

EIA COMPLETES CORRECTIONS TO DRILLING ACTIVITY ESTIMATES SERIES

by William Trapmann and Phil Shambaugh

INTRODUCTION

The Energy Information Administration (EIA) has published monthly and annual estimates of oil and gas drilling activity since 1978. These data are key information for many industry analysts, serving as a leading indicator of trends in the industry and a barometer of general industry status. They are assessed directly for trends, as well as in combination with other measures to assess the productivity and profitability of upstream industry operations. They are major reference points for policymakers at both the Federal and State level. Users in the private sector include financial analysts assessing investment opportunities. Firms with upstream operations also rely on these data in apprising their circumstances relative to those of their competitors. EIA uses these data in its own analytical and modeling work.

EIA does not itself collect drilling activity data. Instead, it relies on an external source for data on oil, gas, and dry well completions. These data are provided to EIA monthly on an *as reported* basis. Due to lags in the reporting of well completions which can (though most do not) range up to several years, EIA must statistically adjust the *as reported* completion data to obtain estimates of the numbers of completions that *would have been reported had there been no reporting lags*. Essentially, this is done by assuming that the pattern of reporting lags observed in the past holds true for the present, and making appropriate upward adjustments to the reported numbers of completions on that basis.

As an integral part of its data gathering function, EIA routinely monitors data quality and periodically conducts work intended to enhance its data systems. During a recent effort to enhance EIA's well completion data system, the detection of unusual patterns in the well completion data as received led to an expanded examination of these data. Substantial discrepancies between the data as received by EIA and correct record counts since 1987 were identified. For total wells by year, the errors ranged up to more than 2,300 wells, 11 percent of the 1995 total, and the impact of these errors extended backward in time to at least the early 1980s.

When the magnitude and extent of the *as reported* well completion data problem were confirmed, EIA suspended its publication and distribution of updated drilling data. EIA staff proceeded to acquire replacement files with the *as reported* records and then revise the statistical portion of its drilling data system to reflect the new information. The replacement files unfortunately also included erroneous data based on the improper allocation of wells between exploration and development. Users of EIA drilling activity data therefore are advised that the drilling activity data that were published or otherwise distributed by EIA prior to August 1998 are not necessarily valid.

EIA has now resolved the two data problems and generated revised time series estimates for well completions and footage drilled. While most industry trends remain consistent with those of the initial, incorrect series, the revised series does exhibit certain differences, chief among which are:

- Total well counts by year in the initial and revised series vary by less than 0.5% until 1996 when the difference is 1.3%.
- Drilling activity did attain its peak level in 1981, but the industry completed an estimated 91,553 wells as opposed to the prior estimate of 90,034.
- The decline in drilling during the mid 1990s was not as steep as previously indicated. The wells in 1995 had been underestimated by 2,384 wells—a difference of 11 percent.
- Success rates, measured as the share of successful gas and oil wells relative to total wells, are similar in the initial and final drilling series, but the improvement in the mid 1990s was not as great as previously indicated.

The remainder of this report presents background on the drilling activity data: what the records are, how they are collected, and the resulting difficulties in developing timely measures of recent drilling activity. This is followed by a discussion of the nature and extent of errors in the raw data files received by EIA. Last, the revised data are presented along with key differences between the prior and revised series and their implications for understanding industry performance.

DRILLING ACTIVITY DATA

About the data. The most widely cited measures of drilling activity consist of summarized information based on individual well records that describe the completion type and status of each well. The individual records contain information regarding the well, including the American Petroleum Institute (API) well number, the well completion date, the well class (exploratory or developmental), the well type (oil, gas, or dry), location data, and measures of the footage drilled. Drilling data traditionally had been compiled and presented as the records are received, a practice which predates EIA's publication and use of these data. (EIA continued this practice as part of its data operation until the mid 1980s.) This approach, however, reflects the reporting activity as measured by the recipient, rather than the industry's activity in the field.

Well completion data by report date ("as-received completions") are not an accurate indicator of actual drilling activity. A preferred measure of drilling activity is a record of well counts and footage drilled by completion date. The completion date marks the point at which the well generally becomes available for production. Drilling measured by completion date is thus more appropriate for industry analysis purposes than drilling measured by report date. Well counts by report date would match counts by completion date *if* wells were reported with no delay, however, that is not the case. In the early 1980s, EIA staff noticed unusual patterns in the as-received completion data, which reflected distortions due to a variable and sometimes very extended reporting lag.

The reporting lag not only creates incorrect magnitudes for peaks, troughs, or changes in drilling activity, it also tends to obscure the timing of these events. The lags are particularly troublesome at times when the drilling trends shift. The historical data provide some especially clear examples of the distortions that can be caused by recording wells by report date. Wells aggregated by completion date now show that total completions peaked in 1981 at 91,553 and had fallen to 84,397 by 1982 (Figure 1). This pattern reflects the effect of oil price levels on drilling activity and is consistent with the data on rotary rigs

running, both of which peaked in 1981. By comparison, total well counts as reported in 1981 and 1982 were 78,538 and 85,795, respectively.¹ The higher count in 1982 reflects an influx of reports of drilling activity that occurred in 1981 and earlier years. The reporting lags were sufficient to cause the 1981 count to be off by roughly 13,000 wells, a 14 percent discrepancy, with the result that the wrong year would be indicated as the peak year for drilling activity. Another critical example of reporting lag distortion is the measure of drilling in 1986, when the industry suffered a collapse in oil and gas prices. The number of wells actually completed in 1986 was about 40,000, while the number of well completion records received was more than 58,000.²

Adjusting for the reporting lag. Concerns about drilling activity data arose at EIA in the early 1980s. These concerns resulted in a plan to develop a procedure that would allow aggregation and reporting of the data on a completion date basis. As noted earlier, EIA does not collect well completion data directly, but relies on a vendor to collect the data, which EIA purchases as a monthly compilation. These are the only such data available, so attempts to convert to a completion date basis must utilize these data. Further, the quality of the EIA's drilling activity estimates depends directly upon the quality of the vendor records as received.

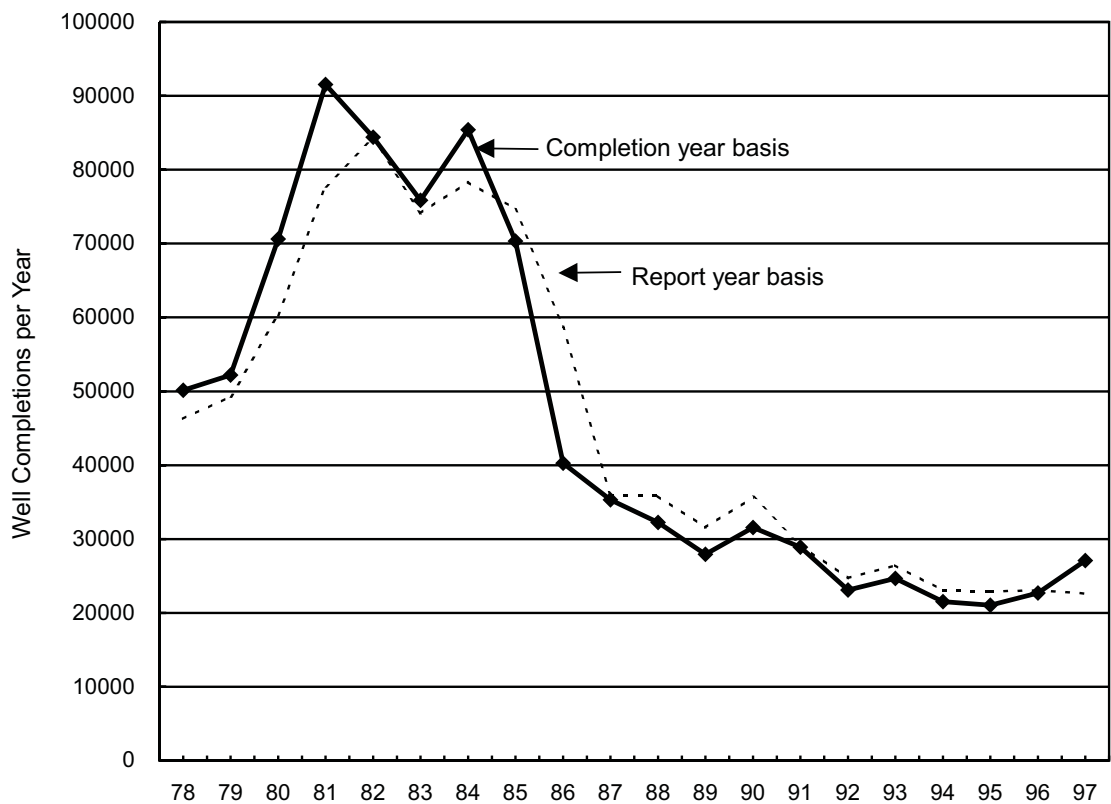
The conversion to completion date required the development of a statistical model to adjust the as-received completion data for incomplete counts due to the time lags between completion and reporting. The reporting lags result in cumulative as-received well counts that are quite incomplete for most recent months. For example, records received in 1995 show that only 16 percent of the wells were reported in the same month that they are completed. More than 13 percent of the well completion records received were for wells completed more than 12 months prior. In fact, 5.5 percent of the records were for well completions more than 5 years old (Figure 2). These reporting lags affect each of the six subcategories (well type by well class), although the effects vary among the subcategories.

EIA's efforts to convert the well completion statistics from a report date to a completion date basis resulted in the Well Completion Estimation Procedure (WELCOM). WELCOM is a system that summarizes

1 *Monthly Energy Review*, EIA (DOE/EIA-0035), January 1984.

2 Reliance on the as-received well counts would have significantly masked the serious impact of the price decline on the industry. By that time, however, EIA had taken steps to convert its drilling activity series to an as completed basis, which avoided this problem.

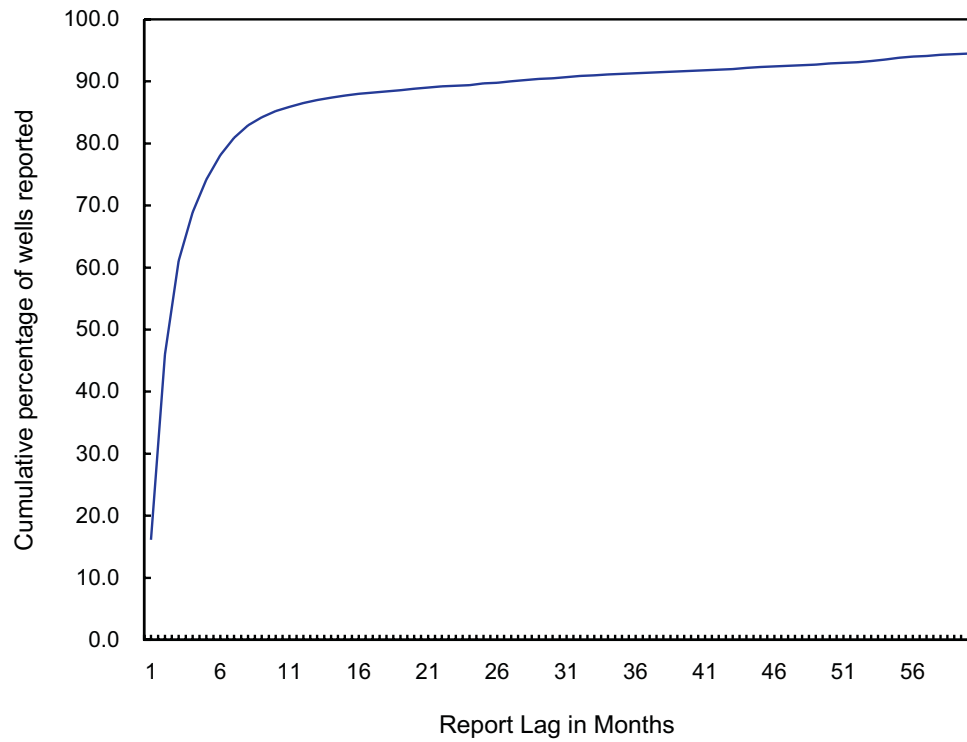
Figure 1. Well Counts by Completion Date and Report Date



Note: Well counts were published by EIA on an as reported basis until March 1985, and on a completion date basis thereafter. Wells by completion year include estimates.

Source: Energy Information Administration, Office of Oil and Gas.

Figure 2. Reporting Lags for Wells Reported in 1995



Source: Energy Information Administration, Office of Oil and Gas.

the historical records and produces estimates for drilling activity in the most recent years. Estimates are monthly at a national level for six major categories: oil-exploratory, oil-development, gas-exploratory, gas-development, dry-exploratory, and dry-development. Until 1982, the files received include few, if any, wells completed more than 5 years earlier. Thus, WELCOM used the cumulative recorded well counts, along with data on rotary rigs running, to construct estimates of the actual numbers of wells completions, and footage drilled, during the most recent 60 months of activity. EIA has used WELCOM since March 1985 to provide the drilling activity estimates published in the *Monthly Energy Review* and the *Annual Energy Review*, as well as other EIA reports. The same drilling data and estimates have also been a part of the foundation underlying numerous analytical efforts including the *Annual Energy Outlook*, the *Short-Term Energy Outlook*, *Performance Profiles of Major Energy Producers*, and *Natural Gas Issues and Trends*.

Problems in the data. Over the years, EIA's operation of WELCOM has included annual reestimation of the system coefficients as well as minor adjustments to the overall implementation of the methodology. The basic system remains essentially the same as that which was developed in 1984. Since then, the patterns and attributes of drilling have evolved as a result of various regulatory reform initiatives and the shift to generally lower oil and gas prices since the mid 1980s. The collection and processing system for the well records has itself changed. Other changes, such as the fact that the data for active rigs now identifies them by whether they are targeted to oil or gas, have offered opportunities to refine the estimation procedure and enhance the precision of the estimates.

A project to enhance the WELCOM procedure was undertaken by EIA in 1997. Its goal was to take maximum advantage of available data, modify the model specification as appropriate, and test alternate statistical approaches to the estimation of the model's coefficients. The associated data work had an unintended impact, however, when curious patterns in the reported well counts were noticed. Well counts can be and often are highly variable between months. Reported well counts for some months in 1995, however, were about half the counts of the prior and succeeding months.

Collaborative examination of the data with the vendor verified the existence of errors and omissions in the data files provided to EIA in these periods. Review of other time periods disclosed errors that were pervasive

in the data over an extensive period. The data files since 1987 were missing some records and contained duplicates of others, updates to many records were not passed along, and records for recompletions—which were not expected—were present in files for some months but not in others. A comparison of the well counts as initially received and the replacement files for 1990 to mid 1997 shows only slight discrepancies, with few exceptions, for the years immediately prior to 1995 when a significant number of records were not received (Figure 3).

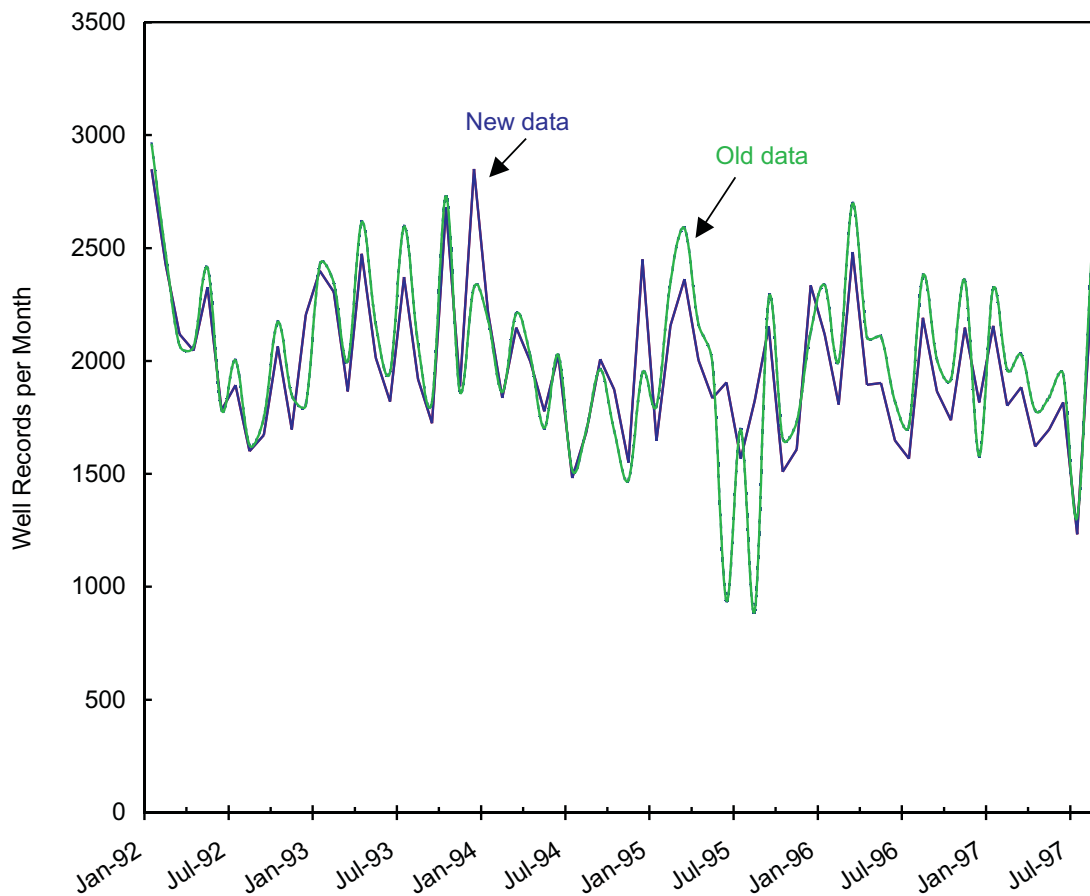
Attempted resolution of this problem required acquisition of replacement files from the vendor, which unfortunately contained data with an improper allocation of wells between exploration and development. The preferred well classification is based on the *final* designation of wells as either exploratory or developmental. The replacement records were identified based on the *initial* well class. The major impact of this slip is that wells initially classified as exploratory should have been classified as developmental, and this occurs more often for successful oil and gas wells than it does for dry holes. The resulting measures reflected exploratory well counts that were too high, developmental wells that were too low, and the oil and gas well counts were distorted more than those for dry wells.

Recognition of the second problem occurred after revised well counts had been published for the five months, February to June 1998. Updated drilling data were not published in July 1998, while EIA staff acquired corrected files from the vendor and re-estimated the WELCOM coefficients. This effort was accomplished in time for EIA to resume publication of drilling activity data in August 1998. The revised series exhibit certain key differences from the prior series, which may alter the perception and understanding of the industry.

COMPARATIVE ASSESSMENT OF THE REVISED AND PRIOR ESTIMATES

Differences between the prior and revised data series. The impact of flawed data files is not limited to the period in which they were received. Due to the extensive reporting lag, many of the records report on much earlier months, affecting earlier records also. For example, even though 1981 is outside the period in which problem data files were received, drilling peaked in that year at 91,553 as opposed to the previous measure of 90,034 (Tables 1.a and 1.b).

Figure 3. Monthly Well Counts as Reported, 1992-1997



Source: Energy Information Administration, Office of Oil and Gas.

Table 1. Oil Wells, Gas Wells, and Dry Holes, Estimates Before and After Data Correction.

1.a. Estimates prior to data correction.

	Exploratory				Developmental				All Wells			
	Oil	Gas	Dry	Total	Oil	Gas	Dry	Total	Oil	Gas	Dry	Total
1978	1191	1792	8054	11037	17874	12613	8537	39024	19065	14405	16591	50061
1979	1335	1920	7478	10733	19368	13250	8560	41178	20703	15170	16038	51911
1980	1781	2094	9035	12910	30497	15129	11302	56928	32278	17223	20337	69838
1981	2667	2533	12297	17497	40176	17374	14987	72537	42843	19907	27284	90034
1982	2470	2168	11346	15984	36672	16776	15036	68484	39142	18944	26382	84468
1983	2113	1660	10271	14044	35086	12896	14065	62047	37199	14556	24336	76091
1984	2335	1599	11482	15416	40250	15413	14315	69978	42585	17012	25797	85394
1985	1879	1282	9445	12606	33142	12970	11763	57875	35021	14252	21208	70481
1986	988	733	5511	7232	17713	7402	7255	32370	18701	8135	12766	39602
1987	859	673	5179	6711	15327	7084	6302	28713	16186	7757	11481	35424
1988	792	663	4766	6221	12530	7575	5476	25581	13322	8238	10242	31802
1989	580	654	4001	5235	9759	8571	4490	22820	10339	9225	8491	28055
1990	628	641	3855	5124	11522	10064	4757	26343	12150	10705	8612	31467
1991	573	542	3393	4508	11335	8910	4521	24766	11908	9452	7914	29274
1992	506	423	2656	3584	8517	7668	3995	20181	9023	8091	6651	23765
1993	485	514	2514	3513	8244	9350	4214	21808	8729	9864	6728	25321
1994	614	777	2203	3594	6166	8200	3070	17436	6780	8977	5273	21030
1995	734	835	1960	3529	6144	6534	2448	15126	6878	7369	4408	18655
1996	822	943	2180	3945	7275	8412	3108	18795	8097	9355	5288	22740
1997	904	856	2352	4112	7134	9424	3656	20214	8038	10280	6008	24326

1.b. Estimates after both data correction efforts.

	Exploratory				Developmental				All Wells			
	Oil	Gas	Dry	Total	Oil	Gas	Dry	Total	Oil	Gas	Dry	Total
1978	1171	1771	7965	10907	18010	12642	8586	39238	19181	14413	16551	50145
1979	1321	1907	7437	10665	19530	13347	8662	41539	20851	15254	16099	52204
1980	1764	2081	9039	12884	30875	15252	11599	57726	32639	17333	20638	70610
1981	2636	2514	12349	17499	40962	17652	15440	74054	43598	20166	27789	91553
1982	2431	2125	11247	15803	36768	16854	14972	68594	39199	18979	26219	84397
1983	2023	1593	10148	13764	35097	12971	14005	62073	37120	14564	24153	75837
1984	2197	1521	11278	14996	40408	15606	14403	70417	42605	17127	25681	85413
1985	1679	1189	8924	11792	33439	12979	12132	58550	35118	14168	21056	70342
1986	1084	793	5549	7426	18013	7719	7112	32844	19097	8512	12661	40270
1987	925	753	5049	6727	15239	7302	6052	28593	16164	8055	11101	35320
1988	855	730	4691	6276	12781	7825	5350	25956	13636	8555	10041	32232
1989	607	707	3924	5238	9597	8832	4264	22693	10204	9539	8188	27931
1990	653	691	3715	5059	11545	10353	4594	26492	12198	11044	8309	31551
1991	593	538	3312	4443	11177	8988	4287	24452	11770	9526	7599	28895
1992	496	424	2510	3430	8261	7785	3608	19654	8757	8209	6118	23084
1993	500	544	2470	3514	7865	9461	3825	21151	8365	10005	6295	24665
1994	566	720	2400	3686	6124	8818	2879	17821	6690	9538	5279	21507
1995	541	567	2198	3306	7086	7770	2877	17733	7627	8337	5075	21039
1996	481	560	2130	3171	7820	8625	3090	19535	8301	9185	5220	22706
1997	418	538	2080	3036	9892	10755	3418	24065	10310	11293	5498	27101

Source: Table 1.a: Estimates prior to data correction are from the Monthly Energy Review (MER), August 1997. Estimates for 1997 are double the figure for the first 6 months. Table 1.b: Estimates after data correction match those published in the MER, August 1998

Another significant difference between the two series arises in the drilling pattern of the mid 1990s. The estimates based on files with missing records showed that the decline in oil and gas prices in 1993 to 1995 contributed to annual well completion declines of 17 percent in 1994, then 11 percent in 1995 to a level of only 18,655 total completions in 1995. The 1996 price surge seemed to drive a 22 percent rise in drilling activity to 22,740 wells. The revised drilling statistics, however, show that the relative worsening of economic conditions resulted in a decline in wells completed from 24,665 to 21,507, followed by a further decline to 21,039 in 1995. The revised series shows 2,384 more wells were drilled in 1995 than previously estimated. The smaller decline in 1995 wells also results in a smaller relative recovery in response to the rise in both oil and gas prices in 1996, 8 percent in contrast to the 22 percent change indicated by the prior series. The 1993-1995 decline of 15 percent in the corrected data is much less than the 26 percent falloff seen in the initial well series. Thus, it appears that the initial estimates overstated the industry reaction in both directions. This casts a much different light on the industry's responsiveness to the economics of the time.

In addition to changes in total well counts, the revised data series contains fewer oil and gas exploratory well completions in recent years than previously estimated, distorting the implied success rates. The shift in recent successful well counts exceeds the corresponding differences in dry holes, so the associated success rates do not increase as much as previously indicated. According to both series, the 1979 peak success rate of 30 percent was almost matched by 1993, and was surpassed in the following year. Thereafter, the initial and revised series exhibit very different trends in exploratory success. In the initial series, success rates increased strongly from 1993, reaching 45 percent in 1996. The final revised series shows improved industry performance in the mid 1990s, but to a lesser degree, with success rates after 1993 ranging between 32 and 35 percent. Thus, the corrected drilling activity data provide a different picture of industry performance and the investment incentives for firms engaged in exploratory activities (Figure 4).

The trends in success rates in both the initial and final series stand in stark contrast to the trend indicated by the exploratory drilling estimates published by EIA from February through June 1998. This intermediate series exhibited sizeable overestimates of the oil and gas exploratory wells for an extended period beginning in the mid 1980s. The mistakenly high successful well counts caused an excessively high exploratory success

rate beginning in 1985 and exceeding the 40 percent level from 1989 forward. Increases in success rates of such a magnitude and at such an early date suggested by these data would be grossly misleading for any analyst relying on these measures. The surge in the trend coincides with the virtual collapse of oil and gas prices. Interpretation of these trends and conditions would likely distort one's perspective of the industry capability to react to changes in economic conditions.

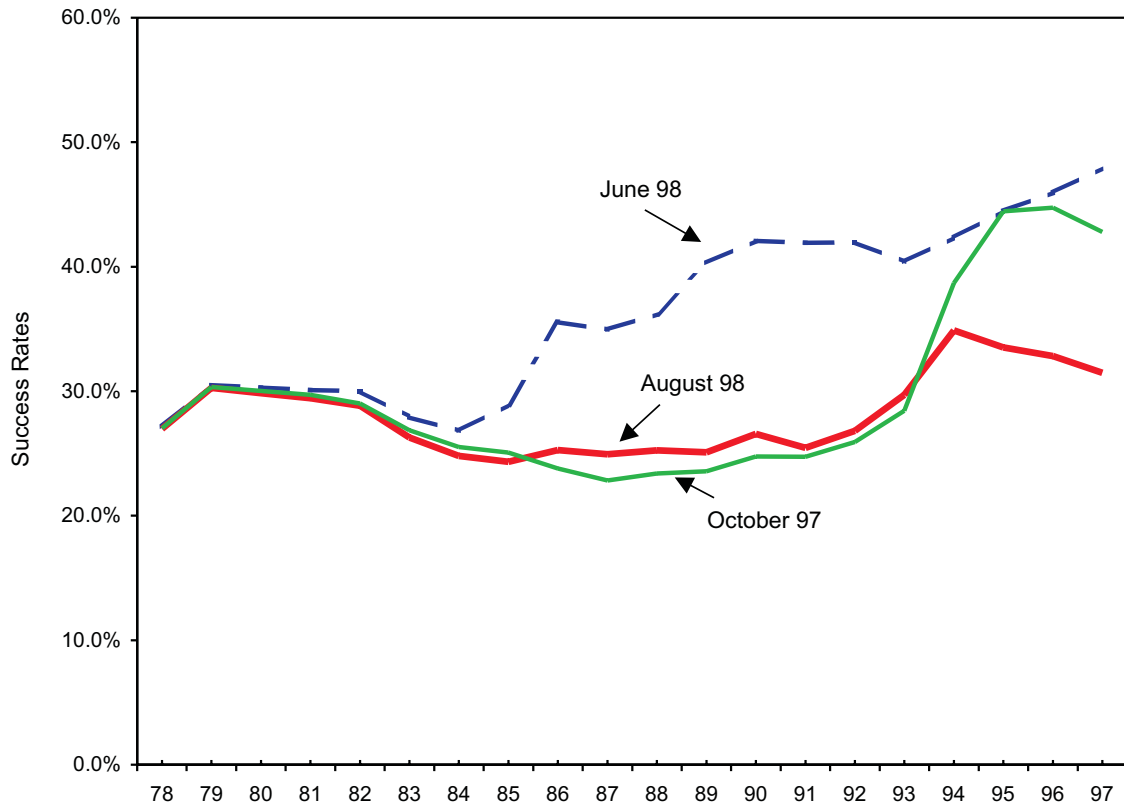
Comparison of the final corrected drilling activity series with the initial erroneous series shows that the overall trends are comparable, with a few significant exceptions. Most aspects of the drilling series are consistent between the initial and final drilling activity data series. However, the implications of the differences in drilling levels for certain years and the new values for exploratory success rates are fundamental to an understanding of the industry.

Implications of the data shift. The significance of the changes is likely to depend on the particular use of the data. EIA itself uses these data to support analysis integral to two prominent information products: the *Annual Energy Outlook (AEO)* and the *Short-Term Energy Outlook*. The drilling data are specifically used in the National Energy Modeling System (NEMS), which is the tool used to produce the integrated energy market projections that are published in the *AEO*.

The productivity of exploratory drilling is represented in the NEMS by finding rates, which are measured as the ratio of oil and gas reserve additions to wells completed. The downward shift in the exploratory well counts lowers the finding rate denominator, so the finding rate will shift upward correspondingly. This productivity increase would raise the available market supplies in future years, all else being equal. However, the corresponding analysis that determines the level of drilling must also be revised, because it is predicated on a high exploratory well count. Thus, the expected drilling response under varying conditions is overestimated in the projection because it is calibrated to an inaccurately high benchmark. When these functional relations are reestimated, drilling response should be less in the future than previously expected. Given that the changes in exploratory drilling primarily affect only the last few years of data, the net result of these two offsetting influences on the outlook is likely to be small.

The long-term supply outlook is also influenced by the degree of success in the search for oil and gas. The lower success rates affect the economic attractiveness of exploratory drilling opportunities. Dry holes represent an unavoidable part of the search for oil and gas that add

Figure 4. Comparison of Exploratory Success Rates by Publication Date



Source: Energy Information Administration, Office of Oil and Gas.

to project cost, lowering the expected profitability of exploratory projects. Lower average success rates imply that the total drilling costs of an average exploration project should be greater, which suggests that the economically recoverable portion of the technically recoverable resource base may not be as large as previously anticipated. The volumetric impact of this shift should be limited, however, because this change will mostly impact oil and gas accumulations that are marginal in size anyway. The economic merit of massive fields that provide the bulk of expected market supplies should remain largely unaffected.

Maintaining data quality in the future. Although EIA conducts a wide range of data quality activities to monitor its data systems, the long-lived problems affecting the drilling data series maintained by EIA are a clear indication that further efforts in this regard are necessary. In a typical EIA data system, ready detection of large discrepancies would be virtually certain, however the well completion data system has a number of attributes that hinder detection. First, the work with the raw data and all initial processing are conducted outside EIA. Such an ‘arms-length’ relationship with the basic data does not facilitate familiarity with the data, and it hampers efforts to investigate concerns about the data. Additionally, the well completion data are inherently quite erratic, so even sizeable fluctuations are not *prima facie* a cause for concern. Yet another factor that impedes ready detection of data problems is the statistically-based processing of the data by EIA. The limited numbers of wells reported early after completion are ‘inflated’ by large factors, which tend to obscure data problems. Last—and arguably paramount—there is no other timely source of information that can be used to validate the data going into or the set of estimates coming out of WELCOM.

Development of new approaches for quality assurance will proceed along a number of paths including closer monitoring of the as-received well completion data and comparison of estimated drilling data by completion month to other data that are available on a timely basis. An indication of the difficulty in establishing any definitive set of guidelines can be seen in the data for wells reported each month. While the well counts in the majority of months vary by 20 percent or less, fluctuations exceeding 30 percent are common. Even relative shifts beyond 40 percent are frequent enough to not be necessarily considered a cause for alarm (Figure 5).

EIA is working to more closely monitor the drilling data through collaboration with its counterparts at the

API. API receives drilling data records from the same vendor as EIA and also generates statistically adjusted estimates for drilling activity by completion date. EIA and API have been exchanging their latest estimates as soon as available to facilitate comparative analysis. In addition to ongoing communication, API and EIA intend to meet periodically to review the status of data and the estimation process. More frequent working meetings will be held on an as-needed basis. While EIA and API may often encounter similar pitfalls in their drilling data efforts, they receive the raw drilling records according to their own specifications and on different reporting cycles. The data files as received are not generated by the same process by the vendor, so future data problems may become apparent early as discrepancies in the data files or estimates are noted.

Drilling-related data are used to evaluate the estimates coming from WELCOM. These data include rotary rigs running, seismic crews, and oil and gas prices. EIA staff is also considering the possibility of obtaining relevant data that may exist at the State level in the major producing States to corroborate the EIA estimates. Drilling-related measures from the major States, while incomplete in themselves, could prove to be useful as a benchmark, since most drilling occurs in a limited number of States.

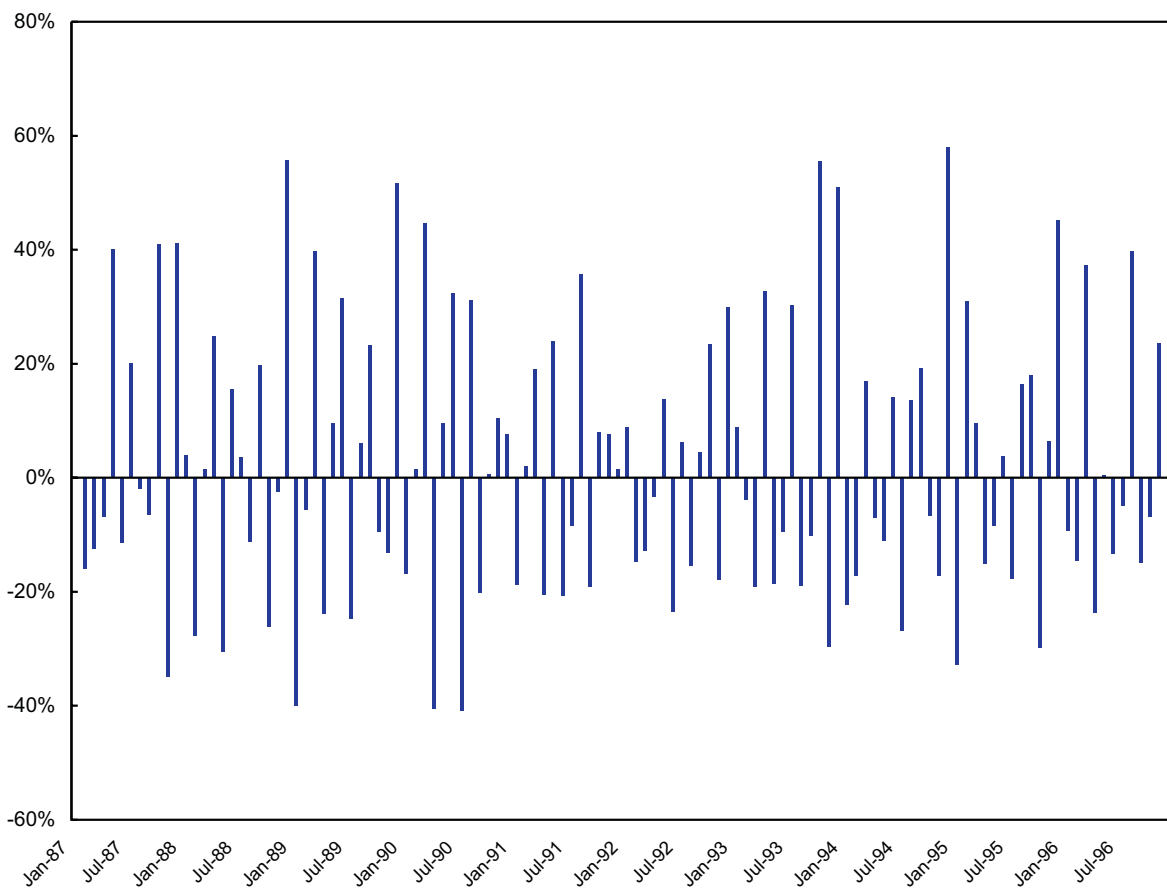
These enhanced monitoring actions will help to avoid future data problems that might require extensive revision of the series.

CONCLUSIONS

EIA has inadvertently been publishing erroneous estimates of oil and gas well drilling activity since the late 1980s. Problems in the raw data obtained from a vendor were detected in late 1997, and again in mid 1998, and have since been resolved. Monthly publication of the corrected EIA drilling activity estimates series resumed with the August 1998 edition of the *Monthly Energy Review*. The major characteristics of the corrected series are comparable in most regards with the initial erroneous series. There are, however a few exceptions which directly bear on fundamental aspects of the industry such as its responses to changing economic conditions and its approach to drilling opportunities.

EIA does not collect the raw data itself, so some data errors may remain extremely difficult to discover, identify, and remedy in a timely manner. However, EIA will make every possible effort to ensure that the same data problem, or a data problem of similar magnitude, does not happen again.

Figure 5. Change in Wells Reported Compared to Prior Month



Source: Energy Information Administration, Office of Oil and Gas.