Volume Title: Fatal Years: Child Mortality in Late Nineteenth-Century America

Volume Author/Editor: Samuel H. Preston and Michael R. Haines

Volume Publisher: Princeton University Press

Volume ISBN: 0-691-04268-3

Volume URL: http://www.nber.org/books/pres91-1

Conference Date: n/a

Publication Date: January 1991

Chapter Title: New Estimates of Child Mortality during the Late Nineteenth Century

Chapter Author: Samuel H. Preston, Michael R. Haines

Chapter URL: http://www.nber.org/chapters/c11542

Chapter pages in book: (p. 49 - 87)

TWO

NEW ESTIMATES OF CHILD MORTALITY DURING THE LATE NINETEENTH CENTURY

THE BASIC PURPOSE of this chapter is to use the public use sample of the 1900 census to construct improved estimates of levels of child mortality in the United States during the last decade of the nineteenth century. The 1900 census asked questions on the number of children that had been born to women and the number of those children who were still living at the time of the census. These data do not provide direct information on such conventional life table measures as the infant mortality rate or the probability of dying before age 5. Instead, these measures must be estimated indirectly from the data, using the extensive procedures that demographers have developed for this purpose and that are described in this chapter.

Our results show a close agreement between the indirect estimates of child mortality for the Death Registration Area (DRA) and the direct estimates that are available from vital registration of deaths in this Area. We also show, however, that child mortality was higher in the DRA than in the nation as a whole for whites and, especially, for blacks. Ironically, the bias in DRA measures is largely offset when both racial groups are combined because blacks represented a very small fraction of population in the DRA, so that their exceptionally high mortality was underweighted.

Previous Estimates of Mortality in the Nineteenth Century

Little is known about trends, levels, and differentials in American mortality in the nineteenth century. It is not altogether clear when or even whether mortality declined in the United States during the period. The official Death Registration Area was not formed until 1900, and even then it only covered ten states and the District of Columbia.[1] As may be seen in Table 2.1, the DRA contained only 26.3 percent of the American population in 1900 and was significantly more urban than the nation as a whole. The percentage of the DRA population that lived in urban areas with at least 2,500 inhabitants was

TABLE 2.1
Comparison of Selected Characteristics of the Original Death Registration States with the United States as a Whole, 1900

	U.S. (%)	*DRA (%)*
Total population	100.0	26.3
Percentage black	11.6	1.9
Percentage urban[a]	39.7	62.9
Percentage foreign-born	13.6	22.4
Percentage of U.S. blacks	100.0	4.4
Percentage of blacks who are urban	20.5	82.0

Source: U.S. Bureau of the Census 1902a, 1975.

Note: The original Death Registration states of 1900 consisted of Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut, New York, New Jersey, Indiana, Michigan, and the District of Columbia. See note 1 in chapter 2 for more information about the DRA.

[a] Population living in incorporated areas with populations of 2,500 and over.

62.9, compared with 39.7 percent for the nation as a whole. The DRA also had a larger proportion of foreign-born residents (22.4 percent versus 13.6 percent for the entire country) and contained only 4.4 percent of the American black population. Of the black population residing in the Death Registration Area, 82 percent were urban, although in the U.S. as a whole only 20.5 percent of blacks lived in urban areas. These differences would not be important if mortality differences along these dimensions had not been pronounced; but, as we will show, there were large differences in mortality by residence, race, and nativity.[2]

Before 1900, official mortality data were limited to selected cities and states and to the imperfect mortality statistics from the decennial federal censuses from 1850 to 1890 that asked questions on household deaths in the preceding year. In 1842, Massachusetts was the first state to institute vital registration, and it is widely cited as a source of information on nineteenth-century American mortality and fertility (Gutman 1956; Vinovskis 1972, 1981). By 1860, the Massachusetts death registration data were quite good, but evidence for years before that date must be sought from other sources such as genealogies, family reconstitutions, and bills of mortality (Vinovskis 1981: app. B). The population of Massachusetts was also more urban and industrial and had a higher percentage born abroad than the population of the country as a whole. Some analysts (e.g., Coale and Zelnick 1963) have been forced to assume that Massachusetts's mortality was representative of that of the United States as a whole, but its representativeness has been seriously questioned (Vinovskis 1978).

The federal census data on mortality have the virtue that they covered the whole nation. But they only provide information on events in the year prior to the census, and they were clearly seriously incomplete in the volume of deaths recorded (see Condran and Crimmins 1979, 1980; Condran and Crimmins-Gardner 1976, 1978; Crimmins and Condran 1983; Haines 1979a; Higgs and Booth 1979; Suliman 1983).

The absence of reliable national-level data has prompted the use of roundabout methods and symptomatic data to estimate mortality trends. Operating from assumptions about the inverse relationships between mortality and income per capita and between mortality and public health adequacy, and assuming a positive relationship between mortality and urbanization, Easterlin (1977:132–40) suggested that the rising effect of income per capita probably outweighed the negative effect of urbanization, with public health playing little or no role before about 1880. He thus posited an increase in expectation of life at birth starting around 1840. This finding contrasts with that of Vinovskis (1981: ch. 2), who suggests that little change occurred in the mortality level in Massachusetts between 1790 and 1860. More recent work by Fogel (1986), using a large genealogical data base, suggests that expectation of life at birth actually declined in the half-century prior to the Civil War, despite evidence of substantial economic growth from 1840 to 1860. One possible explanation for such a decline is that increases in income per capita were accompanied by a poorer income distribution (Williamson and Lindert 1980: ch. 4; Pessen 1973), although extensive data on workers hired by the Army is inconsistent with such a deterioration (Margo and Villafor 1987). Another explanation is that urbanization more than offset the gains from higher income. Such a process can be better documented in England in the first half of the nineteenth century, where urbanization was far more widespread (Woods 1985).

More evidence exists for the postbellum era. Higgs (1973) argues that mortality began its decline in rural America in the 1870s and that the decline took place largely as a result of improvements in diet, nutrition, housing, and general levels of living and without much assistance from public health. Meeker (1972, 1974) contends that mortality improved little if at all before about 1880, and that only after about 1880 was the fall in urban death rates substantially aided by new public-health measures, especially installation of sanitary sewers and pure central water supplies. Both analysts use a variant of intercensal survival analysis—tracking the survivorship of a birth cohort from one census to the next—which produces virtually no information on mortality in early childhood. Meeker's result is supported by work with extant nineteenth-century American life tables and model life table systems, which shows little evidence of sustained mortality reduction before about 1880 (Haines 1979a). Table 2.2 compiles previous estimates of nineteenth-century mortality in the United States. The data are confined to available life table information. On the whole, the results indicate little or no decline before the 1870s, higher mortality in urban areas, and much higher mortality among blacks.

By the 1890s, it is likely that mortality was declining in both rural and urban areas, although the absence of high-quality data of national scope leaves the matter open (Condran and Crimmins 1980). Urban death rates began at a higher level but apparently fell more rapidly, probably pushed by improvements in standards of living as well as advances in public health. Mortality improvements have been linked to specific public-health initiatives in the late nineteenth and early twentieth centuries in New York (Duffy 1974), Baltimore (Howard 1924), Philadelphia (Condran and Cheney 1982), Boston (Meckel 1985), Chicago (Cain 1972, 1974, 1977), and New Orleans (Lentzner 1987). Indeed, detailed studies of individual cities furnish perhaps the best opportunity to study this complex process.

As noted in Chapter 1, several European countries in the late nineteenth century also provide several examples of more rapid mortality decline in urban than in rural areas. Kingsley Davis (1973), focussing especially on Stockholm, demonstrates that urban areas in Europe often had mortality declines that were even more rapid than the much-heralded declines in less-developed countries after World War II. The accumulating evidence calls into question, at least for urban residents, Thomas McKeown's influential studies discrediting the importance of public-health measures as a factor in the nineteenth-century mortality decline (see especially McKeown and Record 1962). Accordingly, his explanatory emphasis on rising standards of nutrition as a factor in the nineteenth-century European mortality decline also appears overdrawn (see also Szreter 1988).

Whatever the progress of the mortality decline in nineteenth-century America, accurate data on mortality levels become available for part of the country with the formation of the Death Registration Area in 1900. Table 2.3 presents data from the DRA for 1900–1902, together with data from other countries during the period 1889–1910. The life table values given are $q(1)$, the probability of dying between birth and exact age 1 (also referred to here and elsewhere as the infant mortality rate); $q(5)$, the probability of dying between birth and exact age 5; and e_o, the expectation of life at birth. The values are calculated for both sexes combined.

In 1900–1902, more than 12 percent of infants in the DRA died before reaching age 1, and more than 18 percent died before their fifth

TABLE 2.2

Estimates of Child Mortality and Expectation of Life at Birth in the United States, 1830–1910

Source	Region	Period	Sex	Child mortality[a]				
				q(1)	*q(2)*	*q(5)*	e_0	e_{10}
Jacobson (1957)	Massachusetts-Maryland	1850	M	.16064	.21394	.27245	40.4	47.8
			F	.13079	.18262	.24122	43.0	48.6
Meech (1898)	United States, whites	1830–60	M	.16195	.21569	.27468	41.0	48.4
			F	.13430	.18752	.24769	42.9	48.8
Elliott (1857)	Massachusetts	1855	Total				39.8	
Vinovskis (1972)	Massachusetts	1859–61	M			.22646	46.4	51.6
			F			.19193	47.3	50.1
Haines (1977)	Seven New York counties	1850–65	M	.14655	.18067	.21268	45.9	49.2
			F	.12389	.15821	.19105	48.9	51.4
			Total	.13549	.16972	.20213	47.4	50.3
Haines (1979a)	United States	1850	M	.24092		.32195	36.5	45.0
			F	.21712		.29845	38.5	46.1
		1860	M	.20210		.27361	40.7	47.1
			F	.19153		.26684	41.2	47.3
		1870	M	.19210		.26007	42.1	47.9
			F	.17724		.24531	43.7	49.0
		1880	M	.22015		.29538	38.7	46.3
			F	.22980		.31019	38.2	46.5
		1890	M	.16334		.22875	43.9	47.9
			F	.15765		.22546	44.5	48.5
		1900	M	.13356		.21252	46.3	48.3
			F	.12476		.18611	47.4	49.2

TABLE 2.2 *(cont.)*

				Child mortality[a]				
Source	*Region*	*Period*	*Sex*	*q(1)*	*q(2)*	*q(5)*	e_0	e_{10}
Haines (1979a)	United States, white	1850	M	.22829		.30697	37.7	45.5
			F	.20596		.28486	39.6	46.6
		1870	M	.18513		.25056	43.1	48.5
			F	.16633		.23114	45.1	49.7
		1880	M	.21436		.28794	39.6	46.7
			F	.21526		.29268	39.6	47.1
		1890	M	.15675		.21914	45.1	48.7
			F	.14490		.20829	46.3	49.4
		1900	M	.12784		.18497	47.6	49.2
			F	.11206		.16781	49.6	50.4
Fogel (1986)	United States	1850–60	M					46.7
Billings[b]	Massachusetts	1878–82	M	.18080	.23250	.28342	41.7	49.9
			F	.15257	.20245	.25408	43.5	50.0
Billings[b]	New Jersey	1879–80	M	.15153	.19398	.24132	45.6	51.6
			F	.13121	.16939	.21217	48.0	52.5
Glover (1921)	Massachusetts	1890	M	.16777	.20851	.25322	42.5	48.4
			F	.14755	.18738	.23415	44.5	49.6
Abbott (1899)	Massachusetts	1893–97	M	.17233	.20726	.24234	44.1	49.3
			F	.14699	.18115	.21593	46.6	50.7
Glover (1921)	DRA, total	1900–1902	M	.13574	.16614	.19452	47.9	50.4
			F	.11267	.14092	.16881	50.7	51.9
			Total	.12448	.15283	.18196	49.2	51.1

Glover (1921)	DRA, whites	1900–1902	M	.13345	.16331	.19136	48.2	50.6
			F	.11061	.13832	.16574	51.1	52.2
Glover (1921)	DRA, blacks	1900–1902	M	.25326	.31098	.35615	32.5	41.9
			F	.21475	.26990	.31944	35.0	43.0
Glover (1921)	DRA, urban	1900–1902	M	.15097	.18683	.22128	44.0	47.5
			F	.12545	.15883	.19195	47.9	50.3
Glover (1921)	DRA, rural	1900–1902	M	.10900	.13065	.15043	54.0	54.4
			F	.08979	.10967	.12983	55.4	54.4
Glover (1921)	DRA, total	1909–11	M	.12495	.15016	.17282	49.9	51.1
			F	.10377	.12743	.14883	53.2	53.3
			Total	.11462	.13908	.16113	51.5	52.2
Glover (1921)	DRA, whites	1909–11	M	.12326	.14799	.17028	50.2	51.3
			F	.10226	.12545	.14651	53.6	53.6
Glover (1921)	DRA, blacks	1909–11	M	.21935	.27155	.31411	34.0	40.6
			F	.18507	.23303	.27232	37.7	42.8
Glover (1921)	DRA, urban	1909–11	M	.13380	.16247	.18815	47.3	49.1
			F	.11123	.13831	.16266	51.4	52.2
Glover (1921)	DRA, rural	1909–11	M	.10326	.12105	.13777	55.1	54.5
			F	.08497	.10119	.11679	57.4	55.5
Haines (1979a)	Suffolk Co., Mass. (Boston)	1859–61	M	.18057		.31389	36.0	44.1
			F	.15970		.29893	43.5	53.5
Haines (1979a)	Suffolk Co., Mass. (Boston)	1874–76	M	.20053		.34800	34.7	45.2
			F	.18403		.32401	37.1	47.1

TABLE 2.2. *(cont.)*

				Child mortality[a]				
Source	*Region*	*Period*	*Sex*	*q(1)*	*q(2)*	*q(5)*	e_0	e_{10}
Billings[b]	Boston, whites	1879–80	M	.21739	.28518	.34218	37.0	47.5
			F	.18873	.25365	.30823	39.1	48.4
Haines (1979a)	Suffolk Co., Mass. (Boston)	1884–86	M	.20171		.32815	36.0	44.9
			F	.17732		.30309	37.9	46.0
Haines (1979a)	Suffolk Co., Mass. (Boston)	1894–96	M	.17884		.29599	37.1	44.1
			F	.15032		.26518	41.0	47.3
Glover (1921)	Boston	1900–1902	M	.15736	.19875	.24002	41.6	46.0
			F	.13548	.16983	.21017	45.1	48.5
Glover (1921)	Boston	1909–11	M	.13527	.16333	.19050	46.0	47.7
			F	.11330	.13851	.16181	50.3	50.9
Condran & Cheney (1982)	Philadelphia	1870	Total	.17400		.29121	39.6	
Condran & Cheney (1982)	Philadelphia	1880	Total	.15970		.26104	42.3	
Condran & Cheney (1982)	Philadelphia	1890	Total	.15290		.24701	43.7	
Glover (1921)	Philadelphia	1900–1902	M	.15027	.18978	.23006	42.5	46.3
			F	.12741	.16369	.20232	46.2	49.1
Glover (1921)	Philadelphia	1909–11	M	.14174	.17456	.20558	45.5	48.1
			F	.11926	.14959	.17796	49.6	51.2

Billings[b]	New York City	1878–81	M	.26278	.35464	.42751	29.0	42.4
			F	.22411	.31513	.38744	32.8	45.3
Billings[b]	New York City, whites	1879–80	M	.23421	.32245	.38085	33.3	44.9
			F	.20427	.28527	.34167	36.8	46.9
Billings[b]	Brooklyn, whites	1879–80	M	.19477	.27036	.33101	37.5	48.1
			F	.16424	.24336	.30545	39.7	49.1
Glover (1921)	New York City	1900–1902	M	.15673	.20308	.24435	40.6	44.9
			F	.13298	.17564	.21542	44.9	48.2
Glover (1921)	New York City	1909–11	M	.13186	.16799	.19907	45.3	47.4
			F	.11405	.14762	.17708	49.5	50.9
Billings[b]	Chicago, whites	1879–80	M	.20526	.27950	.34394	38.1	50.6
			F	.15107	.22919	.29958	41.3	51.6
Glover (1921)	Chicago	1900–1902	M	.12010	.15142	.18191	46.3	47.7
			F	.09762	.12764	.15676	50.8	55.0
Glover (1921)	Chicago	1909–11	M	.13066	.16079	.18980	45.9	51.5
			F	.10431	.13196	.15959	51.7	52.4

Source: Jacobson 1957; Meech 1898; Abbott 1899; Glover 1921; Haines 1977, 1979a; Vinovskis 1972; Fogel 1986: Table 3; U.S. Bureau of the Census 1886 (Billings); Condran and Cheney 1982: Table 1; various Massachusetts vital statistics and census data (Haines 1979a).

[a] $Q(1)$is the probability of dying before reaching age 1. It is the infant mortality rate. $Q(2)$ and $q(5)$ are the probabilities of dying before reaching ages 2 and 5, respectively. E_0 and e_{10} are the expectations of life at birth and at age 10.

[b] From U.S. Bureau of the Census 1886.

TABLE 2.3
Comparison of Published Life Table Values: U.S., 1900–1902 and 1909–11, and Selected Foreign Nations, 1889–1911

Location	*Period*	$q(1)$	$q(5)$	e_0
United States (DRA)	1900–1902	.124	.182	49.2
Whites	1900–1902	.122	.179	49.6
Blacks	1900–1902	.234	.338	33.8
Urban (whites)	1900–1902	.138	.207	45.9
Rural (whites)	1900–1902	.100	.140	54.7
United States (DRA)	1901–10	.117	.166	50.9
	1909–11	.115	.161	51.5
Australia	1890–1900	.110	.151	52.9
	1901–10	.088	.116	57.0
Austria	1900–1901	.230	.321	38.8
Belgium	1891–1900	.156	.224	47.0
	1900	.195	.230	46.8
Bulgaria	1899–1902	.155	.289	40.1
Czechoslovakia	1899–1902	.229	.307	40.3
Denmark	1895–1900	.134	.177	51.7
England & Wales	1891–1900	.156	.234	45.9
	1901–10	.131	.192	50.4
France	1890–92	.172	.250	43.3
	1895–97	.159	.220	46.1
	1898–1903	.150	.209	47.4
Germany	1891–1900	.217	.292	42.2
	1901–10	.187	.243	46.5
Prussia	1891–1900	.203	.282	42.8
Ireland	1890–92	.098	.157	48.8
	1900–1902	.103	.159	50.2
Italy	1891	.190	.339	38.6
	1900–1902	.167	.283	43.0
	1901–10	.160	.266	44.5
Japan	1899	.162	.246	43.4
Netherlands	1890–99	.158	.226	47.6
	1901	.116	.224	49.0
	1900–1909	.129	.186	52.2
New Zealand (whites)	1891–95	.088	.118	58.1
	1896–1900	.080	.104	59.9

TABLE 2.3 (*cont.*)

Location	*Period*	$q(1)$	$q(5)$	e_0
Norway	1891–1900	.096	.149	52.2
	1901–10	.074	.109	56.2
Russia (European)	1896–97	.277	.422	32.4
Scotland	1891–1900	.131	.212	46.0
Sweden	1891–1900	.102	.160	52.2
	1898–1902	.107	.154	52.7
	1901–10	.084	.126	55.7
Switzerland	1889–1900	.151	.199	47.1
	1901–10	.126	.163	50.7

Source: United States: Glover 1921. All other life tables are from the published official life tables used by Coale and Demeny 1966, except those for Belgium (1900), the Netherlands (1901), and Sweden (1898–1902), which are taken from Keyfitz and Flieger 1968; Italy (1891) and Japan (1899), which were taken from Preston, Keyfitz, and Schoen 1972; and Ireland (1890–92 and 1900–1902), which were constructed from data given in Mitchell and Deane 1971 using the Reed-Merrell method (U.S. Bureau of the Census 1971: ch. 15). Male and female life tables were combined assuming a sex ratio at birth of 105 males per 100 females.

Note: $q(1)$ is the probability of dying between birth and exact age 1. It is the infant mortality rate; $q(5)$ is the probability of dying between birth and exact age 5; e_0 is the expectation of life at birth.

birthday. Mortality in the Death Registration Area was considerably better than that achieved in central, eastern, and southern Europe (i.e., Germany, Prussia, Austria, Czechoslovakia, European Russia, Bulgaria, and Italy); but Norway, Sweden, Australia, and New Zealand had superior survivorship. DRA mortality was not greatly ahead of that of Japan, the one non-Western nation represented in Table 2.3. For the largely urban black population of the DRA, mortality was so severe that it approached levels in European Russia in 1896–97, which are the highest mortality rates presented here. The series of life tables from 1900–1902 to 1901–10 to 1909–11 for the Death Registration Area indicates that mortality fell after 1900.

Child Mortality Estimates Based upon the Census Sample

We now turn attention to estimating levels of child mortality for the nation as a whole based upon the enumerators' manuscripts from the 1900 United States Census. Estimates are made separately for the white and black populations; Chapter 3 will describe levels of childhood mortality according to more detailed characteristics.

The Sample

The original schedules of the 1900 census asked questions on the number of children who had been born to women who had ever been married and the number of those children who were still living. Instructions to enumerators indicated that stillbirths were to be excluded (U.S. Bureau of the Census 1979:34). For reasons that are not clear, the returns from these questions were never published or analyzed.[3] Several years ago, a 1-in-750 stratified random sample of households was produced from these manuscripts at the University of Washington (Graham 1980). The data consist of a self-weighted sample of 27,069 households containing 100,438 individuals from all the states and territories of the United States, including Alaska and Hawaii.

A comparison of selected characteristics of the sample with published census data reveals that differences in age, sex, race, residence, and nativity distributions were small and insignificant. Table 2.4 provides a number of these comparisons, including calculations of the singulate mean age at first marriage for females in various race and nativity groups. The latter requires distributions of the population by age, sex, and marital status (Hajnal 1953). As can be seen, the differences from published results are negligible.[4] This sample has been used by a number of other scholars, who have, in some cases, confirmed its representativeness (see Haines and Anderson 1988). For the analysis presented in this chapter, a subfile was created containing a sample of all 32,866 adult women who completed questionnaire information on both children ever born and children surviving and whose responses were legible. Other restrictions on the data analyzed in the chapter are presented below.[5]

Estimation Procedures

The indirect estimation procedures used in this chapter begin with the recognition that the proportion dead among children ever born to a group of women is the joint outcome of a set of age-specific death rates and the distribution of exposure times to the risk of death that were experienced by offspring of those women. For example, if the probability of dying before age 5 is .30 and if all of the women's births occurred exactly 5 years earlier, then the proportion dead among their children should be .30. If all of their births had occurred

TABLE 2.4
Comparison of Selected Population Characteristics in the National Sample of the 1900 U.S. Census and Published Census Results

	Sample	*Published census*
Total persons (*N*)	100,468	75,994,575
Percentage female	48.91	48.92
Percentage black	11.34	11.62
Percentage foreign-born	14.23	13.61
Percentage urban	40.52	39.69
Percentage female at ages 20–24		
Total	46.66	46.66
White	45.03	45.26
Native white	44.36	45.15
Foreign white	47.03	45.91
Black	56.61	54.68
Singulate mean age at first marriage, females		
Total	23.54	23.66
White	23.69	23.86
Native white	23.76	23.88
Foreign white	23.50	23.58
Black	22.81	22.49

Source: Sample of census enumerators' manuscripts, U.S., 1900. U.S. Bureau of the Census 1902a, 1975. Singulate mean age at marriage calculated according to Hajnal 1953.

exactly 2 years earlier, however, then the proportion dead among their children would be less than .30, since some child deaths occur between ages 2 and 5. The aim of indirect estimation techniques is to provide an adjustment for children's exposure to the risk of death that allows the underlying probabilities of death to emerge. In particular, the procedures are based on the following identity (Sullivan 1972; Brass 1975; Trussell 1975; United Nations 1983a: ch. 3):

$$D/B = \int_0^{\alpha} c(a)q(a)da, \tag{2.1}$$

where B is the cumulative number of children born to reporting women; D is the cumulative number of deaths among those children; $c(a)da$ is the proportion of children born to reporting women who were born within period a to $a + da$ years before the census; $q(a)$ is the probability of death before age a for a child born to reporting women a years before the census; and α is the number of years since the birth of the first child born to reporting women.

By the mean value theorem, there must be some age A between 0 and α such that

$$D/B = q(A) \int_0^\alpha c(a)da = q(A);$$

that is, the proportion dead among children ever born to the women must equal the probability of death prior to some age A in the life table pertaining to those children. The briefer the period of the child's exposure to the risk of death, the lower will be A. Short exposure periods can be constructed, for example, by limiting data to women aged 15–19 or to women who have been married less than 5 years. Numerous simulations of mortality and fertility histories (Sullivan 1972; Trussell 1975) have established that $q(1)$ (the probability of dying before exact age 1) is best identified by proportions dead among children born to women aged 15–19, $q(2)$ is best identified by reports of women aged 20–24 or in marital duration category 0–4 years, $q(3)$ by women aged 25–29 or married 5–9 years, and so on. The complete set of these correspondences is presented in Table 2.5 below.

The correspondences are not exact, of course, and conventional estimation procedures provide adjustment factors tailored to a particular application. These adjustment factors are designed to correct the estimates according to the shape of the age-specific fertility function prevailing in the population under study, a shape that determines the time distribution of children's exposure to the risk of mortality. This shape is indexed by the ratio of cumulative average numbers of children ever born in successive age or marital-duration intervals. Clearly, the ratio involves comparisons of cumulative childbearing across cohorts; to apply the methods, it is necessary to assume that the ratios also pertained in the course of childbearing to an actual cohort, which amounts to assuming that fertility has been constant.

An alternative approach to the indirect estimation of child mortality is the surviving-children method (Preston and Palloni 1978). This method involves the backward projection of the age distribution of surviving "own-children" by various levels of mortality within a model life table system to the point where the back-projected number of births equals the number of children reported as ever born by the group of women. A model life table is simply an empirical representation of a "typical" life table for populations at a particular level of mortality. A model life table *system* consists of a set of model life tables that vary systematically in their level of mortality, typically indexed by life expectancy at birth. Various systems of model life tables have been constructed that vary in their input data and in their methods of estimation. Most frequently used, by virtue of their broad data base and careful construction, are the four regional systems of Coale and Demeny (1966). Coale and Demeny observed four different types of relationships among age-specific death rates that prevailed historically in (mainly) European populations and assigned labels to these relationships that correspond roughly to the region of Europe supplying input data for a particular system.

The surviving-children procedure is based on a rearrangement of equation (2.1):

$$B/(B - D) = \int_0^{\alpha} [C_s(a)/(1 - q(a)]da, \qquad (2.2)$$

where $C_s(a)da$ is the proportion of surviving children who were aged a to $a+da$ at the time of the census. Women can be grouped into broad age, marital-duration, or other categories to implement this approach. The census sample provides direct reports on B and D, and $C_s(a)$ can be estimated directly from the age distribution of surviving own-children enumerated with the mother. An "own-child" is not simply any child in the household but one who is identified as, or is surmised to be, the natural offspring of the mother. The matching of mothers and children is done through an examination of information on relationship to head of household, age, surname, place of birth, and order of enumeration in the original census manuscripts for both mother and child. The availability of the age distribution of these own-children is one of the advantages of a sample of original census returns. Given B, D, and $C_s(a)$, the analyst then locates the set of $q(a)$'s within a model life table system that will satisfy equation (2.2). In order that the own-children estimates of $C_s(a)$ not be biased by children having left the home, it is necessary to confine the analysis to younger women. The Coale and Demeny (1966) "West" model life table system is used here to provide values of $q(a)$; and the solution is derived by an iterative procedure built into a model life table generation program (Avery 1981).

The surviving-children method has some advantages over the more conventional estimation procedures based on equation (2.1).

Most important, it is insensitive to recent fertility declines or to irregular patterns of fertility behavior in the past. The history of fertility is explicitly represented in the age distribution of surviving children, whereas fertility must be assumed constant in the conventional approach. Second, the method is flexible with respect to the age or marital-duration groups of women that can be included in the analysis, a feature of particular advantage in dealing with some of the small-sample problems that are encountered here. The procedure is more sensitive than the others, however, to age-selective omissions and misreporting of children's ages. Fortunately, the 1900 U.S. Census appears to have had exceptionally accurate age reporting, probably attributable to the unusual inclusion of questions on both age at last birthday and year of birth (Coale and Zelnik 1963).

Each of the estimation procedures used a set of model life tables. Under the conventional age and marital-duration procedures, these model life tables are embodied in the multipliers that take account of the shape of the fertility history in a particular application. Different sets of multipliers exist for different model life table systems (United Nations 1983a: Tables 47 and 56). In the surviving-children technique, the model life table is imposed directly by the analyst. In neither case, however, are results sensitive to the model life table system chosen. Alternative model life table systems applied to the same set of data will produce identical values of $q(a)$ at some age A^*. The age of child at which this identity pertains for a particular age or marital-duration group of reporting women is usually close to the age shown in Table 2.5. That is, it is around age 1 for women aged 15–19, around age 2 for women aged 20–24 or married 0–4 years, around age 3 for women aged 25–29 or married 5–9 years, etc. The reason that this identity applies is that any pair of solutions to equation (2.2) that are drawn from different model mortality systems must intersect somewhere in the range of ages 0 to α (Preston and Palloni 1978). If they did not intersect—that is, if one $q(a)$ function lay above the other at all ages—then they could not both be solutions. The result is that two $q(a)$ solutions drawn from different model life table systems for the ages shown in Table 2.5 are usually within 1 to 4 percent of one another. For the same reason, there is also an intersection between two solutions, one that is drawn from a model life table system and the other that has an arbitrary time trend in $q(a)$ built into the system. Results of simulations of various types of mortality decline enable the assignment of a "date" to each estimate. The date is the approximate point at which plausible time trends intersect. In this chapter we use the dating equations developed in the United Nations's *Manual X* (United Nations 1983a: ch. 3).[6]

Implementing the Estimation Equations

A number of filters were applied to the census data, particularly for the marital-duration model, to increase the accuracy of estimation.

First, for the surviving-children approach, analysis was confined to women aged 14–34 because of the potential bias resulting from migration of children away from home. In implementing the surviving-children method, we also excluded women whose oldest "own-child" was implied to have been born before the women reached age 14.

Second, when a woman's age was used as the index of her children's exposure to the risk of mortality (the "age model"), all women in the relevant age groups were used in the estimations, with the exception of those for whom an illegible or missing response was given either for children ever born or for children surviving. (As noted earlier, these women were also excluded from the other estimation approaches). Mean parity estimates by age, required for adjustment factors, are based on all women with a legible response on children ever born.

Third, when a woman's marital duration was used as the index of her children's exposure to the risk of mortality (the "marital-duration model"), we attempted to exclude women not in their first marriage, for whom the duration in their *current* marriage—the only information available in the census—would be a very imperfect indicator of their children's exposure to mortality. In particular, we selected only women currently married with husband present who reported no surviving children other than own-children present in their household; whose implied age at marriage (current age minus duration of marriage) was between 10 and 34 years; whose oldest own-child's age was not more than two years greater than duration of current marriage; and whose reported number of children ever born was not more than two greater than duration of current marriage in years.

Despite the efforts to exclude from the marital-duration model women who had borne children prior to their current marriage, it is likely that our procedures have not been completely successful. One of the main clues about remarriage in the census manuscripts (which listed only current marital status and duration of current marriage) is the age of the oldest own-child. But under high-mortality conditions, many of the early births would not have survived to the 1900 census and thus would have left no evidence of the earlier marriage. These same high-mortality conditions would also tend to produce a higher proportion of remarried women in the population because of marriages disrupted by the death of the husband. Thus, remarried women with high child mortality would tend to be located at earlier marital durations and would bias upward estimates of child mortality at early ages. This bias occurs because duration of marriage is being used as a proxy for the time that the children are exposed to the risk of death.

These problems are more acute among blacks, for whom both marital disruption and mortality were high at the turn of the century. The percentage of women who were widowed or divorced was approximately twice as high in the black population as it was in the white population. For example, at ages 35–44, the percentage of black women reported as "widowed or divorced" in the Census of 1900 was 19.6, compared with 8.1 for whites (U.S. Bureau of the Census 1902a: Table 29). This result is, of course, not conclusive, since it is remarriage that is of direct interest. But the census results by age, race, and marital status are suggestive of the higher rates of marital dissolution among the black population of the United States at the turn of the century. Further, evidence from the Death Registration Area for 1900–1902 (Table 2.2) and from other sources (e.g., Condran 1984) indicates that adult male mortality among blacks was substantially above that of black females and well above the average for white males.

Differences among results obtained from using the four different regional sets of Coale and Demeny tables (i.e., North, South, East, West) are usually very small, as expected. In choosing among them, it was noted that Model West fitted well to the 1900–1902 Death Registration Area life table for the total and the white populations (Coale and Zelnik 1963). It is less clear whether any of the Coale and Demeny models fits the age patterns of black mortality well (Zelnik 1969; Condran 1984). Model West, an "average" pattern, was chosen for the black population as a compromise.

For the surviving-children approach, equation (2.2) was solved to provide estimates for all women aged 14–34 and for the subgroups of all women aged 14–24 and 25–34. It should be recognized that what constitutes a mortality level in the surviving-children approach is simply a complete model life table. Although considerable detail by age of child is presented for this method, the estimates for any particular solution are not independent of one another but are constrained to correspond to the same model life table. Depending on the model life table family chosen, different $q(a)$ sequences may result. All of the model life table systems, when applied to women aged 14–34, however, yield very similar results at age 5 because of the tendency for solutions produced by different model life table systems—or by a model life table arbitrarily deformed by different time trends in mortality—to intersect at some age of child. For the total population, the range of $q(5)$'s indicated by the various Coale and Demeny model life table solutions is only .004, whereas it is .024 for $q(1)$ and .030 for $q(20)$ (Preston and Haines 1984). Using a formula presented in Preston and Palloni (1978:84), we estimate that the year to which this robust surviving-children estimate of $q(5)$ pertains is 1896.

Results

The results of the different estimation procedures are given in Table 2.5. The table includes the $q(a)$'s (i.e., the probabilities of death between birth and exact age a), the N's (number of children ever born) for each group, and the level of West Model life table implied by each estimate. In addition, for the age and marital-duration models, Table 2.5 presents the estimated dates to which each of the various $q(a)$ estimates pertains, expressed in terms of years prior to the census of June 1, 1900. The time reference becomes earlier for older women, whose children were, on average, exposed to mortality in more distant periods.

Perhaps the best way to begin summarizing the mass of information in Table 2.5 is by means of a graph. Figure 2.1 presents age-specific estimates of $q(a)$ for the total population, using the three main approaches and, in each instance, using Model West estimation equations. Agreement among the three approaches is close for ages 3, 5, and 10. Beyond ages 5 and 10, the surviving-children estimate is basically an extrapolation using the same model life table identified as pertaining to younger children; because estimation stops with women aged 34, the surviving-children approach contains little or no information on mortality among older children. Nevertheless, the inclusion of the complete surviving-children $q(a)$ function in Figure 2.1 is illuminating because it suggests that the child mortality experience among older women—who are represented in the other two estimation approaches—diverges systematically from Model West level 13.6, which is the surviving-children method estimate for the total U.S. population. If we make the reasonable assumption that the West model life table system pertained in the period 1880–1900 roughly as accurately as it did in 1900, then children of older women were clearly subject to higher mortality conditions than were children of younger women.

These estimates thus suggest that a substantial reduction in child mortality occurred prior to the census of 1900, an implication consistent with some of the research cited earlier in this chapter. The pace

TABLE 2.5
Estimates of Child Mortality by Race: U.S., 1900

	q(1)	*q(2)*	*q(3)*	*q(5)*	*q(10)*	*q(15)*	*q(20)*	*q(25)*	*Level*	*Implied life expectancy at birth*
						Age model				
Age group of women[a]	15–19	20–24	25–29	30–34	35–39	40–44	45–49	—		
q(a)										
Total	.15332	.17664	.16438	.17736	.20662	.21983	.26076	—		
White	.16168	.15176	.15109	.16705	.19512	.20920	.24755	—		
Black	.13090	.26216	.21502	.25164	.27776	.29367	.34327	—		
N's[b]										
Total	382	3,378	6,886	9,123	11,212	10,861	9,760	—		
White	288	2,620	5,740	7,995	9,681	9,534	8,421	—		
Black	93	732	1,079	1,099	1,488	1,275	1,315	—		
Implied level[c]										
Total	11.38	12.19	13.52	13.60	13.06	13.03	12.19	—		
White	10.85	13.50	14.19	14.06	13.53	13.44	12.66	—		
Black	12.88	8.11	11.13	10.45	10.26	10.26	9.28	—		
Years prior to census to which estimates apply[d]										
Total	0.9	2.1	3.9	6.0	8.5	11.2	14.2	—		
White	0.9	2.0	3.7	5.8	8.1	10.8	13.8	—		
Black	0.7	2.1	4.4	7.2	10.2	13.3	16.3	—		

	Marital-duration model							
Duration group of women[a]	—	0–4	5–9	10–14	15–19	20–24	25–29	30–34
q(*a*)								
Total	—	.14722	.15514	.18234	.19496	.21885	.25267	.27768
White	—	.12926	.13949	.17267	.19234	.21101	.24398	.26915
Black	—	.28021	.26441	.25096	.22168	.27879	.32477	.35960
N′s[b]								
Total	—	2,592	6,716	9,088	9,034	8,746	7,222	6,326
White	—	2,261	5,868	8,120	8,224	7,796	6,462	5,744
Black	—	322	811	916	791	924	733	564
Implied level[c]								
Total	—	13.74	13.97	13.38	13.35	13.07	12.48	12.43
White	—	14.79	14.82	13.81	13.64	13.37	12.79	12.72
Black	—	7.36	8.96	10.48	12.47	10.80	9.90	9.68
Years prior to census to which estimates apply[d]								
Total	—	1.3	3.4	5.8	8.2	11.0	14.1	17.1
White	—	1.4	3.5	5.8	8.2	10.8	14.0	17.0
Black	—	1.3	3.2	5.8	8.8	11.8	14.8	17.5

TABLE 2.5 *(cont.)*

	$q(1)$	$q(2)$	$q(3)$	$q(5)$	$q(10)$	$q(15)$	$q(20)$	$q(25)$	*Level*	*Implied life expectancy at birth*
					Surviving-children method[c]					
Women ages 14–34										
Total	.12025	.14906	.16183	.17636	.19218	.20381	.22040	.24234	13.65	50.08
White	.11076	.13658	.14802	.16104	.17561	.18638	.20187	.22255	14.36	51.83
Black	.17034	.21380	.23304	.25496	.27640	.29209	.31304	.34026	10.32	41.83
Women ages 14–24										
Total	.13255	.16566	.18033	.19703	.21441	.22718	.24482	.26802	12.75	47.92
White	.11775	.14576	.15818	.17231	.18780	.19921	.21551	.23718	13.82	50.54
Black	.18525	.23237	.25325	.27701	.29983	.31647	.33853	.36706	9.44	39.68
Women ages 25–34										
Total	.11806	.14617	.15863	.17281	.18834	.19978	.21611	.23782	13.80	50.48
White	.10970	.13518	.14647	.15932	.17375	.18441	.19978	.22030	14.44	52.03
Black	.16612	.20851	.22728	.24866	.26969	.28509	.30571	.33253	10.57	42.46

Sources: Data are from a sample of census enumerators' manuscripts, U.S., 1900. The age and duration methods use equations in United Nations 1983a; ch. 3.

Note: Coale and Demeny (1966) West model life tables are used as the basis for all three methods in the table; $q(a)$ is the probability of death before exact age a. Levels calculated from life tables for both sexes assuming a sex ratio at birth of 1.05.

[a] The age and duration groups of women are those used to estimate the particular q's listed in that column.

[b] N is the number of children ever born to each group of women.

[c] The following expectations of life are associated with these levels in the Coale-Demeny West model system. The total is estimated assuming a sex ratio at birth of 1.05.

Level	*Female e_0^0*	*Male e_0^0*	*Total e_0^0*
7	35.00	32.48	33.71
8	37.50	34.89	36.16
9	40.00	37.30	38.62
10	42.50	39.71	41.07
11	45.00	42.12	43.52
12	47.50	44.52	45.98
13	50.00	47.11	48.52
14	52.50	49.56	50.99

[d] These numbers relate the date, expressed in terms of years prior to the census (June 1, 1900), to which each estimate, on average, applied.

[e] The numbers of children ever born, children surviving, and surviving own-children present for women ages 14–34 used to make those estimates are:

	Children ever born	*Children surviving*	*Children present*
Total	19,292	16,121	15,263
White	16,331	13,882	13,276
Black	2,838	2,156	1,917

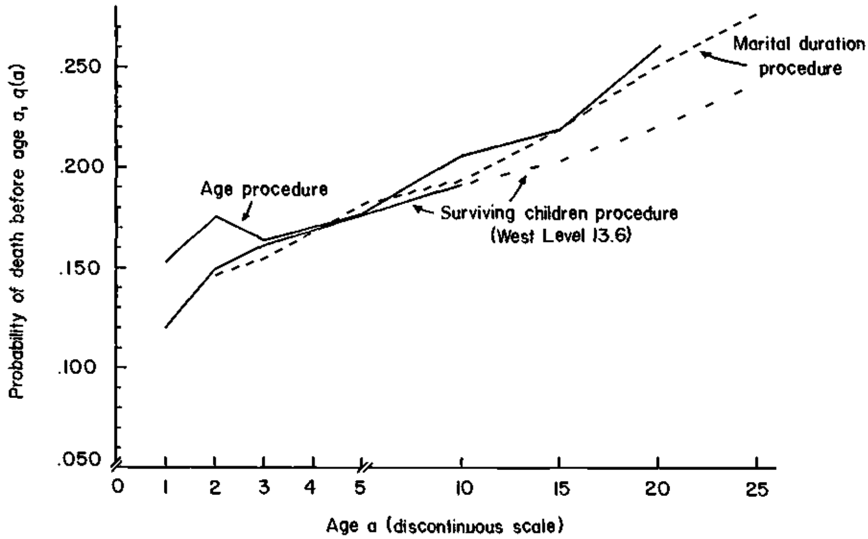

Figure 2.1 Estimates of Cumulative Mortality by Three Indirect Procedures, U.S., 1900

of this reduction can be estimated from Table 2.5. Using the marital-duration estimates, the average Model West level of mortality for marital durations 5–9 and 10–14 years was 13.68 and the average date to which these estimates pertain is 1895.9 (i.e., 1900.5 − (3.4 + 5.8)/2). For women married 25–29 and 30–34 years, the average Model West level was 12.46 and the average date 1884.9. Thus, over the course of 11 years, the improvement in level was 1.22. These estimates translate into a decadal rate of gain in expectation of life at birth of 2.8 years between the mid-1880s and the mid-1890s.

This pace is consistent with Stolnitz's summary of changes in expectation of life at birth in western European countries between the 1880s and the 1900s, which suggested a median decadal rate of gain of 3.05 years (Stolnitz 1955: Table 6). The estimated pace of mortality decline in the U.S. depends, of course, on the suitability of the West model life table system. If post-infant mortality were much higher than assumed in the Model West pattern, some of the divergence shown in Figure 2.1 would be accounted for by this disparity in age patterns. In the extreme, if the North model were appropriate, with its very high tuberculosis death rates and relatively high mortality above age 5, then the mortality improvement over the 11-year period would be only 0.68 levels, roughly half as great as indicated by the

West model. On the other hand, if the East model is used, with its low post-infant mortality, the gain would be 1.81 levels.

Other sources, using different indirect procedures but without direct information on child mortality, have also suggested that mortality declines were occurring in the United States during this period (Higgs 1973, 1979; Meeker 1972; Haines 1979a; D. S. Smith 1983). Direct evidence from registration data (for the limited number of states and cities that had registration systems in place) also points to mortality decline for infants and for children aged 1–4 years in the 1890s (Condran and Crimmins 1980: Table 1). But the child mortality data analyzed here are the strongest evidence yet available, or likely to become available, that child mortality levels were improving for the United States as a whole in the decades before 1900.

Referring again to Figure 2.1, it can be seen that estimates of $q(1)$ and $q(2)$ are much less consistent than those at more advanced ages. In particular, the age model gives relatively high estimates of $q(1)$ and $q(2)$, and the $q(2)$ estimate exceeds all of the estimates of $q(3)$. Such an irregularity could have been produced by a sharp rise in childhood mortality in a short period before the census, but such an event seems unlikely. More plausibly, the explanation lies in data problems. As shown in Table 2.5, the age-model estimates of $q(1)$ are based on relatively few births. Also, age-model estimates of $q(1)$ and $q(2)$ are based disproportionately on first births and births to younger women, births known to be at unusually high risk of death (World Health Organization 1978). This high-risk composition of births is exacerbated by the relatively late age at marriage, with a singulate mean age at marriage of 23.66 years for females in 1900 (Table 2.4).

Before discarding age-model estimates for $q(1)$ and $q(2)$, we experimented with alternative age groupings for women. Despite the fact that the degree of age misreporting in the 1900 census was low relative to previous and subsequent censuses (Coale and Zelnik 1963), there is some indication of age heaping, particularly among blacks. The digital preference appeared largely for ages ending in 0 or 5. This pattern could create a bias if less educated or poorer women, who would also have been more likely to have experienced high mortality among their children, were also more likely to have misstated their ages. In an effort to test whether alternative age groupings would improve age-model estimates, the equations prepared by Hill, Zlotnik, and Durch (1981) were used for the age groups 18–22 (to estimate $q[2]$), 23–27 (to estimate $q[3]$), 28–32 (to estimate $q[5]$), and 33–37 (to estimate $q[10]$). The results (not presented) showed an even less regular pattern than when conventional age groupings were used. Therefore we are inclined to disregard age-model estimates of $q(1)$ and $q(2)$.

To derive a single best estimate of child mortality conditions in the United States near the turn of the century, we amalgamated the $q(3)$, $q(5)$, and $q(10)$ estimates from the three different estimation procedures. The mean West model mortality level corresponding to these variables for the age model is 13.39, and their mean date is 1893.4; for the marital-duration model, the corresponding figures are 13.57 for the date 1894.7. The surviving-children model provides a level of 13.65 and a date of approximately 1896. These are highly consistent with one another and allow for some trend of improved mortality. The grand mean is approximately a level of 13.5 for 1895. At this level of mortality in the West model life table system, $q(5)$ is .180 and the implied expectation of life at birth is 49.8 years.

The estimate of a $q(5)$ of .180 for 1895 is probably the single most robust estimate of childhood mortality that we can make based on the census sample. At this level, American child mortality compared favorably with that in most other Western countries. Among the countries shown in Table 2.3, only Australia, New Zealand, Norway, Sweden, and Ireland had lower childhood mortality in the 1890s, while Denmark's level was nearly identical. This group of countries is preponderantly rural. For the more industrialized countries of western Europe—Belgium, England, France, Germany, and the Netherlands—child mortality was 25–62 percent higher than in the United States, and in southern and eastern Europe the excess was even greater.

It should be noted that selective mortality of mothers could introduce a downward bias into our estimates of mortality for children of older mothers, and hence into estimates of trends. If women who died before 1900 experienced higher mortality among their children than women who survived, which seems likely, then the child mortality experience reported by the survivors in the census of 1900 underestimates that experienced by the cohort of women who began childbearing. Such a correlation could result from household episodes of disease that raise the death risks for both mothers and children; from shared hazards of the birth process; and from social and economic influences that affect the health of all family members.

A rough estimate of the amount of bias that might be introduced into reports of women aged 45–49 can be obtained through the following considerations:

1. The proportion of women reaching age 22 who died before age 47 in the DRA life table of 1900–1902 was 19.2 percent (Glover 1921:60–61).

2. These women can be assumed to have died about halfway through this interval, and so to have contributed about 10 percent of the cohort's births.

3. If mortality among their children was 50 percent higher than average, then child mortality in the original cohort of women would have been 5 percent higher than among the cohort of surviving women.

A differential of 50 percent is much larger than what is implied by social-class differences in child mortality described in Chapters 3 and 4. That is, the clustering of mortality by social class is unlikely to have induced a child mortality differential between living and dying mothers as large as 50 percent, even if all deaths of mothers were confined to the lowest social classes. A differential of 20 percent is more plausible. Nor could deaths of mothers and babies during childbirth create a bias as large as 5 percent. Only 1.3 percent of women surviving to age 20 died of maternal causes in the DRA life table of 1900 (Preston, Keyfitz, and Schoen 1972:727). Even if *all* of their children had died (and again assuming that they bore half as many children by age 50 as surviving women), the downward bias in the $q(5)$ of .180 (based on estimates supplied by surviving women) would only be 3.0 percent.

So 5 percent appears to be close to an upper limit on the extent of bias in child mortality resulting from the selective mortality of women before age 50. And most of our analysis is based upon younger women, among whom the forces of selection would be weaker still. We need to be aware of the potential bias from selective mortality of mothers, but it does not appear to be large enough to have seriously distorted our estimates.

Reliability of the Estimates

Before a discussion and interpretation of the mortality estimates is undertaken, it is useful to conduct tests of the reliability of the data and estimation procedures. Two tests were performed, although they were not entirely independent. The first test involved a comparison to figures contained in the 1900–1902 life tables for states in the Death Registration Area (Glover 1921). To make this comparison, we repeated the foregoing calculations for women in the 1900 census sample who resided in the states that constituted the Death Registration Area of 1900–1902. As has been mentioned above, this area comprised a minority of the population in 1900 (26.3 percent), but its mortality conditions are relatively accurately known by virtue of the

Glover life tables. Table 2.6 presents the basic results of this comparison. The various estimates are graphed in Figure 2.2. It is clear that the surviving-children approach produced a life table (West model level 13.29) in remarkably close agreement with the Glover table. It is important to note, however, that the Glover table pertains to a date some five years later than the surviving-children estimates. Nevertheless, our estimates of $q(a)$ are slightly higher, allowing the possibility of a small downtrend in mortality. Table 2.6 and Figure 2.2 are also instructive regarding the very close conformity of mortality in the DRA to the West model life table system.

As in the total American population, estimates based on the marital-duration and age models diverge systematically from the surviving-children estimates beyond age 5. This divergence occurs because the surviving-children estimates are limited to women below age 35, whereas the others are not. The indication of a downtrend in mortality—higher child mortality conditions for offspring of older women—is even clearer in Figure 2.2 than in Figure 2.1. The age model again produces high estimates for $q(1)$ and, especially, $q(2)$. The $q(a)$ sequences for both the age and duration models are less smooth and regular than for the total American population, and there are larger (but unsystematic) divergences between the two sets of estimates. Both of these traits are plausibly ascribed to the smaller number of observations available in the census sample of the DRA states. For the values believed to be most reliably estimated, $q(3)$, $q(5)$, and $q(10)$, the mean West model level is 13.03 for marital-duration-based estimates and 13.11 for age-based estimates. These levels are, respectively, 0.54 and 0.28 levels below our corresponding estimates for the total United States. The surviving-children estimate of level 13.29 is .36 levels below the estimate for the total United States in Table 2.5. Since each level represents about 2.4 years of life expectancy at birth, it appears that life expectancy at birth in the Death Registration Area was about one year lower than in the United States as a whole at the turn of the century. Note that this conclusion is not based on a comparison of data drawn from different sources but on a comparison of data for different areas from the census sample alone.

For whites, applying the surviving-children method to the census sample yields a level of mortality in the Death Registration Area very similar to that contained in the Glover 1900–1902 Death Registration Area life table: $q(3)$ in the two sources is .167 and .164, $q(5)$ is .182 and .179, and $q(10)$ is .198 and .196. The census sample data and procedures thus receive strong validation for whites and for the total population through comparison to the Glover table.

For blacks, however, there is a larger discrepancy. For $q(3)$, the cen-

TABLE 2.6
Estimates of Child Mortality by Race: U.S. Death Registration Area, 1900

	q(1)	*q(2)*	*q(3)*	*q(5)*	*q(10)*	*q(15)*	*q(20)*	*q(25)*	*Level*	*Implied life expectancy at birth*
	1900 Census sample using West model									
Age model										
Total	0.13010	0.17495	0.16259	0.18398	0.22284	0.24371	0.27520	—		
(*N*)	(56)	(675)	(1,408)	(2,156)	(2,490)	(2,729)	(2,575)	—		
White	0.13145	0.17403	0.16436	0.17947	0.21932	0.24585	0.27521	—		
(*N*)	(55)	(655)	(1,399)	(2,107)	(2,448)	(2,706)	(2,513)	—		
Duration model										
Total	—	0.13855	0.14707	0.19962	0.23267	0.23796	0.28059	0.30224		
(*N*)	—	(609)	(1,620)	(2,174)	(2,062)	(2,026)	(1,810)	(1,561)		
White	—	0.13576	0.14724	0.19472	0.23133	0.23848	0.27854	0.30215		
(*N*)	—	(597)	(1,599)	(2,139)	(2,037)	(2,022)	(1,784)	(1,542)		
Surviving children model women aged 14–34										
Total	0.12532	0.15577	0.16926	0.18462	0.20110	0.21318	0.23019	0.25264	13.29	49.21
White	0.12358	0.15347	0.16671	0.18179	0.19804	0.20997	0.22684	0.24912	13.41	49.51
Black	0.25814	0.32084	0.34862	0.38024	0.40824	0.42834	0.45426	0.48717	5.77	30.69
	Glover's 1900–1902 life tables for the Death Registration Area									
Total	0.12448	0.15383	0.16708	0.18196	0.19948	0.21037	0.22761	0.25232	—	49.24
White	0.12231	0.15112	0.16414	0.17886	0.19616	0.20674	0.22355	0.24785	—	49.62
Black	0.23447	0.29094	0.31561	0.33824	0.36621	0.39008	0.42135	0.45491	—	33.76

TABLE 2.6 *(cont.)*

	$q(1)$	$q(2)$	$q(3)$	$q(5)$	$q(10)$	$q(15)$	$q(20)$	$q(25)$	*Level*	*Implied life expectancy at birth*
	Implied level in West model life table system									
Age model										
Total	12.93	12.28	13.61	13.31	12.42	12.14	11.67	—	—	—
White	12.84	12.33	13.52	13.51	12.56	12.06	11.67	—	—	—
Duration model										
Total	—	14.24	14.40	12.65	12.04	12.35	11.47	11.60	—	—
White	—	14.40	14.40	12.85	12.09	12.33	11.55	11.60	—	—

Source: Sample of census enumerator's manuscripts, U.S., 1900. See text for estimation procedures. Death Registration Area life tables for 1900–1902 are found in Glover 1921.

Note: $q(a)$ is the probability of death before exact age a. Values of $q(a)$ for both sexes combined are derived by combining life tables for males and females assuming a sex ratio at birth of 105 males per 100 females. The value of N for the age and duration models, shown in parentheses, is the number of children ever born. For the surviving-children model, the relevant N's are:

	Children ever born	*Children surviving*	*Children present*
Total	4,261	3,533	3,344
White	4,182	3,478	3,303
Black	70	46	34

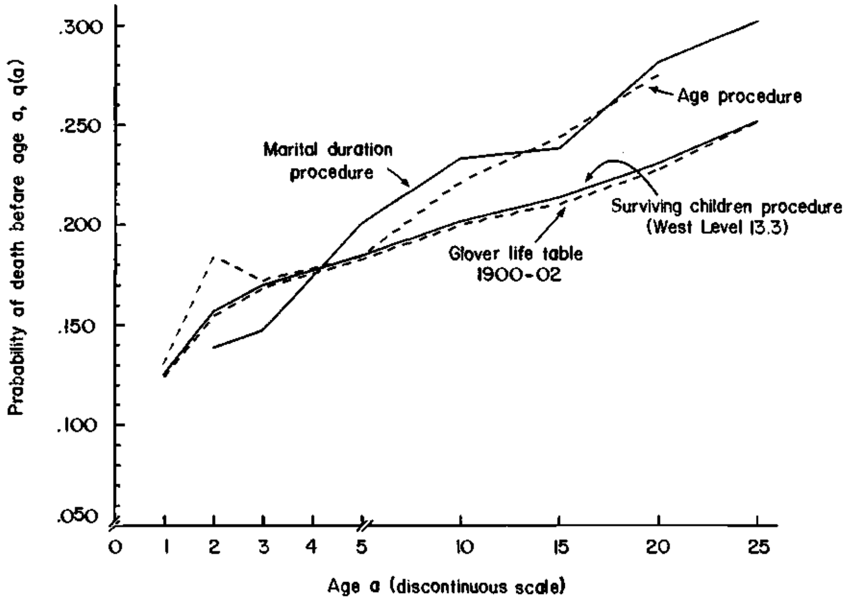

Figure 2.2 Estimates of Cumulative Mortality for Death Registration Area, U.S., 1900

sus sample gives a mortality level of .349, compared with .316 for the Glover life table. The $q(5)$ figures are .380 and .338, and $q(10)$, .408 and .366. The census sample implies even higher mortality in the DRA than does the Glover table. But the surviving-children estimates for blacks in the Death Registration Area are based on only 70 births. When such small numbers are involved, tests of significance are in order. We have assumed that death is a binomial process and that the underlying probability of death before age 5 for blacks in the DRA at this time is .33824, as in the Glover life table. With 70 observations and an "observed" $q(5)$ of .38024, the standard error of the number of deaths is $(.33824*(1 - .33824)*70)^{1/2} = 3.96$.[7] The observed number of deaths, 70*(.38024) = 26.61, is thus within one standard error of the expected number of deaths, 70*(.33824) = 23.68. Hence we conclude that the surviving-children approach gives a mortality level for blacks in the Death Registration Area that is not statistically significantly different from that in the 1900–1902 Glover life table for blacks.

This result is a reassuring indication that surviving—children data for blacks, at least in the Death Registration Area, are in line with other estimates believed to be accurate. It is also reassuring that our mortality estimates are slightly higher, since reporting errors in the census seem more likely to lead to an underestimate than to an overestimate of mortality.

A second test of the reliability of the census data uses a data set consisting of states and territories as the units of observation.[8] For each state or territory, a summary mortality index in the form of a ratio of actual to expected deaths was prepared using the information in the 1900 census sample on children ever born and children surviving for each geographic unit. This index, used extensively in the remainder of the book, is described in detail in Chapter 3. In addition, a death rate for children aged 0–4 was calculated from published data in the census of 1900 referring to mortality in the year prior to the census (June 1, 1899 to May 31, 1900). The census of 1900 included registration mortality data for the states of the Death Registration Area, and for cities in states outside the Death Registration Area whenever such data were available. When registration data were unavailable, responses to a census question on deaths in the household in the year prior to the census were substituted (Condran and Crimmins 1979). The registration data are known to have been more accurate than the "deaths last year" question.

The correlation between our mortality index and the published census "deaths last year" information for children 0–4 would not be expected to be perfect, since they covered different time periods and age groups. But the index was most influenced by young children, and many of those deaths had taken place in the late 1890s. The zero-order correlation between the index and the census death rate for the 45 regional aggregates (see note 8) was, in fact, .649, which is statistically significantly different from zero at a one percent confidence level. So the two independent sources of information on geographic variation in mortality are in reasonably good agreement.

Which data source, the census questions on children born and surviving (providing the "index") or the tabulations of deaths from registration and census reports (the "death rate"), is more accurate? To answer this question, a weighted least squares regression was run with a state's mortality index as the independent variable and the state's 1900 census death rate for children aged 0–4 as the dependent variable. The weights were the number of children ever born in each state. If our data are more accurate, then this simple regression fitted to *all* states and territories should produce positive residuals (i.e., actual values of the census death rate exceeding predicted values) for the DRA states and largely negative residuals (i.e., actual values of the census death rate less than predicted values) for non-DRA states and territories. That is, the death rate should be higher (relative to our index) in states in the DRA than it is in states that are not in the DRA; the regression line itself, of course, reflects the average level of incompleteness in the death rate across states both in and out of the DRA.

Exactly such a result emerges. For the ten DRA states plus the District of Columbia (also in the Death Registration Area), the mean residual was +10.74, and only one state, Maine, had a small negative residual (−1.28). For the other 34 states and territories (or groupings), the mean residual was −4.73, and 26 of the 34 had negative residuals. Of the states and territories in this latter group that had unexpectedly positive residuals, most had substantial registration coverage that was reflected in the census death reports.[9] Thus, the results here strongly support the superiority of the indirect mortality estimates from the census questions on children ever born and surviving relative to direct census mortality data on deaths last year.

Black Mortality

An important modification to received wisdom posed by the new figures relates to the black population. The three basic estimates of black mortality for the entire United States (from Table 2.5) are plotted in Figure 2.3. The age model and the surviving-children procedure give similar results, with the latter series basically representing a smoothed version of the former. The age model gives erratic results for the younger ages, where *N*'s are small. The anomalous series is that pertaining to marital-duration estimates, which declines from $q(1)$ to $q(10)$ before rising sharply. A likely explanation for the irregularity is the high rate of marital disruption and nonmarital unions among the black population (Farley 1970: ch. 6). It appears that the age and surviving-children procedures afford the best estimates of black child mortality. For the whole United States, the surviving-children estimate for black women aged 14–34 is West model level 10.32. The mean West model level corresponding to the $q(3)$, $q(5)$, and $q(10)$ estimates by the age procedure is 10.61 (e_0 = 42.46), and the mean date to which these estimates pertain is 1893.1. The mean of the $q(5)$ values for levels 10.32 and 10.61 is .255.

The $q(5)$ figure of .255 that emerges from the census sample for blacks is far below the figure of .338 appearing in Glover's life table for blacks in the DRA. Though it is possible that errors in one or both sources account for this discrepancy, it is reassuring that the census sample for the Death Registration Area itself implies a level of mortality even higher than in the Glover table, as we have just seen. The

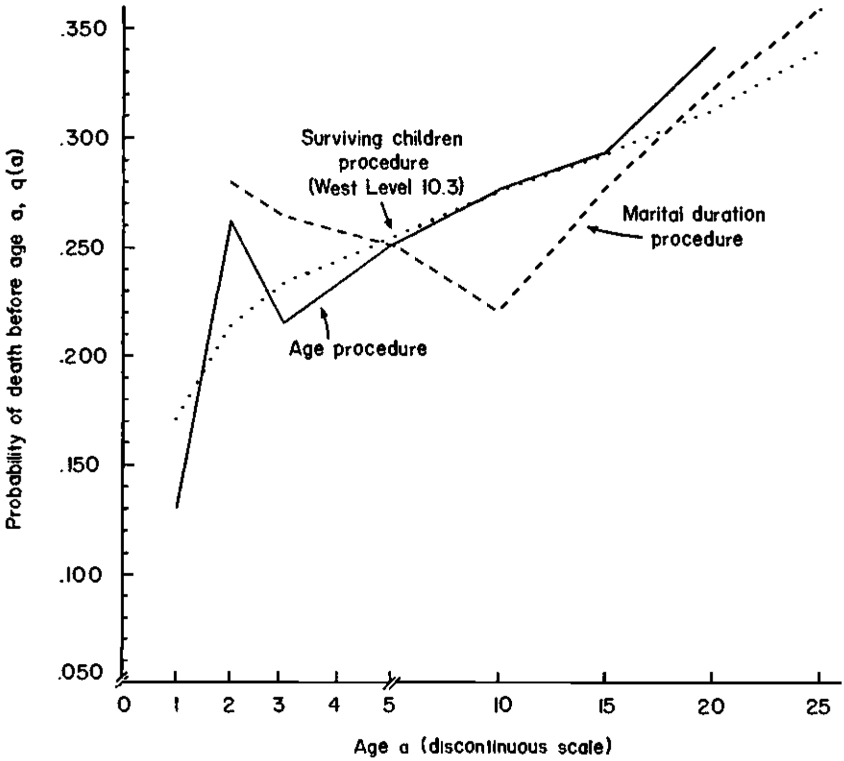

Figure 2.3 Estimates of Cumulative Mortality by Three Indirect Procedures, U.S. Blacks, 1900

most persuasive explanation of the discrepancy is that, before the extensive deployment of public-health measures aimed at communicable diseases during the twentieth century, there was a decisive rural advantage in mortality. This advantage was discussed in Chapter 1, and confirming evidence will be presented subsequently. The highly urbanized blacks in the Death Registration Area seem to have left behind a seriously distorted impression of general black mortality conditions, which has also exaggerated the black/white gap. Instead of a black/white ratio of $q(5)$'s of 1.89 from the Glover life tables (.3382/.1789), our results for the entire United States give a figure of 1.58 (.2550/.1610). Black child mortality appears to have been, both absolutely and in relation to whites, much poorer in the urban industrial states that formed the bulk of the DRA than in the more rural South. This revision of racial mortality differentials around 1900 also implies that less progress has been made during the twentieth century in narrowing the gap between black and white child mortality than is commonly assumed.

Several early warnings were sounded about the likely unrepresentativeness of Death Registration Area figures for blacks. American census officials later considered it highly probable that black mortality was better in the South than in the Death Registration Area (U.S. Bureau of the Census 1918:341). Nevertheless, most modern analysts have accepted as nationally representative the Death Registration Area mortality rates for blacks, or have even considered them too low. One reason why the DRA figures for black children appeared plausible is that *adult* black mortality for the whole United States was extraordinarily high, as revealed by one or another form of intercensal survival analysis or by stable population analysis. Demeny and Gingrich (1967), Farley (1970), and Meeker (1976) used West model life tables to combine adult mortality levels estimated from these procedures with presumed levels of child mortality. The resulting levels of expectation of life at birth for both sexes combined were 32.3 years for 1900–1910 (Demeny and Gingrich 1967), 30.2 years for 1900 (Meeker 1976), and 25.0 years for black females in 1880–1900 (Farley 1970). But the level of expectation of life at birth corresponding in the West model system to the level of black *child* mortality by the surviving-child method is 41.8 years. Coale and Rives (1973), in their reconstruction of black age distributions, used several mortality assumptions that are in the range of $e_0 = 30$ for the period; they suggested that levels of child mortality in the Death Registration Area were actually underestimates for blacks in the nation as a whole, rather than overestimates as we have shown. At a life expectancy level of 30 for 1900, $q(2)$ in the West model life table system is .328 and $q(5)$ is .388. These figures are about 50 percent higher than those which we estimate based on the census sample.

A probable key to the discrepancy is the appropriateness of the West model to black American mortality in the era. The best evidence on this matter is the age pattern of mortality in the Death Registration Area states. Zelnik (1969) has carefully studied this pattern. He demonstrated that the relation between child and adult mortality for blacks was very different from that implied by the West model between 1900–1902 and 1949–51. Mortality below age 10 was very favorable relative to mortality in the adult years, with differences in implied levels of expectation of life at birth (i.e., based on age-specific death rates in combination with West models) as large as 25 years. Moreover, the discrepancies increased as the Death Registration Area expanded to national coverage. Condran (1984) has produced similar findings for Philadelphia in the late nineteenth century. Demeny and

Gingrich (1967) argued that this trend could be explained by poor death registration for southern children, who were successively incorporated into the Death Registration Area, but the required amounts of underregistration are implausibly high. Furthermore, Zelnik introduced a life table of black Metropolitan Life Insurance clients (i.e., based on quite good data) that shows exactly the same age pattern of deviations as the entire United States life table for blacks. Although Zelnik did not speculate on reasons for the pattern of deviations, it is likely that tuberculosis played an important role. This disease was exceptionally common among American blacks (Meeker 1976; Condran 1984) and is capable of heavily distorting age patterns of mortality in the implied direction (Preston 1976).

Eblen (1974) is the only analyst to come close to what now appears to be the correct range of black child mortality. He used a more flexible model life table system that allowed the data (age distributions in successive censuses) to determine, in part, the relation between child and adult mortality. His estimate of $q(1)$ for 1890–1900 was .200, and for $q(10)$, .352. These estimates are only about 15 percent above our own.

So the previously accepted picture of extremely high black child mortality conditions around 1900 appears to have resulted from two distortions that reinforced one another: highly unrepresentative mortality conditions in the urban Northeast, the only area having an appreciable amount of direct vital registration data; and a very peculiar age pattern of mortality for blacks in the nation as a whole, with much better child mortality conditions than are implied by the levels of adult mortality that could be estimated for the nation as a whole by intercensal comparisons.

Just as the inference of child mortality levels from adult mortality levels can, and apparently did, lead to serious error, so can the extrapolation from child levels to adult levels produce distortions. Although we have presented in our tables the life-expectancy estimates corresponding to child mortality levels for blacks as a convenient metric, we caution against using them as valid estimates. They are almost certainly too high because blacks had higher adult mortality, relative to child mortality, than is implied by the West model life table system. But there is every reason to believe that the child mortality estimates for blacks that are presented here are superior to others that have been proposed.

In contrast to results for the total and white populations, the black estimates in Figure 2.3 do not suggest much of a downtrend in child mortality. Only $q(20)$ and $q(25)$ in the marital-duration model are higher than estimates implied by the surviving-children procedure (i.e., by the best-fitting West model). Even this discrepancy could be accounted for by the unusually high adult mortality in the black population relative to the West model life table system. What appears to be a mild decline could simply be the result of age distortions in the model life table used. It is always possible that higher fractions of dead children were omitted by older women and that such omissions are obscuring a true decline. The most we can say is that the census sample data are not consistent with much improvement in black child mortality in the late nineteenth century. The results do not provide much support for the possibility that black life expectancy followed a U-shaped time trend between 1860 and 1900 (Fogel, et al. 1978:78). They are, however, consistent with Ewbank's (1987) recent conclusion that black mortality rates were essentially stagnant in the late nineteenth century.

If our estimates of black child mortality are correct, they imply that the major accounts of black demographic history (e.g., Coale and Rives 1973; Farley 1970) may need revising. In particular, birth-rate estimates for the nineteenth century appear to need downward revision by approximately 8 to 10 percent, since the number of (surviving) children in censuses, on which the reconstructions are primarily based, would require fewer births to produce if child mortality were lower than previously assumed.

Mortality of Whites and of the Total Population

A similar but much smaller bias exists for the white population. Because whites in the Death Registration Area were more highly urban than in the nation as a whole (67 percent versus 43 percent), one might expect that childhood mortality for whites in the DRA was also higher than in the nation as a whole. A comparison of Tables 2.5 and 2.6 confirms this expectation. The surviving-children method applied to whites in the nation as a whole yields a $q(5)$ of .161; but for whites in the Death Registration Area, it is .182, or some 13 percent higher. It is likely that a higher proportion of foreign-born persons in the DRA (22 percent versus 14 percent for the nation as a whole) also contributed to this outcome. Since our results for whites in the Death Registration Area came very close to Glover's life table for whites in 1900–1902 ($q[5]$ = .179 from Table 2.6), we conclude that the Glover life tables also give a somewhat biased view of white mortality in the entire U.S. at the turn of the century. In terms of its implication for expectation of life at birth, as shown in Tables 2.5 and 2.6, the difference between $q(5)$'s of .161 and .182 (using the surviving-children method) amounts to 2.32 years, or expectations of life at birth of 51.83 years (for the nation as a whole) versus 49.51 years (for the Death Registration Area).

Thus the Death Registration Area life tables, the most authoritative and widely cited information on American mortality rates at the turn of the century, present too pessimistic a picture of mortality conditions for whites and, especially, for blacks. Ironically, this bias is sharply attenuated among the total American population. The census sample gives a $q(5)$ of .176 for the whole United States and .185 for the Death Registration Area. The relatively small difference between these figures results from the fact that blacks contributed a much smaller proportion of births in the Death Registration Area than they did in the nation as a whole. Black births used for the surviving-children estimates were 14.7 percent of total births in the United States, but they were only 1.6 percent of births in the Death Registration Area.

Thus, the fact that the Death Registration Area life table provides reasonably good estimates of child mortality for the United States as a whole in 1900–1902 is simply the result of errors that were largely offsetting. Mortality for both blacks and whites appears to have been too high in the Death Registration Area tables, but the upward bias is largely offset by the very low proportion of blacks in the Death Registration Area.

Quantitative Summary

To summarize results of this chapter, we use the surviving-children method because it aggregates over different ages and marital durations of women and appears to work very well, especially for blacks. The basic estimates of the probability of dying before age 5, $q(5)$, are as follows (from Tables 2.2, 2.5, and 2.6):

	Census sample, United States, 1896	*Census sample, Death Registration Area, 1896*	*Vital registration, Death Registration Area, 1900–1902*
White	.161	.182	.179
Black	.255	.380*	.338
Total	.176	.185	.182

* *based upon only 70 births, insignificantly different from .338*

The summary shows clearly that mortality was substantially lower in the nation as a whole than it was in the Death Registration Area for both whites and blacks; that the census sample gives results very close to the vital statistics when confined to the states constituting the Death Registration Area; and that the bias in DRA figures is substantially offset when blacks and whites are combined because such a low percentage (1.9 percent) of the DRA population was black.