THE CHANGING SIZE DISTRIBUTION OF U.S. TRADE UNIONS AND ITS DESCRIPTION BY PARETO'S DISTRIBUTION

John Pencavel

Mainz, June 2012

Between 1974 and 2007, there were 101 fewer labor organizations so that, notwithstanding the drop in membership, the average size of U.S. unions rose:

the number of members per union grew from 114 thousand in 1974 to 180 thousand in 2007.

The changes in the size distribution are linked to the growth of a few very large unions.

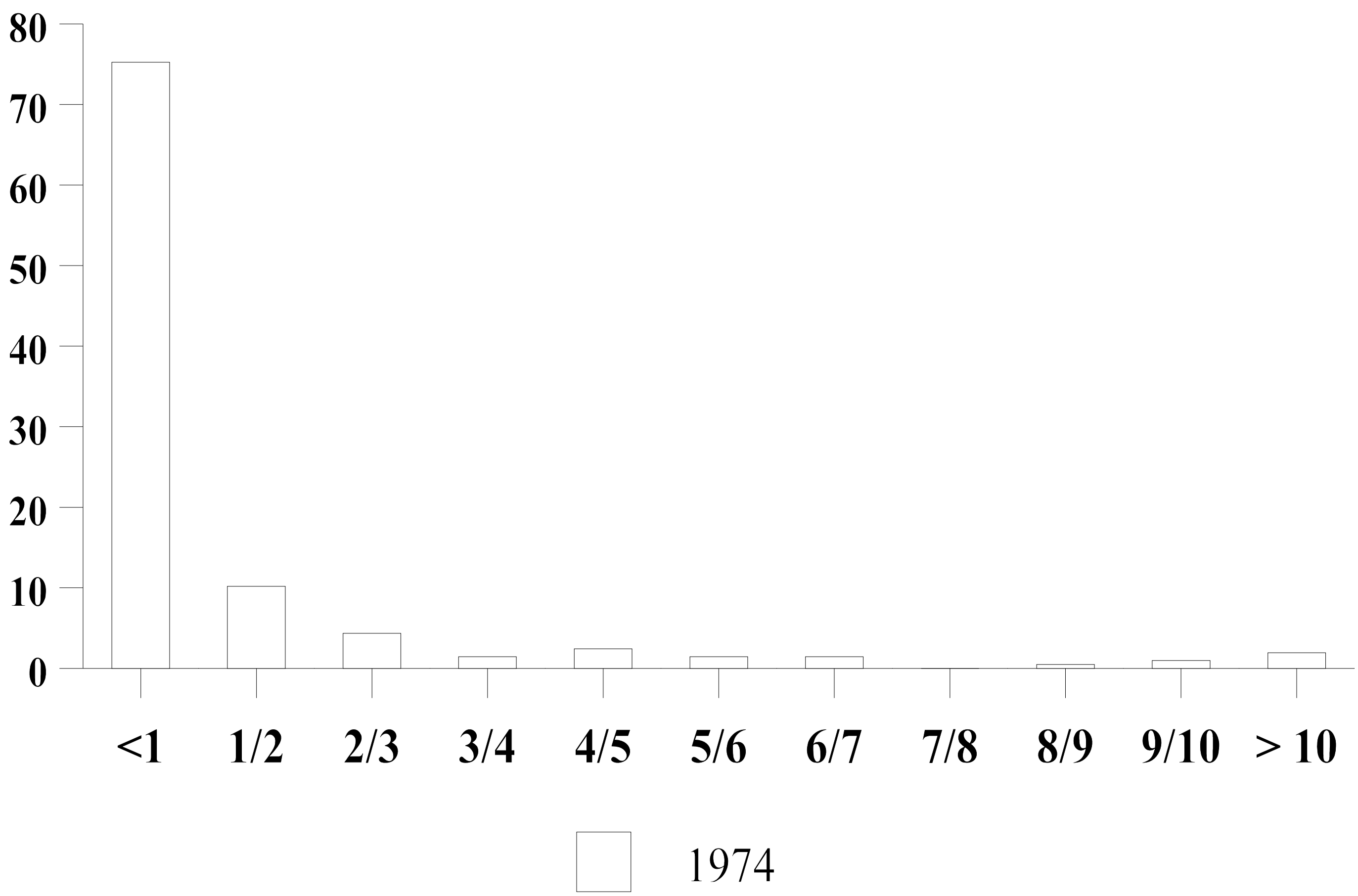
“union” or “labor organization” includes employee associations as well as conventional trade unions

National Unions "The principal locus of political and economic power in the American union movement has long been the national unions" (Rees (1988))

1) The Five Largest Unions in 1974

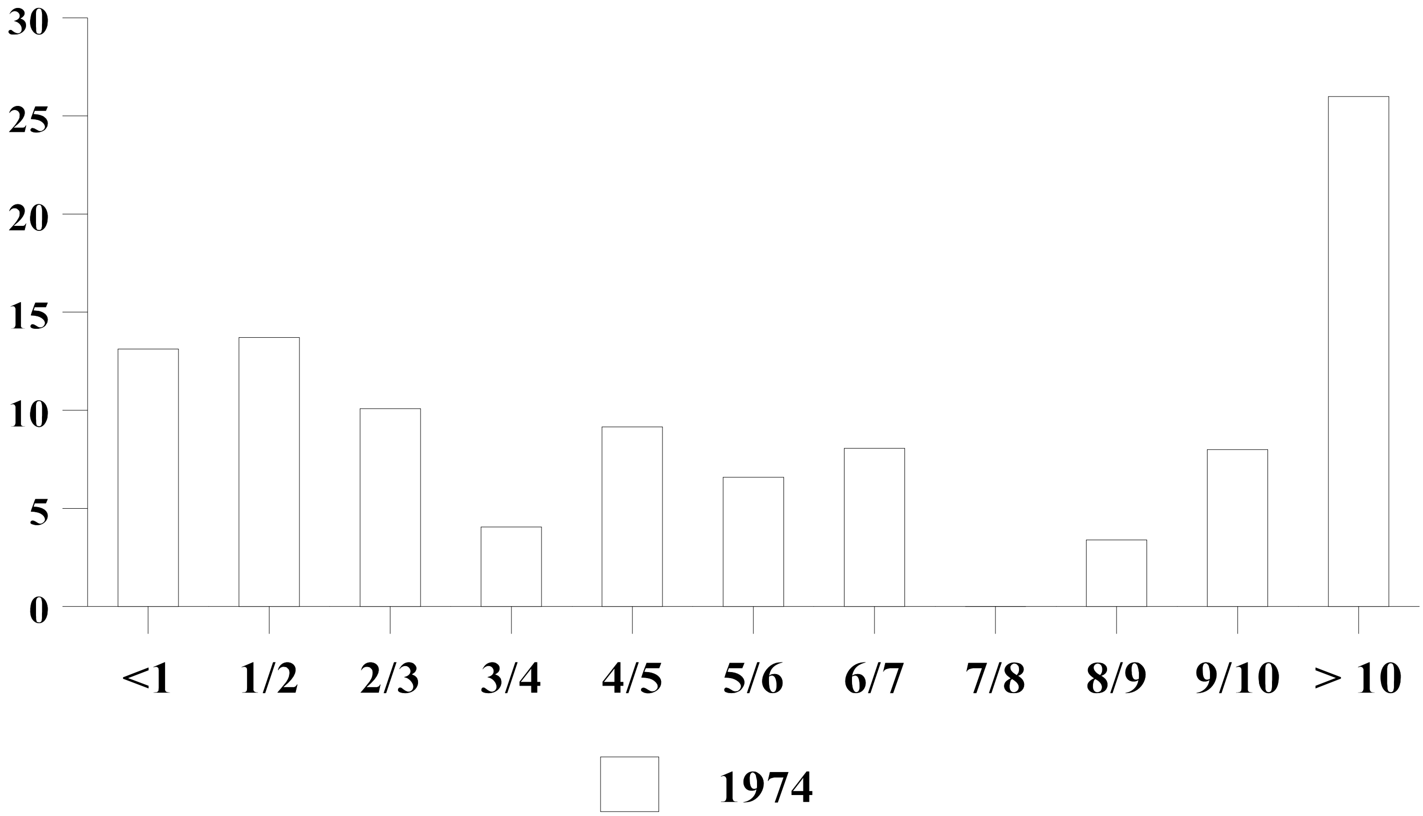
Full Name of Union in 1974	Abbreviation	Membership in Thousands	
		1974	2007
International Brotherhood of Teamsters, Chauffeurs, Warehouse-men & Helpers of America	IBT	1,973.3	1,398.6
International Union of Automobile, Aerospace & Agricultural Implement Workers of America	UAW	1,544.9	538.4
National Education Association	NEA	1,470.2	3,167.6
United Steelworkers of America	USW	1,300.0	730.9
International Brotherhood of Electrical Workers	IBEW	991.2	697.9

2) Percent Distribution of the Number of Unions by Size of Union in 1974: Size Classes in Intervals of One Hundred Thousand Members



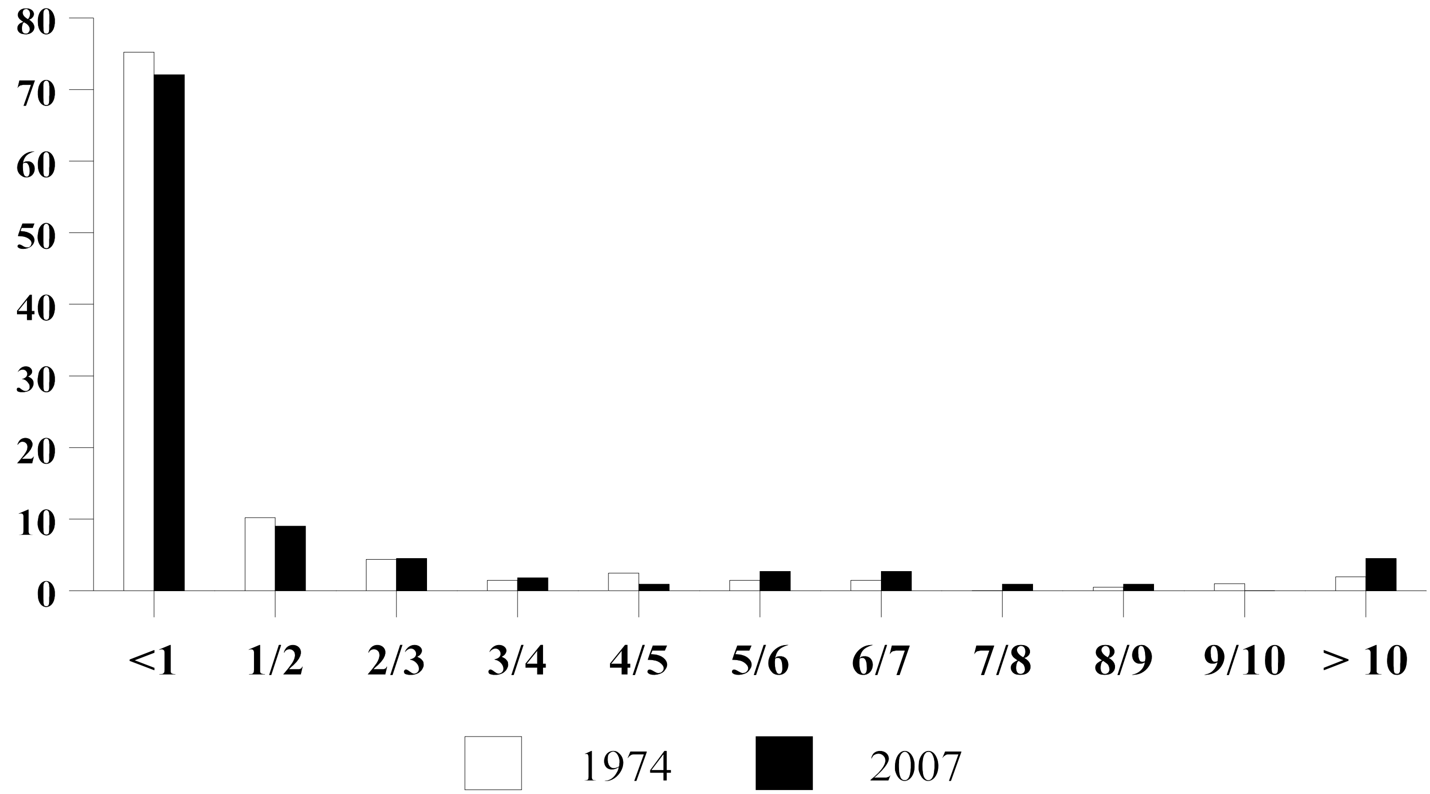
The horizontal axis measures membership in labor organizations by hundred thousand of members. Thus " < 1" means less than 100,000, "1/2" means from 100,000 to 199,999, and so on in 100,000 intervals until the largest class of 1,000,000 or more denoted " > 10 " .
The vertical axis measures the total number of unions in the size class as a percentage of the total number of unions in all size classes.

3) Percent Distribution of Union Membership by Size of Union in 1974: Size Classes in Intervals of One Hundred Thousand Members



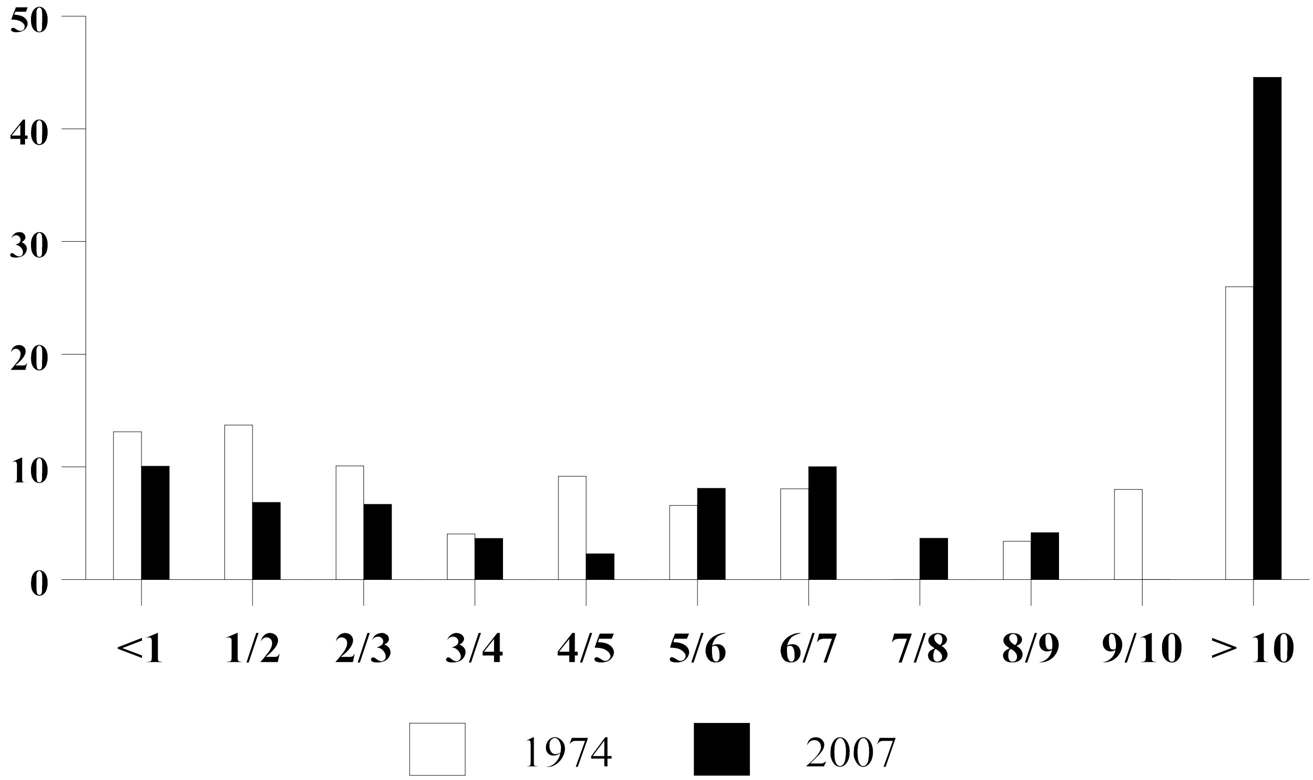
The horizontal axis measures membership in labor organizations by hundred thousand of members. Thus " < 1" means less than 100,000, "1/2" means from 100,000 to 199,999, and so on in 100,000 intervals until the largest class of 1,000,000 or more denoted " > 10 " .
The vertical axis measures the membership in unions in the size class as a percentage of total union membership in all size classes.

4) Percent Distribution of the Number of Unions by Size of Union in 1974 and 2007 : Size Classes in Intervals of One Hundred Thousand Members



The horizontal axis measures membership in labor organizations by hundred thousand of members. Thus < 1 means less than 100,000, 1/2 means from 100,000 to 199,999, and so on. The vertical axis measures the total number of unions in the size class as a percentage of the total number of unions in all size classes. The black columns describe the year 2007 and the white columns describe the year 1974.

5) Percent Distribution of Union Membership by Size of Union in 1974 and 2007: Size Classes in Intervals of One Hundred Thousand Members



The horizontal axis measures membership in labor organizations by hundred thousand of members. Thus “ < 1” means less than 100,000, “1/2” means from 100,000 to 199,999, and so on in 100,000 intervals until the largest class of 1,000,000 or more denoted “ > 10 ” .

The vertical axis measures the membership in unions in the size class as a percentage of total union membership in all size classes.

The black columns describe the year 2007 and the white columns describe the year 1974.

6) The Distribution of Unions by the Number of their Members in 1974 and 2007

	size of union (number of members)	1974		2007	
		percent of all unions	percent of all members	percent of all unions	percent of all members
1	< 1,000	12.7	0.04	9.0	0.02
2	from 1,000 to 9,999	24.1	0.87	22.5	0.50
3	from 10,000 to 99,999	38.2	8.21	40.5	9.54
4	from 100,000 to 499,999	19.3	37.01	16.2	19.45
5	from 500,000 to 999,999	3.8	23.89	7.2	25.93
6	≥ 1,000,000	1.9	25.99	4.5	44.58

7) Changes in the Number of Unions and in the Number of Members (in thousands) between 1974 and 2007 by Size Class

size of union in thousands of members	number in 2007 minus number in 1974	
	number of unions	number of members in thousands
< 1	- 17	- 4.8
from 1 to 4.9	- 20	- 64.8
from 5 to 9.9	- 6	- 46.6
from 10 to 24.9	- 13	- 187.0
from 25 to 49.9	- 19	- 577.1
from 50 to 99.9	- 4	- 282.2
from 100 to 199.9	- 13	- 1,949.6
from 200 to 299.9	- 5	- 1,107
from 300 to 399.9	- 1	- 251.4
from 400 to 499.9	- 4	- 1,756.1
from 500 to 999.9	0	- 593.4
≥ 1,000	+ 1	+ 2,627.8
Total	- 101	- 4,195.1

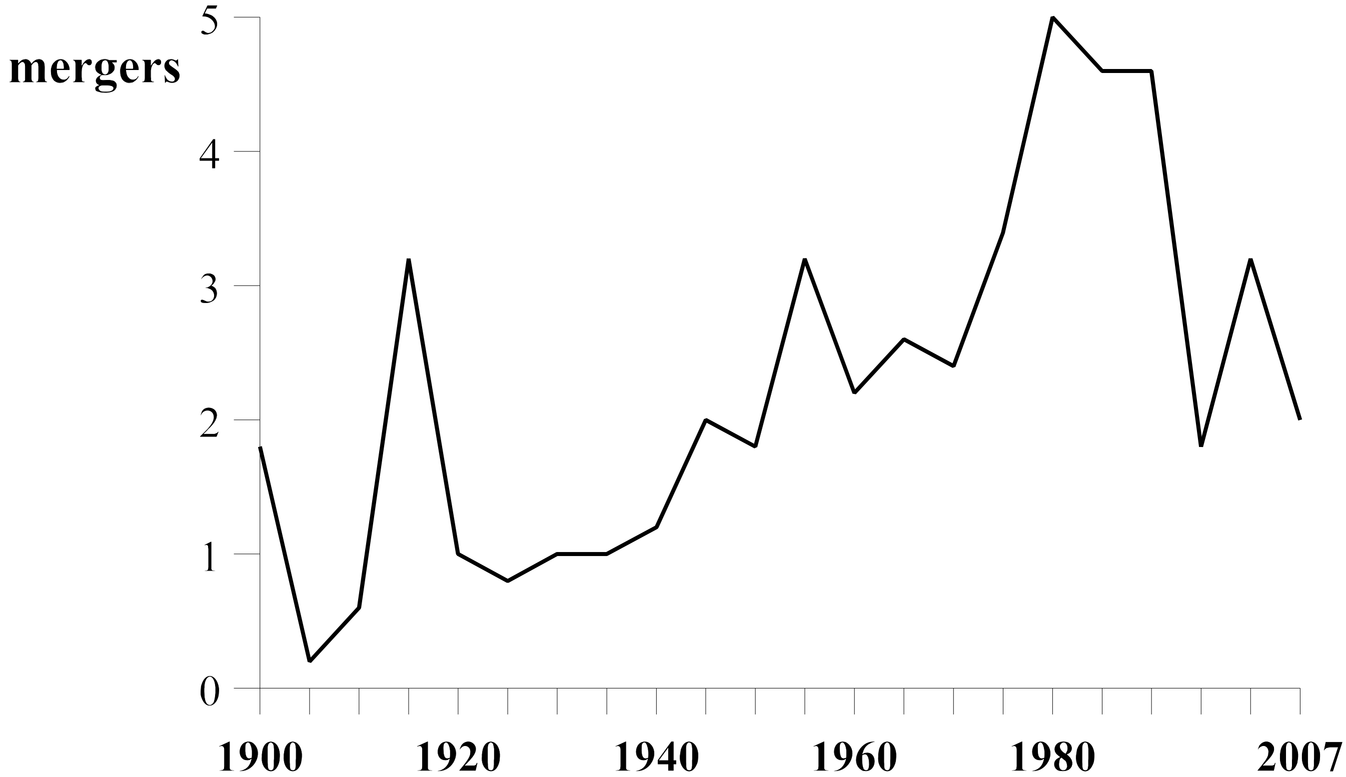
In 1974, those labor organizations with half-a-million members or more made up fewer than 6 percent of all unions but 50 percent of all union members belonged to these unions.

In 1974, three-quarters of all U.S. unions consisted of those where each had fewer than 10,000 members but merely 10 percent of all union members were members of these unions.

By 2007, more than 70 percent of all union members belonged to organizations each of whose membership consisted of half-a-million members or more. The five organizations with one million or more members accounted for 45 percent of all members. Yet these labor organizations each with a million or more members represented less than 5 percent of all labor organizations.

In 2007, seventy-two percent of unions consisted of organizations where each had less than 100,000 members and yet this 72 percent of unions had in aggregate only 10 percent of all members.

8) Average Number of Mergers of Labor Unions per Year, 1900-2007



9) The Five Largest Unions in 2007

Full Name of Union in 2007	Abbreviation	Membership in Thousands	
		in 1974	in 2007
National Education Association	NEA	1,470.2	3,167.6
Service Employees International Union	SEIU	550.0	1,575.5
American Federation of State County, & Municipal Employees	AFSCME	648.2	1,470.1
International Brotherhood of Teamsters	IBT	1,973.3	1,398.6
United Food & Commercial Workers International Union	UFCW	1,175.9	1,304.1

10) Percent of All Union Members Who Are Members of the Ten, Five, and Two Largest Unions, Selected Years 1920 - 2007

year	membership of the ten largest unions as a percent of total union membership	membership of the five largest unions as a percent of total union membership	membership of the two largest unions as a percent of total union membership
1920	43.85	28.70	14.98
1939	36.84	17.88	14.29
1968	43.82	28.79	14.72
1974	45.43	30.09	14.54
1978	47.13	30.46	14.85
1983	48.10	31.60	14.70
2007	62.40	44.58	23.71

Pareto’s Distribution

The concentration of the union structure and the growth of the largest unions may be taken up in a less impressionistic and more organized fashion if the distribution in the sizes of unions followed a compact systematic pattern.......

the lognormal distribution

Pareto’s distribution

high incomes, the size distribution of firms, the size distribution of large cities

Consider the case for the same relationship to describe the membership of large unions.

Start with the data on union membership in 2007.

Union Membership in 2007

Suppose the threshold for inclusion in the set of "large" unions in 2007 are those with at least 200,000 members. There are 21 labor organizations in this set. These unions constituted almost a fifth (precisely18.9 percent) of all the unions and over four-fifths (precisely 83.1 percent) of all the reported membership.

Let M_i denote the number of members belonging to union i and suppose S_i is the percentage of unions with membership greater than union i's membership. The values of S_i constitute the survivor function and it is the complement of the cumulative distribution function of membership. Pareto's Law maintains that, beyond a certain threshold (initially, a level corresponding to membership of 200,000 in 2007), the survivor function of union membership is $S(M_i) = e^{\lambda} (M_i)^{-\alpha}$ or, in logarithms,

(1) $$\ln S_i = \lambda - \alpha \ln M_i \quad ,$$

where $\lambda > 0$ and $\alpha > 0$ are parameters. α is Pareto's coefficient.

The special case of $\alpha = 1$ is Zipf's Law or the "rank-size" rule.

In the study of the distribution of high incomes, α is often estimated between 1.5 and 2.5;

in its application to the size distribution of large firms, α is estimated between 0.89 and 1.06; and

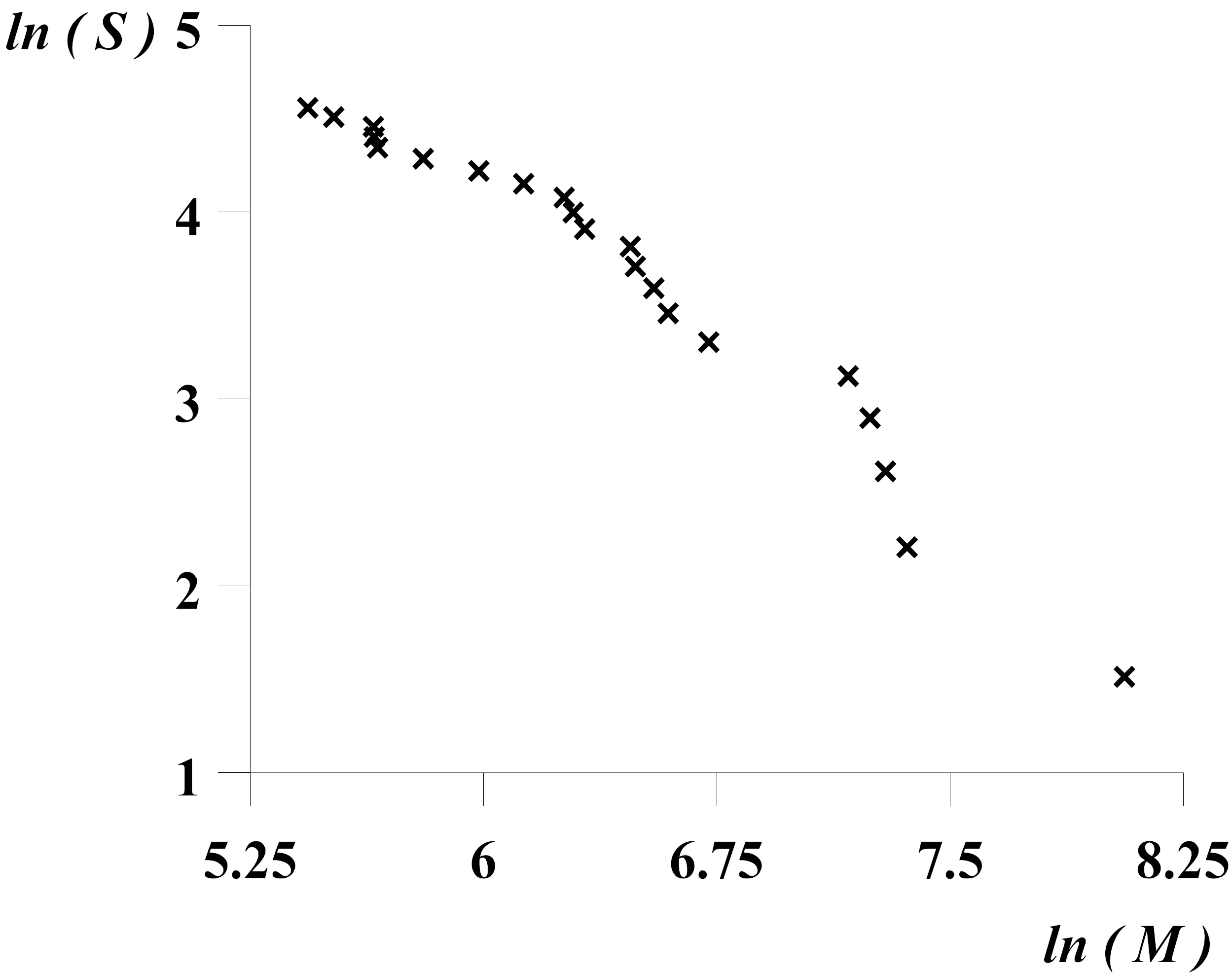
in its application to the distribution of large cities in a country, α is estimated between 0.73 and 1.96.

The value of α provides an indicator of the degree of concentration of the values of the variable among those observations to which Pareto's distribution is applied: higher values of α denote less concentration because the larger the value of α, the steeper the decline of the survivor function and the larger the range of values of the survivor function for a given difference in union membership.

A useful attribute of Pareto's distribution is that the average membership of those unions equal to and above a minimum value, say, M_N, is $\alpha (\alpha - 1)^{-1} M_N$.

Or the ratio of the average membership to M_N is $\alpha (\alpha - 1)^{-1}$.

Pareto's Law for Labor Unions with ≥ 200,000 Members in 2007 ?



The negative relationship between $ln\ S_i$ and $ln\ M_i$ does not appear to be linear although are the deviations from linearity random and inconsequential? To what extent is this relationship satisfactorily described as linear? Is this a good fit by some criterion? Are the deviations from the fitted relationship randomly distributed? These issues will be taken up by estimating the unknown parameters of Pareto's distribution by the method of least-squares.

Random departures from precise log-linearity may be accommodated by adding a stochastic term:

(1.1) $$ln\, S_i = a_0 + b_0\, ln\, M_i + e_{0\,i}$$

where a_0 and b_0 are parameters to be estimated with the interpretation that $a_0 = \lambda$ and $b_0 = -\alpha$ and where $e_{0\,i}$ represents the value of i's residual from the fitted line.

Estimates of Equation (1.1) to Largest Unions in 2007

	size of unions	number of unions	estimated a_0 (s.e.)	estimated b_0 (s.e.)	implied α	R^2
1	> 200,000	21	10.894 (0.395)	-1.126 (0.061)	1.13	0.947
2	> 100,000	31	8.890 (0.284)	-0.883 (0.047)	0.88	0.923
3	> 50,000	46	7.873 (0.167)	-0.784 (0.030)	0.78	0.938
4	> 25,000	60	7.239 (0.120)	-0.723 (0.023)	0.72	0.943
5	> 10,000	76	6.407 (0.101)	-0.615 (0.021)	0.62	0.919

As the membership threshold is lowered, so Pareto's coefficient is estimated to be lower ⇒

the concentration of union members into the largest unions appears more salient as smaller and smaller unions are added to the data

Pareto's coefficient is sensitive to the set of observations to which it is applied and, in particular, to the threshold value that defines inclusion into the set.

When the estimated residuals from this regression are ordered by the size of the union, the null hypothesis that they display no first-order serial correlation is rejected at conventional levels of significance by familiar tests.

Are different inferences about Pareto's coefficient would follow from applying some procedure that recognizes this pattern in the fitted residuals. For this, the residuals were assumed to follow a first-order autoregressive pattern and Cochrane and Orcutt's familiar iterative technique was applied to equation (1.1). Using the same 21 largest unions in 2007, the consequences for the estimated parameters are shown in Table 10. The estimates of Pareto's coefficient are similar - though not identical - to those that do not recognize such serial correlation.

Estimates of Equation (1.1) After Allowing for First-Order Serial Correlation in the Residuals

		CORC		OLS	
	size of unions	estimated b_0 (s.e.)	implied α	estimated b_0 (s.e.)	implied α
1	> 200,000	-1.149 (0.098)	1.15	-1.126 (0.061)	1.13
2	> 100,000	-0.967 (0.090)	0.97	-0.883 (0.047)	0.88
3	> 50,000	-0.890 (0.069)	0.89	-0.784 (0.030)	0.78
4	> 25,000	-0.852 (0.055)	0.85	-0.723 (0.023)	0.72
5	> 10,000	-0.795 (0.054)	0.80	-0.615 (0.021)	0.62

Perhaps the serial correlation of the residuals indicates functional form mis-specification. Fit a non-Pareto function:

(3.1) $$\ln S_i = a_2 + b_2 \ln M_i + c_2 (\ln M_i)^2 + e_{2i} \quad ,$$

Estimates of Equation (3.1) to the Largest Unions in 2007

$$(3.1) \qquad \ln S_i = a_2 + b_2 \ln M_i + c_2 (\ln M_i)^2 + e_{2i} \quad ,$$

	size of unions	number of unions	estimated a_2 (s.e.)	estimated b_2 (s.e.)	estimated c_2 (se)	R^2
1	> 200,000	21	0.331 (2.401)	2.112 (0.732)	-0.245 (0.055)	0.975
2	> 100,000	31	0.344 (0.862)	1.987 (0.287)	-0.235 (0.023)	0.983
3	> 50,000	46	2.666 (0.480)	1.117 (0.173)	-0.167 (0.015)	0.984
4	> 25,000	60	3.744 (0.290)	0.655 (0.112)	-0.128 (0.010)	0.984
5	> 10,000	76	4.086 (0.125)	0.440 (0.055)	-0.109 (0.006)	0.987

With $b_2 > 0$ and $c_2 < 0$, the relationship is concave from below. Although the relevance of the second-order term constitutes a strict rejection of log-linearity, a less drastic reaction to this result is to view Pareto's log-linearity as a useful first-order approximation to the relationship and not to discard Pareto's rule unreservedly.

Union Membership in Earlier Years

Having estimated Pareto's distribution to the 21 unions with at least 200,000 members in 2007, consider how these estimates contrast with those fitted to a comparable set of unions in other years and consider what these estimates in other years imply.

What constitutes a "comparable" set of unions?

estimate Pareto's distribution to large unions in earlier years by specifying

(i) the same number of unions in these years as in 2007 (namely, the 21 largest unions) or
(ii) the same minimum threshold level of union members as in 2007 (namely, 200,000 members) or
(iii) the same percentage of all unions as in 2007 (namely, the largest 18.9 percent of all unions).

These three selection criteria result in a different number of unions being defined as "large" in a given year.

Estimates of $ln\, S_i = a_0 + b_0\, ln\, M_i + e_{0i}$

year	21 largest unions		unions with ≥ 200,000 members		largest 18.9% of unions	
	implied α	R^2	implied α	R^2	implied α	R^2
2007	1.13	0.947	1.13	0.947	1.13	0.947
1978	1.28	0.907	1.16	0.917	1.10	0.944
1968	1.43	0.958	1.13	0.954	1.23	0.964
1939	1.46	0.948	1.70	0.882	1.32	0.967
1920	1.55	0.955			1.51	0.962

In all instances but one, the goodness of fit (R^2) statistic exceeds ninety percent.

The estimate of Pareto's coefficient tends to fall as the equation is fitted to more recent years implying more concentration among large unions in recent years.

What is implied by $\alpha = 1.53$ compared with $\alpha = 1.13$?
Use the property of Pareto's distribution that, provided $\alpha > 1$, for any union with membership M^*, the average membership of those unions whose membership equals or exceeds M^* is
$[\alpha (\alpha - 1)^{-1}] M^*$.

This implies that a lower value of α increases the gap between M^* and the average-sized union with membership above M^*.

If M^* equals 200,000, then the average size of unions with membership equal to or greater than 200,000 is 577,359 when α = 1.53

and the average union size is three times this number (precisely, 1,738,462) when α = 1.13.

This difference in α reflects a considerable difference in inequality among these large unions. Apparent “small” differences in α have “large” implications for concentration.

Two Extensions to the Use of Pareto's Distribution Describing Trade Unions

First, how do the estimates of Pareto's coefficient change if the size of unions is measured by something other than the number of members?

Second, because the history and structure of U.S. unionism are singular, how does Pareto's distribution fare as a description of the size of larger unions in other countries?

Assets

Compare the estimates for Pareto's coefficient when applied to union membership with the estimates when applied to the net assets of unions where K_i denotes the net assets of union i :

(1.2) $$ln\, S_i = a_3 + b_3\, ln\, K_i + e_{3i}$$

(3.2) $$ln\, S_i = a_4 + b_4\, ln\, K_i + c_4\, (ln\, K_i)^2 + e_{4i}$$

Britain

British unionism has been characterized by the dominance of a few "general" unions that draw their membership from many different industries and that account for a large fraction of total union membership. In this respect, the growth of certain very large unions in the U.S.A. such as the Service Employees International Union, the Teamsters, and Steel Workers Union follows the pattern set earlier by British unionism.

Another similarity between American and British unionism is the decline in the extent of unionism in both countries in the last few decades.

U.S.A.

Membership and Net Assets of the 21 Largest U.S. Unions in 1982 (Net Assets) and in 1983 (Membership)

(1.1) $$\ln S_i = a_0 + b_0 \ln M_i + e_{0i}$$

(1.2) $$\ln S_i = a_3 + b_3 \ln K_i + e_{3i}$$

membership equation (1.1)				net assets equation (1.2)			
estimates of.....		implied α	R^2	estimates of.....		implied α	R^2
a_0 (s.e.)	b_0 (s.e.)			a_3 (s.e.)	b_3 (s.e.)		
11.600 (0.540)	-1.278 (0.087)	1.28	0.920	13.086 (0.387)	-0.848 (0.035)	0.85	0.969

(3.1) $$\ln S_i = a_2 + b_2 \ln M_i + c_2 (\ln M_i)^2 + e_{2i} \quad ,$$

(3.2) $$\ln S_i = a_4 + b_4 \ln K_i + c_4 (\ln K_i)^2 + e_{4i}$$

membership equation (3.1)				net assets equation (3.2)			
estimates of.....			R^2	estimates of.....			R^2
a_2 (s.e.)	b_2 (s.e.)	c_2 (s.e.)		a_4 (s.e.)	b_4 (s.e.)	c_4 (s.e.)	
-10.356 (3.096)	5.768 (0.990)	-0.560 (0.079)	0.979	5.187 (5.261)	0.532 (0.917)	-0.060 (0.039)	0.973

Britain

Membership (1984) and Net Worth (1985) of the 20 Largest British Unions

(1.1) $ln\ S_i = a_0 + b_0\ ln\ M_i + e_{0\,i}$

(1.2) $ln\ S_i = a_3 + b_3\ ln\ K_i + e_{3\,i}$

membership equation (1.1)				net worth equation (1.2)			
estimates of.....		implied α	R^2	estimates of.....		implied α	R^2
a_0 (s.e.)	b_0 (s.e.)			a_3 (s.e.)	b_3 (s.e.)		
10.313 (0.281)	-1.153 (0.048)	1.15	0.9 68	12.691 (0.747)	-0.955 (0.079)	0.96	0.891

(3.1) $ln\ S_i = a_2 + b_2\ ln\ M_i + c_2\ (\ ln\ M_i\ )^2 + e_{2\,i}$,

(3.2) $ln\ S_i = a_4 + 4_5\ ln\ K_i + c_4\ (\ ln\ K_i\ )^2 + e_{4\,i}$

membership equation (3.1)				net worth equation (3.2)			
estimates of.....			R^2	estimates of.....			R^2
a_2 (s.e.)	b_2 (s.e.)	c_2 (s.e.)		a_4 (s.e.)	b_4 (s.e.)	c_4 (s.e.)	
3.451 (2.224)	1.162 (0.747)	-0.192 (0.062)	0.979	-15.892 (3.103)	5.125 (0.658)	-0.321 (0.035)	0.982

Inferences from these Estimates

1. Assets v. Membership

In both the U.S. and Britain in the 1980s, there was more inequality in the assets of unions than in their membership

2. U.S. v. Britain

As measured by Pareto's coefficient in the 1980s, there was more inequality in union membership among large British unions than among large U.S. unions and there was more inequality in the assets of unions among large U.S. unions than among large British unions

3. Judgment about Pareto's distribution

a) Pareto's distribution provides a good <u>first-order</u> fit to the distribution of membership and of assets among large U.S. and British unions in the 1980s

b) The statistical significance of the (negative) quadratic terms in both the membership and the asset equations in Britain suggests a departure from strict log-linearity. The same statement holds for membership though not for assets in the U.S.

4. The U.S. over time

There has been an increase in concentration of membership in national unions since the 1970s and perhaps longer. This concentration has come about as the heart of U.S. unionism has shifted from the private to the public sector. The “labor conglomerates” are situated principally in the public sector.

How important in this concentration? Cf with product market concentration

Does this trend toward concentration matter?