



The Logistic Curve and the Census Count of 1940

Raymond Pearl; Lowell J. Reed; Joseph F. Kish

Science, New Series, Vol. 92, No. 2395. (Nov. 22, 1940), pp. 486-488.

Stable URL:

<http://links.jstor.org/sici?sici=0036-8075%2819401122%293%3A92%3A2395%3C486%3ATLCATC%3E2.0.CO%3B2-T>

Science is currently published by American Association for the Advancement of Science.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/aaas.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

S. H. Caldwell, the Massachusetts Institute of Technology.
 E. U. Condon, Westinghouse Electric and Manufacturing Company.
 Donald Cooksey, University of California.
 C. S. Draper, the Massachusetts Institute of Technology.
 E. W. Engstrom, Radio Corporation of America.
 W. C. Evans, Westinghouse Electric and Manufacturing Company.
 T. C. Fry, Bell Telephone Laboratories.
 L. O. Grondahl, Union Switch and Signal Company.
 William W. Hansen, Stanford University.
 G. R. Harrison, the Massachusetts Institute of Technology.
 Thomas H. Johnson, Franklin Institute.
 M. J. Kelly, Bell Telephone Laboratories.
 Paul Klopsteg, Central Scientific Company.
 Ernest O. Lawrence, University of California.
 Frank D. Lewis, Loomis Laboratories.
 George Metcalf, General Electric Company.
 J. P. Molnar, the Massachusetts Institute of Technology.
 Philip McCord Morse, the Massachusetts Institute of Technology.
 Edward J. Poitras, the California Institute of Technology.
 Louis A. Turner, Princeton University.
 Warren Weaver, Rockefeller Foundation.
 H. Hugh Willis, Sperry Gyroscope Company.
 C. L. Wilson, Research Corporation.

Serving on committees reporting directly to the chairman or otherwise assisting the chairman are the following:

J. W. Beams, University of Virginia.
 Gregory Breit, University of Wisconsin.
 Lyman J. Briggs, director, National Bureau of Standards.
 T. H. Dillon (Colonel), Carnegie Institution of Washington.

Ross Gunn, Naval Research Laboratory.
 John H. Howard, the Massachusetts Institute of Technology.
 G. B. Pegram, Columbia University.
 H. C. Urey, Columbia University.

Lloyd Sutton is serving on a Patent Advisory Committee with Commissioner Coe.

With the exception of six who have obtained leave of absence from their regular positions and who are devoting full time to the work of the committee, all have volunteered their services on a part-time basis without compensation. The committee has completed or has under negotiation approximately sixty contracts with industrial laboratories and universities for carrying out research on specific problems.

According to Dr. Bush, the scientists and engineers of the United States have been most enthusiastic in the offer of their services. In addition, they have understood that it is not possible on short notice to make the most effective use of the skill and talent of thousands of persons, and they have been sympathetic with the program of the committee in accepting the services of individual scientists only as problems were brought into a form where their talents could most effectively be used.

The task assigned to the committee by the Council of National Defense is that of correlating and supporting scientific research on mechanisms and devices of warfare. It does not extend to such matters as food or medicine and health. For that reason the men selected to aid the committee will continue to be drawn largely from the fields of chemistry, physics and engineering.

SPECIAL ARTICLES

THE LOGISTIC CURVE AND THE CENSUS COUNT OF 1940¹

EARLY in 1920 Pearl and Reed² published the results of fitting a logistic curve to the census counts of the population of the United States from 1790 to 1910, inclusive. At the time the computations were made in 1919 the results of the census of 1920 were not available, and therefore were not and could not be included in the calculations. Soon after the original paper was published the notation was improved, and the curve took the definitive form, still, however, *without* the use of the 1920 count:

$$y = \frac{197.27}{1 + 67.32e^{-0.0313x}} \quad (1)$$

¹ From the department of biology and the department of biostatistics (Paper No. 215) of the School of Hygiene and Public Health, Johns Hopkins University.

² R. Pearl and L. J. Reed, *Proc. Nat. Acad. Sci.*, 6: 275-288, 1920.

where y denotes calculated population in millions, and x time, in base units of one year. This curve (1) will be hereinafter referred to as Logistic I. It is depicted graphically in Fig. 1 with the actual census counts of 1920, 1930 and 1940 (preliminary figure) inserted as crossed circles connected by a dash line to indicate that in the derivation of equation (1) no census count after that of 1910 was used.

Thirty years have elapsed since the last *datum* (the census count of 1910) available when the curve was calculated. It seems justifiable now to make a further examination of how the case goes, as was done in 1930.³ The following tabulation gives the facts.

It thus appears that the forecast made by Logistic I missed the counted population by 16 parts in a thousand in excess in 1920; by 2.5 parts in a thousand in defect in 1930, and by 37.3 parts in a thousand in excess in 1940. The error in 1940 is of perhaps about

³ R. Pearl and L. J. Reed, *SCIENCE*, 72: 399-401, 1930.

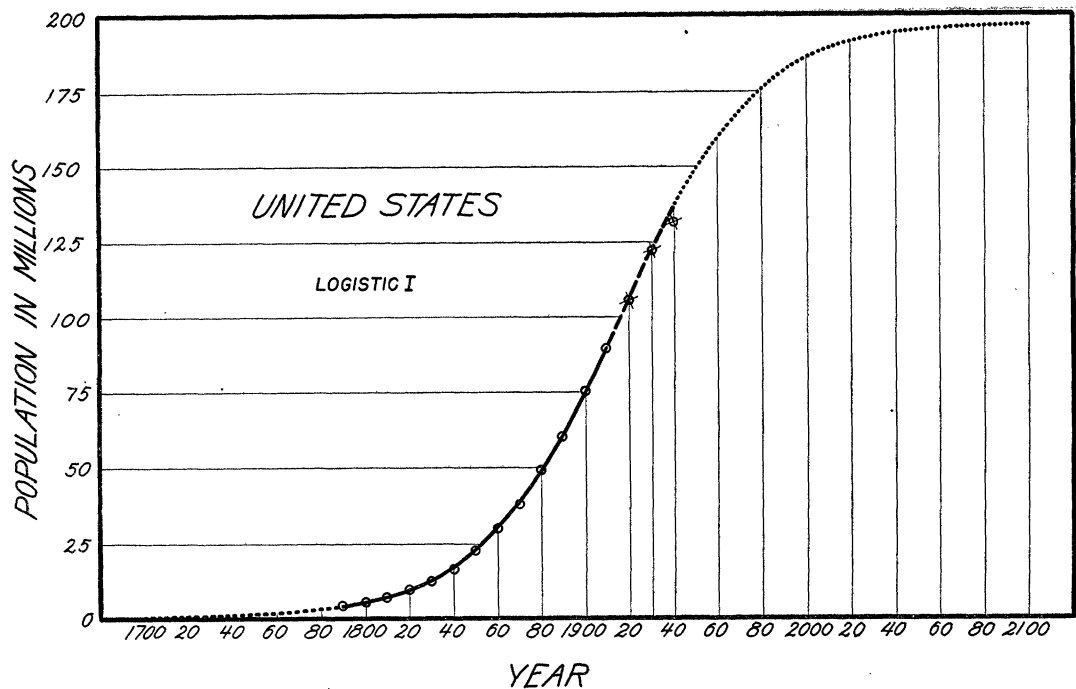


FIG. 1. The census counts of the population of the United States from 1790 to 1940, inclusive (given by circles). The smooth curve is the logistic of equation (1) fitted to the census counts from 1790 to 1910 inclusive. The broken lines show the extrapolation of the curve beyond the data to which it was fitted. The dash portion from 1910 to 1940 is the part of the extrapolation which has been tested by census counts (crossed circles) which have been made since the logistic was originally fitted. The dotted line shows the further extrapolation of the same curve.

TABLE I POPULATION OF CONTINENTAL UNITED STATES (IN MILLIONS)		
	A. As forecast by Logistic I (data of 1790-1910)	B. As counted by the Census Bureau
1920	107.4	105.7
1930	122.4	122.7
1940	136.3	131.4

the same order of magnitude as that probably inherent in the count itself. But it is considerably larger than that made by the curve in either 1920 or 1930.

In our earlier work with the logistic curve we have repeatedly emphasized (a) that it is a basic postulate of the logistic theory of population growth that any particular population can be expected to continue to follow in its later growth the same logistic curve that it has followed in its earlier growth *only* if there has been no serious or cataclysmic alteration of the conditions (climatic, geological, biological, economic or social) under which its earlier growth has taken place; and (b) that if, and when, evidence appeared that indicated that substantial alteration of the conditions of growth had occurred a new logistic should, and would, be derived to take into account the new evidence. The census counts of 1920 and 1930 indicated no such evidence. Does the count of 1940? We do

TABLE II THE POPULATION OF THE UNITED STATES: (A) AS COUNTED IN OFFICIAL CENSUSES; (B) ACCORDING TO LOGISTIC II; (C) ACCORDING TO LOGISTIC I. (ALL POPULATIONS ARE IN MILLIONS OF PERSONS)					
Year	A Observed	B Logistic II	C Logistic I	B-A	C-A
Lower asympt000	.000
1700210	.239
1720399	.446
1740758	.833
1760	1.438	1.553
1780	2.718	2.887
1790	3.929	3.730	3.929	-.199	.000
1800	5.308	5.108	5.336	-.200	+.028
1810	7.240	6.974	7.228	-.266	+.012
1820	9.638	9.487	9.757	-.151	+.119
1830	12.866	12.840	13.109	-.026	+.243
1840	17.069	17.260	17.506	+.191	+.437
1850	23.192	22.997	23.192	-.195	.000
1860	31.443	30.294	30.412	-1.149	-1.031
1870	38.558	39.342	39.372	+.784	+.814
1880	50.156	50.209	50.177	+.053	+.021
1890	62.948	62.775	62.769	-.173	-.179
1900	75.995	76.684	76.870	+.689	+.875
1910	91.972	91.352	91.972	-.620	.000
1920	105.711	106.053	107.394	+.342	+1.683
1930	122.775	120.054	122.397	-2.721	-.378
1940	131.410	132.756	136.318	+1.346	+4.908
1950	143.779	148.678
1960	152.986	159.230
1970	160.430	167.945
1980	166.295	174.941
1990	170.820	180.437
2000	174.256	184.678
2020	178.751	190.341
2040	181.205	193.509
2060	182.522	195.249
2080	183.221	196.337
2100	183.590	196.681
Upper asympt	184.000	197.274

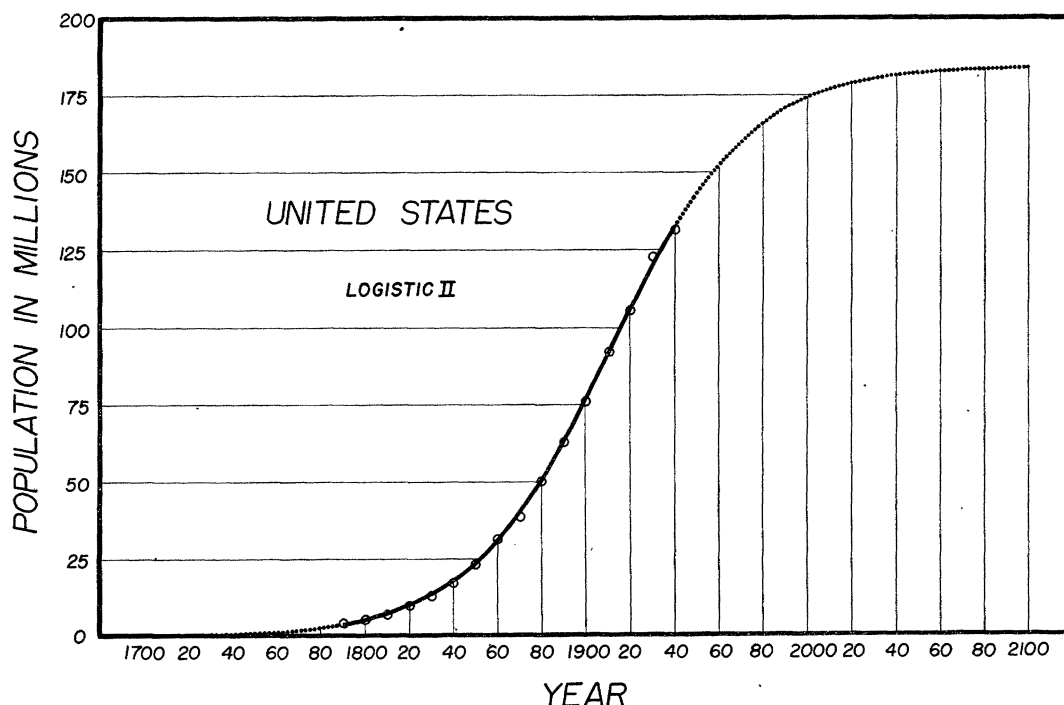


Fig. 2. Logistic curve II (smooth curve) fitted to the census counts (circles) of the population of the United States from 1790 to 1940 inclusive.

not know, and shall not know until the count of 1950 gives some indication as to whether the discrepancy between theory and observation in 1940 is merely a minor fluctuation that will be compensated for in the next ten years, or marks instead the beginning of a different trend of the curve.

But in the meantime, in part as a token of good faith relative to our past promises, we have computed a new logistic curve (Logistic II), using the method of successive least square approximations,⁴ to the history of the growth of the population of the United States, including all the recorded data from 1790 to 1940 inclusive. The results are set forth in Table II, and Fig. 2. The equation of this new logistic (Logistic II) is:

$$y = \frac{184.00}{1 + 66.69e^{-0.0322x}} \quad (2)$$

The constants K and C in (2) are *smaller* than in (1) by amounts that are respectively 6.7 and 0.9 per cent. of their values in (1). The constant r is *larger* in (2) than in (1) by 2.9 per cent. of its value in (1). From Table II the following root mean square deviations are computed

$$\sqrt{\frac{\sum (B-A)^2}{16}} = \sqrt{\frac{12.408}{16}} = .8805$$

$$\sqrt{\frac{\sum (C-A)^2}{16}} = \sqrt{\frac{29.853}{16}} = 1.3659$$

⁴ R. Pearl, "Introduction to Medical Biometry and Statistics." Third Edit. Chapter XVIII *passim*. Philadelphia: W. B. Saunders Company, 1940.

It is plain that either of the two logisties fits the 16 observed populations very well, considering that each is only a three constant curve. But as to which makes the better forecast of the population to be counted in 1950 we are not prepared to say at this time. We hope to make a statement on the point in 1950.

RAYMOND PEARL⁵

LOWELL J. REED

JOSEPH F. KISH

SCHOOL OF HYGIENE AND PUBLIC HEALTH,
THE JOHNS HOPKINS UNIVERSITY

THE HEPARINS OF VARIOUS MAMMALIAN SPECIES AND THEIR RELATIVE ANTI-COAGULANT POTENCY

ON the basis of studies of impure beef heparin Jorpes¹ postulated that heparin is a mixture of mucosin polysulfuric acids, the anticoagulant action of which is dependent on the sulfate groups, the potency increasing with increasing sulfur content. The isolation of the crystalline barium salt of heparin by Charles and Scott² has made it possible to prepare large amounts of the pure anticoagulant. Working with samples of this crystalline material isolated from different tissues of the ox, Charles and Todd³ could

⁵ Professor Pearl died on November 17.

¹ Jorpes, *Biochem. Jour.*, 29: 1817, 1935; *ibid.*, 33: 47, 1938.

² Charles and Scott, *Biochem. Jour.*, 30: 1927, 1936.

³ Charles and Todd, *Biochem. Jour.*, 34: 112, 1940.