STATISTICAL APPENDIX

This is the statistical appendix to "Will Most of Us be Working for Giant Enterprises by 2028?" in the Journal of Economic Behavior and Organization.

A. Data on Enterprises

As noted in the article, before 1992 Enterprise Statistics does not have statistics completely covering four sectors and, therefore, it is necessary to make estimates. For 1992 comparisons of total employment in 1992 from Enterprise Statistics with the employment data from the NIPA statistics [U.S. Department of Commerce, Bureau of Economic Analysis, 1998, Table 6.4] reveal relatively complete coverage for seven sectors; for services, however, the former source covers only about four fifths of all employees. I made, however, no corrections for this shortfall either in 1992 or in previous years.

For the period 1977 - 1992 I have been able to link unpublished data on the employment size of enterprises estimated by the U.S. Small Business Administration to the Census data so as to estimate a series covering all sectors. For the period from 1958 through 1977, the estimation procedures are more problematic.

Using sectoral employment data from the NIPA statistics I constructed an "employment elasticity" for each size class of enterprise for the period 1978 - 1982. This statistic showed the percentage change in employment of any size class for each of the four sectors when total employment in the sector increased by 1 percent. These elasticities were then used to estimate employment in each size class for each sector back to 1963.

Assuming that the percentage change in the employment size of enterprises in each size class for the sectors not fully included in Enterprise Statistics was the same as for the core

sectors, I then combined these estimates with the estimates of total employment in each employment size to calculate the number of enterprises.

Although I present only the data for all sectors in the text, I also calculated the same tables for only the core sectors for which data are complete for all years. Both sets of data reveal similar trends.

Some data sets on the size distribution of enterprises are available for years earlier than 1958, but they were estimated using much different sources and methods than the current Enterprise Statistics series. The lack of comparability between the data in earlier and later years appeared sufficiently great that it did not appear worthwhile to include these earlier data in this discussion [1]

B. Interpolation Procedures

Since the size distribution data are grouped, it is necessary to interpolate and extrapolate these data in order to calculate the Florence median and the share of employment accounted for by the largest number of firms. Given the uncertainty in the type of distribution with which I was dealing, I only used the tail of the distribution (for enterprises, only those with 500 or more employees; for establishment, only those with 50 or more employees) for the curve fitting. For calculating the various interpolations and extrapolations, it proved useful to fit two types of curves.

[1] Such data are available for 1945 through 1956 (Churchill [1954] or U.S. Department of Commerce [1965, p. 80, series V20-29] and, except for services, appear to cover most of the labor force in the private sector. Making adjustments for this undercount and comparing these estimates for 1956 with the estimates used in Table 2 reveals certain differences, for instance, in the former set of statistics the number of enterprises was larger and the covered labor force was smaller. Various measures of size are also sufficiently different that the two series do not seem comparable. The major cause for these inconsistencies lies with my estimate of services for the early series, which were necessarily crude because there was no year for which the coverage of services was relatively complete.

For determining the size of the firm with a given number of workers, the Pareto distribution proved to have the best fit. That is, the coefficient of determination (as well as the standard errors of estimate) obtained by regressing the log of the lower limit of the size distribution against the log of the accumulated number of firms was higher than any other method with which I experimented, and it is from these calculations that the results in Appendix Table A-1 are derived. For determining the total number of employees of a given number of firms, another type of interpolation is necessary. Experimenting with 1992 data I found empirically that regressing the log of the cumulative number of firms against the cumulative number of employees provided the best fit.

Appendix Table A-1 about here.

Of the various measures of enterprise and establishment size in the tables below, no estimates were required for calculation of the arithmetic average and the percent of employees in firms with more than 999 and 9,999 employees. Given these estimating procedures and the small number of points with which to fit the curves used in interpolating and extrapolation, the other three statistics are more problematic and, in descending order of reliability, are: the Florence median, which only required interpolations; the average number of employees in the largest 1,000 enterprises, which required relatively small extrapolations; and the average number of employees in the largest 100 enterprises, which required extrapolations far beyond the lowest limit of the highest size category (over 9,999 employees).

C. Establishment Size

As noted in Table 4, the statistics on establishment size come from various issues of County Business Patterns. Appendix Table A-2 presents various measures of average employment size when the sector composition of employment is held constant.

Appendix Table A-2 about here.Table A-1: Pareto Coefficients ( ) for the Size Distribution of Enterprises

Years		Pareto coefficients
Census data for grouped data of all domestic enterprises over 500 employees		Data from Fortune of largest 100 U.S. employers in the industrial sector where both employees at home and abroad are counted
1958	-0.98	-0.68
1963	-0.95	-0.67
1968	-0.91	-
1972	-0.90	-0.63
1977	-0.92	-
1982	-0.94	-0.60
1987	-0.98	-
1992	-0.99	-0.66
1997	-	-0.66

Notes: Using Gibrat's ideas about a dynamic stochastic of enterprise growth, Ijiri and Simon [1977, p. 13] argue that the Pareto coefficient is a more theoretically based assessment of changing aggregate concentration (sometimes called agglomeration) of enterprises. The closer the Pareto coefficient is to zero, the greater the concentration.

Table A-2: Various Measurements of Establishment Size Holding Employment Structure Constant

Year	Establishments> 19 employees		Florence median		Percent of employees in establishments with:					
	Arithmetic average		Number of employees		less than 20 employees		more than 500 employees		more than 1000 employees	
Employment weights										
	1962	1995	1962	1995	1962	1995	1962	1995	1962	1995
1962 data	101.3	93.1	99	66	27.0%	31.7%	27.6%	24.2%	18.7%	16.3%
1995 data	91.1	84.9	97	76	22.5	25.5	22.2	20.0	14.1	13.0

Note: Eight major sectors are used in the constant-employment-structure indices. For sources of data and other information, see note for Table 4 in the original article.

.