ESTIMATION OF SUMMARY MEASURES OF INCOME SIZE DISTRIBUTION FROM GROUPED DATA

Emmett Spiers, U.S. Bureau of the Census

The impetus for this research springs from a decision by the Census Bureau to update its historical series on the trends in the income of families and persons [1]. Since such an undertaking involved the handling of truly massive amounts of grouped income data, it was necessary to employ methods for calculating summary distribution measures which were inexpensive as well as reasonably accurate. In the present paper we will discuss the methods finally chosen and compare them to some of the alternatives considered.

Organizationally, the paper is divided into four sections. The first of these provides a brief overview of available techniques and describes the properties we will require for our application. Sections 2 and 3 discuss some numerical comparisons made between various alternative estimation procedures. Section 4 provides a few concluding remarks.

1. PROPERTIES DESIRED AND ALTERNATIVES CONSIDERED

As a preliminary to the work discussed in this paper, a number of desired properties were set down as requirements. There were four general criteria imposed:

(1) The method should fit the given points exactly (no curve fitting).

(2) Some bias in the estimates can be allowed providing it is consistent; i.e., the estimation technique should not introduce spurious trends into the data.

(3) Simple and efficient methods are best, if possible.

(4) All the summary measures from the grouped data (quantiles, income shares, Gini ratios, etc.) should be consistent with one another and with income distribution theory (i.e., the distribution functions and Lorenz curves obtained should always be nondecreasing).

Since the entire historical series to be updated comes from the Current Population Survey (CPS), several more criteria were imposed that were tailored specifically to that survey:

(5) The data should be "smoothed" somewhat to allow for rounding in the CPS [2].

(6) Because of the widths of the upper income intervals, methods consistent with the theory of income distribution [e.g., 3] are preferable.

(7) The method should be able to handle unusually shaped income distributions; e.g., the method will be used for doctors and surgeons, as well as for service workers.

(8) Since the mean incomes per income interval generally are unavailable for the major portion of the series, the method has to be one which does not depend on this information.

Some of the best known interpolation procedures for income data are precluded by these requirements. In particular, the techniques suggested by Gastwirth-Glauberman [4] and Budd [5] both employ knowledge of the mean income in each interval. A number of general purpose interpolation techniques, unless modified, also lack one or more of the above properties. Two, for instance, that we examined and which proved unsatisfactory were cubic spline interpolation [6] and Akima's method of Local Procedures [7,8]. 1/

From a companion paper by Oh [9] we did have available a general purpose interpolation scheme, Karup-King osculatory interpolation, which had been modified to handle income data. 2/ In the next section we will compare Oh's procedure with the combination of Pareto and linear interpolation we suggest here. The Hermite interpolation technique advocated by Gastwirth-Glauberman will also be considered, even though it cannot always be used in the CPS.

2. ESTIMATING INCOME QUANTILES IN THE CPS

In this section we will examine three different methods for estimating income quantiles from the CPS. The three methods are--

(1) Actual quantiles--The "actual" quantiles from the ungrouped CPS data were calculated by sorting the CPS microdata files and picking the income representing each of the quantiles selected for comparison (i.e., the 20th, 60th, 80th and 95th percentiles). This was done separately for families (table 1) and unrelated individuals (table 2) for each income year 1958-1974.

(2) Pareto-linear--The Pareto-linear estimates were developed assuming uniform distributions in the lower income intervals and Pareto distributions in the upper income intervals. The starting point of the calculations was annual Census Bureau CPS income reports. Each interval was interpolated separately. Pareto interpolation was used whenever the absolute value of Pareto's slope parameter was greater than 1. Usually this condition occurred in the income intervals above the median. The absolute value of this parameter is generally greater than 2, in the top interval, and decreases as income decreases. Pareto interpolation could have been used

even after the parameter became less than one; however, we did not use it, because the estimates derived from Pareto interpolation were frequently less accurate than those derived from linear interpolation.

(3) Karup-King osculatory interpolation--The third method used was Karup-King osculatory interpolation, modified as necessary for use with income data [9]. For the comparisons in this paper, we first converted the income and frequency information to a log scale, in order to better graduate the distributions in the longer intervals in the upper tail. Basically, the procedure consisted of deriving the cumulative distribution function in the interval [b, c) by examining the interval just before it, say [a, b), and just after it, say [c, d). Two quadratic equations were then fit through the points $\{a, b, c\}$ and $\{b, c, d\}$. These two quadratic equations were then weighted in such a way as to force a smooth nondecreasing cumulative distribution through b and c. Moreover, the procedure had to fit a, b, c, and d exactly. An extra point was provided in the top open-end interval by fitting a Pareto distribution to the interval preceeding the open interval and estimating the frequency above $100,000.

Now that we have outlined the three methods to be looked at, it is appropriate to turn to the actual (numerical) comparisons in tables 1 and 2.3/ Several observations are possible:

(1) Relatively speaking, income quantiles can be more accurately estimated for families than for unrelated individuals (i.e., both interpolation procedures tend to be relatively closer to the ungrouped data for families than for unrelated individuals).

(2) The pattern of accuracy is also different for families than for unrelated individuals. For families, the data are better for the lower quantiles than for upper quantiles, while the reverse is true for unrelated individuals. Undoubtedly, this pattern occurs for families because the income intervals used to calculate higher quantiles are much broader than for lower quantiles. However, for unrelated individuals, lower quantiles fall in the extreme bottom intervals, where the size of the interval is still large relative to the magnitude of the estimate being attempted.

(3) The CPS data follow the Pareto law rather closely in the upper tail of the income distribution, as has been mentioned, especially if one fits the CPS to a Pareto which can change from interval to interval, as is done here. This is one of the main reasons the Pareto-linear interpolation works so well.

(4) The Pareto-linear procedure seems to provide more accurate measures more of the time than does the Karup-King. This was in some sense unexpected because Karup-King, as employed in this paper, essentially represents a refinement to a simple log-log (Pareto type) interpolation procedure. I suspect that the Karup-King might have been better had we accumulated the data from higher intervals to lower intervals and then applied the osculatory interpolation formulas.

3. LORENZ CURVE ESTIMATION IN THE CPS

We now turn from the interpolation of income quantiles to obtaining selected Lorenz curve measures (income shares and Gini ratios). Again, we will make comparisons (in tables 3 and 4) between three methods:

(1) Actual values--For each year we calculated the aggregate income received by each percentile of the population. This was done separately for families and unrelated individuals from CPS microdata files sorted by amount of income. In table 3 we look at just families over the period 1967-1974 so as to be consistent with [4]. In table 4 we examine the entire time series.4/

(2) Pareto-linear--To obtain Lorenz curve values using this method, the aggregate income in each size class had to be derived. We did this by assuming Pareto distributions in each income interval for the higher intervals and assuming a uniform distribution in the lower intervals. The same decision rule as before was used for switching from one method to the other. In the top open-end interval, the frequency with income above $100,000 was estimated from a Pareto distribution fitted to the previous interval. An assumed mean of $100,000 was assigned to units with income over $100,000. The closed interval form of the Pareto mean income estimation formula was used for the remaining units in the open-end interval (see [10] for full details). The Gini index was estimated by splitting the given Lorenz curve into 100 intervals, each of one percent, and using Simpson's rule for approximate integration.

(3) Hermite--Gastwirth and Glauberman [4] employed Hermite interpolation to develop Lorenz curve measures from the CPS for the years 1967-1974. We have reproduced these here, in part, because Karup-King estimates were not available in time for the presentation at the session.

At least two overall observations seem in order for the comparisons in the tables:

(1) The Hermite interpolation procedures of Gastwirth and Glauberman assume that mean income per income interval is known. For this reason,we expected their estimates to be better than the Pareto-linear ones,since,for the latter,the actual means in each interval are not used. However, the results seem to indicate that the Pareto-linear method is slightly more accurate than Gastwirth-Glauberman's. I suspect, though, that data for all families do not represent an adequate test. It is my opinion that Hermite interpolation might be better than Pareto-linear for unrelated individuals or for race data.

(2) For Gini indexes, the Pareto-linear differs from the ungrouped data by, at most, .004, while Gastwirth-Glauberman differs by, at most, .006. Both methods tend to underestimate the Gini index slightly. However, neither method appears to introduce a spurious trend. Similar closeness to ungrouped data is indicated for shares of aggregate income.

4. CONCLUSIONS

This paper examines several methods for estimating summary measures of income distributions from grouped data. Of those considered in detail, it would seem that the Pareto-linear is best suited for our application to the Current Population Survey historical income series. The advantage of the method grows when one considers its simplicity and ease of use. In fact, the Census Bureau has adopted Pareto interpolation for calculating published CPS medians when these fall in intervals of more than $1,000 in length. This will be fully implemented for the annual 1976, series P-60, income report.

ACKNOWLEDGEMENTS AND FOOTNOTES

The author would like to thank Fritz Scheuren and H. Lock Oh of the Social Security Administration for their many helpful comments and suggestions. Thanks are also due to Professor Gastwirth of the George Washington University for his stimulating discussion of an earlier draft of this paper. Editorial assistance was provided by Mary Henson and Gordon Green. The typing was done by Rubye Ellis.

1/ It is possible that we were not patient enough in applying these methods; even as trivial a modification as converting to logs before interpolating may well have yielded acceptable results. However, given the comparisons made with Karup-King Osculatory Interpolation [9], we suspect that these methods would generally not be better than the simpler (Pareto-linear) technique actually adopted.

2/ Oh's procedure satisfies all our requirements with the exception of perhaps number 6.

3/ Very little work has been done so far to estimate the standard errors of the differences among the several interpolation methods presented. Sampling error is not, however, likely to be a serious limitation on the comparisons in the tables, since each of the methods was applied in turn to exactly the same data sets, the March CPS's from 1959 to 1975 (i.e., income years 1958-1974, respectively).

4/ This paper does not represent the first appearance of these ungrouped figures in print. Most of them were originally prepared by me several years ago and published in Series P-60 beginning with report No. 90.

REFERENCES

[1] U.S. Bureau of the Census, "Trends in the Income of Families and Persons in the U.S., 1947-1964," Technical Paper No. 17.
U.S. Bureau of the Census, "Trends in the Income of Families and Persons, 1947-1959," Technical Paper No. 8.

[2] Knott, J., "An Analysis of the Effect of Income Rounding in the Current Population Survey," 1971 American Statistical Association Proceedings, Social Statistics Section, 1972.

[3] Bjerke, K., "Income and Wage Distributions - Part I: A survey of the literature," Review of Income and Wealth, Series 16, No. 3 (Sept. 1970), pp. 235-252.

[4] Gastwirth, J.L., and Glauberman, M., "The Interpolation of the Lorenz Curve and Gini Index from Grouped Data," Econometrica, Vol. 44, pp. 479-483, May, 1976.

[5] Budd, E.C., "Postwar Changes in the Size Distribution of Income in the U.S.," American Economic Review, 60 (May 1970), pp. 247-260.

[6] Sperry-Rand Corporation, Univac Division, Math-Pack, Programmers Reference, UP7542, 1967, section 2, pp. 68-74.

[7] Akima, H., "A New Method of Interpolation and Smooth Curve Fitting Based on Local Procedures," Journal of the Association for Computing Machinery, Vol. 17, No. 4 (October, 1970), pp. 589-603.

[8] Akima, H., "Interpolation and Smooth Curve Fitting Based on Local Procedures," Collected Algorithms from Communications of the Association for Computing Machinery, Algorithm 433, March 1, 1972.

[9] Oh, H.L., "Osculatory Interpolation with a Monotonicity Constraint," 1977 American Statistical Association Proceedings, Statistical Computation Section.

[10] Spiers, E., "Some Notes on the Derivation of Computation Formulas Assuming a Pareto Distribution," (Unpublished Working Paper), 1976.

Table 1.--COMPARISON OF DATA ON SELECTED INCOME QUANTILES FOR FAMILIES BY TOTAL MONEY INCOME IN 1958 TO 1974, BY TYPE OF ESTIMATION METHOD, FOR THE UNITED STATES

Year	Ungrouped Data (1)	Pareto-Linear (2)	Percent Difference $\frac{(2)-(1)}{(1)}$x100 (3)	Karup-King Log-Log Scale (4)	Percent Difference $\frac{(4)-(1)}{(1)}$x100 (5)	Ungrouped Data (6)	Pareto-Linear (7)	Percent Difference $\frac{(7)-(6)}{(6)}$x100 (8)	Karup-King Log-Log Scale (9)	Percent Difference $\frac{(9)-(6)}{(6)}$x100 (10)
	TWENTIETH PERCENTILE					SIXTIETH PERCENTILE				
1974	$6,500	$6,551	0.8	$6,552	0.8	$14,916	$14,944	0.2	$14,948	0.2
1973	6,081	6,141	1.0	6,143	1.0	14,000	13,883	-0.8	14,012	0.1
1972	5,612	5,668	1.0	5,671	1.1	12,855	12,816	-0.3	12,932	0.6
1971	5,211	5,275	1.2	5,277	1.3	11,826	11,850	0.2	11,873	0.4
1970	5,100	5,148	0.9	5,154	1.1	11,299	11,337	0.3	11,404	0.9
1969	5,000	5,005	0.1	5,005	0.1	10,800	10,799	-	10,883	0.8
1968	4,544	4,598	1.2	4,610	1.5	9,960	9,968	0.1	9,970	0.1
1967	4,097	4,164	1.6	4,172	1.8	9,000	9,129	1.4	9,137	1.5
1966	3,935	3,950	0.4	3,951	0.4	8,563	8,644	0.9	8,664	1.2
1965	3,500	3,508	0.2	3,508	0.2	7,910	7,982	0.9	7,991	1.0
1964	3,250	3,288	1.2	3,288	1.2	7,500	7,574	1.0	7,601	1.3
1963	3,096	3,150	1.7	3,156	1.9	7,134	7,223	1.2	7,244	1.5
1962	3,000	3,018	0.6	3,019	0.6	6,800	6,851	0.8	6,863	0.9
1961	2,800	2,827	1.0	2,831	1.1	6,560	6,631	1.1	6,663	1.6
1960	2,784	2,795	0.4	2,799	0.5	6,364	6,423	0.9	6,451	1.4
1959	2,677	2,715	1.4	2,719	1.6	6,081	6,176	1.6	6,194	1.9
1958	2,530	2,558	1.1	2,556	1.0	5,720	5,774	0.9	5,817	1.7
Average absolute % difference			0.9		1.0			0.7		1.0
Maximum % diff			1.7		1.9			1.4		1.9
Number times better			9		1			14		1
Maximum less minimum			1.6		1.8			2.4		1.8
Year	EIGHTIETH PERCENTILE					NINETY-FIFTH PERCENTILE				
1974	$20,445	$19,894	-2.7	$20,968	2.6	$31,948	$31,957	-	$30,562	-4.3
1973	19,253	18,658	-3.1	19,596	1.8	30,015	30,296	0.9	28,970	-3.5
1972	17,760	17,418	-1.9	18,058	1.7	27,836	28,152	1.1	27,072	-2.7
1971	16,218	16,119	-0.6	16,370	0.9	25,325	25,520	0.8	25,310	-0.1
1970	15,531	15,538	-	15,633	0.7	24,250	24,342	0.4	24,597	1.4
1969	14,751	14,783	0.2	14,815	0.4	22,703	22,757	0.2	23,435	3.2
1968	13,400	13,434	0.3	13,556	1.2	20,590	20,664	0.4	21,168	2.8
1967	12,270	12,395	1.0	12,432	1.3	19,025	19,171	0.8	19,124	0.5
1966	11,640	11,721	0.7	11,743	0.9	18,000	18,297	1.7	17,858	-0.8
1965	10,800	10,876	0.7	10,948	1.4	16,695	17,071	2.3	16,806	0.7
1964	10,201	10,415	2.1	10,465	2.6	15,788	16,088	1.9	15,924	0.9
1963	9,969	9,980	0.1	9,981	0.1	15,144	15,400	1.7	15,315	1.1
1962	9,500	9,504	-	9,558	0.6	14,900	14,928	0.2	14,950	0.3
1961	9,035	9,120	0.9	9,169	1.5	14,600	14,676	0.5	14,756	1.1
1960	8,800	8,796	-	8,849	0.6	13,536	13,756	1.6	13,983	3.3
1959	8,380	8,393	0.2	8,424	0.5	12,800	13,057	2.0	13,255	3.6
1958	7,800	7,776	-0.3	7,864	0.8	12,000	12,165	1.4	12,206	1.7
Average absolute % difference			0.9		1.2			1.1		1.9
Maximum % diff			-3.1		2.6			2.3		3.6
Number times better			13		-			11		6
Maximum less minimum			5.2		2.5			2.3		7.1

- Rounds to zero.

SOURCE: CURRENT POPULATION SURVEY
U.S. BUREAU OF THE CENSUS

Table 2.--COMPARISON OF DATA ON SELECTED INCOME QUANTILES FOR UNRELATED INDIVIDUALS BY TOTAL MONEY INCOME IN 1958 TO 1974, BY TYPE OF ESTIMATION METHOD, FOR THE UNITED STATES

Year	Ungrouped Data (1)	Pareto-Linear (2)	Percent Difference $\frac{(2)-(1)}{(1)}$x100 (3)	Karup-King Log-Log Scale (4)	Percent Difference $\frac{(4)-(1)}{(1)}$x100 (5)	Ungrouped Data (6)	Pareto-Linear (7)	Percent Difference $\frac{(7)-(6)}{(6)}$x100 (8)	Karup-King Log-Log Scale (9)	Percent Difference $\frac{(9)-(6)}{(6)}$x100 (10)
	TWENTIETH PERCENTILE					SIXTIETH PERCENTILE				
1974	$2,095	$2,120	1.2	$2,124	1.4	$5,636	$5,749	2.0	$5,737	1.8
1973	1,872	1,883	0.6	1,891	1.0	5,160	5,242	1.6	5,251	1.8
1972	1,596	1,604	0.5	1,608	0.8	4,660	4,698	0.8	4,680	0.4
1971	1,461	1,472	0.8	1,473	0.8	4,332	4,422	2.1	4,401	1.6
1970	1,368	1,361	-0.5	1,366	-0.1	4,100	4,191	2.2	4,174	1.8
1969	1,235	1,247	1.0	1,255	1.6	3,895	3,906	0.3	3,900	0.1
1968	1,180	1,185	0.4	1,193	1.1	3,600	3,667	1.9	3,651	1.4
1967	1,000	1,015	1.5	1,016	1.6	3,128	3,249	3.9	3,240	3.6
1966	998	998	-	998	-	3,000	3,095	3.2	3,108	3.6
1965	900	870	-3.3	856	-4.9	2,995	2,995	-	2,995	-
1964	839	769	-8.3	748	-10.8	2,654	2,740	3.2	2,727	2.8
1963	792	709	-10.5	685	-13.5	2,400	2,421	0.9	2,407	0.3
1962	775	749	-3.4	756	-2.5	2,340	2,367	1.2	2,350	0.4
1961	695	664	-4.5	664	-4.5	2,340	2,379	1.7	2,364	1.0
1960	650	644	-0.9	645	-0.8	2,400	2,408	0.3	2,399	-
1959	600	568	-5.3	568	-5.3	2,080	2,148	3.3	2,136	2.7
1958	550	559	1.6	559	1.6	2,040	2,128	4.3	2,121	4.0
Average absolute % difference. .			2.6		3.1			1.9		1.6
Maximum % diff . .			-10.5		-13.5			4.3		4.0
Number times better			9		3			2		14
Maximum less minimum			12.1		15.1			4.3		4.0
Year	EIGHTIETH PERCENTILE					NINETY-FIFTH PERCENTILE				
1974	$9,296	$9,384	0.9	$9,395	1.1	$15,658	$15,849	1.2	$15,815	1.0
1973	8,802	8,853	0.6	8,860	0.7	15,000	15,216	1.4	15,192	1.3
1972	8,000	8,045	0.6	8,050	0.6	13,500	13,710	1.6	13,775	2.0
1971	7,500	7,528	0.4	7,555	0.7	12,900	12,918	0.1	12,953	0.4
1970	7,200	7,254	0.8	7,281	1.1	12,270	12,435	1.3	12,428	1.3
1969	6,635	6,717	1.2	6,743	1.6	11,800	11,909	0.9	11,917	1.0
1968	6,250	6,375	2.0	6,405	2.5	10,770	10,937	1.6	10,990	2.0
1967	5,593	5,727	2.4	5,756	2.9	9,840	9,925	0.9	9,928	0.9
1966	5,200	5,320	2.3	5,350	2.9	9,200	9,352	1.7	9,372	1.9
1965	5,101	5,260	3.1	5,297	3.8	8,727	8,842	1.3	8,847	1.4
1964	4,996	4,997	-	4,998	-	8,160	8,338	2.2	8,343	2.2
1963	4,675	4,710	0.7	4,748	1.6	8,000	8,074	0.9	8,076	1.0
1962	4,560	4,603	0.9	4,615	1.2	7,800	7,824	0.3	7,834	0.4
1961	4,300	4,373	1.7	4,382	1.9	7,200	7,315	1.6	7,280	1.1
1960	4,181	4,261	1.9	4,277	2.3	6,611	6,753	2.1	6,761	2.3
1959	3,891	3,929	1.0	3,936	1.2	6,492	6,597	1.6	6,605	1.7
1958	3,800	3,836	0.9	3,848	1.3	6,300	6,481	2.9	6,495	3.1
Average absolute % difference. .			1.3		1.6			1.4		1.5
Maximum % diff . .			3.1		3.8			2.9		3.1
Number times better			15		-			11		3
Maximum less minimum			3.1		3.8			2.8		2.2

- Rounds to zero.

SOURCE: CURRENT POPULATION SURVEY
U.S. Bureau of the Census

Table 3.--GINI INDEXES AND SELECTED PERCENTAGE SHARES OF AGGREGATE MONEY INCOME IN 1967 TO 1972, FOR ALL FAMILIES, BY TYPE OF ESTIMATION METHOD, FOR THE UNITED STATES

Year	Un-grouped Data (1)	Pareto-Linear (2)	Difference (2)-(1) (3)	Hermite[1] (4)	Difference (4)-(1) (5)	Un-grouped Data (6)	Pareto-Linear (7)	Difference (7)-(6) (8)	Hermite[1] (9)	Difference (9)-(6) (10)
	Gini Index					Lowest 20 Percent				
1972..	.360	.357[2]	-.003	.359	-.001	5.4	5.5	0.1	5.6	0.2
1971..	.356	.355[2]	-.001	.352	-.004	5.5	5.5	-	5.7	0.2
1970..	.354	.353[2]	-.001	.349	-.005	5.4	5.5	0.1	5.7	0.3
1969..	.349	.347[2]	-.002	.345	-.004	5.6	5.6	-	5.8	0.2
1968..	.348	.344[2]	-.004	.342	-.006	5.6	5.7	0.1	5.9	0.3
1967..	.348	.347[2]	-.001	.344	-.004	5.5	5.6	0.1	5.8	0.3
Year	60 TO 80 PERCENT					TOP 5 PERCENT				
1972..	23.9	23.8	-0.1	23.7	-0.2	15.9	15.9	-	16.2	0.3
1971..	23.8	23.8	-	23.7	-0.1	15.7	15.9	0.2	15.6	-0.1
1970..	23.8	23.8	-	23.7	-0.1	15.6	15.8	0.2	15.4	-0.2
1969..	23.7	23.7	-	23.7	-	15.6	15.6	-	15.4	-0.2
1968..	23.7	23.8	0.1	23.7	-	15.6	15.3	-0.3	15.4	-0.2
1967..	23.9	23.8	-0.1	23.8	-0.1	15.2	15.3	0.1	15.4	0.2

\- Rounds to zero.

1 Gastwirth and Glauberman, "On the Interpolation of the Lorenz Curve and Gini Index", Unpublished Paper.

2 Gini Index calculated using Simpson's rule for approximate integration after splitting the Lorenz Curve into 100 equal intervals.

SOURCE: CURRENT POPULATION SURVEY
U.S. BUREAU OF THE CENSUS

Table 4.--GINI INDEX AND PERCENTAGE SHARE OF AGGREGATE MONEY INCOME IN 1958 TO 1974 RECEIVED BY THE TOP 5 PERCENT OF FAMILIES AND UNRELATED INDIVIDUALS, FOR THE UNITED STATES

Year	FAMILIES						UNRELATED INDIVIDUALS					
	Gini Index			Top 5 Percent			Gini Index			Top 5 Percent		
	Un-grouped Data	Pareto-Linear[1]	Difference	Un-grouped Data	Pareto-Linear	Difference	Un-grouped Data	Pareto-Linear[1]	Difference	Un-grouped Data	Pareto-Linear	Difference
1974..	.356	.352	-.004	15.3	15.4	0.1	.448	.446	-.002	19.3	19.4	0.1
1973..	.357	.355	-.002	15.5	15.8	0.3	.460	.463	.003	20.0	20.9	0.9
1972..	.360	.357	-.003	15.9	15.9	-	.478	.474	-.004	21.4	21.3	-0.1
1971..	.356	.355	-.001	15.7	15.9	0.2	.473	.471	-.002	20.5	20.5	-
1970..	.354	.353	-.001	15.6	15.8	0.2	.478	.478	-	20.8	20.9	0.1
1969..	.349	.347	-.002	15.6	15.6	-	.481	.478	-.003	20.7	20.6	-0.1
1968..	.348	.344	-.004	15.6	15.3	-0.3	.480	.478	-.002	20.8	20.4	-0.4
1967..	.348	.347	-.001	15.2	15.3	0.1	.490	.491	.001	21.1	21.2	0.1
1966..	.349	.348	-.001	15.6	15.6	-	.484	.488	.004	21.2	21.4	0.2
1965..	.356	.356	-	15.5	15.7	0.2	.486	.487	.001	20.0	20.1	0.1
1964..	.361	.356	-.005	15.9	15.4	-0.5	.512	.508	-.004	22.9	22.3	-0.6
1963..	.362	.359	-.003	15.8	15.6	-0.2	.500	.504	.004	20.1	21.0	0.9
1962..	.362	.362	-	15.7	15.9	0.2	.502	.497	-.005	20.8	21.0	0.2
1961..	.374	.373	-.001	16.6	16.8	0.2	.510	.508	-.002	21.6	22.4	0.8
1960..	.364	.366	.002	15.9	16.6	0.7	.506	.490[2]	-.016[2]	20.2	19.9[2]	-0.3[2]
1959..	.361	.360	-.001	15.9	16.1	0.2	.522	.524	.002	22.1	24.4	2.3
1958..	.354	.354	-	15.4	15.6	0.2	.519	.505	-.014	21.6	21.3	-0.3

\- Rounds to zero.

1 Gini Index calculated using Simpson's rule for approximate integration after splitting the Lorenz Curve into 100 equal intervals.

2 Pareto invalid in top interval. Assumed Pareto Alpha = 2.85.
Mean for $25,000 and over = $37,500, mean for $15,000 to $25,000 = $20,000.

SOURCE: CURRENT POPULATION SURVEY
U.S. BUREAU OF THE CENSUS