

Effects of Individual Site Characteristics

The RPMs' responses to the questions on fast and slow sites provide important information about the types of problems that lead to slow cleanups and the advantages that promote fast cleanups, but they do not go very far in identifying the types of sites at which these problems or advantages are likely to occur. Outlined below are findings (for nonfederal facilities) on a site's likelihood of being fast or slow, and of being slow because of inherent site problems or enforcement and legal problems, as a function of specific site characteristics.¹⁵ These findings cannot be considered definitive because any characteristic examined in isolation may be serving as a proxy for more fundamental causes, but they can help suggest important areas for further study.

Effects on Site Duration. Of the universe of 1,124 nonfederal sites, the RPMs identified 25 percent as slow and 40 percent as fast, as shown in Table 6; the speed of the remaining 35 percent was either intermediate or unknown.¹⁶ To summarize the findings, an above-average chance of slow cleanup is found among sites proposed before 1984; sites with more than two operable units, 200 acres, or 50,000 cubic yards of waste; sites involving dry cleaning, chemical manufacturing, mining land, or waste disposal to a lake or river; sites with more than 10 PRPs; and sites located in New England. Fast sites are disproportionately found among sites proposed after 1983, sites with fewer than 10 PRPs, and orphan sites.

As expected, given the average distributions presented in Table 2, sites proposed for the NPL between 1981 and 1983 and those with three or more operable units have greatly above-average chances of being slow--41 percent and 48 percent, respectively. These characteristics may have a direct influence on the duration of cleanup, if EPA's inexperience and start-up problems delayed the progress of early sites and if dividing a site into several operable units itself leads to a lengthier cleanup schedule. However, they are also likely to be serving, at least in part, as proxies for other characteristics leading to early listing or subdivision into multiple units. Again, the correlation with the year of listing on the proposed NPL may also be partly or wholly an artifact of excessive RPM optimism.

15. Sites owned by the federal government (typically military bases or Department of Energy facilities) are excluded from the analysis because their problems of cleanup length are often different from those encountered at nonfederal sites.

16. The RPMs' classification of the 125 federal facilities is 31 percent slow sites, 35 percent fast sites, and 34 percent sites of medium or unknown cleanup time.

Other, more exogenous characteristics related to the size of a site and the complexity of its contamination problem also show correlations with the distribution of cleanup times. The acreage of a site is highly relevant, with almost half of the sites exceeding 200 acres classified as slow. Estimated waste quantity has a noticeable but less dramatic effect: the percentage of slow sites is roughly 7 points higher for sites with more than 50,000 cubic yards of waste than for sites with less than that amount. The data on acreage and waste quantity used here are from EPA's NPL Characterization Project, which contains information collected as sites were screened for placement on the NPL. Data on waste quantity for many sites were not available from the Characterization Project, which explains why the figures for the sites with data on waste do not average out to those for all 1,124 nonfederal sites.

TABLE 6. DISTRIBUTIONS BY DURATION OF NONFEDERAL NPL SITES WITH SELECTED CHARACTERISTICS

| | Number of Sites | Duration Distribution (In percent) | | |
|--|-----------------|------------------------------------|------|-------------------------|
| | | Slow | Fast | Intermediate or Unknown |
| All Nonfederal Sites | 1,124 | 25.1 | 39.9 | 35.0 |
| Proposed During 1981-1983 | 538 | 41.1 | 22.9 | 36.1 |
| Proposed During 1984-1992 | 584 | 10.4 | 55.7 | 33.9 |
| Three or More Operable Units | 145 | 48.3 | 20.7 | 31.0 |
| Site Area | | | | |
| More than 200 acres | 89 | 47.2 | 21.3 | 31.5 |
| Not more than 200 acres | 919 | 23.8 | 41.3 | 34.8 |
| Waste Quantity | | | | |
| More than 50,000 cubic yards | 157 | 31.8 | 33.8 | 34.4 |
| Not more than 50,000 cubic yards | 691 | 25.2 | 39.9 | 34.9 |
| Past or Current Site Use Includes | | | | |
| Dry cleaners, laundries | 27 | 51.9 | 22.2 | 25.9 |
| Chemical manufacturing | 151 | 39.1 | 36.4 | 24.5 |
| Metal mining | 30 | 33.3 | 26.7 | 30.0 |
| Mixed-waste landfills | 127 | 26.8 | 31.5 | 41.7 |
| Contamination Included Lake/River Disposal | 47 | 42.6 | 17.0 | 40.4 |

(Continued)

The type of land uses associated with a site are sometimes implicated in the length of its cleanup. As Table 6 shows, 52 percent of the sites used partly or entirely by dry cleaners and other laundry businesses are slow, more than twice the national average. The small number of sites involved, however, suggests special caution in interpreting this finding. In fact, the very large areas of contamination in the San Fernando and San Gabriel valleys in California, which are divided into four NPL sites each and received contamination from large numbers of sources, account for 8 of the 14 slow sites at which the land use included dry cleaners. Sites involving chemical manufacturing and metal mining are also more likely to be slow; the sample of mining sites is small, however. Mixed-waste landfills--that is, landfills containing both industrial waste and municipal solid waste--are similar to

TABLE 6. CONTINUED

| | Number of Sites | Duration Distribution (In percent) | | |
|---|--|---------------------------------------|------|-------------------------|
| | | Slow | Fast | Intermediate or Unknown |
| Contamination Includes | | | | |
| Dioxins, PCP | 54 | 29.6 | 40.7 | 29.6 |
| Asbestos | 42 | 28.6 | 35.7 | 35.7 |
| PCBs | 210 | 27.6 | 36.7 | 35.7 |
| Metals | 677 | 25.0 | 40.3 | 34.7 |
| Lead | 382 | 20.7 | 42.4 | 36.9 |
| Number of PRPs | | | | |
| Zero to 10 | 656 | 20.6 | 47.9 | 31.6 |
| More than 10 | 400 | 31.3 | 28.0 | 40.8 |
| Orphan Sites | 88 | 10.2 | 59.1 | 30.7 |
| Number of PRPs Contributing Less Than 1 Percent of Waste | | | | |
| Zero to 10 | 532 | 22.6 | 44.4 | 33.1 |
| More than 10 | 113 | 31.9 | 31.0 | 37.2 |
| Located in EPA Region 1 | 77 | 39.0 | 19.5 | 41.6 |
| SOURCE: | Congressional Budget Office. | | | |
| NOTES: | NPL = National Priorities List; PCP = pentachlorophenol; PCBs = polychlorinated biphenyls; PRP = potentially responsible party; EPA = Environmental Protection Agency. | | | |

all nonfederal sites in their percentage of slow sites, but are less likely to be fast and more likely to be in the remaining "intermediate or unknown" duration category.

As for the types of contamination involved at a site, the distribution by duration category of the 47 sites at which some contamination was caused by disposal to a lake or river is tilted sharply toward slow sites and away from fast sites compared with the national average. The data do not show any significant correlations between the presence of dioxins or PCPs and a site's duration category; the same is true for asbestos, PCBs, and metals in general. Sites involving lead contamination appear somewhat less likely to be slow.¹⁷

Another marked correlation exists between the distribution of durations and the number of potentially responsible parties. Sites with more than 10 PRPs are half again more likely to be slow and 40 percent less likely to be fast than sites with zero to 10 PRPs. Orphan sites--that is, sites with no identifiable PRPs capable of contributing to the cost of cleanup--show the strongest tilt away from slow durations and toward fast durations of any group of sites examined in the table. The number of potential *de minimis* parties does not seem, however, to have a distinct effect: sites at which more than 10 PRPs each contributed less than 1 percent of the waste show roughly the same distribution as all sites with more than 10 PRPs.

Finally, the RPMs in EPA's Region 1 (New England) expect a much higher fraction of slow cleanups at their sites than do all RPMs nationwide (39 percent as against 25 percent) and a much lower fraction of fast cleanups (19 percent as against 40 percent).¹⁸ This correlation could be an artifact of the data, if Region 1's sites are atypical in other ways that affect duration or if its RPMs simply took a less optimistic view of Superfund's future progress.

Interestingly, however, Region 1 reported a disproportionate number of funding and staffing constraints. The region accounts for less than 7 percent of all nonfederal NPL sites but 20 percent of the sites at which RPMs cited these constraints as a factor leading to slow cleanup (16 cases out of a total of 82) and 43 percent of the sites for which RPMs identified either constraint as the primary factor (3 of 7 funding problems and 6 of 14 staffing problems).

17. One characteristic not reported in Table 6 is the presence of groundwater contamination. Because such contamination is so common, having been identified at 81 percent of all nonfederal sites, it has little power to explain why some sites take longer than average.

18. For all other regions, the share of long sites was 31 percent or less. At the other end of the spectrum from Region 1, Region 6 (Arkansas, Louisiana, New Mexico, Oklahoma, and Texas) and Region 7 (Iowa, Kansas, Missouri, and Nebraska) reported 15 percent and 14 percent long sites and 65 percent and 68 percent short sites, respectively, on the basis of 66 and 56 total sites.

Without the nine slow sites involving primary funding or staffing problems, the share of slow sites in Region 1 would be similar to the national average.

Effects on Intrinsic Site Difficulties. EPA's cleanup managers frequently cited intrinsic site difficulties, such as size and complexity, as factors contributing to slow site cleanups. As Table 7 shows, the RPMs identified such difficulties as primary or secondary explanations at 45 percent of all 282 slow nonfederal NPL sites, and as primary explanations at 31 percent.

Two site characteristics that were shown above to be correlated with slow cleanups are nonetheless not positively associated with the perceived difficulty of slow sites. Slow sites proposed for the NPL between 1981 and 1983 are no more likely to be judged difficult than their later counterparts. This result suggests that site difficulty per se is at best part of the reason why a higher share of early sites have long cleanup times; again, other possible reasons include EPA inexperience and start-up problems in the program's initial years and underestimation by RPMs of the length of sites added after 1983. As Table 7 also shows, slow sites in EPA's Region 1 are distinctly less likely to be considered difficult than those in other regions; this result is expected, since Region 1 RPMs attributed more of their slow sites to funding and staffing problems.

In the remaining cases, the expected pattern generally prevails: physical and technical site characteristics that were found to be strongly correlated with slow cleanups are also correlated with a slow site's perceived difficulty. Slow sites with three or more operable units are much more likely to be considered intrinsically difficult, as are those with more than 200 acres or 50,000 cubic yards of waste. Site difficulty is either a primary or secondary explanation at *all* slow sites involving dry cleaners and mining, and a primary explanation in almost all the cases. Again, however, the samples of sites included in these categories are small. A larger sample suggests that chemical manufacturing is correlated with the difficulty of slow sites, not just with site duration itself. Conversely, mixed-waste landfills, which are not associated with slow cleanups, also show no association with the difficulty of slow sites.

Site difficulty is almost uniformly considered a primary or secondary explanation for slow cleanups in the relatively few cases involving waste disposal to a lake or river. Two other indicators of surface-water hazards are introduced here: the presence of widespread sediment contamination is strongly correlated with the difficulty of slow sites, but the more general existence of an impact on surface water is not. The small number of slow sites with contamination by dioxins or PCP are more likely than the average slow site to be difficult, even though these substances did not show a major



**TABLE 7. CORRELATIONS BETWEEN INTRINSIC DIFFICULTY AND
SELECTED CHARACTERISTICS OF SLOW SITES**

| | Number of Sites | Percentage with Intrinsic Difficulties | Percentage with Primary Intrinsic Difficulties |
|---|--------------------|--|---|
| All Slow Nonfederal Sites | 282 | 45.0 | 30.9 |
| Proposed During 1981-1983 | 221 | 44.8 | 30.3 |
| Proposed During 1984-1992 | 61 | 45.9 | 32.8 |
| Three or More Operable Units | 70 | 80.0 | 60.0 |
| Site Area | | | |
| More than 200 acres | 42 | 85.7 | 71.5 |
| Not more than 200 acres | 219 | 36.1 | 21.9 |
| Waste Quantity | | | |
| More than 50,000 cubic yards | 50 | 60.0 | 48.0 |
| Not more than 50,000 cubic yards | 174 | 36.8 | 22.4 |
| Past or Current Site Use Includes | | | |
| Dry cleaners, laundries | 14 | 100.0 | 85.7 |
| Chemical manufacturing | 59 | 67.8 | 45.8 |
| Metal mining | 10 | 100.0 | 90.0 |
| Mixed-waste landfills | 34 | 38.2 | 29.4 |
| Contamination Included Lake/River Disposal | 20 | 95.0 | 70.0 |
| Widespread Sediment Contamination | 14 | 85.7 | 50.0 |
| Surface Water Impacts | 181 | 44.8 | 31.5 |
| Contamination Includes | | | |
| Dioxins, PCP | 16 | 62.5 | 56.3 |
| Asbestos | 12 | 41.7 | 33.3 |
| PCBs | 58 | 41.4 | 27.6 |
| Metals | 169 | 43.8 | 30.2 |
| Lead | 79 | 44.3 | 26.6 |
| Located in Region 1 | 30 | 33.3 | 23.3 |

SOURCE: Congressional Budget Office.

NOTES: PCP = pentachlorophenol; PCBs = polychlorinated biphenyls.



correlation with site duration itself. The presence of asbestos, PCBs, metals in general, or lead has no bearing on a slow site's perceived difficulty.

Effects on Enforcement and Legal Problems. Enforcement problems are cited nearly as often as intrinsic site difficulties as primary explanations at slow nonfederal NPL sites (29 percent of the time), and even more often as primary or secondary explanations (57 percent of the time). As Table 8 shows, many factors appear to predispose slow sites to having enforcement problems: the existence of more than 10 PRPs (or even more so, more than 50 PRPs), more than 10 PRPs that contributed less than 1 percent of the site's waste, the share of waste from off-site orphan parties exceeding 12 percent, past or current land use including mixed-waste landfills or chemical manufacturing, municipal solid waste exceeding 50 percent, and expected PRP costs exceeding \$15 million. Dry cleaning and mining land uses have little bearing on the probability that a slow site has significant enforcement problems and a negative impact on the likelihood that such problems are the primary explanation. This result is expected, given that intrinsic difficulties are often identified as the primary explanation at these sites.¹⁹

The characteristics examined in Table 8 highlight the limitations of single-factor analysis, in that many of the investigated characteristics are likely to be correlated with each other. For example, sites with many PRPs are more likely to have many PRPs who contributed small shares of waste, and sites with a majority of municipal solid waste are likely to be mixed-waste landfills. Given such correlations, multivariate regression techniques are required if analysis is to distinguish the truly causal factors from the indirect proxies.

Multiple Regression Analysis of Site Duration

In essence, regression analyses attempt to find the mathematical line or curve of a given type that best fits the given data. By examining multiple explanatory factors at the same time, such analyses can estimate the independent effects of each individual factor when all other factors remain constant. CBO has conducted only the most preliminary regression analysis. Nonetheless, the available results do offer some qualified support for the findings of the single-variable correlations and suggest some potentially important topics for further study.

19. Waste type and quantity, among the characteristics examined in previous tables, are omitted here on the theory that they are unlikely to affect a site's contentiousness except through their impact on cleanup costs.

dichotomous data.²⁰ Other regressions analyzed estimated durations themselves, using the ordinary least squares (OLS) technique. The two measures of duration are not fully consistent: 66 of the 282 slow nonfederal sites identified by the RPMs appear to fit the definitions of medium or fast sites, based on the available listing dates, completion dates, and numbers of operable units. Time constraints precluded a site-by-site reconciliation of the two types of data.

Most of the factors that were shown in Table 6 to be individually correlated with the distribution of sites by duration also appear to be statistically significant--that is, to be very likely to have non-zero impacts--in one or both types of regression analysis (see Table 9). Both the OLS and logit regressions indicate that orphan sites tend to be faster and that sites proposed early, divided into several operable units, large in area, or involving dry cleaning or chemical manufacturing land uses are more likely to be slow. Again, because of the limited sample, the results on sites involving dry cleaning should be considered particularly tentative. The OLS analysis also suggests that larger waste volumes and the presence of more than 10 PRPs increase a site's duration, and the logit analysis suggests that the presence of widespread sediment contamination increases a site's chance of being slow.

One finding of note is that sites proposed for the NPL through 1983 and those with several operable units still appear significantly different in duration, even after correcting for site area, waste quantity, existence of surface water contamination, number of PRPs, and other factors. This finding has three possible explanations, one of which applies to the NPL proposal year but not to the number of operable units.

First, these two characteristics could still be serving as proxies for other factors or combinations of factors not included in the regressions.²¹ Second, as discussed above, the connection with the proposal year could be an artifact of the data, resulting from the PRPs' underestimating the difficulties facing Superfund in the future. Third, a site's year of proposal and number of units could truly have direct impacts on its cleanup duration. To the extent that dividing a site into several operable units itself lengthens the time required to reach construction completion, the tension between EPA's current strategy

20. That many of the explanatory variables considered in CBO's analysis, like the slow/not-slow outcome variable, are also dichotomous suggests that discriminant analysis might produce better estimates than those from the logit regressions. Further research is necessary to answer this question.

21. When the variable indicating whether a site's proposal year is "early" or "late" is removed from the equation, the variable indicating the presence of contaminated surface water becomes more significant. This result suggests that many early sites had surface water problems, but does not explain why earliness itself appears to be a better predictor of long cleanup duration.

of dividing cleanup into short-term response work and long-term remediation, on the one hand, and the goals of reducing cleanup duration and achieving site completions, on the other hand, may be greater than expected.

Two variables that are individually correlated with the distribution of site durations--disposal to a lake or a river as a source of contamination and mining as a past or current land use--do not appear to be significant in the multivariate regressions. The lack of impact of the mining variable suggests that it was serving as a proxy for site size in the individual correlations.

TABLE 9. REGRESSION RESULTS ON CLEANUP DURATION

| OLS Regressions on Duration in Years | Logit Regressions on Probability That Site Is "Slow" |
|--|--|
| Characteristics Found to Be Significantly Associated with Long Cleanups | |
| Early Proposal Year | Early Proposal Year |
| Several Operable Units | Several Operable Units |
| Large Size (Acreage) | Large Size (Acreage) |
| Large Waste Volume | |
| | Widespread Sediment Contamination |
| Land Use Including Dry Cleaning | Land Use Including Dry Cleaning |
| Land Use Including Chemical Manufacturing | Land Use Including Chemical Manufacturing |
| More Than 10 PRPs | |
| Viable PRPs (Not orphan site) | Viable PRPs (Not orphan site) |
| Characteristics Generally Found Not to Be Significant | |
| Waste Disposal to Lake or River | |
| Contamination Affecting Surface Water | |
| Land Use Including Mining | |
| Many PRPs Contributing Less Than 1 Percent of Waste | |
| Large Estimated Orphan Share from Off-Site Parties | |
| SOURCE: | Congressional Budget Office. |
| NOTES: | OLS = ordinary least squares; PRP = potentially responsible party. |

The regressions investigated the differences between each EPA region and Region 10 (arbitrarily chosen as the default category) rather than the difference between each region and the national average. In qualitative terms, both OLS and logit results suggest that sites in Regions 1, 4, and 5 (corresponding roughly to the New England, southeastern, and Great Lakes states) take longer than average, while sites in Regions 6, 7, and 8 (the south-central, central plains, and Rocky Mountain and Dakota states) take less time than average. CBO cannot identify which of these differences, if any, are statistically significant.

Further research into the explanations of cleanup duration could benefit from the use of other statistical techniques, more variables, and a closer scrutiny of the underlying data. Such research would help not only to confirm or refute the present qualitative findings of significance, but also to quantify the size of the effects. Additional analysis could also help clarify the characteristics that make a site slow because of intrinsic difficulties or enforcement problems.

