table using a particular characteristic. For example, once the unit records had been sorted by race, sex, and nativity, those subgroups could be reused for multiple tabulations, with subsequent counts dividing these basic groups by detailed categories of such variables as age, birthplace, or occupation. Under the traditional tally system, each new table began from scratch, which meant duplicating the same work over and over again.

To choose between the three systems, Porter appointed a committee to conduct a contest. Each contestant was required to transfer the information on 10,491 residents of St. Louis from the 1880 census onto their chips, slips, or cards, and make a set of tables. Hollerith was the clear winner. As shown in Table 1, the punched cards were significantly faster to prepare than either the chips or slips.[35] The real advantage of the Hollerith system, however, came in the tabulation phase: the punched cards were ten times faster than the slips, and eight times faster than the chips.

Table 1. Data processing hours: 1889 Census Office Contest

	Transcription	Tabulation	Total
Pidgen chips	110.9	44.7	155.6
Hunt slips	144.4	55.4	199.8
Hollerith cards	72.5	5.5	77.9

The Hollerith system was faster in the tabulation phase because the counting was done electrically. Each card had room for 288 punched holes, and the position of each hole identified its meaning. The central element of the machine was the circuit-closing press, which looked a little like a waffle iron (see Figure 5). The top part of the press contained 288 spring-loaded pins corresponding to the positions of the punched holes. The bottom part had 288 small cups filled with mercury (see Figure 6). The operator placed a punched card in the press, and pulled the handle to lower the pins. Most of the pins were pressed upwards, but wherever there was a hole in the card the pins went through and into the mercury-filled cups, creating an

[35] Truesdell, *The Development of Punch Card Tabulation In the Bureau of the Census, 1890-1940*, 40-41, 142.

electrical connection.[36]

The electrical connections activated electromagnets that advanced a set of dials. The machine could be set up to count individual holes or combinations of holes. For example, a dial might be advanced when the holes for “white,” “female,” and “native-born” were all punched. Up to forty characteristics or combinations of characteristics could be counted with each pass through the circuit-closing press. The counters could store up to 9,999 cases for each characteristic; when that limit was reached, the operator transcribed the readings from all the dials onto paper, and reset the dials to zero.

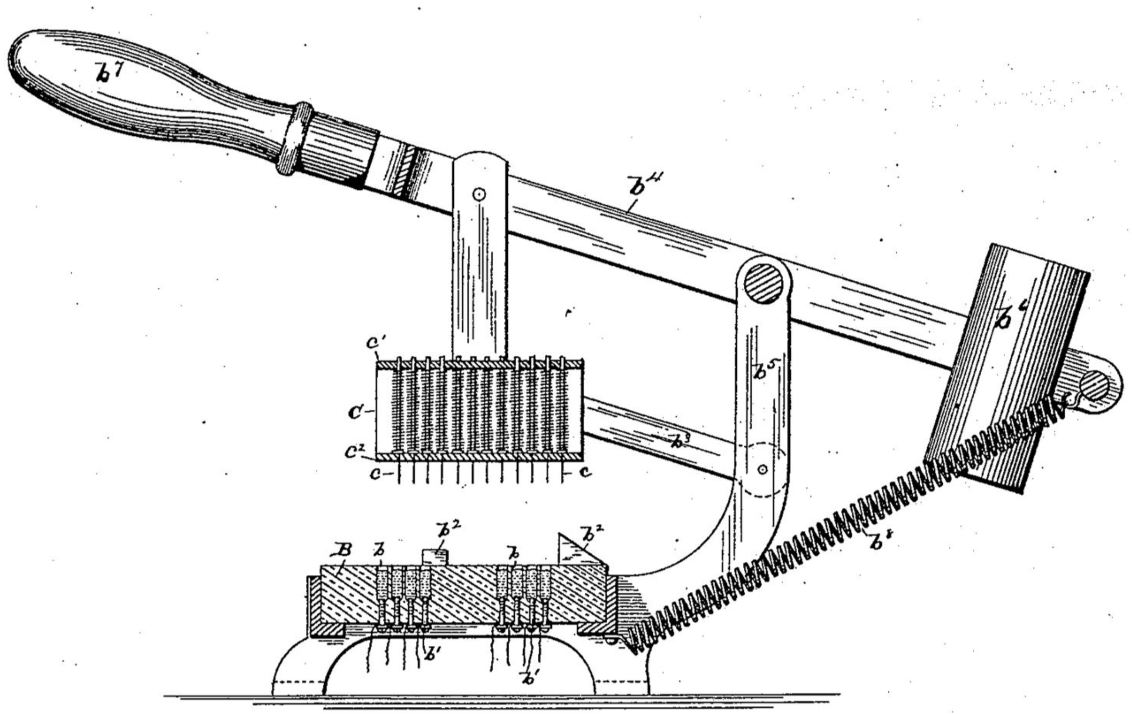
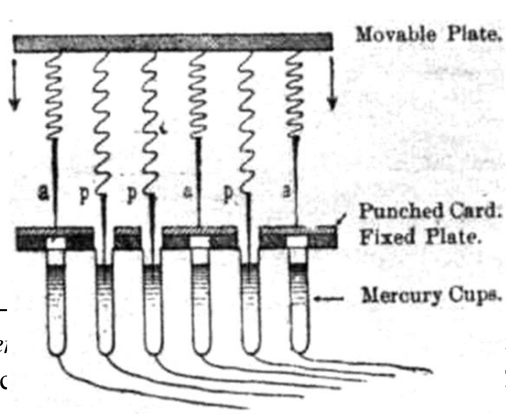


Figure 5. Hollerith circuit-closing card press
Source: H. Hollerith, US Pat. No. 395,781, "Art of Compiling Statistics," Jan. 8, 1889.



[36] Glen Fleck, ed., *A Compute* ambridge, Massachusetts, 1973), 26. Herman Hollerith, “An Elec olumbia University School of Mines, Vol.X No.16 (April

.

Figure 6. Operation of mercury-filled cups
Source: IBM Archive.

In addition to the counting function, Hollerith's machine also had a sorting function. The press activated a sorting box, which consisted of twenty-four compartments with spring-loaded lids. When the press was activated, an electromagnet activated the catch on one of the lids, springing the compartment open. The system could be set so that each compartment represented a census characteristic or combination of characteristics. Using the sorting box and the counters simultaneously, the cards could be sorted and counted in one operation. The sorted cards could then be fed through the machine again, allowing for complex cross-classifications of census characteristics.[37]

Clerks using Hollerith machines to tabulate the 1890 census were able to process an average of 7,000 to 8,000 punched cards per day.[38] The Census Office estimated that the system would save $580,000 in wages; Hollerith was paid $230,390 in rental fees for 56 machines for four years, or about 40% of the savings.[39] The machines were returned to Hollerith when tabulation was completed in 1894.

In preparation for the 1900 census, the Census Office again held a contest for tabulation systems in June 1889. This time, the only competitors were Hollerith and Pidgin. Pidgin offered three new systems: the "Automatic Mechanical Tabulation System," the "Pin Board Electrical Tabulation System," and the "Electrical Typewriter Tabulator."[40] Hollerith was so nervous about the new devices that he hired a Pinkerton

[37] Heide, *Punched-Card Systems and the Early Information Explosion, 1880-1945*, 25-26.

[38] According to *Scientific American*, "a single operator can dispose of 50,000 names in a day." *Scientific American*, Vol.63, no. 9 (August 30, 1890): 132.

[39] Heide, *Punched-Card Systems and the Early Information Explosion, 1880-1945*, 24 and 39.

[40] Heide, *Punched-Card Systems and the Early Information Explosion, 1880-1945*, 51. Geoffrey Austrian, *Herman Hollerith: Forgotten Giant of Information Processing* (New York, 1982), 122.

detective to go up to Boston to conduct industrial espionage. He need not have worried; in the end the Hollerith machines—which were virtually unchanged from the ones used in 1890—completed the census tabulation contest in less than half the time of the best Pidgin machine.[41] Census Director Merriam negotiated a contract with Hollerith for an annual rental of $1,000 per tabulator.

During the course of tabulating the 1900 Census, Hollerith introduced several innovations. The most important of these was the "Automatic Tabulating Machine," in which "the work of separately placing each card beneath the pin box, depressing the pin box, and removing the card is performed automatically," thus speeding tabulation by a factor of six.[42] To process the Census of Agriculture, he developed an "Adding Tabulator" that not only counted the number of farms, but also summed up their acreage and production. Finally, Hollerith came out with a new keypunch and an automatic sorting machine. Taken together, these new devices represented dramatic improvements over the technology used in 1890. In all, Hollerith supplied 311 tabulating machines, twenty automatic sorters, 1,021 punches for preparing the cards, for which he was paid $428,239 in rental fees.[43] The tabulation was fast; the time elapsed from the census day to the publication of the population volume was the shortest since 1820.[44]

From 1790 to 1900 the Census Office was a temporary organization. Each decade, a Census Office was established, the census was taken, and then the office was closed. This system starkly contrasted with European countries, virtually all of which had established permanent central statistical offices by the mid-nineteenth century.[45] The inefficiency of shutting down and reopening the Census Office every decade was obvious, and most superintendents of the census since the 1850s—including Joseph Kennedy, Francis Walker, Robert Porter, and Carroll Wright—vigorously advocated for the creation of a permanent agency.[46]

---

[41] Austrian, *Herman Hollerith: Forgotten Giant of Information Processing*, 122.

[42] Edward W. Byrn, "The Mechanical Work of the Twelfth Census," *Scientific American* vol. 86 no. 16 (April 19, 1902), 275. Austrian, *Herman Hollerith: Forgotten Giant of Information Processing*, 176.

[43] Austrian, *Herman Hollerith: Forgotten Giant of Information Processing*, 128.

[44] Merriam, "The Evolution of American Census Taking," 838.

[45] Peter Flora, "Historical Sources of Statistics," *Current Sociology*, 23:2 (1975), 113-139.

[46] Wright and Hunt, *The History and Growth of the United States Census*, 79-83. "Permanent Census

Census Director Merriam finally succeeded, partly because President Theodore Roosevelt wanted more extensive statistical facts to undergird social legislation. On July 1, 1902, the Census Office became a permanent agency and soon after was renamed the "Bureau of the Census."[47]

Merriam resigned as Census Director in 1903, and a year later he became president of Hollerith's Tabulating Machine Company. The cozy relationship between the Census Bureau and the Hollerith Company ended under the new Census Director, Simon N.D. North. Hollerith had lobbied against North's appointment, and North felt that the rental fees for the Hollerith machines were exorbitant. The conflict led to the removal of Hollerith machines from the Census Bureau, and for a brief period the adoption of the slower Pidgin devices.[48]

North reasoned that with the establishment of a permanent Census Bureau, it made sense to have permanent equipment under Bureau control. Hollerith's original patents were set to expire on January 8, 1906, freeing anyone to develop similar tabulators.[49] In 1905, North successfully applied to Congress for $40,000 to develop tabulating machinery.[50] He used these funds to establish the Census Machine Shop in 1907, enabling the Census Bureau to build and maintain its own tabulating equipment.[51] North wrote in his

---

Office," *Publications of the American Statistical Association*, Vol. 7, No. 56 (December 1901), 1-62.

[47] Merriam, "The Evolution of American Census Taking," 832. The Act of March 6, 1902 (32 Stat. 51) established a permanent Census Office.

[48] "W.R. Merriam to Resign," *New York Times*, March 1, 1903, p. 9. "New York Confirms Report that Mr. Merriam Has Resigned," *The Washington Post*, March 3, 1903, p. 5. Austrian, *Herman Hollerith: Forgotten Giant of Information Processing*, 227-228. Truesdell, *The Development of Punch Card Tabulation In the Bureau of the Census, 1890-1940,* 119-120. Austrian, *Herman Hollerith: Forgotten Giant of Information Processing,* 225-228 and 236. Bureau of the Census, *Report of the Director to the Secretary of Commerce and Labor, Concerning the Operations of the Bureau for the Year 1907-08* (Washington, 1908), 20-21.

[49] Bureau of the Census, *Report of the Director to the Secretary of Commerce and Labor, Concerning the Operations of the Bureau for the Year 1904-05* (Washington, 1905), 19. Truesdell, *The Development of Punch Card Tabulation In the Bureau of the Census, 1890-1940*, 121.

[50] Truesdell, *The Development of Punch Card Tabulation In the Bureau of the Census, 1890-1940,* 120. Bureau of the Census, *Report of the Director to the Secretary of Commerce and Labor, Concerning the Operations of the Bureau for the Year 1905-06* (Washington, 1906), 24.

[51] W. Stull Holt, *The Bureau of the Census, Its History, Activities and Organization* (Washington, 1929), 51. According to Holt, the "force employed was not large, consisting of an expert on patents of the type

annual report, "For the small sum . . . the Census experts have succeeded in devising a tabulating apparatus, along lines entirely novel, which infringes no patents and which is a marvel of simplicity, of accuracy, and of rapid manipulation of punched cards. . . . The possibilities of saving, in the costs of compiling future censuses . . . are enormous."[52] Further, "With this machine shop, the building and repair of machinery can be done much more effectively and economically by the Bureau than by contract with private manufactures."[53] The Census Bureau recruited four disaffected Hollerith engineers to carry out the work, along with James Powers, a Russian immigrant who had invented machinery for several other companies.[54] To attract talent, the engineers were allowed to take out private patents on their discoveries, as long as their inventions remained freely available for government use.[55]

By 1907, the Census Bureau Machine Shop developed prototypes that had some advantages over Hollerith machines. The new tabulator printed results instead of showing them on a dial, eliminating the time-consuming step of transcribing the information on the dials onto paper. Powers developed a new kind of electric keypunch with 240 keys, one for each hole on a punch card, which promised to double the pace of data entry compared with the Hollerith punch. In addition, the Machine Shop was working on a fully-automatic tabulator; like Hollerith's latest design, these machines were designed to process a stack of punch cards with no intervention by the operator.[56]

---

involved...a foreman and assistant foreman, and seven or eight mechanicians and electricians." Joseph W. Duncan and William C. Shelton, *Revolution in United States Government Statistics, 1926-1976* (Washington, 1978), 118.

[52] Bureau of the Census, *Report of the Director to the Secretary of Commerce and Labor, Concerning the Operations of the Bureau for the Year 1905-06* (Washington, 1906), 24.

[53] Bureau of the Census, *Report of the Director to the Secretary of Commerce and Labor, Concerning the Operations of the Bureau for the Year 1906-07* (Washington, 1907), 18.

[54] Austrian, *Herman Hollerith: Forgotten Giant of Information Processing*, 271-72. Heide, *Punched-Card Systems and the Early Information Explosion, 1880-1945*, 75-76.

[55] Heide, *Punched-Card Systems and the Early Information Explosion, 1880-1945*, 75.

[56] Bureau of the Census, *Report of the Director to the Secretary of Commerce and Labor, Concerning the Operations of the Bureau for the Year 1906-07* (Washington, 1907), 18. Heide, *Punched-Card Systems and the Early Information Explosion, 1880-1945*, 76-78.Truesdell, *The Development of Punch Card Tabulation In the Bureau of the Census, 1890-1940*, 121-122, 126. Duncan and Shelton, *Revolution in*

Hollerith was furious. Convinced that his patents were being infringed upon, he launched a lobbying campaign with the Secretary of Commerce and President Roosevelt. The anti-monopolist Roosevelt was unsympathetic, but when Taft became president in 1909 Hollerith managed to get North fired.[57] The new Director, Edward Durand, nevertheless opted to continue with the technology developed in the Census Bureau Machine Shop to process the 1910 census. Less than three months before the 1910 count was to begin, Hollerith filed suit against Durand, alleging patent infringement on card-sorting machines that were being altered by the Machine Shop. Hollerith got a restraining order, halting work on the machines, but the District of Columbia court overturned the order, whereupon the lawsuit fizzled out.[58]

While Hollerith's lawsuit was still going on in 1911, James Powers resigned from his position at the Census Bureau and formed the Powers Tabulating Machine Company. Powers introduced a full line of punches, tabulators, and sorters that competed directly with the products of Hollerith's Tabulating Machine Company. The Powers Company joined with the Remington Typewriter Company and the Rand Kardex office supply firm to form Remington Rand in 1927. Meanwhile, in 1911 Hollerith's company merged with a commercial scale company and two makers of employee time-clocks to form the Computing-Tabulating-Recording Company (CTR), and in 1924 CTR changed its name to the International Business Machines Corporation (IBM). IBM, Remington Rand, and the Census Bureau Machine Shop dominated data processing until the late twentieth century.[59]

From a twenty-first century perspective, the Census Bureau Machine Shop is extraordinary. The shop was a government manufacturing establishment, explicitly set up to compete with a private-sector

---

United States Government Statistics, 1926-1976, 118.

[57] Austrian, "The Growing Impasse," chapter in *Herman Hollerith: Forgotten Giant of Information Processing*, 277-295.

[58] Truesdell, *The Development of Punch Card Tabulation In the Bureau of the Census, 1890-1940*, 129. Heide, *Punched-Card Systems and the Early Information Explosion, 1880-1945*, 79. Austrian, "Tabulating Machine Company v. Durand," chapter in *Herman Hollerith: Forgotten Giant of Information Processing*, 296-305.

[59] Heide, *Punched-Card Systems and the Early Information Explosion, 1880-1945*, 82, 89, 105-107. Austrian, *Herman Hollerith: Forgotten Giant of Information Processing*, 305.

vendor. The Machine Shop even recruited talent from that vendor and barely skirted (or possibly infringed) its patents.[60] It was, however, an era of progressive reform. Although Hollerith's company was not a massive trust, it did have a complete monopoly on tabulating equipment. Hollerith complained repeatedly that the government should not compete with private enterprise, but his objections fell on deaf ears. The key players in the Bureau and the executive branch felt that Hollerith had been holding the government hostage, since no alternative suppliers of tabulating equipment existed. Further, the Census Machine Shop had already demonstrated its ability to innovate existing machinery and develop machinery in response to its own needs.

According to A. Ross Eckler, former Director of the Census, "From 1910 to 1950, equipment built by the Bureau's machine shop was substantially more productive for census work than equipment available commercially."[61] In the director's report from 1912, Durand pushed for increased funding for the Machine Shop, "It is desirable that an appropriation . . . should be made to enable the Bureau to continue the employment of as large a part of its force of patent experts, inventors, and mechanics as possible. Should the bureau lose the services of all or the greater part of its present mechanical force, it would be almost impossible to secure competent men for the expert mechanical work" which these men had "long and successful experience."[62]

---

60 Austrian, *Herman Hollerith: Forgotten Giant of Information Processing*, 261-262 and 267-268. *Annual Report of the Director of the Census, To the Secretary of Commerce and Labor for the Fiscal Year Ended June 30, 1912* (Washington, 1913), 14.

61 A. Ross Eckler, *The Bureau of the Census* (New York, 1972), 112.

62 *Annual Report of the Director of the Census, To the Secretary of Commerce and Labor For the Fiscal Year Ended June 30, 1912* (Washington, 1913), 14-15. Heide, *Punched-Card Systems and the Early Information Explosion, 1880-1945*, 80. *Annual Report of the Director of the Census, to the Secretary of Commerce and Labor for the Year 1906-1907* (Washington, 1907), 18. *Annual Report of the Director of the Census, To the Secretary of Commerce and Labor for the Year 1909-1910* (Washington, D.C., 1911), 43. *Annual Report of the Director of the Census, to the Secretary of Commerce and Labor for the Fiscal Year Ended June 30, 1916* (Washington, 1916), 20-21. *Annual Report of the Director of the Census, to the Secretary of Commerce and Labor for the Fiscal Year Ended June 30, 1918* (Washington, 1918), 22-23. *Annual Report of the Director of the Census, to the Secretary of Commerce and Labor for the Fiscal Year Ended June 30, 1919* (Washington, 1919), 34-35. *Annual Report of the Director of the Census, to the Secretary of Commerce and Labor For the Fiscal Year Ended June 30, 1921* (Washington, D.C., 1921), 28. *Annual Report of the Director of the Census, to the Secretary of Commerce and Labor for the Fiscal Year Ended June 30, 1922* (Washington, 1922), 24-25. *Annual Report of the Director of the Census, to the Secretary of Commerce and Labor for the Fiscal Year Ended June 30, 1923* (Washington, 1923), 27.

The 1910 census was processed mainly on machines developed in the Census Machine Shop.[63] The Powers punches frequently jammed, so the early card-punching was done on old Hollerith punching machines owned by the Census Bureau. The Machine Shop eventually reduced the problem of jamming, and about two-thirds of the 1910 cards were ultimately punched on the Powers machines. Most of the tabulation was carried out on semi-automatic tabulators designed by the Machine Shop, and the Census Bureau introduced fully automatic self-feeding tabulators after the tabulation was underway.[64]

For the next four decades, the Census Bureau relied mainly on equipment designed and maintained by the Mechanical Laboratory (formerly the Machine Shop), purchasing or renting supplemental equipment when necessary. According to Eckler, maintaining equipment purchased from outside vendors "proved to be definitely economical," but also "paved the way for Bureau engineers to introduce modifications in line with special needs."[65] For example, the 1920 population census was processed with the self-feeding automatic

Annual Report of the Director of the Census, to the Secretary of Commerce and Labor for the Fiscal Year Ended June 30, 1924 (Washington, 1924), 23-24. *Annual Report of the Director of the Census, to the Secretary of Commerce and Labor for the Fiscal Year Ended June 30, 1925* (Washington, 1921), 25-26. *Annual Report of the Director of the Census, to the Secretary of Commerce and Labor for the Fiscal Year Ended June 30, 1928* (Washington, 1928), 5-6. *Annual Report of the Director of the Census, to the Secretary of Commerce and Labor for the Fiscal Year Ended June 30, 1930* (Washington, 1930), 22-23.

[63] The semi-automatic electric tabulating machines were designed by the Census Bureau but manufactured by the Sloan & Chase Manufacturing Co. Bureau of the Census, Report *of the Director to the Secretary of Commerce and Labor, Concerning the Operations of the Bureau for the Year 1909-10* (Washington, 1910), 44.

[64] In 1918 the Bureau conducted a remarkable experiment with field-punched cards in a special census of Okmulgee County, Oklahoma. Enumerators were issued "a modified (enlarged) conductor's punch" to punch 11 ¼ by 6 ¼ inch cards to capture "standard population data" in the field. No official report on the results of the experiment is extant, but "oral tradition" suggests "a machine was devised" to convert the field-punched cards into tabulation cards Unfortunately, the machine "did not come up to expectations" and the field punch data capture method was not repeated. It was not until 1960 that the Census Bureau again attempted enumerator data capture in the field, this time with "bubble" sheets used to record population characteristics by enumerators going from door to door. Truesdell, *The Development of Punch Card Tabulation In the Bureau of the Census, 1890-1940*, 128-129. Bureau of the Census, Report of the Director to the Secretary of Commerce and Labor, Concerning the Operations of the Bureau for the Year 1910-11 (Washington, 1911), 23.

[65] Eckler, *The Bureau of the Census*, 118. Truesdell, *The Development of Punch Card Tabulation In the Bureau of the Census, 1890-1940*, 128-129 and 139. Duncan and Shelton, *Revolution in United States Government Statistics, 1926-1976*, 118.

tabulators that were introduced during the 1910 tabulation process and improved versions of the sorting machines used in 1900 and 1910. The keypunch equipment was more problematic; the Census Bureau decided that the Powers punch used for most of the cards in the 1910 census was too unreliable, and reverted to the older style of punch based on the original Hollerith designs. For the 1920 agricultural census, the Bureau rented equipment from the Tabulating Machine Company, since summing quantities of crops required an adding tabulator, a device never successfully constructed by the Bureau's Mechanical Laboratory.[66]

The Census Mechanical Laboratory made substantial advances in the 1940s and 50s, with the development of a multicolumn sorter and major improvements to the tabulator. [67] Mechanical Laboratory engineer Anthony Berlinsky, the "Thomas Edison of the Census Bureau," was reportedly responsible for numerous mechanical improvements for data processing.[68] In 1950, *Popular Mechanics* enthused that "census employees have the fanciest collection of complicated machinery the mind of many has been able to devise." "These machines are masterpieces of ingenuity" and it was Berlinsky who was "busily at work making them more ingenious." Berlinsky told the magazine that "When we finally got a machine to handle 40 columns of facts or 'holes' in a card those guys upstairs had to go think of 20 more questions they wanted to ask on the next census so everything is obsolete. This time I'm rebuilding the multi-column sorters to combine the sorting and tabulating operations. They'll handle 80 columns of facts. I hope it holds 'em for a while. It probably won't, though." Berlinsky also "conceived the idea for constructing an automatic microfilming machine, which was designed, tested, modified as necessary, and constructed in its final form under his immediate direction."[69]

---

[66] Truesdell, *The Development of Punch Card Tabulation In the Bureau of the Census, 1890-1940*, 141-142.

[67] Robert Jenkins, *Procedural History of the 1940 Census of Population and Housing* (Wisconsin, 1983), 82. Morris B. Ullman, *The 1950 Censuses—How They Were Taken: Procedural Studies of the 1950 Censuses, No. 2 Population, Housing, Agriculture, Irrigation, Drainage* (Washington, 1955), 36.

[68] Robert Voight, "Conrad Taeuber Oral History," transcript of an oral history conducted April 12, 1989 by Robert Voight, United States Census Bureau, 9; Ann Herbert Scott, *Census, U.S.A.* (New York, 1968), 58.

[69] Richard F. Dempewolff, "Count off, Americans," *Popular Mechanics*, Vol. 92, no. 2 (February 1950),

IBM hired the Census Bureau's Chief of Machine Tabulation Lawrence Wilson in the late 1940s. Soon after, IBM came out with a new unit record machine, Model 101, which combined the latest innovations in sorting and tabulation that had been developed at the Mechanical Laboratory. Described in *Popular Science* as a "super-dooper census gadget," the Census Bureau leased 32 of these machines from IBM for the 1950 census. In an ironic turnaround of the Bureau's appropriation of Hollerith technology in 1907, for the 1950 census the Bureau paid the descendant of the Hollerith Company for technology that had been developed by the Bureau.[70]

106-110 and 260. Eckler, *The Bureau of the Census*, 114.

[70] John F. Loosbrock, "How Machines Turn You Into Statistics," *Popular Science* (April 1950), 152-155. Eckler, *The Bureau of the Census*, 115. Duncan and Shelton, *Revolution in United States Government Statistics, 1926-1976*, 119.

Figure 7. Census Mechanical Laboratory, 1950s
Source: U. S. Census Bureau,
https://broadcast.census.gov/pio/photos/machines_1950s_and_1960s/machines_1950_08023-hi.jpg.

UNIVAC

By 1950, punch card tabulation was a mature technology. Refinements continued to be made and data processing speeds continued to improve. As the tabulation technology was perfected, the biggest bottleneck became the punch card itself. The population and housing components of the 1950 census required about twenty-two gigabytes of data storage. These data were stored on 282 million eighty-column cards weighing some 600 tons that if piled up would make a stack thirty-one miles tall. The cards were fussy and fragile, and high humidity or mishandling often made them unreadable. As processing of punched data was increasingly automated, keypunching remained highly labor-intensive. Moreover, the storage, retrieval, and organization of the massive physical collection of data began to consume a growing share of resources.[71]

Electronic computers offered a solution to the punch card bottleneck. The first electronic programmable computer was the Electronic Numerical Integrator And Computer (ENIAC), built from 1943 to 1946 by John W. Mauchly and J. Presper Eckert at the University of Pennsylvania Moore School of Engineering. ENIAC was a wartime project designed with the goal of calculating ballistics for artillery. Mauchly was convinced that computers would eventually have civilian applications, and he approached Morris Hansen, Statistical Assistant to the Director of the Census Bureau, to discuss Bureau needs for data processing. Beginning in 1944 and continuing over the next two years, Mauchly had multiple meetings with Hanson and other census experts "to discuss the applicability of computers to Census problems."[72]

---

[71] Steven Ruggles et al. "Frozen Film and FOSDIC Forms: Restoring the 1960 U.S. Census of Population and Housing," *Historical Methods* (April-June 2011), 69-78. For a discussion of the number of cards punched for the 1950 census of population enumeration, see Morris B. Ullman, *The 1950 Censuses—How They Were Taken: Procedural Studies of the 1950 Censuses, No. 2 Population, Housing, Agriculture, Irrigation, Drainage* (Washington, 1955), 29-38.

[72] Duncan and Shelton, *Revolution in United States Government Statistics, 1926-1976*, 120-121 and 124. Nancy Stern, "ENIAC Research and Development," chapter in *From ENIAC To UNIVAC: An Appraisal of the Eckert-Mauchly Computers* (Digital Equipment Corporation, 1981), 24-47.

The University of Pennsylvania asserted intellectual property rights over ENIAC in 1946. According to Mauchly, he and Eckert were given the choice of signing over their patents to the University or resigning, and they chose the latter. Morris Hansen, meanwhile, himself now convinced of the potential of the electronic computer for census applications, arranged for the National Bureau of Standards (NBS) to study the feasibility of computerizing the census. Backed by Census Bureau funding, in October 1946 Eckert and Mauchly received a $75,000 study contract from NBS to draw up plans for a census computer.[73] With preliminary funds in hand, in December 1947 the inventors established the Eckert-Mauchly Computer Corporation. The following June, the new company received the full contract worth some $300,000 to build the world's first commercial computer for the Census Bureau.[74] Their machine, The Universal Automatic Computer (UNIVAC) was designed for data processing. A key innovation was the use of magnetic tape for data storage, which meant that the punch cards could be discarded once their information was transferred to tape. UNIVAC was completed in March 1951, too late to contribute significantly to the 1950 census tabulation.[75] The Bureau conducted extensive testing, however, and concluded that the machine had the

---

[73] Duncan and Shelton, *Revolution in United States Government Statistics, 1926-1976*, 125. Eckler, *The Bureau of the Census*, 116.

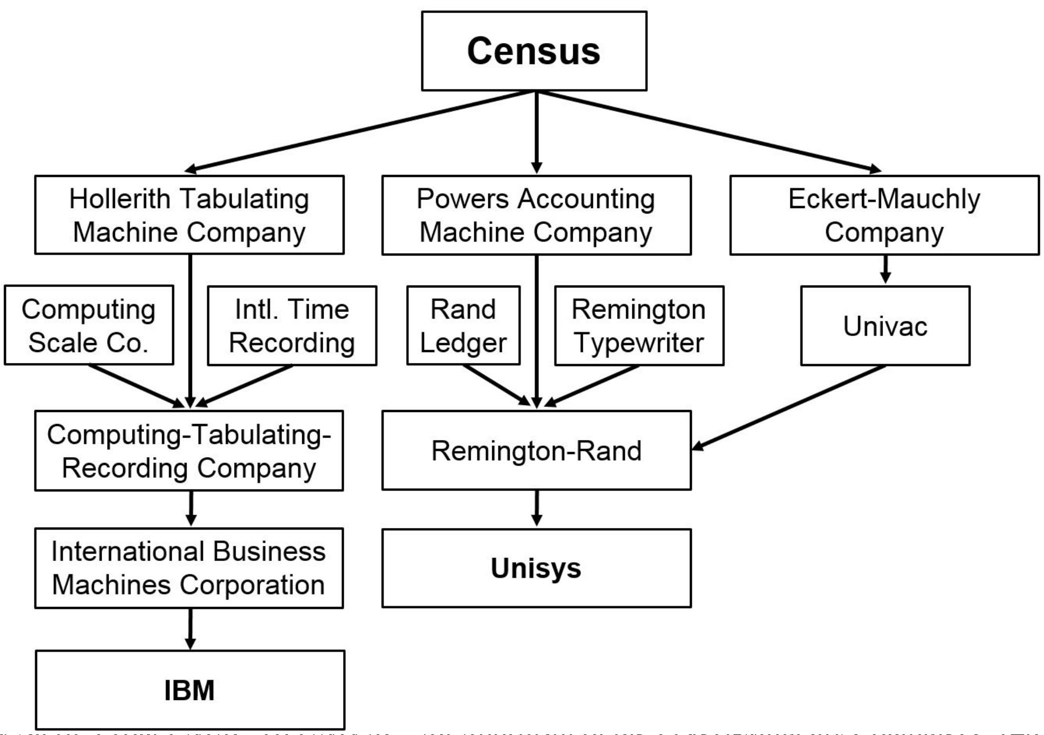
[74]Lars Heide, "Making Business of a Revolutionary New Technology: the Eckert Mauchly Company, 1945-1951." *History of Nordic Computing 3, Third IFIP WG 9.7 Conference, HiNC3*. (Springer, Heidelberg, 2011), 207-214. Nancy Stern, *From ENIAC to UNIVAC: An Appraisal of the Eckert-Mauchly Computers* (Digital Equipment Corporation, 1981), 91-92. Arthur L. Norberg, *Computers and Commerce: A Study of Technology and Management at Eckert-Mauchly Computer Company, Engineering Research Associates, and Remington Rand, 1946-1957* (Massachusetts Institute of Technology, 2005), 73. Duncan and Shelton, *Revolution in United States Government Statistics, 1926-1976*, 126. Norberg, *Computers and Commerce: A Study of Technology and Management at Eckert-Mauchly Computer Company, Engineering Research Associates, and Remington Rand, 1946-1957*, 92-95. Nancy Stern, "Transition to the Private Sector," chapter in *From ENIAC To UNIVAC: An Appraisal of the Eckert-Mauchly Computers* (Digital Equipment Corporation, 1981), 87-115. The NBS was established in 1901 as an agency of the Commerce Department to give aid to "manufacturing, commerce, the makers of scientific apparatus, the scientific work of the government, of schools, colleges and universities." Stern, 101.

[75] Duncan and Shelton, *Revolution in United States Government Statistics, 1926-1976* (U.S. Department of Commerce, Washington, 1978): 126-127. Bryan Hargis and Donald Larson, *Scientific and Technological Development Activities of the Bureau of the Census, Technical Paper 29* (U.S. Department of Commerce, September 1973), 3.

potential to cut data processing costs in half.[76]

By the time the first UNIVAC was delivered, the Eckert-Mauchly Computing Corporation had been purchased by Remington-Rand, itself a descendent from the Powers Tabulating Machine Company. Soon the UNIVAC faced competition. Until 1946, IBM's leadership had no interest in electronic computing, but the Census Bureau contract with Eckert-Mauchly was a wake-up call.[77] By 1949, IBM had established a substantial electronics research staff, and the company unveiled its first computer in 1953. Largely because of the strength of the IBM marketing and service divisions, IBM's computer rentals grew rapidly By 1955 the number of installed IBM machines exceeded the number of Univacs.[78]

Figure 8 presents a genealogy of data processing. The Hollerith Tabulating Machine Company and the Powers Tabulating Company, the two main producers of unit record machines in the early twentieth century, were both Census Bureau spin-offs started by former Census employees. The Eckert-Mauchly Computer Company was established with start-up funding from the Census Bureau. After various mergers



---

[76] Eck… 0 *Cen… Housing, Agr… ment Stat…*

[77] Nancy Stern, *From ENIAC To UNIVAC: An Appraisal of the Eckert-Mauchly Computers*, 148, 151.

[78] Duncan and Shelton, *Revolution in United States Government Statistics, 1926-1976*, 129-130.

Figure 8. Genealogy of Data Processing

and acquisitions, Hollerith became IBM and Powers became Remington-Rand. The Eckert-Mauchly Computer Corporation was renamed Univac, and Remington-Rand purchased Univac. Thus, the two biggest computer companies of the second half of the twentieth century could both trace their roots to the census.

FOSDIC and TIGER, 1960-1990

Two transformational technological developments in the second half of the twentieth century addressed problems associated with transforming census responses into machine-processable data: the Film Optical Sensing Device for Input to Computers (FOSDIC) and Topographically Integrated Geographic Encoding and Referencing (TIGER).

In the late 1940s the Census Bureau concluded that punch cards represented the biggest bottleneck in data processing.[79] At operational peak in processing the 1950 census, the Bureau employed nearly 2,000 key punch operators for over fourteen months at a cost of nearly six million dollars to record population and housing data on punch cards.[80] As soon as plans for developing the UNIVAC were confirmed, Census Bureau engineers realized its use would be limited without a new method for data input. In 1949—two years before the first UNIVAC was operational—the Census Bureau began collaborating with the National Bureau of Standards to design and build the first high-speed Optical Mark Recognition (OMR) system.[81]

---

[79] Duncan and Shelton, *Revolution in United States Government Statistics, 1926-1976*, 129. Hargis and Larson, *Scientific and Technological Development Activities of the Bureau of the Census, Technical Paper 29*, 3.

[80] Martin H. Weik, "A Third Survey of Domestic Electronic Digital Computing Systems," Ballistic Research Laboratories, Report No. 1115 (March 1961), 258-261.

[81] Hargis and Larson, *Scientific and Technological Development Activities of the Bureau of the Census,*

The UNIVAC made it possible to discard punch cards once the information had been transferred to magnetic tape, but the computer did not immediately solve the punch card bottleneck. Initially, the only method for getting data from census form to magnetic tape was via punch cards, using a card-to-tape converting machine. Accordingly, to input the data clerks still had had to prepare hundreds of millions of punch cards, using manual methods that had only marginally improved over the previous half century.

The idea of making a machine to "read" marks on paper was not entirely new. In 1938, IBM came out with an electrographic test-scoring machine that used the conductivity of special graphite pencils to interpret responses on what they called "mark sense" cards and convert them into punch cards. The mark sense system was used to score SAT college admission tests, and it was adapted for data capture in the 1951 Census of Canada. The Census Bureau initially considered using the same system for the U. S., but determined it was impractical because of the larger scale of the U.S. census.[82]

By 1951, the National Bureau of Standards (NBS) had settled on a purely optical sensing system that used a photoelectric cell to read marks onto magnetic tape. NBS engineers determined that precision could be maximized by taking negative microfilms of marked paper forms, so that the blackened responses would appear as transparent spots on the negatives. They used a cathode-ray tube to send a beam of light to each position on the form in sequence; when the light penetrated the microfilm, it was sensed by a photoelectric cell, and the response was encoded on magnetic tape.[83] The basic principle was essentially the same as Hollerith's original punch card, with an electric eye substituting for Hollerith's mercury-filled cups. Microfilm was initially adopted as an intermediate stage because it simplified the sensing of the marks, but it

au

Figure 9. Keypunch operator, 1950 Census
Source: U. S. Census Bureau, https://broadcast.census.gov/pio/photos/process_1950/process_1950_08010_hi.jpg.

immediately became apparent that film offered two practical advantages over paper. First, it was extremely compact; once filmed, the paper forms could be discarded and the microfilm could be preserved for long-term storage. Second, reels of microfilm are much easier to handle than sheets of paper, so microfilm greatly simplified automated feeding of the machine at high speed.

The first functional Film Optical Sensing Device for Input to Computers (FOSDIC) was delivered to the Census Bureau early in 1954 (see Figure 10). The device could read 2,000 marks per second, and it was successfully used for digitizing several surveys. Work on an improved version designed for the 1960 census began in 1956. The NBS built a prototype, and the Census Mechanical Laboratory constructed four identical units to be used for the census.[84]

[84] Eckler, *Bureau of the Census, Factfinder for the Nation*, 49-50. Duncan and Shelton, *Revolution in United States Government Statistics, 1926-1976*, 119 and 131. U.S. Bureau of the Census, *U.S. Census of Population and Housing: 1970 Procedural History* (Washington, 1976), 8-7.

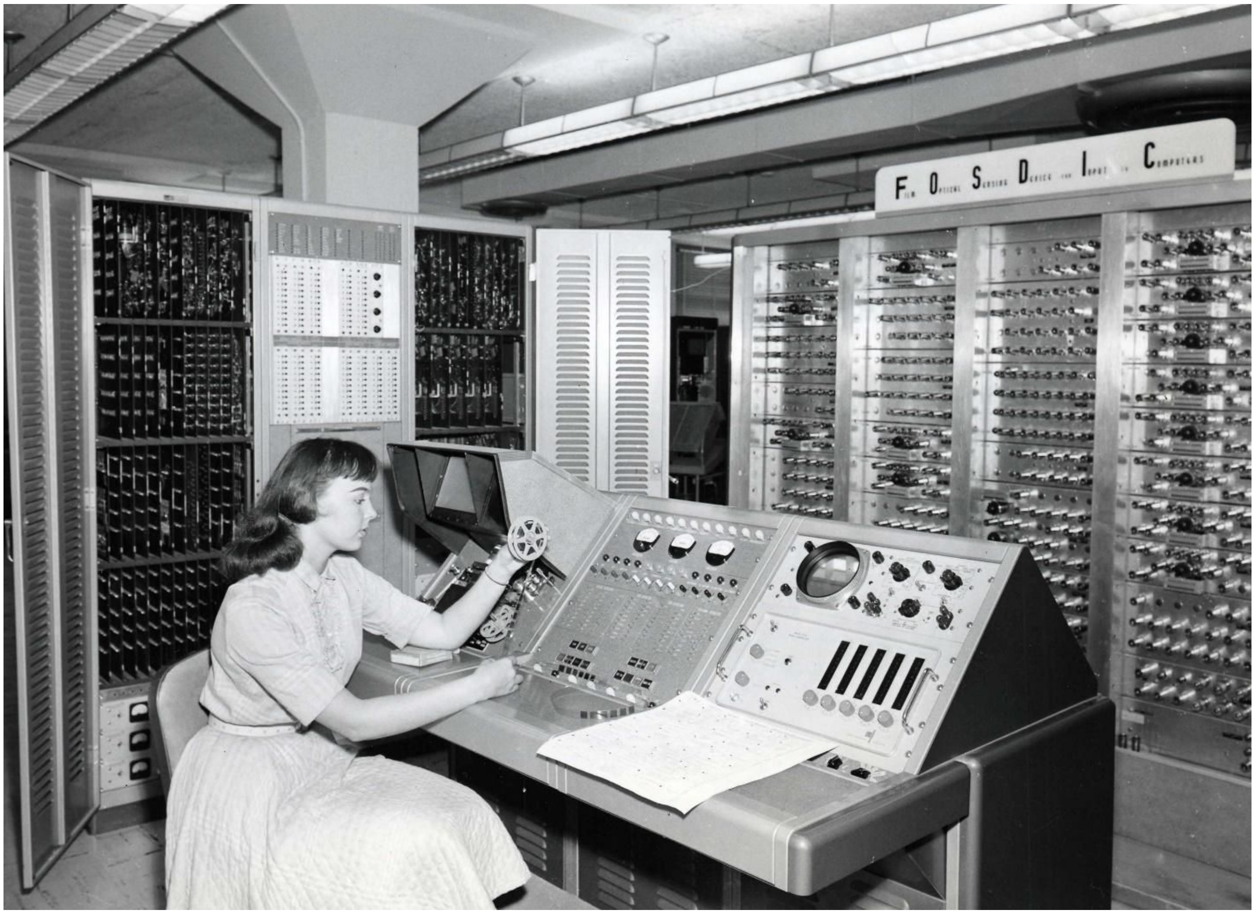


Figure 10. Film Optical Sensing Device for Input to Computers (FOSDIC)
Source: U. S. Census Bureau, https://www. census. gov/library/photos/fosdic_2. html.

The bubble sheets used to record information in the 1960 census (Figure 11) were filled in by enumerators going from door to door. This meant that the most labor-intensive part of data capture was carried out during the face-to-face interview. From 1850 to 1950, individual characteristics were transcribed in longhand on census forms, which then had to be centrally processed by hand. From 1850 to 1880, this hand processing was done by tally-marks; from 1890 to 1950, the forms were transcribed onto punch cards, so they could be tallied by machine. In a sense, the decentralization of data capture in 1960 was a throwback to the first data capture system of 1790-1840, since in both eras the front-line enumerators were the ones converting the data into a form suitable for processing.

The 1960 census made unprecedented use of sampling. All households were required to respond to the "short form" questionnaire, which included just seven questions per person. At every fourth household canvassed by enumerators, they dropped off a "long-form" questionnaire, with forty-six housing questions and twenty-nine additional questions pertaining to individual characteristics. In most of the country, respondents were instructed to fill out and mail back the long-form questionnaire, which was then copied onto FOSDIC bubble sheets.[85] The 1970 census further automated data capture. The Census Bureau constructed an address list that included most of the nation's households, and mailed out FOSDIC forms. The mailings instructed respondents to fill in the bubbles themselves and mail back the form. Households were randomly selected to receive a short form or a long form. Over 87% of households that received a FOSDIC form filled it out and sent it back, dramatically reducing data capture costs.[86]

The 1970 census also benefited from new and improved machinery designed and built by the Census Engineering Development Laboratory (formerly known as the Census Machine Shop and the Census Mechanical Laboratory). Microfilming was automated with a device that automatically fed the paper schedules. The long-form schedules were multi-page booklets, and the machine turned the pages using "a vacuum head mounted on a rotating arm" and filmed each page. The FOSDIC-70 machine was also greatly improved and could read 450 microfilmed pages a minute, a four-fold improvement over 1960.[87] Both automatic microfilming and optical sensing of the microfilm continued to advance over the next two decades. By 1990, the FOSDIC machine could read 40,000 pages per minute, about 350 times as fast as in 1960.[88]

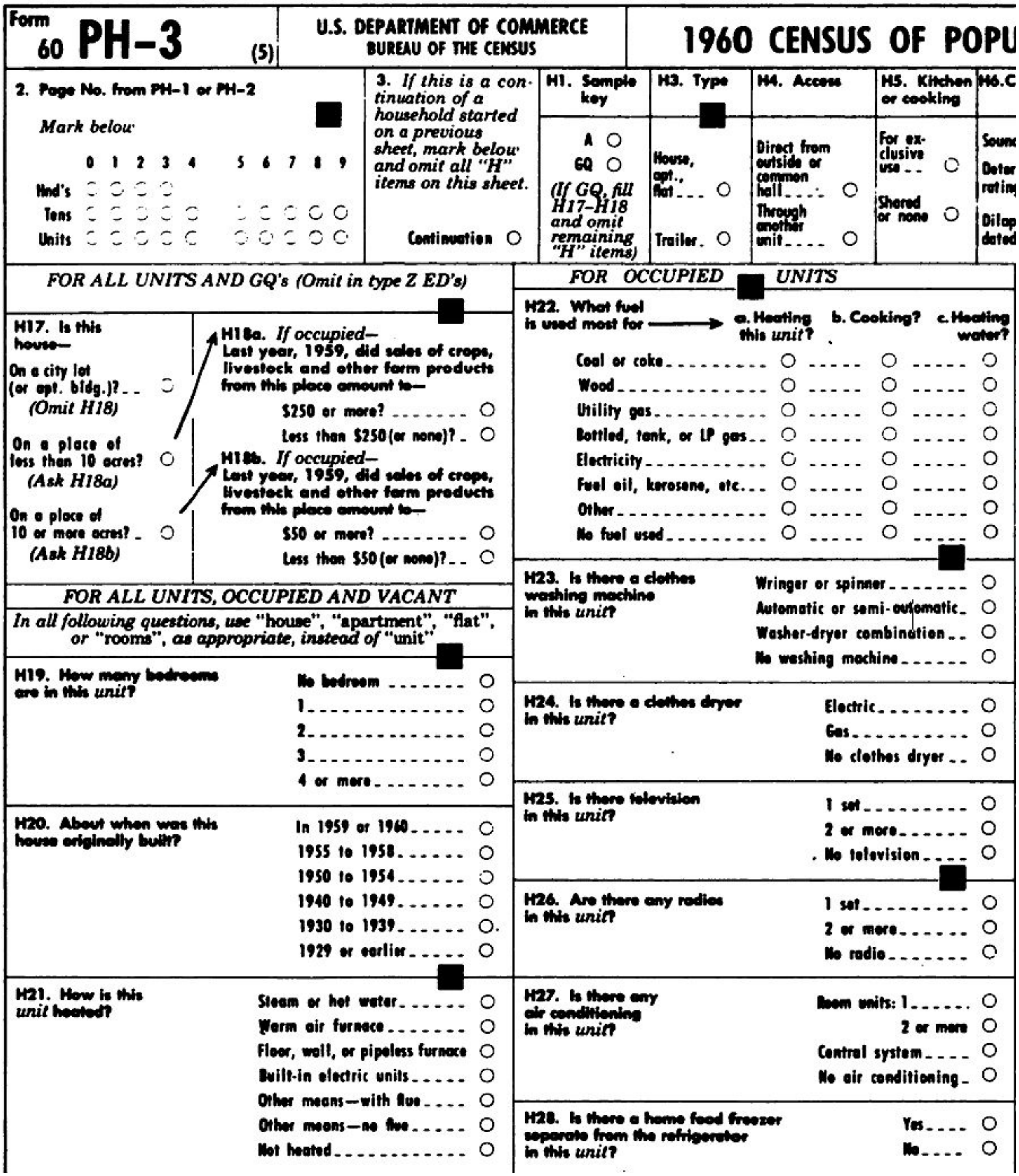
---

[85] The 1940 and 1950 censuses made some use of sampling, but the entire population answered most questions. U.S. Bureau of the Census, *1960 Censuses of Population and Housing: Procedural History*, Chapter 3, "Sampling and Estimation" (Washington, March 1965), 17-24.

[86] U.S. Bureau of the Census, *U.S. Census of Population and Housing: 1970 Procedural History* (Washington, 1976), 1-6.

[87] Duncan and Shelton, *Revolution in United States Government Statistics, 1926-1976*, 132. U.S. Bureau of the Census, *U.S. Census of Population and Housing: 1970 Procedural History* (Washington, 1976), 8-4-8-8.

[88] *1990 Census of Population and Housing—History* (Washington, U.S. Department of Commerce, October 1993), Chapter 1 p. 29.



Form 60 PH–3 (5)

U.S. DEPARTMENT OF COMMERCE
BUREAU OF THE CENSUS

1960 CENSUS OF POPU

2. Page No. from PH–1 or PH–2

Mark below

0 1 2 3 4 5 6 7 8 9

Hnd's
Tens
Units

3. *If this is a continuation of a household started on a previous sheet, mark below and omit all "H" items on this sheet.*

Continuation

H1. Sample key

A
GQ
(If GQ, fill H17–H18 and omit remaining "H" items)

H3. Type

House, apt., flat
Trailer

H4. Access

Direct from outside or common hall
Through another unit

H5. Kitchen or cooking

For exclusive use
Shared or none

H6.C

Soun
Deter ratin
Dilap dated

FOR ALL UNITS AND GQ's (Omit in type Z ED's)

H17. Is this house—

On a city lot (or apt. bldg.)?
(Omit H18)

On a place of less than 10 acres?
(Ask H18a)

On a place of 10 or more acres?
(Ask H18b)

H18a. *If occupied—* Last year, 1959, did sales of crops, livestock and other farm products from this place amount to—

$250 or more?
Less than $250 (or none)?

H18b. *If occupied—* Last year, 1959, did sales of crops, livestock and other farm products from this place amount to—

$50 or more?
Less than $50 (or none)?

FOR ALL UNITS, OCCUPIED AND VACANT

In all following questions, use "house", "apartment", "flat", *or* "rooms", *as appropriate, instead of* "unit"

H19. How many bedrooms are in this *unit*?

No bedroom
1
2
3
4 or more

H20. About when was this house originally built?

In 1959 or 1960
1955 to 1958
1950 to 1954
1940 to 1949
1930 to 1939
1929 or earlier

H21. How is this *unit* heated?

Steam or hot water
Warm air furnace
Floor, wall, or pipeless furnace
Built-in electric units
Other means—with flue
Other means—no flue
Not heated

FOR OCCUPIED UNITS

H22. What fuel is used most for → a. Heating this *unit*? b. Cooking? c. Heating water?

Coal or coke
Wood
Utility gas
Bottled, tank, or LP gas
Electricity
Fuel oil, kerosene, etc.
Other
No fuel used

H23. Is there a clothes washing machine in this *unit*?

Wringer or spinner
Automatic or semi-automatic
Washer-dryer combination
No washing machine

H24. Is there a clothes dryer in this *unit*?

Electric
Gas
No clothes dryer

H25. Is there television in this *unit*?

1 set
2 or more
No television

H26. Are there any radios in this *unit*?

1 set
2 or more
No radio

H27. Is there any air conditioning in this *unit*?

Room units: 1
2 or more
Central system
No air conditioning

H28. Is there a home food freezer separate from the refrigerator in this *unit*?

Yes
No

Figure 11. FOSDIC Bubble-Sheet for 1960 Census (Detail)
Source: U.S. Bureau of the Census, *United States Census of Population and Housing 1960: Principal Data-Collection Forms and Procedures* (Washington, 1961), 21.

The most consequential Census Bureau data capture innovation of the late-twentieth century was the development of computerized digital street mapping. As noted, after 1960, the Bureau decided to abandon door-to-door enumeration in many places, substituting a mail-out mail-back enumeration strategy. This made development of a complete computerized address list essential. In 1967, researchers working on a pretest for the 1970 census in New Haven, Connecticut, decided that a digital map of the city was needed to ensure complete coverage.[89] Census Bureau mathematician James Corbett had the idea of representing streets and intersections as vectors, and Census software developers created a protocol they called Geographic Base File-Dual Independent Map Encoding (GBF-DIME) to store the information.[90] During the 1970s, the Bureau developed digital maps for all U. S. cities based on the GBF-DIME technology, and shared the maps and the associated software with local planning agencies so they could make updates, corrections, and adaptations to meet local planning needs.[91]

In 1983 the Census Bureau reached a data-sharing agreement with the U. S. Geological Survey, which provided scanned maps to form the basis for a comprehensive national digital map including all streets and roads, as well as other geographical features, such as rivers, political boundaries, railroads, and locations of each housing unit. The new system—Topologically Integrated Geographic Encoding and Referencing (TIGER)—was far more accurate and comprehensive than GBF-DIME, and was the first nationwide general-purpose spatial dataset.[92] TIGER was used extensively in data capture for the 1990 census. After the

---

[89] J. Sperling, "Development and Maintenance of the TIGER Database: Experiences in Spatial Data Sharing at the US Bureau of the Census," in *Sharing Geographic Information*, ed. H.J. Onsrud and G. Rushton (Center for Urban Policy Research: New Brunswick, NJ, 1995), 377-396. Margo Anderson, *The American Census: A Social History*, 253-257.

[90] "James Corbett," United States Census Bureau, https://www.census.gov/history/www/census_then_now/notable_alumni/james_corbett.html. "DIME Underwent Lots of Testing Too," *Census Bulletin*, vol. XVIII, no. 18 (September 6, 1968), https://www.census.gov/history/pdf/1968censusbulletin-dime.pdf.

[91] Shoreh Elhami and Robert A. Lamacchia, "Geographic Information Systems GIS," in *Encyclopedia of the U.S. Census: From the Constitution to the American Community Survey* (Washington, D.C., 2012), 248-249.

[92] Donald F. Cooke, "Topology and TIGER: The Census Bureau's Contribution," chapter in *The History of Geographic Information Systems: Perspectives from the Pioneers*, Timothy W. Foresman, editor (New

census, the Bureau released the TIGER files to the public. The availability of a national spatial database free of any licensing restrictions stimulated development of a Geographic Information Systems industry.[93] Since 1990, the Census Bureau has continuously revised and improved the TIGER system, and it is now an indispensable element of national data infrastructure, used by planning agencies, nonprofits, and commercial entities for an extraordinary range of applications.[94]

Privatization, 1990-2020

For most of the twentieth century, the Federal Government expanded the scope of its activities. The first phase of the New Deal programs in the Roosevelt Administration built massive dams and bridges, provided insurance for banks, subsidized farmers, and developed new electric power systems. The second phase introduced vast new insurance programs covering unemployment and old-age support. After World War II, the Federal Government built the immense Interstate Highway System and made unprecedented investments in health, education, and anti-poverty initiatives, culminating in the Great Society programs of the 1960s.

The expansion of government came under attack in the 1980s. President Reagan believed that the private sector was inherently more efficient than the government, and campaigned to privatize government functions. In his first inaugural address, Reagan asserted that "government is not the solution to our problem; government *is* the problem."[95] The Reagan Administration defined privatization as "a strategy to shift the production of goods and services from the Government to the private sector in order to reduce Government

---

Jersey, 1998), 47. United States Census Bureau, "25th Anniversary of TIGER," http://census.maps.arcgis.com/apps/MapJournal/index.html?appid=2b9a7b6923a940db84172d6de138eb7e.

93 Donald F. Cooke, "Topology and TIGER: The Census Bureau's Contribution," 49. Elhami and Lamacchia, "Geographic Information Systems (GIS)," 248.

94 Shoreh Elhami and Robert A. Lamacchia, "Geographic Information Systems (GIS)," 249. United States Census Bureau, "25th Anniversary of TIGER," http://census.maps.arcgis.com/apps/MapJournal/index.html?appid=2b9a7b6923a940db84172d6de138eb7e.

95 Ronald Reagan, "Inaugural Address," January 20, 1981, http://www.presidency.ucsb.edu/ws/?pid=43130.

expenditures and to take advantage of the efficiencies that normally result when services are provided through the competitive marketplace."[96] With a Democratic majority in both houses of Congress committed to preserving the achievements of the New Deal and Great Society, Reagan's privatization efforts had only modest consequences.

President Clinton had a much bigger impact. He campaigned in 1992 on a "third way" platform of shrinking the federal government; he proposed to "Cut 100,000 unnecessary bureaucratic positions through attrition and mandate three percent across-the-board savings in every federal agency."[97] Once in office, his "reinventing government" initiative expanded this goal to a reduction of 272,900.[98] With the support of Congress, every department in the executive branch began a program of outsourcing, and hundreds of thousands of jobs formerly performed by federal employees were transferred to private contractors. Figure 12 shows the number of executive branch employees per million Americans. The drop in federal jobs began during the Carter Administration, with almost a 10% reduction. The trend leveled off during the Reagan and first Bush administrations, but then began a precipitous drop during the Clinton years, when the ratio of federal jobs to population fell by almost 30%. Since the Clinton era, there has been little change, except for a temporary bump upwards during the Great Recession.

---

[96] Office of Management and Budget, Budget of the United States Government, Fiscal Year 1988 Budget, supra note 4 at Supp. 2-38; Michal Laurie Tingle, Privatization and the Reagan Administration: Ideology and Application. Yale Law and Policy Review 6:1 (1988): 229-257.

[97] "Bill Clinton for President 1992 Campaign Brochures," http://www.4president.org/brochures/billclinton1992brochure.htm. The American Presidency Project, "Executive Order 12839--Reduction of 100,000 Federal Workers," February 10, 1993, http://www.presidency.ucsb.edu/ws/index.php?pid=61535.

[98] "A Brief History of the National Performance Review," https://govinfo.library.unt.edu/npr/library/papers/bkgrd/brief.html. Dale V. Jones, *Downsizing the Federal Government* (Armonk, NY: M.E. Sharpe, 1998): 4.

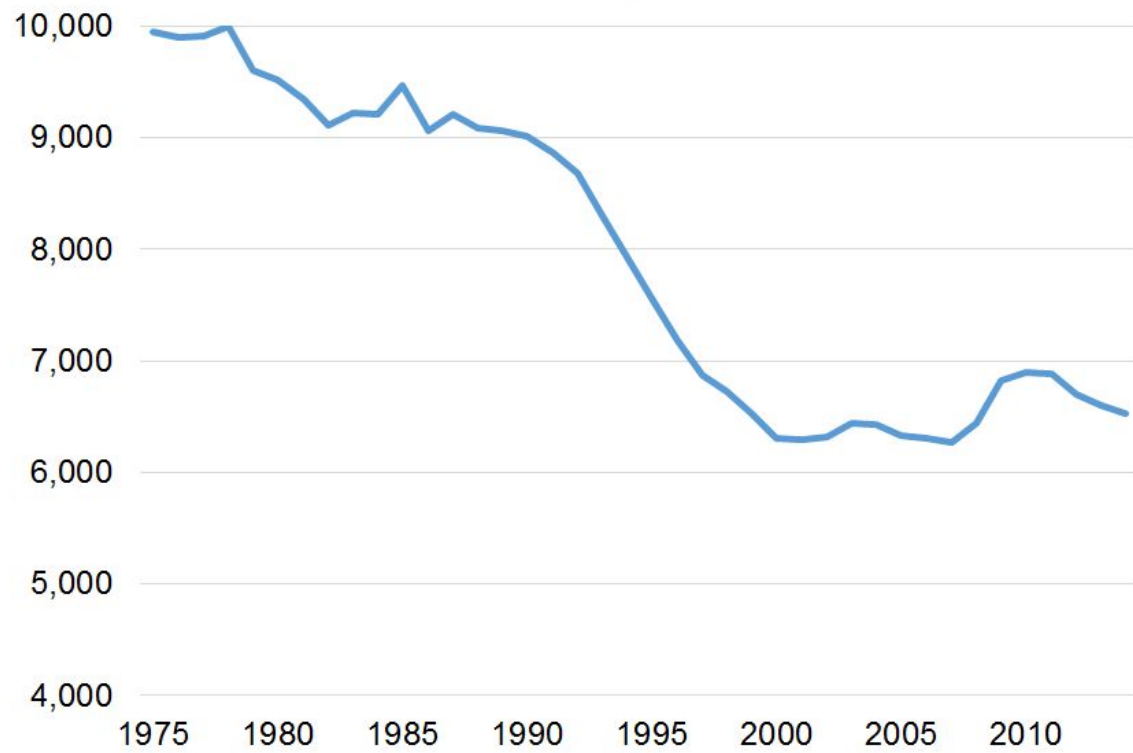


Figure 12. Executive Branch Employment per Million Population: 1975-2014
Source:
https://www.opm.gov/policy-data-oversight/data-analysis-documentation/federal-employment-reports/historical-tables/executive-branch-civilian-employment-since-1940.

In 1996, Clinton announced in his State of the Union Address that "the era of big government is over."[99] By that year, executive branch employment had already dropped by 300,000 workers. Clinton signed the Federal Activities Inventory Reform Act (FAIR) in 1998, which required federal agencies to "review their activities and define them as either inherently governmental or potentially subject to privatization," and the number of federal employees fell by another 150,000.[100] At the same time that the federal workforce was

[99] William J. Clinton, "Address Before a Joint Session of the Congress on the State of the Union," January 23, 1996, *The American Presidency Project*, http://www.presidency.ucsb.edu/ws/index.php?pid=53091.

[100] Louis Jacobson, "Clinton signs privatization bill," October 22, 1998, *Government Executive*, https://www.govexec.com/federal-news/1998/10/clinton-signs-privatization-bill/4725/. Public Law 105-270-October 19, 1998, 112 Stat. 2382, "Federal Activities Inventory Reform Act of 1998," https://www.congress.gov/105/plaws/publ270/PLAW-105publ270.pdf. Data, Analysis, and Documentation: Federal Employment Reports, "Historical Federal Workforce Tables," https://www.opm.gov/policy-data-oversight/data-analysis-documentation/federal-employment-reports/historical-tables/executive-branch-civilian-employment-since-1940/.

shrinking, the responsibilities of the federal government continued to grow. The inevitable consequence was a massive increase in the number of private contractors, who began doing work previously carried out by federal employees.

The Census Bureau was transformed by privatization. As planning for the 2000 census ramped up in the early 1990s, there was intense pressure to identify census functions that could be outsourced to private vendors. In 1993, the Clinton Administration's privatization mandate was compounded by a hiring freeze and budget reductions, which restricted the ability of the Bureau to hire experts to investigate the potential for privatizing data capture operations. Accordingly, the Bureau outsourced the job of investigating the feasibility of outsourcing by hiring the Rochester Institute of Technology Research Corporation (RITRC) to investigate requirements for outsourcing data capture. RITRC developed a set of technical specifications for replacing the FOSDIC infrastructure with digital imaging directly from paper. Building on the RITRC work, in early 1996 the Census Bureau commissioned Advanced Research Technologies, Inc (ARTI) to conduct a benefit-cost analysis comparing the FOSDIC system with digital imaging from paper. ARTI recommended that the Bureau switch to digital imaging of paper forms. The report concluded that digital imaging would be the least expensive solution, but also noted that it was high risk, since the technology was untested and had never been used at the scale needed for the census.[101]

In February 1996, John H. Thompson, then Chief of the Decennial Management Division of the Census Bureau, wrote that based on the ARTI benefit cost analysis, "I am recommending that the Census Bureau require the use of imaging technology to perform the data capture function for the 2000 Census."[102] The recommendation was approved, and in August 1996, the Bureau invited bids for data capture for the 2000 Census. The same year, the Bureau permanently shut down the Technical Services Division, the direct descendant of the machine shop established by Census Director North in 1907, signaling the end of the

[101] U.S. Census Bureau, *History: 2000 Census of Population and Housing*, 274-276.

[102] Census 2000 Decision Memorandum No. 1, 2000 Decennial Census: DMD to DIR Memorandum No. 9609. 21 February 1996.

Census Bureau's century-long effort to develop new technologies for data capture. Vice President Al Gore gave the Census Bureau a Hammer Award, consisting of a "$6.00 hammer with a little red, white, and blue ribbon," which was "the Vice President's symbolic answer to the $400.00 hammer of yesterday's government."[103]

The decision to outsource data capture to the private sector was not a major concern for Census Bureau leadership. The Census Bureau historian conducted oral history interviews with key leaders of the era, explicitly asking them to identify the most important developments at the Bureau during their tenure; none of the leaders of the period from 1989 through 1998 mentioned privatization or outsourcing.[104] The Clinton administration took almost two years to get a census director in place, and during that critical period an acting director authorized the initial contracts to study outsourcing of data capture.[105] Director Martha Riche, confirmed in November 1994, had little experience in managing contracts. Privatization was implemented by Michael Longini, Chief of the Census Bureau Decennial Systems and Contracts Management Office, a 25-year veteran of the Bureau. Longini recruited the assistance of Commerce Department staff to identify requirements, draft requests for proposals, and evaluate bids. Despite continuing resistance from some career staff, Longini prevailed.[106]

---

[103] U.S. Census Bureau, *History: 2000 Census of Population and Housing*, 274. Appendix I: List of Hammer Awards as of August 1, 1996, https://govinfo.library.unt.edu/npr/library/nprrpt/annrpt/vp-rpt96/appendix/hammer.html

[104] Charles D. Jones, Associate Director for the Decennial Census, 1987-1994, interviewed March 27, 1997. https://www.census.gov/history/pdf/charles_jones_oral_history.pdf; William P. Butz, Associate Director for Demographic Programs, 1987-1995, interviewed Sept. 8, 1998 https://www.census.gov/history/pdf/Butz_Oral_History.pdf; Barbara Everett Bryant, Director 1989-1993, interviewed January 19, 1993 https://www.census.gov/history/www/reference/oral_histories/barbara_everitt_bryant.html; Martha Farnsworth Riche, Director, 1994-1998; interviewed January 28, 1998. https://www.census.gov/history/pdf/Riche_Oral_History.pdf

[105] Prepared statement of Harry A. Scarr, Hearing of the Subcommittee on Census Statistics and Postal Personnel, House of Representatives, 103rd Congress, May 27, 1993. Serial No. 103-12.

[106] Brian Freil, "Relationships Count." Government Executive, August 15, 2001. https://www.govexec.com/magazine-acquisition-awards/magazine-acquisition-awards-census-bureau/2001/08/relationships-count/9760/. U.S. Census Bureau, *History: 2000 Census of Population and Housing*, 283.

The two major contracts for data capture in the 2000 census were won by the aerospace contractors Lockheed Martin and TRW. Lockheed Martin won the $49 million contact for Data Capture System 2000 (DCS 2000), to provide the hardware and software needed to process paper census forms through scanning, Optical Mark Recognition (OMR), and Optical Character Recognition (OCR). OMR had been used by the Census since the introduction of FOSDIC in 1960; OCR, however, was new, and involved the automatic interpretation of open-ended handwritten census responses, reducing the need for hand-keying data. TRW won a $188 million contract for setting up and operating regional Data Capture Centers where the paper forms would be processed and converted to electronic format.[107]

There were many unanticipated costs. The contractors did not consult with the Census Bureau personnel who had institutional knowledge and experience processing millions of paper forms. The Bureau's lack of experience with contractors led to inefficiencies. Requirements were poorly documented, resulting in frequent changes. There were major philosophical differences between the contractors and the Bureau, especially in the area of quality assurance. Misunderstanding led to change orders, which increased costs.[108]

The digital imaging and OCR systems did not work as well as anticipated. The scanning machines were far slower than the FOSDIC machines they replaced, so the number of scanning machines grew dramatically. The 1990 census used eight active FOSDIC machines, with several additional machines in reserve in case of breakdowns.[109] DCS 2000 replaced the eight FOSDIC machines with 162 Kodak Digital Science Scanner 9500 machines operating at one-thirtieth the speed of the old technology.[110] The big advantage of the new system was its OCR capability for reading hand-written responses, but in practice the

---

[107] Colleen O'Hara, "Census counts on Lockheed to build $49M imaging system," *FCW The Business of Federal Technology*, March 30, 1997, https://fcw.com/Articles/1997/03/30/Census-counts-on-Lockheed-to-build-49M-imaging-system.aspx. Colleen O"Hara, "TRW wins $187M data center contract," *FCW The Business of Federal Technology*," February 1, 1998, https://fcw.com/Articles/1998/02/01/TRW-wins-187M-data-center-contract.aspx.

[108] U.S. Census Bureau, *History: 2000 Census of Population and Housing*, 274.

[109] U.S. Census Bureau, *1990 Census of Population and Housing History* (Washington, 1996), 1-29.

[110] U.S. Census Bureau, *History: 2000 Census of Population and Housing*, 286.

OCR proved much less reliable than had been hoped, so a high proportion of the forms had to be manually keyed.

As census day approached, the General Accounting Office (GAO) warned repeatedly that DCS 2000 was not ready. The system had inadequate testing and had not been demonstrated to provide the speed and accuracy needed to meet statutory deadlines for production of census results. In the 1998 dress rehearsal for the census, the data capture system crashed repeatedly "due to flaws in the system software, which could not handle the workload." The lenses on the scanners quickly accumulated dust, so every fifteen minutes the system had to be shut down to clean them. DCS 2000 was further plagued by missing data, misinterpretation of responses, sorter jams, and a high error rate for write-in responses.[111]

The Census Bureau had planned to use a sampling strategy to improve accuracy and reduce the cost of following up on people who had not responded to the census, but in January 1999 the Supreme Court blocked the use of this approach. That meant more cost. Compounding the challenges facing the census, the technology still had serious problems. An operational test in Pomona, California failed in November 1999. This end-to-end test of data capture revealed a much higher rate of manual keying of open-ended responses than anticipated, along with a manual keying rate only half as fast as projected. This bottleneck would slow the entire system to unacceptable levels. With six weeks to go before census day, the GAO reported that "the Bureau faces formidable challenges in performing critical data capture operations." DCS 2000 was supposed to be fully operational by October 1999; because of system requirement changes, the target date slipped to February 25, 2000, just two weeks before data capture operations were scheduled to begin.[112]

The operational testing of DCS 2000 showed that the system was too slow to digitize the entire census by the statutory deadline of December 31, 2000, when the statistics needed for reapportionment had to

---

[111] U.S. Census Bureau, *History: 2000 Census of Population and Housing*, 278. Colleen O'Hara, "GAO Questions Success of Census Test," *FCW The Business of Federal Technology*, August 9, 1998 https://fcw.com/articles/1998/08/09/gao-questions-success-of-census-test.aspx.

[112] General Accounting Office, *2000 Census: New Data Capture System Progress and Risks* (GAO/AIMD-00-61, Washington, 2000).

be delivered to Congress.[113] At the eleventh hour, the Census Bureau and the contractors redesigned the workflow to meet the deadline. Instead of digitizing the entire census as it came in to the processing centers, the Bureau adopted a “two pass” approach to data capture. During the first pass the contractors would process only the items needed for the apportionment counts. Once that work was complete, the contractors would begin work on capturing additional fields, such as income and education. This late fix required new software and hardware, as well as a substantial extension of the overall processing time.[114]

The 2000 census narrowly averted disaster. A post-mortem study concluded that the Census Bureau had failed to capitalize on the institutional knowledge and experience of Bureau professional staff, relying instead on military contractors with no prior experience in census operations. Some census career professionals felt that continuing use of the FOSDIC system would have been far more cost-effective than outsourcing data capture. Bureau in-house content and processing experts profoundly disagreed with TRW’s decisions on data capture procedures. Subject matter experts within the Bureau complained that they could not even understand how DCS 2000 worked because the documentation was written in technical jargon.[115] Poor understanding of data capture requirements led to frequent changes. The contracts were “cost-plus,” meaning that any increase in expenses of contractors could be passed along to the Census Bureau. In the end,

---

[113] General Accounting Office, Testimony Before the Subcommittee on the Census Committee on Government Reform, House of Representatives, *2000 Census: Status of Key Operations, Statement of J. Christopher Mihm, Associate Director, Federal Management and Workforce Issues, General Government Division*, GAO/T-GGD/AIMD-00-91, February 15, 2000, https://www.gao.gov/new.items/g100091t.pdf. U.S. Census Bureau, *History: 2000 Census of Population and Housing*, 281. GAO, United States General Accounting Office, Report to the Subcommittee on the Census, Committee on Government Reform, House of Representatives, “2000 Census, Contingency Planning Needed to Address Risks That Pose a Threat to a Successful Census,” (GAO/GGD--00-6), December 1999, https://www.gao.gov/assets/230/228556.pdf.

[114] U.S. General Accounting Office, correspondence, “2000 Census: Update on Data Capture Operations and System,” September 29, 2000, http://www.gao.gov/assets/100/90103.pdf.

[115] Titan Systems Corporation/System Resources Division, “Census 2000 Data Capture System Requirements Study,” Census 2000 Evaluation R.3.d, August 23, 2002. Donald Kline. Census 2000 Testing, Experimentation, and Evaluation Program Topic Report No. 3, TR-3, Census 2000 Data Capture, U. S. Census Bureau, Washington, 2004. https://www.census.gov/pred/www/rpts/TR3.pdf.

the cost for the Lockheed Martin contract jumped from $49 million to $220 million, and the contract to TRW went from $188 million to $314 million.[116]

Planning for the 2010 census began before the processing of the 2000 census was complete. A key innovation for the 2010 census was elimination of the long form census questionnaire. From 1940 to 2000, the Census Bureau asked a subset of the census respondents an extra set of detailed questions. In 2000, for example, one-in-six households received a long census form with eighty-one questions, ranging from the amount of their electric bill to the educational attainment of each household member. The other five-sixths of the population received a short form with just seven questions. The use of both the long form and short form reduced the burden of filling out the census form for most people, but still provided sufficient information to produce detailed statistics for individual communities. In 2010, the long form was replaced by the American Community Survey (ACS). The ACS is similar to the 2000 long form, but instead of asking the questions as part of the decennial census, the ACS is spread out over the course of the decade. Each year, about three percent of the population is surveyed; by adding up the responses for several years, the ACS can provide community-level statistics similar to those from the decennial long form.

The elimination of the long form from the decennial census promised dramatic savings in data capture. In Census 2000, the long form accounted for 63% of census responses, even though it went out to only a sixth of the population. Other new technologies promised additional savings. The Census Bureau planned an internet response option, so that people could fill out the census using a web-based form. For in-person interviews of people who did not initially respond by mail, internet, or phone, the Bureau planned to use hand-held devices to capture the data at the time of the interview, along with the precise GPS

---

[116] U.S. Census Bureau, *History: 2000 Census of Population and Housing*, 283-284. An enumerator for Census 2000 noted a "small example of the problem with using private contractors with lowest cost bids for the Census: The follow-up enumerators work kits included mechanical pencils with poor quality lead that quickly broke. The size of leads used by the pencils were not sold in any major office supply stores, so the enumerators had to buy their own number 2 pencils to fill out the forms. Perhaps the provider put in the lowest bid, but the material didn't work so the actual cost was absorbed by individual enumerators supplying a substitute." Personal communication, Miriam L. King, July 3, 2018.

coordinates of the housing unit. Both the internet response option and the handheld devices would produce tabulation-ready digital data, greatly reducing the burden of processing paper forms.

The reliance on outsourcing was even greater in 2010 than it had been in 2000.[117] The biggest data capture contract was the Decennial Response Integration System (DRIS), which included both the paper data capture operations and the development of an internet response option. The DRIS contract went to Lockheed Martin; the initial award in 2005 was for $500 million, a dramatic jump from the size of the contracts awarded for the 2000 census. Six months later, the Census Bureau turned to another defense contractor, the Harris Corporation, awarding a $600 million contract for the Field Data Collection Automation Project. The goal of the Harris contract was development of a hand-held electronic device (Figure 13) for in-person census enumeration.[118] Problems quickly arose. Between 2004 and 2008, the GAO issued nine warnings that the census was at risk because of mismanagement of the contractors.[119] These warnings were prescient.

The internet response option was promptly abandoned. There had been an internet option for the 2000 Census, but it was not widely publicized and only garnered 63,000 responses. Nevertheless, a subsequent Census Bureau evaluation judged the experiment an operational success, and made the recommendation that this response mode should be better publicized in future censuses. The Bureau conducted successful usability tests of internet response in 2003 and 2005. Despite this early progress, the DRIS contractor Lockheed Martin announced in 2006 that it could not provide an internet response facility in time for the 2008 dress rehearsal, which meant that a large-scale test would be impossible before Census day.

---

[117] U.S. Government Accountability Office, "2010 Census: Census Bureau Generally Follows Selected Leading Acquisition Planning Practices, but Continued Management Is Needed to Help Ensure Success," GAO-06-277, May 18, 2006. https://www.gao.gov/assets/260/250136.pdf.

[118] Harris had also received a $209 million contract in 2002 to improve the accuracy of the Census Bureau's address list and maps. U.S. Government Publishing Office, "2010 Census: Census Bureau Generally Follows Selected Leading Acquisition Planning Practices, but Continued Management Is Needed to Help Ensure Success," May 18, 2006.

[119] Rep. Henry A Waxman, *Fact Sheet: Chronology of Warnings about the Census Bureau's Field Data Collection Automation System* (Washington, 2008), https://democrats-oversight.house.gov/sites/democrats.oversight.house.gov/files/migrated/20080409140918.pdf.

Census Director Kincannon immediately made the surprise decision to cancel internet response, citing the risk of "phishing" and denial of service attacks, as well as the Lockheed Martin failure to deliver the system in time for large-scale testing.[120]

The handheld device contract with the Harris Corporation was an even greater disaster. The software did not function correctly, the work fell behind schedule, and the projected cost more than doubled to $1.3 billion. In 2008 the Census Bureau abruptly canceled the plan to use the handheld devices for non-response follow-up and reverted to entirely paper-based processing. The last-minute change further increased the cost of the 2010 census by up to three billion dollars, making the Harris debacle one of the most expensive failed software systems in history.[121]

Figure 13. Handheld Harris Device for 2010 Enumeration
Source: U. S. Census Bureau http://calleam. com/wp-content/uploads/2-fdca_handheld-device_high_resolution1. jpg.

[120] "Apportionment in the Balance: A Look Into the Progress of the 2010 Decennial Census," Hearing before the Subcommittee on Federalism and the Census of the Committee on Government Reform, House of Representatives, 109th Congress, 2nd Session, March 1, 2006, https://archive.org/stream/gov.gpo.fdsys.CHRG-109hhrg27016/CHRG-109hhrg27016_djvu.txt. National Research Council. *Envisioning the 2020 Census,* Committee on National Statistics, Division of Behavioral and Social Sciences and Education (Washington, 2010).

[121] U.S. Census Bureau, "2010 Census Planning Memoranda Series, No. 190," April 30, 2012, https://www.census.gov/2010census/pdf/2010_Census_NRFU_Operations_Assessment.pdf.

In the wake of the cancellation of the two key innovations designed to modernize the 2010 census and reduce the cost of data capture, there was bipartisan congressional condemnation of the Census Bureau. Senator Coburn blamed "the failure, mismanagement, and incompetence of the Census Bureau." Senator Mikulski said the fiasco "borders on scandal," Representative Maloney called it "a statistical Katrina," and Representative Waxman pronounced it a "colossal failure."[122]

Other countries had much better success in modernizing their censuses. The Brazilian 2010 census was conducted entirely by in-person interviews using handheld devices for data capture. The Brazilian census had one of the longest questionnaires in the world, with thirty-seven short-form questions and 107 long-form questions, compared with just ten questions on the 2010 U. S. census. The cost for the 190,000 devices used in the Brazil enumeration was less than $100 each, a total of under $20 million.[123] Multiple countries have used the internet to collect census data. Canada, for example, began offering an internet response option for its 2006 census, and 18% of the population took advantage of it. By the next census in 2011, 54% of Canadians responded by internet, driving the overall self-response rate up to a remarkable 85%.[124]

After the cancellation of both the handheld devices and the internet response option, the 2010 U.S. census had to be conducted entirely on paper. The cost ballooned; the Lockheed-Martin contract doubled to $1.02 billion, Harris received $1.06 billion, and the other major contractors were paid $580 million.[125] The

---

[122] "Statement of Sen. Thomas Coburn on the Supplemental Appropriations Act of 2008." 110th Congress, 2nd Session. Congressional Record Vol 154 No. 107, S6234-6237. "Opening Statement of Rep. Henry A. Waxman, Committee on Oversight and Government Reform, Joint Hearing on 2010 Census: Field Data Collection Automation Contract (FDCA) Contract," 110th Congress, House of Representatives, Committee on Oversight and Government Reform, April 9, 2008.

[123] Economic Commission for Europe, "The 2010 Brazilian Population Census: innovations and lessons learned," (April 30, 2012), https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2012/38-SP_Brazil.pdf.

[124] By 2016, Canada's internet response rate was up to 68%, and the overall self-response rate up to 89%. Statistics Canada, 2016 Census of Population collection response rates. http://www12.statcan.gc.ca/census-recensement/2016/ref/response-rates-eng.cfm.

[125] Daniel H. Weinberg, "Management Challenges of the 2010 U.S. Census," July 28, 2011, https://www.census.gov/history/pdf/weinberg-managing2010.pdf.

privatization of data capture led not to the promised efficiencies, but rather to rapidly escalating costs, reduced productivity, and near catastrophic failure of both the 2000 and 2010 censuses.[126]

Figure 14 describes the costs of the census from 1940 to 2010 in constant 2010 dollars. Panel A on the left shows the cost per household enumerated, since the data are collected on a household-by-household basis. By this measure, there was a 40% decline in census costs from 1940 to 1960. This savings was partly a consequence of the growing use of sampling in that era, which meant that there were fewer data items to be captured, despite the continuing population increase. From 1970 to 1990 census costs per household escalated rapidly, and census officials and outside critics pointed to this increase as evidence of the need for outsourcing. Privatization of census data capture in 2000, however, came with a heavy price tag, as costs per household rose by an unprecedented 79%, followed by another 30% increase in 2010.

Panel B on the right side of Figure 14 shows an alternative measure of census costs: the expense of the census per item of information gathered. Since 1950, costs per item have increased steadily, but the pace of increase accelerated dramatically with privatization. From 1990 to 2000, census costs per item of data collected doubled. With the elimination of the long form from the decennial census on 2010, the amount of information collected dropped in half, and the cost per item of information collected almost tripled.

The 2020 census will be almost entirely reliant on private contracts, and the scale of the work to be outsourced is greater than ever. Costs may not go up as much in 2020 as in 2000 or 2010, but there is still likely to be a significant increase. Even if the Census Bureau wanted to carry out some of the data capture operations in-house, after twenty-five years of attrition there are few employees left with the skills and

---

[126] *2020 Census Operational Plan: A New Design for the 21st Century*, Issued November 2015, Version 1.1, pp. 5-7. U.S. Census Bureau, *History: 2000 Census of Population and Housing* (Washington, 2009). Michael Krigsman, "Billion-dollar IT failure at Census Bureau" *Beyond IT Failure* (March 20, 2008). http://www.zdnet.com/article/billion-dollar-it-failure-at-census-bureau/. GAO U.S. Government Accountability Office, *High Risk Series, Progress on Many High Risk-Areas, While Substantial Efforts Needed on Others*, GAO-17-317, February 2017. U.S. General Accounting Office, *2000 Census: New Data Capture System Progress and Risks* (GAO/AIMD-00-61, Washington, 2000): 4. U.S. General Accounting Office, Report to the Subcommittee on the Census, Committee on Government Reform, House of Representatives, "Headquarters Processing System Status and Risks," October 2000.

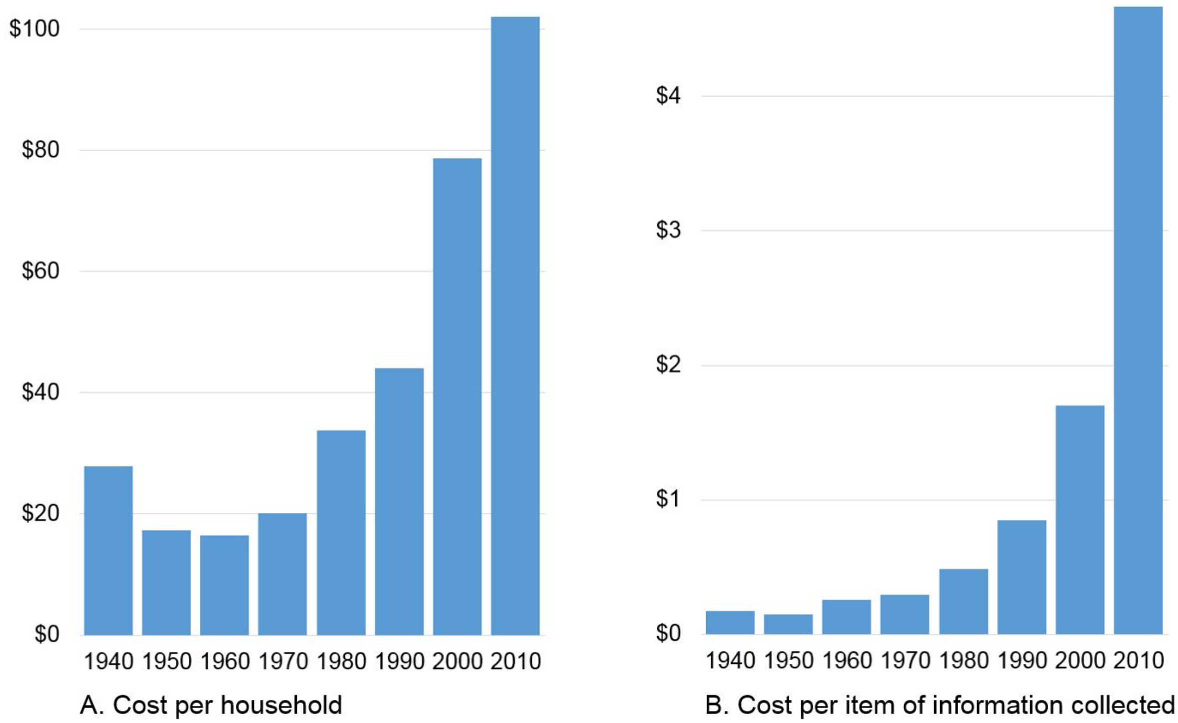


Figure 14. Cost of Decennial Censuses per household and per item collected, 1940-2010
Sources: Costs in 1940-1990: U.S. Census Bureau, *Measuring America: The Decennial Censuses from 1790 to 2000* (Washington, 2002), A1; Costs in 2000: U.S. Census Bureau, *History, 2000 Census of Population and Housing* (Washington, 2009), 28; Costs in 2010: Government Accountability Office, 2020 Census: Progress Report on the Census Bureau's Efforts to Contain Enumeration Costs (GAO-13-857T) (Washington, 2013), https://www.gao.gov/assets/660/657732.pdf; 2010 cost adjustment to remove the cost of the ACS: Government Accountability Office, 2020 Census: Census Bureau Should Take Action to Improve the Credibility and Accuracy of Its Cost Estimate (GAO-08-554),(Washington, 2008), https://www.gao.gov/assets/280/276782.pdf; Number of households and items collected: authors' calculations based on Steven Ruggles, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. *Integrated Public Use Microdata Series: Version 7.0* [dataset]. Minneapolis: University of Minnesota, 2017. https://doi.org/10.18128/D010.V7.0. Prices, 1980-2010: CPI Research Series Using Current Methods (CPI-U-RS), U.S. Bureau of Labor Statistics https://www.bls.gov/cpi/research-series/allitems.pdf; Prices 1940-1970, U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: All Items (CPI-U), retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CPIAUCNS.

expertise to carry out the work. As they did in 2000 and 2010, the GAO has designated the 2020 census as "High Risk," raising frequent alarms about the progress of Census contractors.[127] The greatest concern for the 2020 census is the potential for information technology failure.

Census Bureau contractors are developing all-new software to replace thousands of legacy systems developed within the Bureau over several decades. The new software is designed for all phases of data collection and processing for surveys as well as the Decennial Census. This is the largest and most complex software project ever undertaken at the Census Bureau, with total estimated costs of $5 billion, about a third of the estimated total cost of conducting the census.[128] The largest contract, with a life cycle cost estimate of $1.28 billion, went to T-Rex Solutions LLC. T-Rex is a privately-held small business which had just $12.8

---

[127] U.S. Government Accountability Office, 2020 Census: Additional Actions Could Strengthen Future Census Test Designs. GAO 14-26. Washington, 2013. https://www.gao.gov/assets/660/658574.pdf. U.S. Government Accountability Office, 2020 Census. Prioritized Information Technology Research and Testing Is Needed for Census Design Decisions. GAO-14-389. Washington, 2014. https://www.gao.gov/assets/670/662286.pdf. U.S. Government Accountability Office, 2020 Census: Census Bureau Needs to Improve its Life-Cycle Cost Estimating Process. GAO-16-628. Washington, 2016. https://www.gao.gov/assets/680/678117.pdf. U,S, Government Accountability Office, 2020 Census: Additional Actions Could Strengthen Field Data Collection Efforts. GAO-17-191. Washington, 2016. https://www.gao.gov/assets/690/682308.pdf. U.S. Government Accountability Office, Information Technology: Better Management of Interdependencies between Programs Supporting 2020 Census Is Needed. GAO 15-623. https://www.gao.gov/assets/680/678905.pdf. U.S. Government Accountability Office, 2020 Census: Sustained Attention to Innovations, IT Systems, and Cost Estimation Is Needed. GAO-17-584T. Washington, 2017. https://www.gao.gov/assets/690/684462.pdf. U.S. Government Accountability Office, 2020 Census: Actions Needed to Mitigate Key Risks Jeopardizing a Cost-Effective Enumeration. GAO-18-215T. Washington, 2017. https://www.gao.gov/assets/690/688060.pdf. U.S. Government Accountability Office, 2020 Census: Preparations for the 2010 Census Underway, but Continued Oversight and Risk Management Are Critical. GAO-07-1106T. Washington, 2017. https://www.hsgac.senate.gov/imo/media/doc/d071106t.pdf. https://www.hsgac.senate.gov/imo/media/doc/GoldenKuffTestimony.pdf. U.S. Government Accountability Office, 2020 Census: Continued Management Attention Needed to Oversee Innovations, Develop and Secure IT Systems, and Improve Cost Estimation. GAO-18-141T. Washington, 2017. https://oversight.house.gov/wp-content/uploads/2017/10/GAO_Goldenkoff-and-Powner_Testimony_2020-Census.pdf.

[128] U.S. Census Bureau, 2020 Census Life-Cycle Cost Estimate Executive Summary, Version 1.0. December 21, 2017. https://www2.census.gov/programs-surveys/decennial/2020/program-management/planning-docs/2020-cost-estimate1.pdf.

million in revenue the year prior to receiving the Census Bureau contract.[129] To develop the census software, T-Rex is subcontracting with a dozen other vendors, including Lockheed Martin and General Dynamics. The GAO warns that IT development is behind schedule, and many critical components are incomplete. Costs have been escalating rapidly, and available funding has not been sufficient to carry out the full set of end-to-end testing that was initially planned.[130]

The Census Bureau is now in turmoil, and many experienced staff are departing. In 2017, shortly after Census Director Thompson asked Congress for an immediate injection of $310 million in new funding to pay for cost overruns, he suddenly and unexpectedly resigned, leaving the 2020 census in the hands of an acting director who may lack the influence needed to persuade Congress of the need for additional funds.[131] The Bureau is projecting large cost increases, and it is unclear whether Congress will cover them.[132] Adding to the confusion is the last-minute insertion of a divisive question on citizenship, a transparently political decision that came too late to be tested and which has the potential to substantially reduce response rates and further increase costs. The 2020 census may limp across the finish line, but it is likely to be expensive and have high undercount of the immigrant population.

---

[129] James B. Treat, "2020 Census Update." In Census Bureau FOIA response 20180611, 336-432. T-Rex Revenue in FY2014 is from "Inc. 5000 2015: The Full List," https://www.inc.com/inc5000/list/2015; T-Rex is number 1917.

[130] U.S. Government Accountability Office, High-Risk Series: Progress on Many High-Risk Areas, While Substantial Efforts Needed on Others GAO-17-317. https://www.gao.gov/highrisk/2020_decennial_census/.

[131] Doreen McCallister, "Census Bureau Director Resigns as Agency Faces Funding Debate," *The Two-Way: Breaking News from NPR*, May 10, 2017. Haley Sweetland Edwards, "The Head of the Census Resigned. It Could Be as Serious as James Comey," *time.com,* May 12, 2017. Tara Bahrampour, "U.S. Census director resigns amid turmoil over funding of 2020 count," *The Washington Post*, May 9, 2017. Tara Bahrampour, "Census watchers warn of a crisis if funding for 2020 count is not increased," The *Washington Post*, April 18, 2017.

[132] U.S. House of Representatives, "Prepared Statement of John H. Thompson, Director U.S. Census Bureau, Before the Appropriations Committee's Subcommittee on Commerce, Justice, Science and Related Agencies," May 3, 2017, https://docs.house.gov/meetings/AP/AP19/20170503/105896/HHRG-115-AP19-Wstate-ThompsonJ-20170503.pdf.

The Politics of Data Capture

For more than two centuries, changes in census data capture techniques were driven by a combination of politics and practical necessity. Madison's original design for the census efficiently decentralized data capture, making it feasible to gather information on millions of people with virtually no clerical staff in Washington. That system collapsed with scandalous errors in 1840, as the expanding scope of the census made Madison's system impractical. With the reorganized individual-level census of 1850, the work of data capture shifted from enumerators in the field to the Census Office in Washington, which was soon overwhelmed by the task. The bottlenecks of tallying census forms led to a century of extraordinary technical innovation, including punch card tabulation, the first practical application of an electronic computer, and the first large-scale optical mark recognition system.

In the progressive era, the Census Bureau developed in-house engineering capabilities that directly competed with private-sector firms. Census officials felt that the Hollerith Tabulating Machine Company was making extortionate profits from their monopoly control of the tabulating business, and that a small public investment in a machine shop would be highly cost-effective. That calculation proved correct, as the Census Bureau Machine Shop produced and maintained increasingly sophisticated unit record machines for the next half century. Even after computers replaced mechanical tabulating equipment, the Census remained at the forefront of innovation in large-scale data capture by producing new technology for the electronic era, such as FOSDIC and TIGER. Prior to the 2000 Census, data capture operations were never outsourced, and by developing and maintaining equipment in-house the census directly competed with private companies that had been originally established as Census Bureau spin-offs.

The self-reliance of the Census Bureau ended abruptly in the 1990s. New political pressure forced the Census Bureau to turn outward for solutions to its data capture problems rather than relying on the experience and expertise of in-house employees.[133] Privatization was never about saving money. Hiring

---

[133] U.S. Census Bureau, *History: 2000 Census of Population and Housing*, 274-280. James Gleick, "The

defense contractors to undertake clerical work and data processing was an expensive proposition. Even the wildly optimistic cost-benefit analysis used to justify outsourcing data capture did not promise significant cost savings. The real agenda of privatization was ideological: The goal was to prove that a Democratic administration could implement Reagan's agenda of shrinking government. Census officials portrayed data capture in the 2000 and 2010 censuses as success stories, but with respect to risks and costs the privatization project has been a dismal failure.

From the late nineteenth century to the late twentieth century, the U.S. Census was the world leader in the development and application of large-scale data capture technology. The Bureau's research and development efforts had numerous spin-offs with transformational impact on broader society, including punch-card business accounting, the two largest computer companies of the twentieth century, and digital street maps. In a broader perspective, the costs of privatizing census data capture are greater than just the increased expense of operating the census; we have also lost a valuable element of our shared institutional capital. The political tides will likely shift again, but at the Census Bureau and across the federal government it will be hard to rebuild the capabilities that the storms of privatization have swept away.

Census: Why We Can't Count," *The New York Times*, July 15, 1990.

Notes