

Business Estimates from the Office of Advocacy: A Discussion of Methodology

A working paper by

Brian Headd
Office of Advocacy
U.S. Small Business Administration



Release Date: June 2005

The statements, findings, conclusions, and recommendations found in this study are those of the authors and do not necessarily reflect the views of the Office of Advocacy, the United States Small Business Administration, or the United States Government.

Abstract

The Office of Advocacy creates estimates of the number of employer and nonemployer businesses for recent years, since the newest available data have around a two-year lag. Advocacy estimates the number of nonemployers using growth rates from Internal Revenue Service's sole proprietor count estimates. Advocacy estimates the number of employers using the net difference in employer births and deaths from the Employment and Training Administration. Analysis shows that historical estimates have been close to the actual figures, as employers and nonemployers exhibited slow steady growth during the period of analysis. Nonemployers, however, tended to be slightly underestimated as sole proprietors have consistently grown faster than nonemployers in recent years. As more data become available, more sophisticated techniques could be employed to create current estimates of the number of businesses.

Introduction

The Office of Advocacy is charged in its authorizing statute to "examine the role of small business in the American economy." Since 1976, the business, academic, and policy communities have come to rely on the Office of Advocacy's small business data. One elusive statistic that is much in demand is the current number of small businesses.¹ The Census Bureau's employer and nonemployer statistics are the most inclusive measures of the number of businesses. Hence, the best and most complete measure of the number of firms (excluding farms) can be found by adding the number of employers from Statistics of U.S. Business (SUSB), published by the Bureau of the Census, to the number of nonemployers from the Census's Nonemployer Statistics.² However, this number is only available with a lag time of about two and a half years. This working paper explains Advocacy's methodology for creating estimates of the number of businesses.³ In a nutshell, Advocacy applies current changes in proxy business counts to historic U.S. Census Bureau levels of employer and nonemployer businesses separately, then adds them together for an estimate of the total number of businesses for each of the previous two years. The reasons for this and the methods used are addressed in the following discussion.

Data Sources

There are a few ways to count the number of businesses. Adding the number of employers and nonemployers from the U.S. Census Bureau's nonemployer database and SUSB is a comprehensive figure that is becoming the gold standard. Unfortunately, both data sources have a lag of a few years. Fortunately, figures are available from the underlying data sources from which they are created. Rates of change based on these more current data sources can be used to try to bring the lagging data sources up to date.

Employers and nonemployers tend to have differing trends with regard to the business cycle, so estimating them separately can create more useful information and result in a strong total business estimate.⁴ In 2002, the most recent year for which actual figures are available, the total number of businesses was 23.3 million. Of these,

¹ For the past several years, the Office of Advocacy has estimated the number of businesses for the latest two years in its publication, *Small Business Economic Indicators*. See www.sba.gov/advo/stats/sbei03.pdf. Note that estimates for some prior year figures calculated in this paper to better evaluate the methods differ from figures listed in *Small Business Economic Indicators* because the actual 2002 nonemployer figure was uniquely available because of the Economic Census.

² Note, however, that the number of employers and nonemployers in existence at any point during the year will be lower than these Census figures, since Census counts all businesses that existed at any time during the year, including those that ceased operations during the year.

³ Comments on these procedures are encouraged to improve the quality of information available on businesses. Helpful comments from Jim Spletzer and Jason Faberman of the U.S. Department of Labor, Bureau of Labor Statistics were received, incorporated, and appreciated.

⁴ A suggestion was received to create estimates for disaggregated figures. In using ETA figures, state estimates would be possible and will be evaluated for future estimates.

17.6 million, or 75 percent consisted of nonemployers, and the remaining 5.7 million (25 percent of businesses) were employers.⁵

Nonemployer Data. According to the Census Bureau, the underlying source for its nonemployer business number is nonfarm “annual business income tax returns filed with the Internal Revenue Service (IRS) and maintained in the Census Bureau’s Business Register.” Census nonemployers have no annual payroll and a minimum of \$1,000 in annual receipts.⁶ The Census Bureau edits and reviews the tax returns to develop tabulations of nonemployer businesses. Of the nonemployers in 2002, 87 percent were sole proprietors.⁷ Because the nonemployer statistics are based primarily on IRS sole proprietorship data, it makes sense to use the latest IRS sole proprietorship data to estimate the percentage change from prior year data.⁸

Employer Firm Data. The Census Bureau, with funding from the Office of Advocacy, has produced annual statistics on nonfarm employer businesses by size of firm beginning with the 1988 data year, under the Statistics of U.S. Business (SUSB) program. SUSB uses employer establishments (or business locations) from Census’s County Business Patterns as a base and uses other data sources to identify which establishments belong to the same firm. The underlying data for the County Business Patterns come from the Census Bureau’s Company Organization Survey and Economic Censuses; administrative records from the IRS, the Social Security Administration, and the Bureau of Labor Statistics; as well as various other surveys.⁹

Year-to-Year Rates of Change. How much does the number of businesses grow from year to year? Counts of business openings and closings are not available from the underlying data sources mentioned above. But openings and closings are available from administrative unemployment insurance records for employers through the Employment and Training Administration (ETA). These data are collected by state, and are available quarterly with a lag of about three months. The underlying data sources for SUSB and figures from ETA should mirror each other with regard to employer openings and closings, and ETA trends are used as a basis for estimates.¹⁰

Because the unemployment insurance tabulations reflect employers within a state, and some firms have establishments in more than one state, the number of new employers and employer closures in these figures is overstated. However, the percent change in the state totals should reflect trends similar to those for the United States as a whole, as the share of establishments belonging to multi-state firms is not believed to vary much from year to year. The quarterly ETA data are added together to produce annual data, which helps “smooth out” unusual peaks and valleys reflecting occasional processing and cleaning of state files. ETA data include three categories: new firms, terminations, and successor firms; Advocacy’s estimates make use of the data on new firms and terminations.

Other measures have more problematic implications:

- The number of IRS tax returns overstates the actual number of businesses, as businesses can and do file more than one return.

⁵ Only 17,000 companies, or 0.3 percent of all employers, had 500 or more employees; thus, the estimate of the number of employers is virtually an estimate of small employers.

⁶ For more detail, see www.census.gov/epcd/nonemployer/view/covmeth.htm.

⁷ One important distinction between nonemployers and sole proprietors is that sole proprietors can have employees.

⁸ It is important to note here that the latest two years of IRS nonfarm sole proprietor data are projected. IRS develops estimates from economic modeling and adjustments using “recent filing patterns, economic and demographic trends, enacted legislation, and administrative changes.” See www.irs.gov/pub/irs-soi/04proj.pdf for more detail.

⁹ For details, see www.sba.gov/advo/research/rs190tot.pdf.

¹⁰ Census’ SSEL (renamed the Business Register) and BLS’ Business Establishment List or BEL, are similar. A study of 1994 data found matches of 95 percent among establishments in both sources (this excludes a large number of SSEL units that did not match and had no payroll). Parker, Spletzer, and Searson, *The Business Establishment List – Standard Statistical Establishment List Comparison Project*, prepared for the Federal Economic Statistics Advisory Committee, June 15, 2000. See www.bls.gov/ore/pdf/st010030.pdf.

- Using self-employment as a measure of nonemployer businesses is inaccurate. Self-employment is an indicator of labor force status. Businesses can have more than one owner (sole proprietor). This too results in multiple filings for individual businesses and overstates the actual amount.
- The number of establishments measures the number of business locations, not the number of firms. (A parent firm can own multiple establishments.) But using establishment openings and closings from the Bureau of Labor Statistics' new Business Employment Dynamics program could be a compelling source in the future. Unfortunately, even this data source, having a time lag of nine months, would be problematic.

Estimating the Number of Businesses

Annual nonemployer data have been available only since 1997; employer data have been available since 1988. An econometric model created from these few available data points would be weak at best. Therefore, Advocacy's more basic approach uses current available data from the IRS and the ETA to estimate year-to-year percentage changes. These rates of change are applied to Census totals of employers and nonemployers. In addition, this approach makes the process more transparent.

The Census Bureau's nonemployers are similar to the population of sole proprietors documented by the IRS, and employers are similar to the employer population covered in state unemployment insurance data provided to the ETA. Estimates of changes in the IRS and ETA data applied to the more comprehensive Census employer and nonemployer data are used to give an approximation of recent year totals.

For nonemployers, the percentage change rates in sole proprietors (as measured by the IRS) is applied to the most current Census nonemployer data. To arrive at figures for employer businesses, estimated employer firm births and closures are calculated and applied to the previous year's number of employers. For employer firm births and closures, the percentage change rates from ETA's state unemployment insurance figures for new firms and terminations are applied to the most current Census employer birth and death data.

Following are examples of recent estimates for 2002 and 2003.¹¹

Nonemployer Firms, 2002 and 2003. The Census nonemployer firm count from 2001, 16,979,498, is increased by the nonfarm sole proprietor annual percentage change from 2001 to 2002 of 2.41 percent, resulting in a rounded nonemployer firm estimate for 2002 of 17,388,800.¹² Increasing this figure by the sole proprietor annual percentage change from 2002 to 2003 of 1.90 percent, results in a rounded nonemployer firm estimate for 2003 of 17,718,500.

(2001 nonemployers)	x	(2001 to 2002 sole proprietor change)	=	(rounded 2002 nonemployer estimate)
16,979,498	x	1.0241	=	17,388,800
(2002 nonemployer estimate)	x	(2002 to 2003 sole proprietor change)	=	(rounded 2003 nonemployer estimate)
17,388,800	x	1.019	=	17,718,500

Employer Firm Births. Census's employer firm birth count from 2001, 585,140, is increased by the ETA state firm birth annual percentage change from 2001 to 2002, 0.77 percent. This yields a rounded employer firm birth

¹¹ Note that listed percentages are rounded.

¹² Figures here differ from the published figures as the 2002 actual nonemployer figure was uniquely available because of the Economic Census.

estimate for 2002 of 589,700. Decreasing this figure by the ETA's annual percentage change in state firm births from 2002 to 2003 (-2.85), results in a rounded firm birth estimate for 2003 of 572,900.

(2001 Census employer births)	x	(2001 to 2002 state birth change)	=	(rounded 2002 employer birth estimate)
585,140	x	1.0077	=	589,700
(2002 employer birth estimate)	x	(2002 to 2003 state birth change)	=	(rounded 2003 employer birth estimate)
589,700	x	(1 - 0.0285)	=	572,900

Employer Firm Closures. Census's employer firm closure count from 2001, 553,291, is increased by the ETA's annual percentage change in state firm closures from 2001 to 2002 (2.84 percent). This results in a rounded employer firm closure estimate for 2002 of 569,000. Decreasing this figure by the ETA's annual percentage change in state firm closures from 2002 to 2003 (-2.49) yields a rounded firm closure estimate for 2003 of 554,800.

(2001 Census employer closures)	x	(2001 to 2002 state closures change)	=	(rounded 2002 employer closure estimate)
553,291	x	1.0284	=	569,000
(2002 employer closure estimate)	x	(2002 to 2003 state closure change)	=	(rounded 2003 employer closure estimate)
569,000	x	(1 - 0.0249)	=	554,800

Employer Firms, 2002 and 2003. Census's employer firm count for 2001 (5,657,774) is increased by Advocacy's estimate of the net difference in firm births and closures in 2002 (20,700). This results in a rounded estimate for 2002 of 5,678,500. Increasing this figure by the 2003 net difference in firm births and closures (estimated by Advocacy at 18,100), results in a rounded estimate for 2003 of 5,696,600.

(2001 Census employers)	+	(2002 employer birth estimate)	-	(2002 employer closure estimate)	=	(rounded 2002 employer estimate)
5,657,774	+	589,700	-	569,00	=	5,678,500
(2002 employer estimate)	+	(2003 employer birth estimate)	-	(2003 employer closure estimate)	=	(rounded 2003 employer estimate)
5,678,500	+	572,900	-	554,800	=	5,696,600

Total Businesses, 2002 and 2003. Adding the employer and nonemployer estimates for 2002 results in a total business estimate of 23,067,300 for 2002. Adding the employer and nonemployer estimates for 2003 results in a total business estimate of 23,415,100 for 2003.

(2002 nonemployer estimate)	+	(2002 employer estimate)	=	(2002 total business estimate)
17,388,800	+	5,678,500	=	23,067,300
(2003 nonemployer estimate)	+	(2003 employer estimate)	=	(2003 total business estimate)
17,718,500	+	5,696,600	=	23,415,100

Table 1 shows the actual and estimated figures for other years.

Results Evaluation

With only seven years of nonemployer data available—1992 and 1997 to 2002—the comparisons that can be made between actual and estimated figures are limited. Attributing statistical significance to so few data points would be of limited value, so charts are employed and attached to the end of the working paper to visualize the estimates.

The first thing to note from Table 1 is that the actual number of employers and nonemployers continually reflect slight annual increases, as there were actual increases in all of the periods reviewed. The increases ranged from 1.7 percent to 3.9 percent for nonemployers and from 0.1 percent to 2.0 percent for employers.

Charts 1, 2 and 3 show that the estimates were quite close to the actual figures, which was most likely the result of the limited natural variation of the figures from year to year. But Chart 3 does show that nonemployer estimates were consistently lower than actual figures. The percentage increases in nonemployers over the measurable period of 1997 to 2002 were larger than the percentage increases in the number of nonfarm sole proprietor tax returns, so the estimation method resulted in a low estimate of the number of nonemployers. Chart 6 shows this undercount more clearly through the annual changes in nonemployers. As more nonemployer data points become available, this imperfect relationship between nonemployer and sole proprietor figures could be quantified and addressed in future nonemployer projections. Because nonemployers constitute most—75 percent—of all businesses and because nonemployer estimates tended to be lower than nonemployer actuals, total business estimates were slightly lower than total business actuals.

Also, Charts 5 and 6 do not show a consistent improvement from an estimate made two years in the future to one year in the future. So having one additional year of data did not yield figures that were better able to capture the change in the numbers of employers and nonemployers.

With nonemployers and employers continually rising throughout the evaluation period, an opportunity to evaluate the estimates' abilities to capture sign changes was not available. And declines predicted by the employer estimates for 2001 did not materialize.

The similarities in estimates and actuals in Charts 1, 2 and 3 lead one to believe that estimates have value for discussion purposes. However the discrepancies in estimates and actual changes in Charts 5 and 6 show that the estimates are less useful for quantitative analysis.

Conclusion

Judging from the few data points where estimate and actual figures are available, estimates from the Office of Advocacy are close to the actual numbers and would be quite adequate for many practical purposes. However, slight variations in the rates of change in the number of employers and nonemployers makes creating an estimate that follows the trend of the annual change in employers and nonemployers quite difficult. Therefore, using these estimates for academic models could present problems.

Comparisons between actual numbers and estimates show that the estimates are close enough to answer more questions than they raise. So at this juncture, the Office of Advocacy feels that applying current changes in proxy business counts to historic U.S. Census Bureau levels is the preferred method of estimation, and we await the views of others on this approach.

As about five more data points (or years) become available in the nonemployer data, the performance of the Office of Advocacy's methods to estimate the number businesses will need to be readdressed and the inaccuracies of the methods described here could be improved through econometric techniques.

List of Illustrations

Table 1: Comparison of Actual Business Figures and Estimates, 1992-2003

Chart 1: Comparison of Estimated and Actual Number of Total Businesses, 1997-2003

Chart 2: Comparison of Estimated and Actual Number of Employers, 1992-2003

Chart 3: Comparison of Estimated and Actual Number of Nonemployers, 1997-2003

Chart 4: Difference in Actual and Estimated Number of Total Businesses, 1998-2003

Chart 5: Difference in Actual and Estimated Number of Employers, 1993-2003

Chart 6: Difference in Actual and Estimated Number of Nonemployers, 1998-2003

Chart 7: Annual Change in Business Measures, 1993-2003

Table 1. Comparison of Actual Business Figures and Estimates, 1992-2003

	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Number of businesses												
Totals												
Actual	19,420,356	--	--	--	--	20,981,527	21,287,904	21,760,347	22,182,499	22,637,272	23,343,821	--
2004 Est.	--	--	--	--	--	--	--	--	--	--	23,067,300	23,415,100
2003 Est.	--	--	--	--	--	--	--	--	--	22,474,151	22,895,182	--
2002 Est.	--	--	--	--	--	--	--	--	21,983,976	22,228,855	--	--
2001 Est.	--	--	--	--	--	--	--	21,522,300	21,723,300	--	--	--
Nonemployers												
Actual (a)	14,325,000	--	--	--	--	15,439,609	15,708,727	16,152,604	16,529,955	16,979,498	17,646,062	--
2004 Est. (b)	--	--	--	--	--	--	--	--	--	--	17,388,800	17,718,500
2003 Est. (b)	--	--	--	--	--	--	--	--	--	16,844,551	17,299,982	--
2002 Est. (b)	--	--	--	--	--	--	--	--	16,332,376	16,588,555	--	--
2001 Est. (b)	--	--	--	--	--	--	--	15,885,500	16,062,300	--	--	--
Employers												
Actual (a)	5,095,356	5,193,642	5,276,964	5,369,068	5,478,047	5,541,918	5,579,177	5,607,743	5,652,544	5,657,774	5,697,759	--
2004 Est. (b)	--	--	--	--	--	--	--	--	--	--	5,678,500	5,696,600
2003 Est. (b)	--	--	--	--	--	--	--	--	--	5,629,600	5,595,200	--
2002 Est. (b)	--	--	--	--	--	--	--	--	5,651,600	5,640,300	--	--
2001 Est. (b)	--	--	--	--	--	--	--	5,636,800	5,661,000	--	--	--
Employer births (a)												
Actual (a)	544,596	564,504	570,587	594,369	597,792	590,644	589,982	579,609	574,300	585,140	--	--
2004 Est. (b)	--	--	--	--	--	--	--	--	--	--	589,700	572,900
2003 Est. (b)	--	--	--	--	--	--	--	--	--	545,400	550,100	--
2002 Est. (b)	--	--	--	--	--	--	--	--	604,500	574,500	--	--
2001 Est. (b)	--	--	--	--	--	--	--	588,900	613,900	--	--	--
Employer closures (a)												
Actual (a)	521,606	492,651	503,563	497,246	512,402	530,003	540,601	544,487	542,831	553,291	--	--
2004 Est. (b)	--	--	--	--	--	--	--	--	--	--	569,000	554,800
2003 Est. (b)	--	--	--	--	--	--	--	--	--	568,300	584,500	--
2002 Est. (b)	--	--	--	--	--	--	--	--	560,600	585,800	--	--
2001 Est. (b)	--	--	--	--	--	--	--	531,300	560,600	--	--	--
Nonfarm sole prop. (c)	15,495,000	15,848,000	16,154,000	16,424,000	16,955,000	17,176,000	17,183,700	17,377,100	17,570,500	17,904,900	18,336,500	18,684,200
State employers (d)	5,791,214	5,900,156	6,035,538	6,105,470	6,190,907	6,330,117	6,494,520	6,621,421	6,766,206	6,871,683	6,952,838	7,044,924
State births (d)	736,773	757,529	806,855	819,477	842,357	885,416	886,235	884,604	922,092	875,714	882,473	857,295
State closures (d)	819,363	802,701	803,290	863,699	849,839	857,073	869,886	854,858	880,230	921,581	947,737	924,164

Table 1. Comparison of Actual Business Figures and Estimates, 1992-2003 (continued)

	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Percent changes												
Nonemployers	--	--	--	--	--	--	1.7	2.8	2.3	2.7	3.9	--
Nonfarm sole prop.	--	2.3	1.9	1.7	3.2	1.3	0.0	1.1	1.1	1.9	2.4	1.9
Census employers	--	1.9	1.6	1.7	2.0	1.2	0.7	0.5	0.8	0.1	--	--
State employers	--	1.9	2.3	1.2	1.4	2.2	2.6	2.0	2.2	1.6	1.2	1.3
Census births	--	3.7	1.1	4.2	0.6	(1.2)	(0.1)	(1.8)	(0.9)	1.9	--	--
State births	--	2.8	6.5	1.6	2.8	5.1	0.1	(0.2)	4.2	(5.0)	0.8	(2.9)
Census closures	--	-5.6	2.2	(1.3)	3.0	3.4	2.0	0.7	(0.3)	1.9	--	--
State closures	--	-2.0	0.1	7.5	(1.6)	0.9	1.5	(1.7)	3.0	4.7	2.8	(2.5)

(a) U.S. Census Bureau, Statistics of U.S. Business and Nonemployer Businesses

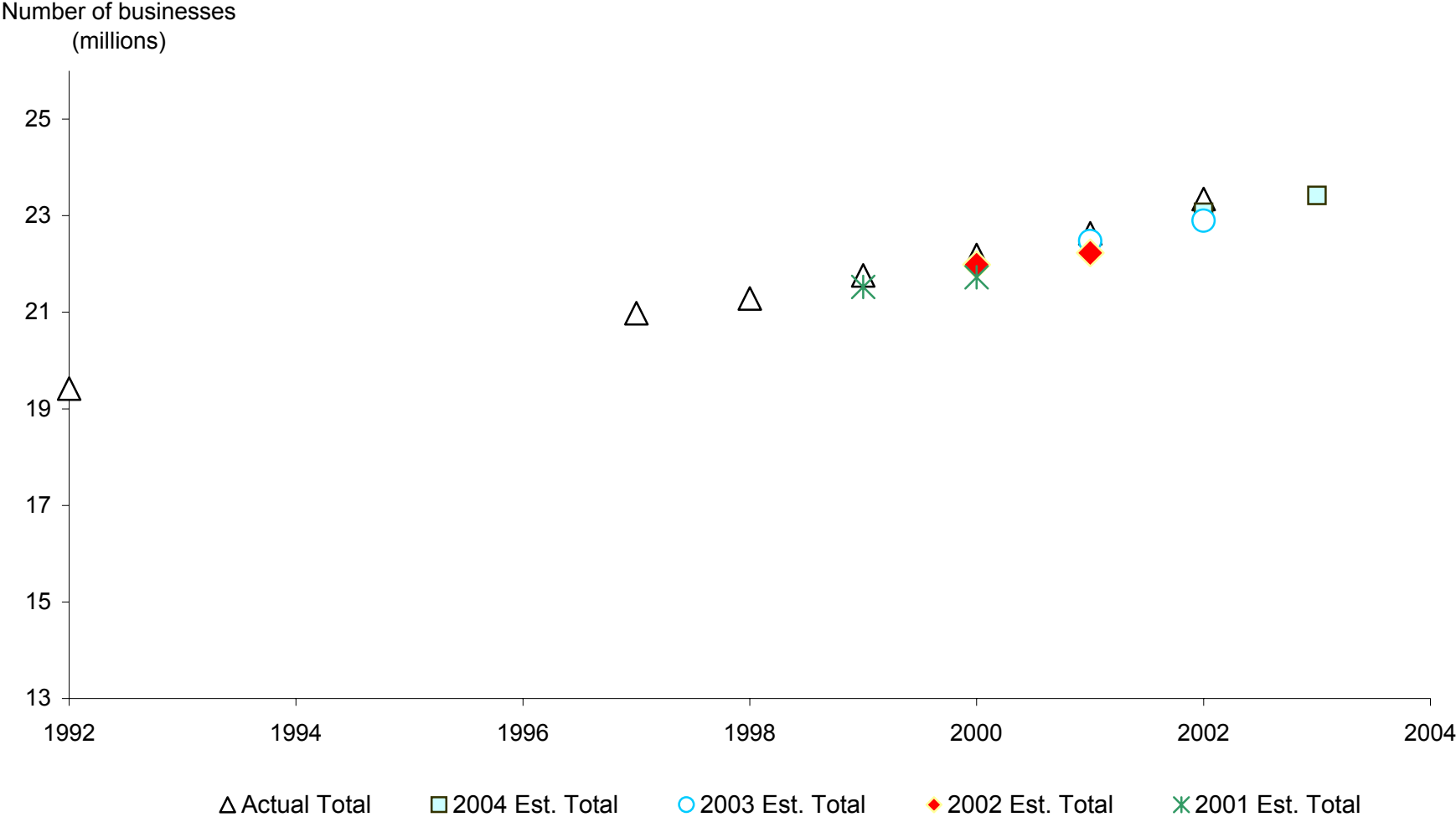
(b) U.S. Small Business Administration, Office of Advocacy estimates

(c) U.S. Department of Treasury, Internal Revenue Service with estimates for the two latest years

(d) U.S. Department of Labor, Employment and Training Administration addition of state figures

Source: U.S. Small Business Administration, Office of Advocacy, from data provided by the sources mentioned above and Advocacy estimates.

Chart 1: Comparison of Estimated and Actual Number of Total Businesses, 1997-2003



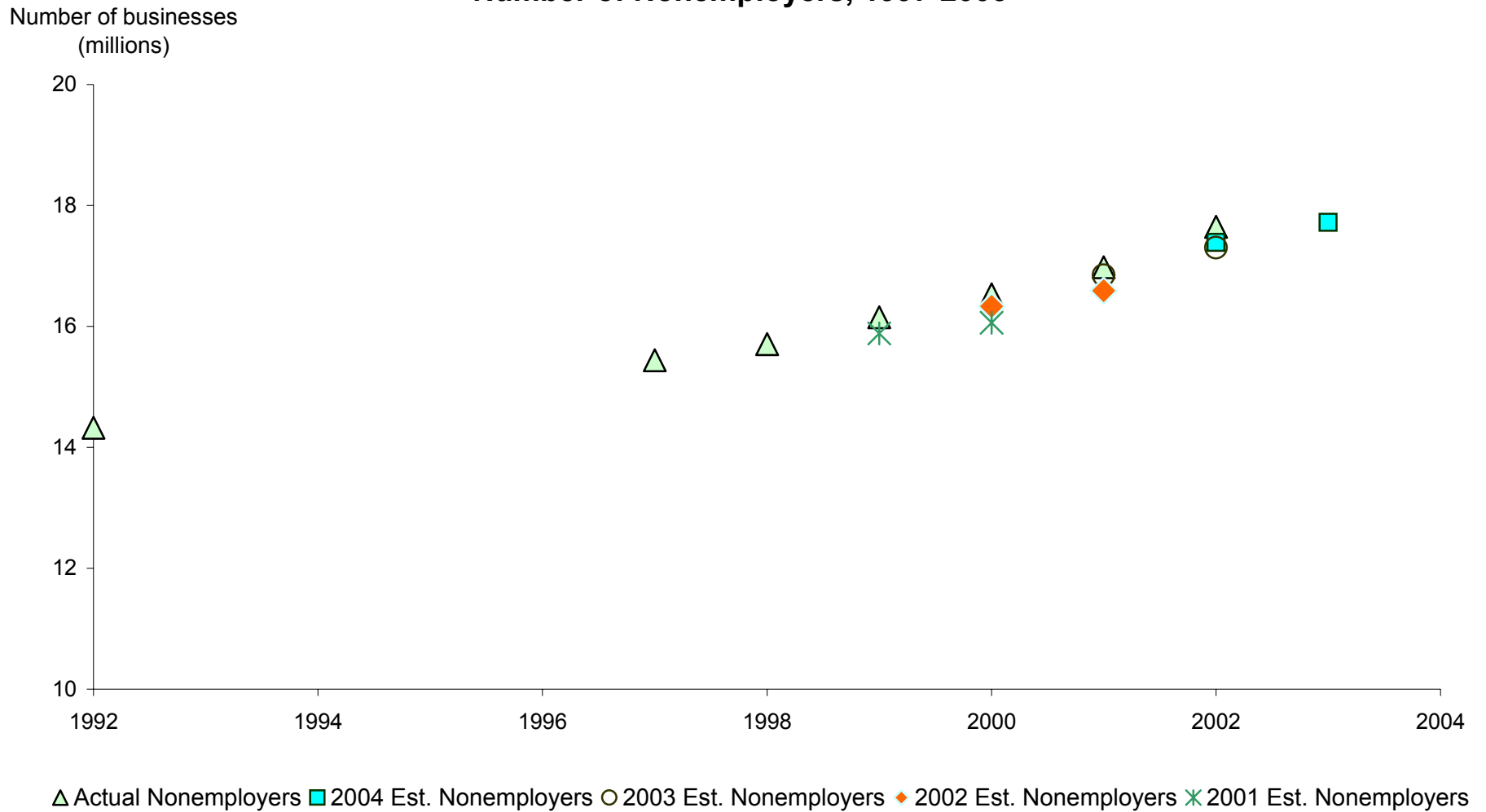
Source: Office of Advocacy, U.S. Small Business administration from U.S. Census Bureau actuals figures and Advocacy

Chart 2: Comparison of Estimated and Actual Number of Employers, 1992-2003



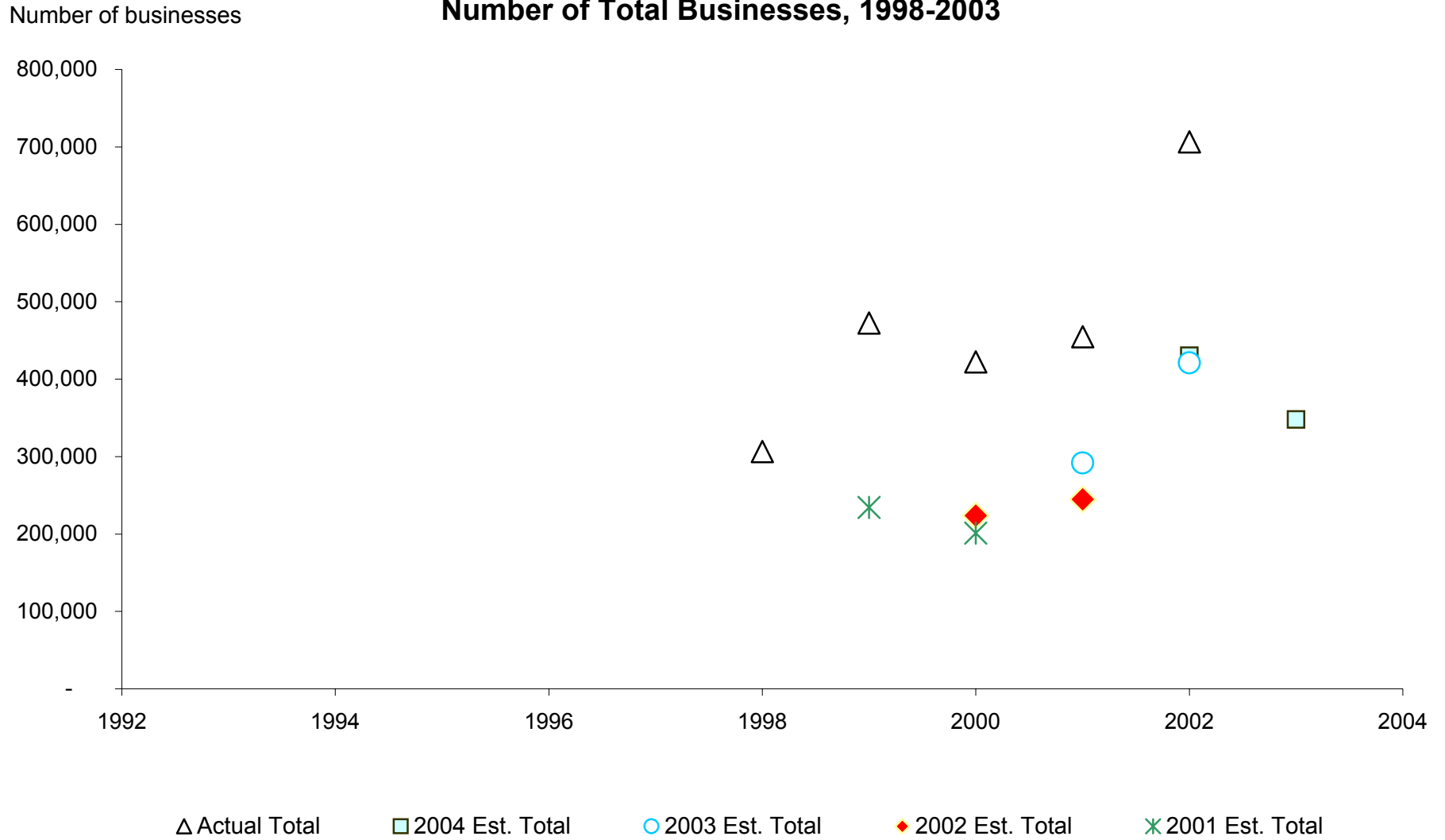
Source: Office of Advocacy, U.S. Small Business administration from U.S. Census Bureau actuals figures and Advocacy

Chart 3: Comparison of Estimated and Actual Number of Nonemployers, 1997-2003



Source: Office of Advocacy, U.S. Small Business administration from U.S. Census Bureau actuals figures and Advocacy

Chart 4: Difference in Actual and Estimated Number of Total Businesses, 1998-2003



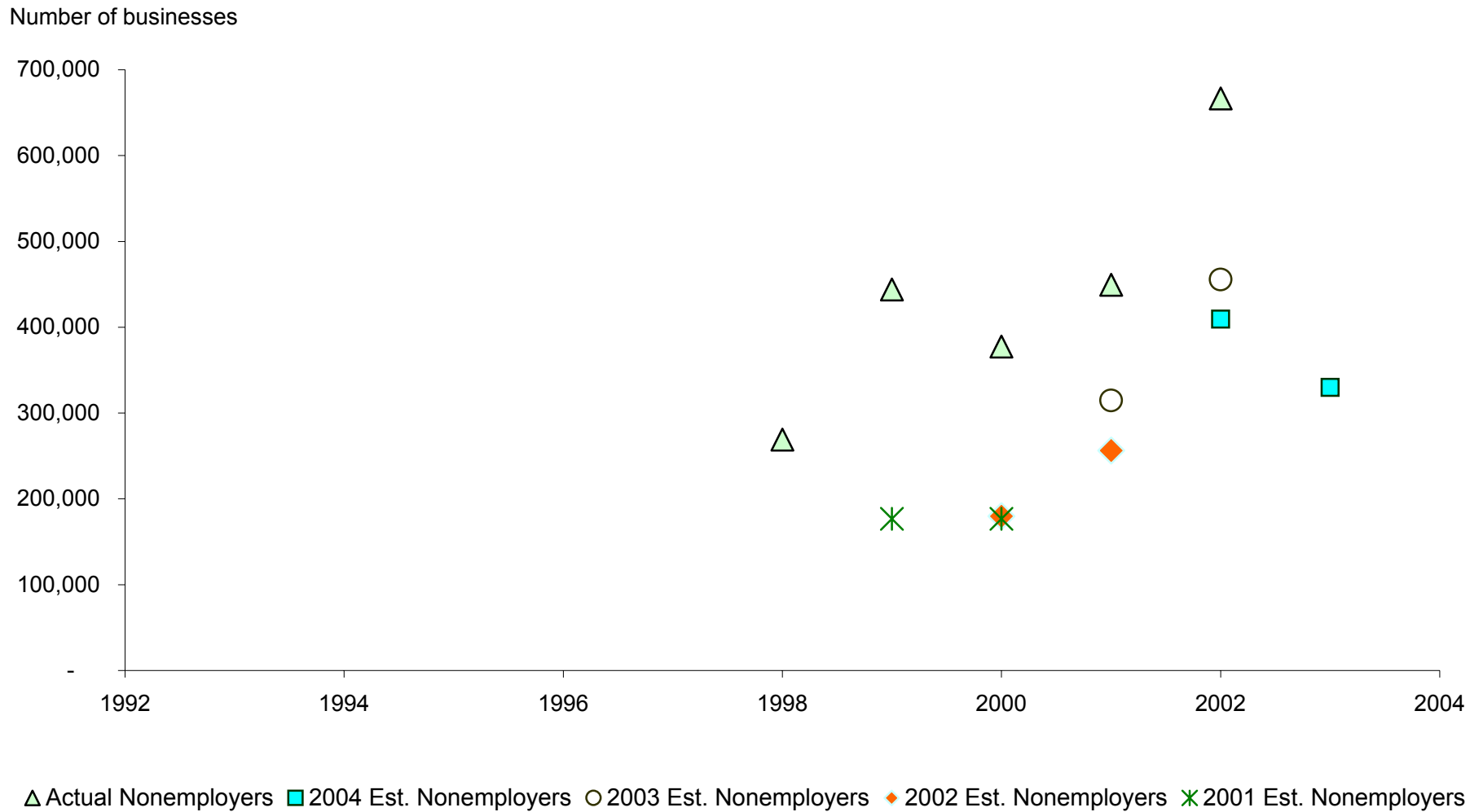
Source: Office of Advocacy, U.S. Small Business administration from U.S. Census Bureau actuals figures and Advocacy

Chart 5: Difference in Actual and Estimated Number of Employers, 1993-2003



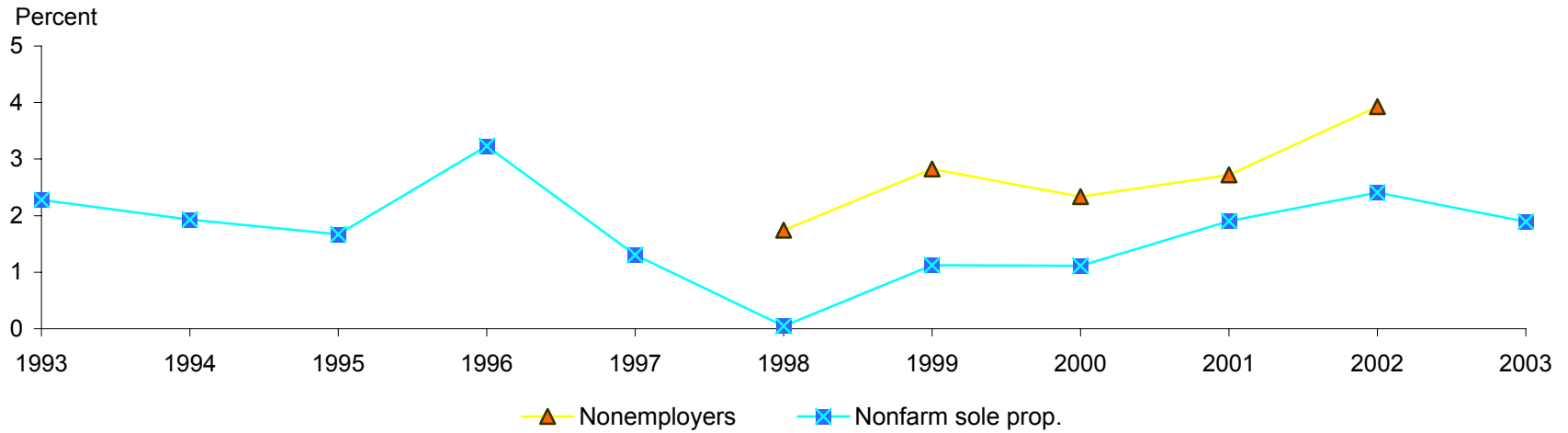
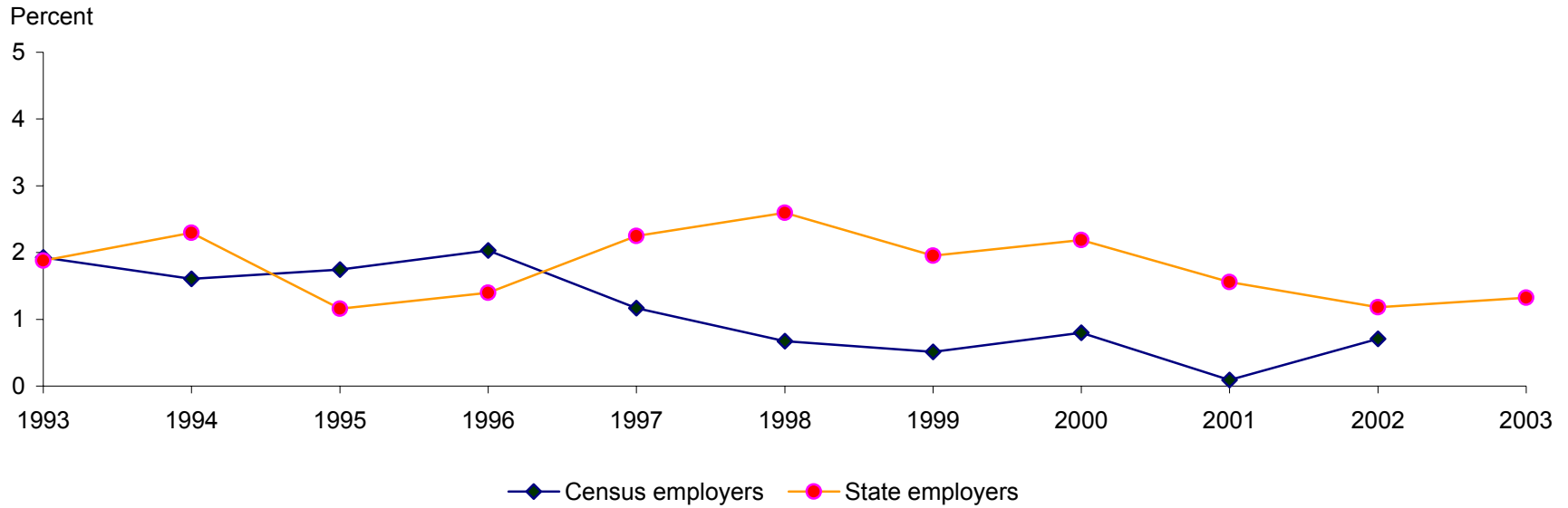
Source: Office of Advocacy, U.S. Small Business administration from U.S. Census Bureau actuals figures and Advocacy

Chart 6: Difference in Actual and Estimated Number of Nonemployers, 1998-2003



Source: Office of Advocacy, U.S. Small Business administration from U.S. Census Bureau actuals figures and Advocacy

Chart 7: Annual Change in Business Measures, 1993-2003



Source: Office of Advocacy, U.S. Small Business administration from U.S. Census Bureau actuals figures and Advocacy