NBER WORKING PAPER SERIES

THE PACE OF PROGRESS AT SUPERFUND SITES: POLICY GOALS AND INTEREST GROUP INFLUENCE

Hilary Sigman

Working Paper 7704
http://www.nber.org/papers/w7704

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2000

I am grateful to Kerry Knight and Andres Lerner for research assistance. I would like to thank Trudy Ann Cameron for her guidance and Howard Chang, Kate Probst, Kip Viscusi, and anonymous referee, and seminar participants at the University of Houston, Columbia, the National Bureau of Economic Research, Resources for the Future, and Yale for comments. This research was supported in part by grant R82-2368 from the Office of Exploratory Research of the U.S. Environment Protection Agency and a faculty research fellowship from the UCLA Center for American Politics and Public Policy. I conducted much of this research while a Gilbert White Fellow at Resources for the Future. The views expressed herein are those of the author and not necessarily those of the National Bureau of Economic Research.

The Pace of Progress at Superfund Sites: Policy Goals and Interest Group Influence
Hilary Sigman
NBER Working Paper No. 7704
May 2000
JEL No. Q2, K32

ABSTRACT

Bureaucracies may set priorities for their workload according to social goals or the desires of concentrated private interests. This paper explores bureaucratic priorities empirically by studying Superfund, the federal program for cleaning up contaminated sites. It examines the amount of time that sites on Superfund's National Priorities List require to complete three states from listing to cleanup, using an econometric method for multiple sequential durations. The empirical results provide little evidence that the EPA prioritizes sites according to their harms. By contrast, concentrated private interests, such as liable parties and local communities, play an important role in the EPA's priorities. Delays caused by liable parties may reduce net benefits of cleanup by 8%. This result suggests a benefit from funding provision of environmental quality and other public goods through diffuse sources, such as broad-based taxes, to avoid the detrimental effects of such concentrated interests.

Hilary Sigman
Department of Economics
Rutgers University
75 Hamilton Street
New Brunswick, NJ 08901-1248
and NBER
sigman@econ.rutgers.edu

Bureaucracies, like other organizations, must solve the problem of how to prioritize their workload. They may chose priorities that reflect social goals. However, concentrated private interest groups may also manage to manipulate agency agendas, as the capture theory articulated by Stigler (1971) and Peltzman (1976) suggests. In addition, bureaucracies may respond to pressures imposed by legislators (Weingast and Moran, 1983).

To study bureaucratic priorities, this paper uses activities the Environmental Protection Agency (EPA) under Superfund, the federal program from cleanup of abandoned contaminated sites. There are a few reasons to choose the Superfund to study these issues. First, the Superfund program requires the agency to make site-by-site decisions at each of over a thousand sites. Thus, we have a large number of comparable observations of the agency's behavior to study econometrically. Second, one can easily connect private interest groups and political pressures to each site. At most sites, private parties pay the costs of the site through legal liability. This funding approach creates a concentrated interest group with a high stake in the Agency's decisions. In addition, activities at Superfund sites are distinctly local. Thus, it is possible to identify another important interest group, the local community, and to associate a Congressional representative with each site.

As a measure of the EPA's priorities for its work, the paper studies the speed with which the EPA moves sites through the Superfund. There are both direct and indirect motivations for focusing on this particular aspect of Superfund decision-making. The direct motivation is that the speed of Superfund is a perennial policy concern. By early 1997, 16 years after Congress enacted the legislation, only 11% of the sites on Superfund's National Priorities List (NPL) had been declared to be clean and therefore deleted from the NPL. President Clinton touted more rapid Superfund progress in his 1997 State of the Union Address, but the GAO (1997) assesses the program's recent pace as poor.

Focusing on the speed with which the EPA moves sites through Superfund also has analytical advantages. Speed provides a measure of the bureaucratic output, rather than an input, so it is closely related to social and interest-group goals. It also provides an indicator of the distribution of agency work effort that the Agency cannot readily manipulate for the sake of appearances. In addition, the Agency has considerable discretion in choosing which sites to expedite. The EPA has set targets for the number of sites to reach certain stages by various dates, but there is no official policy on which sites to prioritize.

This paper uses data on the record of progress at NPL sites through January 1997.[1] I estimate an econometric model of the length of time for sites to complete various stages of the Superfund process, using a model of multiple sequential durations. The paper focuses on three transitions: (1) proposed listing of the site on the National Priorities List, which marks the selection the site for federal cleanup; (2) signing of a Record of Decision, which reflects the completion of decision-making about a remedy; and (3) construction completion, which marks the end of most cleanup activities at the site. The model allows unobserved cross-site heterogeneity, perhaps associated with the complexity of the site. In the presence of unobserved heterogeneity, joint estimation of the duration of the various stages addresses the selection problem that arises because sites cannot make later transitions if they have not completed earlier stages.

The empirical results provide evidence that the EPA responds to interest groups in prioritizing its resources. Both the liable parties and the local communities get weight in the agency's priorities. Liable parties appear to delay progress at their sites. Sites without viable liable parties experienced 29% faster decision-making than sites with viable PRPs. Sites with deep-pocketed liable parties (where the private financial stake in the site is likely to be large) have slower cleanup. Similarly, powerful communities manage to expedite progress. Sites in communities with higher voter turnout received faster cleanup of their sites, while higher income communities receive faster listing.

This interest group influence may have a small but significant impact on social welfare. A rough calculation suggests that delays caused by liable parties may decrease the Superfund's net benefits by 8% (about \$.8 million per site or \$1 billion in total). Thus, the results support current Congressional proposals to reduce the role of liability financing (Reisch,

[1]NPL sites are chosen based on risk assessments and political factors (Hamilton and Viscusi, 1999a). Although NPL sites differ systematically from all contaminated sites, the NPL sites are the relevant universe for an analysis of large-scale Superfund cleanups.

1998), although other effects of liability funding also need to be taken into consideration for a full assessment (see the discussion in section 5 below). More generally, the results provide an argument for funding provision of environmental quality and other public goods from diffuse sources, such as broad-bases taxes or general appropriations, to avoid the detrimental effects of interest group politics.

The Agency does not appear to use broader social goals, such as health risks, to prioritize its workload. Sites' rate of progress is not materially affected by their hazardousness. In addition, sites in densely populated areas, where there may be greater human exposure, do not progress faster than other sites. Although contrary to social welfare, the irrelevance of exposure is not surprising because Superfund policy generally focuses on individual risk levels (Viscusi and Hamilton, 1999).

Despite the direct influence of interest groups such as the liable parties and local communities, the legislature does not have observable influence on cleanup speeds. Little evidence supports the frequent contention that powerful legislators (identified either by seniority or by serving on the Superfund authorizing subcommittee) significantly speed sites in their districts.

Previous studies by Beider (1994), Hamilton and Viscusi (1999a), and Hird (1994) include econometric analyses of the speed of progress among other outcomes at NPL sites. A broader literature examines determinants of other attributes of EPA's decision-making at NPL sites, such as the attributes of selected remedies, the target risk levels for the remedies, and spending per site.[2] Results from these studies are discussed in Section 1.

This paper differs from earlier work on the pace of Superfund in several ways. First, it studies multiple stages of the process through completion of cleanup and uses a joint model of the successive stages with explicit consideration of site-specific heterogeneity across stages. Second, this paper emphasizes the influence of liable parties, using an original data set on PRP financial characteristics to explore their role. Finally, unlike previous studies,

[2]These studies include those mentioned in the text and also Barnett (1994), Gupta et al (1995, 1996), Sigman (1998), Stratmann (1998), and Viscusi and Hamilton (1999).

the analysis includes time-varying explanatory variables, including the level of Superfund funding and the nature of Congressional representation.

The paper proceeds as follows. Section 1 presents a simple model of EPA decision-making to suggest hypotheses about the reasons for differential progress across Superfund sites. Section 2 provides background on the important milestones in progress toward cleaning up Superfund sites. Section 3 outlines the econometric model that jointly estimates parameters for the three transitions. Section 4 presents econometric results for models without unobserved heterogeneity and with parametric unobserved heterogeneity. A final section concludes by summarizing the empirical findings.

1 Determinants of progress

This section presents a simple model of the EPA's problem in deciding how to prioritize its workload. It identifies several potential influences and then discusses the empirical representations of those influences.

Suppose the EPA has exogenously-determined resources to move sites through the Superfund process. Let b represent the total available resource budget and r_i represent resources devoted to site i. When the Agency devotes more resources to a site, the length of time, t_{ij}, that site i remains in stage j declines. However, the site's technical complexity may also affect the speed. Thus, resources translate into faster progress through an Agency production function, $a_j(r_i, S_i)$, where S_i represents the technological site characteristics. The function may vary with the stage of progress, j, reflecting different technical constraints as the process proceeds.

The EPA might deploy these resources to accomplish several goals. It could address health consequences from Superfund sites. Ideally, it would consider three factors: the hazardousness of the site, h_i; the exposed population e_i; and the time until completion of cleanup, which is the sum of the times in the various stages, j. Multiplying these factors together yields an expression for cumulative health harms for the site, $h_i e_i \sum_j t_{ij}$.

In addition, the Agency may respond to political and interest group pressures to speed progress. To summarize these pressures, suppose there is a cost to the Agency, c_{ij}, that depends on the amount of time the site is in a given stage. The cost depends on pressures from liable parties, represented by L_i, the local community, M_i, and political oversight, P_i. To reflect these pressures, $c_{ij} = c_j(t_{ij,} L_i, M_i, P_i)$.

Thus, a general null hypothesis is that the EPA seeks to minimize total health and political costs, $h_i e_i \sum_j t_{ij} + c_j(t_{ij}, L_i, M_i, P_i)$. The Agency minimizes the sum of these costs across all sites, subject to its resource constraint and the function that converts resources into speed,

$$\min_{r_i} \sum_{i=1}^{N} h_i e_i \Sigma_j t_{ij} + c_j(t_{ij}, L_i, M_i, P_i) \qquad \text{s.t. } t_{ij} \geq a_j(r_i, S_i) \tag{1}$$

$$\sum_{i=1}^{N} r_i \leq b \tag{2}$$

$$r_i \geq 0$$

where N is the total number of sites.

The solution for r_i depends on the aggregate administrative budget, b, and site attributes, h_i, e_i, L_i, M_i, and P_i . It may also depend on the vector of the technical characteristics, S_i, if the marginal productivity of administrative resources varies with these characteristics.

The input r_i is not directly observable and is less relevant for policy prescriptions than the output t_{ij}, the length of time sites spend in the various stages. Thus, this analysis focuses on reduced form relationships between t_{ij} and the variables above:

$$t_{ij} = A_j(b, h_i, e_i, L_i, M_i, P_i, S_i). \tag{3}$$

The remainder of this section discusses the empirical representations of these variables. The variables include: (1) Superfund program resources, b; (2) the social costs of delay, h_i and e_i; (3) liable party pressures, L_i; (4) community pressures, M_i; (5) political oversight, P_i; and (6) technical complexity, S_i. Table 1 presents the means and standard deviations of these variables for the sites analyzed.

Superfund resources. Superfund resources, b, are represented by the aggregate program funding, which has varied considerably over time. When Congress allowed the program's authorization to lapse in 1985–86, funding for the program virtually disappeared. Funding levels peaked in the early 1990s and have declined more recently. Tax contributions to the Superfund Trust Fund ceased in the end of 1995 when their authorization lapsed, but funding remains from recent appropriations and cost recovery from liable parties. To represent this wavering commitment, the equations include the annual appropriation for Superfund divided by the number of sites on the NPL. Table 1 reports that funding has averaged \$1 million per site in 1994 dollars. Because sites reach various stages in the process at different times, these funding levels may affect the relative speeds of progress across sites.

Health-risk priorities. The most important social reason for expediting cleanup is avoided health harms, resulting from the intrinsic hazard, h_i, and the exposed population, e_i. Studies of Superfund by Gupta, Van Houtven, and Cropper (1995) and Hamilton and Viscusi (1995, 1999a, 1999b) use a direct measure of assessed health hazards at the site: estimated risk levels for individuals through various pathways. These risk levels are not used here because they are calculated as part of the Remedial Investigation that precedes the Record of Decision. Thus, they cannot be used in the first two stages. In addition, the risk levels are not available for many sites with early Records of Decision.

Instead, I use the site's Hazard Ranking System (HRS) score as a measure of hazards. These scores reflect a number of characteristics, such as number of people potentially exposed, quantities and types of wastes, and likelihood of migration of contaminants. To create a total HRS score, the EPA scores several different exposure pathways, including groundwater, surface water, and air. Because it sometimes stops evaluating additional pathways once the total score exceeds the number required for NPL listing, the aggregate HRS may not be a good measure of hazards. Thus, the equations include HRS scores separately by pathway, with an indicator variable to reflect whether the pathway was scored. As Table 1 indicates,

Table 1: Means and standard deviations of explanatory variables

	Mean	Standard Deviation
Budget (b):		
Authorization per NPL site (million 1994 dollars)	1.137	.121
Costs of delay (e_i and h_i):		
Population density in census tract (thousands per square km)	.434	.690
Hazard Ranking System scores:		
Groundwater scored	.968	–
Value of groundwater HRS score	63.8	20.4
Surface water scored	.686	–
Value of surface water HRS score	19.2	20.8
Air scored	.138	–
Value of air HRS score	55.8	15.6
Liability funding (L_i):		
Orphan site	.074	–
Large publicly-traded PRP	.285	–
PRP financial stability (Altman's Z for 1980)	2.97	.98
Community characteristics (M_i):		
Voter turnout in county for 1984 Presidential election	.562	.078
Median household income in census tract (thousand 1990 $)	32.3	11.97
Fraction black in census tract	.082	.170
Fraction Hispanic (any race) in census tract	.058	.127
Political oversight (P_i):		
Representative on Superfund authorizing subcommittee	.037	–
Representative's seniority (years)	8.66	1.99
Techological complexity (S_i):		
Size of site in acres	48.97	238.2
Facility type:		
Landfill	.300	–
Surface impoundment	.070	–
Chemical plant	.113	–
Other manufacturing plant	.221	–
Wood preserving site	.046	–
Mine	.041	–
Wellfield	.066	–
Other facility	.1445	–
Contaminants:		
Acids and bases	.213	–
Dioxins	.051	–
Metals	.495	–
Radioactive materials	.020	–
Other organic contaminants	.742	–
Other inorganic contaminants	.288	–

the EPA almost always scores groundwater, often surface water, and less frequently air exposure.[3] Groundwater exposure typically yields the highest absolute value for this risk measure.

The HRS score combines both intrinsic hazard, h_i, and the exposed population, e_i. However, I also include a separate measure of exposed population, the density of the site's census tract, so that the weighting of exposure is not restricted to the HRS weights. If the coefficients on the HRS scores are statistically significant, failing to find a statistically significant coefficient on population density does not necessarily imply that exposed population does not matter; it could imply that the HRS score captures the effects of e_i. However, population density may have a statistically significant coefficient if the weight the HRS score attributes to population density is not the true coefficient on population density in the duration equation. In addition, population density in the census tract represents a broader definition of the exposed population than typically used in HRS scores, so it may add to the information contained in the score.

Superfund's regulations specify that cleanup is required when individual cancer risks exceed a certain threshold. This regulatory guidance does not include a role for the exposed population. Viscusi and Hamilton (1999) find evidence of a "scope effect" in which only the risk level, rather than the number of people exposed, determines the EPA's remedy selection. Given this lack of attention to exposed population, it will not be surprising if this variable does not affect the speed of progress either.

Liable party pressure. The Superfund program relies heavily on financing through *ex post* liability rather than more conventional tax financing. Potentially Responsible Parties (PRPs) who may be held liable for cleanup costs include generators and transporters who contributed hazardous materials to a site and past and present site owners and operators. The EPA and PRPs may reach agreements for the PRPs to undertake study or cleanup of Superfund sites. Alternatively, the EPA may attempt to recover costs from the PRPs after

[3]The EPA added a fourth HRS pathway, soil exposure, in the 1980s. This pathway is rarely scored so it is not included in the analysis.

using its own funds to clean up a site.

Liable parties may have an important role in shaping the EPA's priorities. PRPs may have incentives to delay progress in order to postpone their costs, lowering the effective cost to them. In addition, Rausser et al. (1998) build a model in which information assymetries between EPA and PRPs contribute incentives for delay. Finally, PRPs' expected costs may change with delay either because contamination diffuses making cleanup more difficult or because contamination naturally attenuates making cleanup cheaper. With these conflicting factors, the direction of any PRP influence, as well as its existence, is an empirical issue.

In the empirical analysis, several measures reflect private pressures from liable parties, L_i. First, the equations include a variable that reflects whether viable PRPs exist at the site. As Table 1 reports, 7.4% of the sites in the data set are "orphan" sites where no viable PRPs have been found. For these sites, Superfund functions like a traditional public works program, funded by tax revenues and appropriations. There should be no pressure from liable parties at these sites.

We can extend this logic to hypothesize that, when PRPs are present, the more likely they are to bear Superfund costs (or the higher share of costs they expect to pay), the stonger their incentive to exert influence on the EPA. As a measure of the likelihood that PRPs will pay for the site, this analysis uses the depth of pockets of the PRPs. This approach is taken because courts have interpreted Superfund to leave PRPs very little way out of liability, except by being judgement proof.

Under Superfund's joint and several liability rule, any PRP at the site can be held liable for the full costs of cleanup, regardless of its share of liability for harms. For this reason, the presence of PRPs with deep pockets raises the likely private stake. PRPs with deep pockets may not end up bearing a large share of costs if their liability for harms (often related to their share of the wastes at the site) is low and other parties, such as smaller firms and individuals, pay their shares. However, even if other parties are judgement proof, the EPA will still secure payment at sites with deep pocket PRPs. Thus, sites with PRPs with great financial resources are likely to be sites where at least some private parties have a high stake.

By contrast, without at least one deep pocket PRP, the EPA is more likely pay out of its own funds and thus the private party stake in the site is lower.

To test this measure of private stakes in the site, I developed an original data set on PRP financial characteristics.[4] PRPs named at each site were manually coded with CUSIP numbers and then matched to the Primary Industrial, Supplementary Industrial, and Tertiary (PST) COMPUSTAT tapes for PRP financial data. From this merge, 29% of sites have at least one PRP with financial data available as indicated in Table 1. The low match rate reflects the abundance of small firms and individuals among PRPs.

The estimated equations use two variables from this merge. First, a match indicates the presence of at least one large publicly-traded PRP. Second, the equations include the average value of Altman's Z index for matched PRPs in 1980. A higher value of this index should indicate a lower likelihood of bankruptcy for each PRP (Altman, 1983).[5] Both PRP financial variables measure the depth of PRPs pockets and thus the likelihood that they will bear substantial costs from the site. To avoid potential endogeneity of firms' financial structure, the equations use 1980 values, prior to any of the outcomes in this study.

A survey of EPA's site managers suggests that PRPs may play an important role in the pace of progress (Beider, 1994). The managers consider unusually cooperative PRPs to be among the most important characteristics of sites with rapid progress. They also report negotiation with PRPs as one of the most common sources of delay at sites with unusually slow progress.

A few previous studies have examined the influence of PRPs on decisions other than

[4]Matching PRP financial data to sites required first creating a list of PRPs by site. The official data set for this purpose is the Site Enforcement Tracking System (SETS), which contains lists of firms and individuals that have been sent letters notifying them of their PRP status. Unfortunately, SETS is incomplete. Therefore, these lists were supplemented by the defendants associated with any NPL site in the EPA Docket, a data set of civil and administrative cases referred by the EPA to the Department of Justice. Even with this combined list of PRPs, the numbers of identified PRPs do not always correspond with the number of PRPs reported by site managers (EPA, 1995). However, there does not appear to be any systematic bias in the differences between the two sources. In addition, the list of PRPs sometimes exceeds the number reported by the site managers, so their reports also contain some noise.

[5]The Z index, based on a discriminant analysis of bankruptcies in the U.S., is 1.2(working capital/total assets)+ 1.4(retained earnings/total assets)+ 3.3(earnings before interest and taxes/total assets)+ .6(market value of equity/book value of total liabilities)+ 1(sales/total assets).

speed.[6] Stratmann (1998) finds that when PRPs are uncooperative (measured by whether there is litigation or ongoing negotiation at the site) the government spends more public funds at the site and that orphan sites are also associated with higher public spending. Sigman (1998) finds that EPA chooses more extensive remedies at sites with a lower likely PRP stake (measured by the sites' orphan status and PRP financial attributes), providing evidence that the EPA responds to PRP pressures.

Community influence. Communities may influence the Superfund process through pressure on public officials to expedite progress. To test for such community influence, the analysis includes several variables that measure the political power of the local community.

The first such measure is voter turnout, which directly represents the strengths of the community's influence on electoral politics. In addition, Hamilton (1993) argues that this variable may indicate a community's ability to engage in collective action. The estimated equations include county-level turnout of eligible voters during the 1984 Presidential election as a nationally-consistent turnout measure.

In addition to voter turnout, the socioeconomic characteristics of the community may determine its political muscle. The equations therefore include several socioeconomic variables for the sites' census tract (a more disaggregated level than voter turnout). Been (1997) argues that the census tract is the appropriate level of aggregation for "environmental equity" analysis because census tracts correspond to perceived community boundaries and thus to the relevant neighborhood for community action. Using a Geographic Information System, sites were matched to census tracts based on their latitude and longitude. For each census tract, the 1990 Census of Population provided three socioeconomic variables, median household income, the percent of the population that is black, and the percent that is Hispanic (of any race). Communities with higher household income may have greater political strength

[6] When previous studies consider PRP influence, they typically include a variable for PRP funding of study or clean-up at the site. However, this approach may fail to capture the true influence of PRPs for two reasons. First, by participating the public oversight process and in negotiations with the EPA, PRPs may influence the pace at sites even when they do not agree to fund the process. Second, the PRPs' decision to participate may be endogenous to the speed and nature of progress at the site.

because of their ability to provide campaign contributions. Communities with higher minority populations may command less influence because of the exclusion of these groups from the political process.

Hamilton and Viscusi (1999a) conduct a thorough study of Superfund "environmental equity" issues, which includes an analysis of the relationship between socioeconomic characteristics of the community and decision-making speeds. Although they find evidence that sites in minority communities (defined as population in a 1 mile concentric ring around the site) get less complete cleanup, they conclude that there is not much evidence of slower decision-making at minority sites. Similarly, they report that household income in a one-mile ring and county-level voter turnout do not affect the length of the decision-making stage. The current study reexamines these effects of these variables in the multistage model estimated here.

Legislative oversight. In addition to private group pressures, legislative oversight might influence the EPA's decision-making. Legislators have incentives to pressure the EPA to direct resources to sites in their district because voters may reward them for these resources.

Legislator's power may come two sources. First, legislators have direct influence over the Agency's activities and budget if they serve on the Superfund authorizing subcommittee (Weingast and Moran, 1983), so membership on this subcommittee is included as an explanatory variable. Second, legislators have more indirect influence through their ability to cooperate or oppose initiatives and funding priorities of the administration. The EPA's political bosses may encourage the agency to direct resources toward sites in the districts of generally powerful legislators. The legislator's seniority provides a measure of their general value to the executive. I focus on House members because they are more likely to be swayed by local concerns, such as Superfund sites, than senators. Unlike previous research, this paper uses congressional representation measures that vary over time.

Previous literature has shown mixed results on legislative influence. Hird (1994) finds that relevant subcommittee assignments for representatives from a site's district have a statistically significant effect on the site's progress, but Hamilton and Viscusi (1999a) find neither this effect nor an effect of the ideological stance of the legislators. Studies that examine other Superfund decision-making also sometimes find an effect of legislative oversight. For example, Stratmann (1998) finds that the seniority of a state's representatives on the authorizing committee increases public spending at sites in that state.

Technical complexity. Superfund sites differ considerably in their technical complexity. Some sites involve a small set of standard contaminants from a single manufacturing operation, whereas other sites may involve diverse contaminants from multiple facilities (for example, in a landfill or other waste management facility). Several variables are included in the equations to capture site complexity. The variables include dummy variables for the type of facility that caused the contamination and for the types of contamination. The site's size in acres, from verbal site descriptions, indicates the extent of the problem to be solved. As Table 2 reports, the average site was 49 acres with a very large variance.

2 Superfund milestones

As political pressure has developed to accelerate Superfund progress, the EPA has formalized and documented a number of stages of progress between initial discovery of a site and the ultimate conclusion of its cleanup (Acton, 1989). Thus, there are numerous indicators of progress prior to a site's removal from the National Priorities List.

NPL listing. The first stage of the process involves choosing sites to receive Superfund resources. Sites identified as potential Superfund sites are included in the EPA's inventory of contaminated sites and studied to determine if they are sufficiently hazardous to merit inclusion on the National Priorities List. As of January 1997, about 40,000 potentially contaminated sites had been identified. On the basis of these preliminary studies, the EPA may propose to list a site on the National Priorities List or may categorize it as "No Further Remedial Action Planned" (NFRAP). Only NPL sites undergo extensive "remedial action,"

Table 2: Progress at nonfederal facility sites on the National Priorities List through January 1997

	Number of sites	Percent of proposed NPL sites	Average years since discovery
Proposed listing on the NPL	1251	100%	3.4
Final listing on the NPL	1145	92%	4.5
First Record of Decision (ROD) signed	1013	81%	8.2
First Remedial Action begun	731	58%	9.4
First Remedial Action completed	504	40%	10.4
Construction completed	407	33%	11.6
Proposed for deletion from the NPL	121	10%	11.0
Deleted from the NPL	118	10%	11.3

Notes: Average number of years is conditional on reaching each stage and begins at the site's discovery date or the beginning of Superfund, whichever is more recent. Federal facility sites and sites in U.S. territories are excluded.
Sources: CERCLIS, SCAP, and Construction Completion List.

although all sites may undergo less costly "removal actions" to address contamination.

As Table 2 reports, 1251 non-federal facilities sites had been proposed for the National Priorities List in the fifty states and District of Columbia by January 1997.[7] This process took an average of 3.4 years and thus is a substantial share of elapsed time for some sites.[8] Virtually all sites proposed for the NPL are finally listed, but with some time delay. Final

[7]The analysis excludes federal facility sites because different institutions govern clean-up at these sites. It excludes territorial sites because they lack covariates such as Census and political variables. The totals include sites that had been deleted from the NPL, except for 3 sites that were deleted because they were judged ineligible for the program.

[8]This listing stage begins at either the site's discovery date or December 11, 1980, the date at which Superfund began. Many sites have discovery dates that well precede the beginning of Superfund. The priority-setting process could not begin on these sites until the program came into effect. Thus, I assign these sites a beginning date of December 1980 to create a duration in this first stage analogous to the time between discovery and listing for sites discovered after the program began. The results were not sensitive to the choice of starting date.

listing had occurred for 1145 sites or 92% by early 1997. The analysis uses time until proposed rather than final listing as the end of the first stage because proposal for the NPL effectively completes the site selection process and marks the time when remedy selection can commence.

Decision-making. Once the EPA names a site to the NPL, decision-making about possible remedies begins. The EPA may decide to address all of the contamination at the site with one cleanup process or break the site down into different "operable units." Operable units may be different areas of the site or different contaminated media (e.g., groundwater versus surface water). Most sites (57%) have only one operable unit, but a few sites have many operable units with a maximum of 29.

A variety of types of remedies must be considered for each operable unit. The EPA or PRPs conduct studies to assess the feasibility, cost, and environmental consequences of several types of remedial alternative. After they reach a decision about which remedy to pursue, the EPA and state authorities sign a Record of Decision (ROD). The ROD contains a justification for the chosen remedy and a summary of information on the costs and consequences of its alternatives. As Table 2 shows, at least one ROD had been signed at 1013 sites, 81% of the proposed sites, by January 1997. The analysis defines the decision-making stage as lasting between the proposed listing of the site on the NPL and signing of the first Record of Decision for the site.[9]

Cleanup. After a remedy has been selected and reported in a Record of Decision, work on the remedy commences. The first stage of work on this remedy consists of "remedial design," the engineering planning for the selected remedy. Following remedial design, the actual cleanup process, the "remedial action," commences. Remedial actions vary considerably in their ambitiousness. They may involve simply limiting migration of contaminants by

[9] Although some sites have multiple Records of Decision for different operable units at the site, the analysis focuses on the time until the first ROD only because of the possible endogeneity in the numbers of RODs at a given site.

capping the site or building retaining walls. They may also attempt the eliminate the source of the contamination by treating contaminants before or after excavating them from the site. Remedial action on at least one operable unit had begun at 731 sites by January 1997.

After completion of a remedial action, the site may remain on the National Priorities List for remedial actions on other operable units or for further cleanup at the same unit if the first cleanup has not been effective. Alternatively, the EPA may decide that the site qualifies as "construction complete." These sites may require continued maintenance activities, such as additional pumping of contaminated water or maintaining fences and security, but require no further major work, such as construction or excavation. As Table 2 reports, EPA judged construction complete at 407 sites by January 1997. The Agency may propose the site for deletion from the NPL and delete it following a comment period. Few sites have reached this stage. Only 121 sites had been proposed for deletion from the NPL and 118 sites actually deleted by January 1997. Average realized times for sites to reach these stages currently exceed 11 years.

In the analysis, the duration of cleanup lasts from the signing of the first ROD until construction completion. For many sites, construction completion follows physical cleanup more immediately than deletion from the NPL and thus captures the true policy goal more accurately.

3 Econometric model

This section discusses the econometric implementation of the model described above. Although the model focuses on the determinants of expected cleanup time, the econometric approach relies on a hazard rate specification. This approach allows the estimation of parameters on time-varying covariates, particularly the funding level, which could not be estimated properly in a model of expected duration.

The model distinguishes three stages of progress: selection of sites, remedy choice, and cleanup. This distinction allows the exogenous variables to affect different aspects of the process; for example, the liable parties may delay the remedy decision but not influence the physical cleanup process. Unlike previous studies of Superfund process, this paper estimates a model of the various stages jointly rather than piecemeal. Heckman and Walker (1990) pioneered the use of this joint approach. The joint model yields several gains if there is unobserved heterogeneity (such as differences in the complexity or contentiousness of sites) that affects multiple stages. Time-varying covariates in later stages are not exogenous because they depend on the completion time in earlier stages. In addition, some observations may not be at risk for later stages because they have not completed earlier stages. The joint model provides appropriate treatment for the selected sample in later durations.

The approach requires specification of a hazard rate function, a function which describes the instantaneous probability of completing a stage at time t, conditional on not having yet completed the stage by that time. In particular, this function may depend on t itself, a phenomenon known as "duration dependence." Sites may be more or less likely to leave a stage if they have been in that stage for a long time. The form of this duration dependence is difficult to specify *a priori* and Heckman and Singer (1984) show that this choice can have important effects on the coefficient estimates. Thus, the estimation uses the flexible functional form for duration dependence suggested by Heckman and Singer. Appendix A provides a technical overview of the construction of the likelihood function.

One further econometric issue that requires consideration is the selection of sites for the model. Sites are included in the data if they were listed on the NPL by January 1997.[10] There are several reasons that this sample may be appropriate for policy predictions, despite selecting on the completion of the first stage. Congress temporarily suspended addition of new sites to the NPL in 1996 unless requested by a state governor. Several current Congressional proposals would limit the number of new NPL sites or close the list. As a result,

[10] Although we attempted to fill in most missing data from the descriptions of the sites in RODs, 118 sites had to be dropped because of missing orphan status and 6 because of missing site characteristics. Thus, the data set analyzed contains 1127 of the 1251 sites in Table 2. A test for the difference in means shows that dropped sites were proposed for the NPL significantly later than the remaining sites, so the exclusion is of some concern. However, when the equations were reestimated with the sites with missing orphan status, the results were not substantially different.

a selection model based on the current regime may not predict the ultimate composition of Superfund any more accurately than the no-new-sites model implied here.[11] Growth of the NPL has also slowed to an average of 21 sites (under 2% of cumulative proposed sites) per year over the last five years. Thus, even in the absence of a regime change, the current NPL may approximate the NPL relevant for forecasting policy changes.

4 Estimated equations

Table 3 shows the estimated duration dependence parameters and coefficients on the covariates. If the estimated coefficient on a covariate is positive, the hazard rate increases with the variable; a positive coefficient thus reflects faster progress and a *shorter* expected duration for that stage. Each column presents the parameters for a different transition.

The model in Table 3 allows unobserved heterogeneity with a normal distribution.[12] The first row in the Table 3 shows the estimated factor loading on the unobserved heterogeneity in each of the three stages. Unobserved heterogeneity that delays progress in the first two stages speeds it in the final stage. An explanation for this pattern of heterogeneity is that especially complex sites pass more rapidly through the first two stages, but present thornier cleanup problems and hence remain longer in the final stage.

The results in Table 3 present estimates of the hazard function parameters. To assist in interpreting the magnitude of these effects, Table 4 translates the parameter estimates into changes in the expected duration of the various stages. The specification does not yield closed form expressions for the marginal effects of the parameters. For this reason, I calculate the survival rate numerically at sample mean covariates and use these survival rates to find

[11] A selection model for the current process might feature "competing risks" model in which inventory sites would transit either onto the NPL or to "No Further Remedial Action Planned" status. However, this model is difficult to estimate because it involves over 40,000 sites and because there is virtually no data on the attributes of non-NPL sites. Even this approach would not resolve the selection issue; new sites may be added to the inventory over time.

[12] The model was also estimated without unobserved heterogeneity and with lognormal heterogeneity. In both cases, the estimated coefficients on the covariates were very similar to those presented in Table 3. Differences in the results with these other distributional assumptions are noted in footnotes.

Table 3: **Maximum likelihood estimates of multiple duration model with normal unobserved heterogeneity**

	NPL listing		ROD signing		Const. compl.	
Factor loading (c_j)	0.201	(.072)	0.911	(.278)	-0.532	(.274)
Duration dependence:						
Linear parameter (γ_j)	0.320	(.052)	1.056	(.117)	0.652	(.143)
Box-Cox parameter (σ_j)	-0.476	(.152)	0.290	(.086)	0.420	(.188)
Superfund budget authorization	0.612	(.103)	0.104	(.124)	0.801	(.445)
Hazards:						
Hazard Ranking Scores (HRS):						
Groundwater scored	0.862	(.265)	0.282	(.288)	-0.376	(.506)
Value of groundwater HRS	0.029	(.183)	0.164	(.253)	-0.706	(.339)
Surface water scored	0.787	(.100)	0.036	(.125)	-0.089	(.176)
Value of surface water HRS	-0.022	(.180)	0.522	(.278)	-1.545	(.448)
Air scored	-0.054	(.237)	0.582	(.602)	0.753	(.715)
Value of air HRS	0.882	(.390)	-0.034	(1.05)	-1.849	(1.29)
Population density	0.045	(.059)	-0.145	(.066)	-0.345	(.120)
Liability:						
Orphan site	0.202	(.127)	0.897	(.235)	0.130	(.233)
Large publicly-traded PRP	0.102	(.082)	0.438	(.130)	-0.377	(.151)
PRP financial stability (Z)	-0.001	(.079)	0.001	(.101)	-0.167	(.115)
Community influence:						
Voter turnout in 1984	0.558	(.503)	-0.601	(.694)	2.835	(.919)
Median household income	0.797	(.303)	-0.127	(.438)	-1.407	(.543)
Fraction black	-0.004	(.218)	0.468	(.322)	-0.392	(.455)
Fraction Hispanic	0.370	(.341)	1.009	(.458)	0.848	(.489)
Political oversight:						
House authorizing subcommittee	0.336	(.223)	0.216	(.192)	0.165	(.266)
House seniority	0.042	(.047)	0.001	(.059)	-0.050	(.074)
Technical complexity:						
Log(Site acreage)	-0.032	(.024)	-0.039	(.036)	-0.234	(.045)
Facility type:						
Landfill	0.228	(.142)	0.386	(.216)	-0.368	(.256)
Surface impoundment	-0.098	(.126)	0.045	(.167)	-0.516	(.233)
Chemical plant	-0.150	(.106)	0.151	(.147)	-0.038	(.190)
Other manufacturing plant	-0.269	(.196)	-0.039	(.266)	-1.004	(.393)
Wood preserving site	-0.173	(.188)	0.180	(.305)	-0.047	(.415)
Mine	1.054	(.144)	0.338	(.216)	-0.626	(.297)
Wellfield	-0.015	(.107)	0.238	(.167)	-0.491	(.221)
Other facility (reference)	–		–		–	
Contaminants:						
Acids and bases	0.258	(.088)	-0.003	(.130)	-0.187	(.160)
Dioxins	-0.069	(.185)	0.721	(.271)	-0.277	(.293)
Metals	-0.176	(.073)	-0.075	(.103)	0.045	(.136)
Radioactive materials	-0.022	(.255)	-1.321	(.385)	-0.778	(.872)
Other organic	0.032	(.082)	-0.025	(.114)	0.315	(.162)
Other inorganic	-0.184	(.083)	-0.023	(.114)	0.252	(.143)
Intercept	-3.706	(.404)	-4.383	(.660)	-4.310	(.899)
Number completing spell	1127		986		403	

Note: Standard errors in parentheses.

the mean durations presented in Table 4.

The first row of Table 4 presents the expected duration with sample mean covariates. Expected durations are 2.49, 6.00, and 10.98 years for the three stages respectively. The remaining rows in Table 4 illustrate the effects of changes in the covariates. For binary covariates, Table 4 shows the percentage difference in expected duration when the variable equals one relative to when it equals zero. For continuous covariates, the table shows the effects of an increase of one standard deviation of the sample population. The table includes predicted effects only for covariates with coefficients that are statistically significant at the 10% level in Table 3.

4.1 NPL listing

The first column of results in Table 3 corresponds to the time from discovery or commencement of Superfund to proposed NPL listing. The duration dependence parameters in the first two rows of the table indicate hazard rates that fall with time. This negative duration dependence suggests that sites that linger in the unlisted state lose momentum. Including unobserved heterogeneity increases the estimated duration dependence (makes it less negative) than when the model is estimated without unobserved heterogeneity. This effect is expected: failure to account for heterogeneity tends to make the observed duration dependence lower because the representation in the risk set of observations with lower hazard rates declines over time.

The remaining rows in Table 3 contain estimates of the parameters for the covariates discussed in section 1. The first variable is the annual per-site Superfund budget authorization. Higher funding speeds listing, a result that is statistically significant. Table 4 shows the estimated substantive contribution of funding: a one standard deviation increase over average public funding ($140 million per year for the studied sites) would reduce listing times by 5.7%. This result supports the view that sites' duration in this stage depends on the EPA's allocation of its scarce administrative resources.

The next rows reflect the extent of the hazard. When EPA chooses to score groundwater

Table 4: Estimated effects of covariates on expected durations

	NPL listing	**ROD signing**	**Construction completion**
Expected duration at sample means (years)	2.49	6.00	10.98
Increase budget authorization 1 S.D.	-5.70%	–	-3.65%
Hazard Ranking Scores:			
Groundwater scored	-50.32%	–	–
Increase groundwater HRS 1 S.D.	–	–	1.48%
Surface water scored	-46.62%	–	–
Increase surface water HRS 1 S.D.	–	-0.90%	2.74%
Increase air HRS 1 S.D.	-3.28%	–	–
Increase population density 1 S.D.	–	0.92%	2.24%
Liable parties:			
Orphan site	–	-29.17%	–
Large publicly-traded PRP	–	-15.21%	15.24%
Community influence:			
Increase voter turnout 1 S.D.	–	–	-2.00%
Increase household income 1 S.D.	-1.79%	–	1.54%
Increase share Hispanic 1 S.D.	–	-1.14%	-0.97%
Increase site area 1 S.D.	–	–	7.05%
Facility types:			
Landfill	–	-13.52%	–
Surface impoundment	–	–	21.10%
Chemical plant	–	–	–
Other manufacturing plant	–	–	44.83%
Wood preserving facility	–	–	–
Mine	-55.18%	–	25.95%
Wellfield	–	–	20.00%
Contaminants:			
Acids and bases	-18.34%	–	–
Dioxins	–	-24.12%	–
Metals	14.91%	–	–
Radioactive materials	–	60.48%	–
Other organic contaminants	–	–	-11.18%
Other inorganic contaminants	15.68%	–	-9.17%

Notes: S.D. = Standard deviation of sample population.
Covariates with coefficient estimates that are statistically significant at 10% level only.

and surface water, listing occurs faster; as discussed above, however, the decision to score these pathways says little about their risks. The level of these scores (which would indicate the actual degree of hazard) does not affect the listing time. The level of the air HRS score does statistically significantly speed listing, but the effect is small: listing is only 3.3% faster when this score increases by one standard deviation. Exposed population, which is a component of HRS scores, also does not appear to affect listing speeds when entered separately. Thus, there is little reason to believe that more hazardous sites receive attention faster than other sites.

The estimates do not suggest a role for liable party influence, L_i, in listing. None of the coefficients on these variables are individually statistically significant. Public policy discussion has not suggested PRP manipulation of the listing process; my results are consistent with the silence on this point.

By contrast, there is evidence of community influence on listing. Higher income communities do appear to get statistically significantly faster listing. Based on this point estimate, a one standard deviation increase in income (about a 40% increase) reduces listing time by 1.8%. There is also a positive point estimate on voter turnout, although it is not statistically significant. The race and ethnicity variables enter with opposite signs but are not statistically significant.

The results do not point to a role for legislative influence on listing times. The coefficient on the House member's seniority is positive, but not statistically significant at the 10% level. It is also not possible to reject the hypothesis that sites in the districts of members of the Superfund authorizing subcommittee have no faster listing.[13]

The remaining variables in Table 3 indicate the technical characteristics of the site that may affect the urgency and the complexity of cleanup. Large sites also do not receive faster listing, contrary to the belief expressed in a survey of EPA officials (Beider, 1994); in fact,

[13]Earlier versions of the equations used the final NPL listing date rather than proposal date as the conclusion of the first stage. Although these estimates generally did not differ much from those discussed here, they did find a strong role for subcommittee membership: if its House member served on the Superfund authorizing subcommittee, the site received final listing 31% faster. This difference may indicate more Congressional influence on the official procedures rather than the EPA's more fundamental decision-making.

the point estimate on this variable is negative. Mines receive the fastest listing among facility types, whereas wood preserving and nonchemical manufacturing facilities receive the slowest listing.

4.2 ROD signing

The second column of Table 3 presents parameters for the decision-making stage, which occurs between NPL listing and the signing of a Record of Decision. Unlike the previous stage, the estimates of duration dependence parameters indicate that the probability of completion rises the longer the site has been in this stage.

In this second stage, Superfund funding continues to enter with a positive coefficient. Unlike the previous stage, this coefficient is not statistically significant, however, so it is not possible to conclude that the EPA's resource constraints determine progress at the stage.

As in the previous stage, there is little evidence that the degree of hazard matters. Only the surface water score enters with a positive coefficient that is statistically significant at the 10% level.[14] As with the air score in the previous stage, the coefficient suggests a minimal influence on actual speeds: times decrease only .9% for a one standard deviation increase in this score.

Unlike the previous stage, PRPs do play a role in the duration of the decision-making stage. Orphan sites appear to complete decision-making more rapidly than other sites. As reported in Table 4, decision-making at orphan sites occurs in 29% less time than at sites with viable PRPs. Thus, the presence of PRPs contributes to cleanup delays. This result suggests both that PRPs do exert influence on the EPA's priorities and that on balance they choose to use that influence to delay progress.

The role of PRP financial characteristics appears to be more complicated. If at least one PRP at the site is in COMPUSTAT, the site has a higher hazard rate and thus shorter expected decision-making time. Such large PRPs are more likely to be active in the studies and public oversight process that precede remedy selection (Sigman, 1998). Thus, the

[14] It is not statistically significant if the model is estimated without unobserved heterogeneity.

positive and statistically significant coefficient on this variable may indicate that such PRP involvement in decision-making helps to expedite the process, reducing cleanup times by 15% according to the estimates in Table 4. The net effect of PRPs is still to slow progress, but these particular PRPs mitigate this delay somewhat.

Community influence, which expedited the first stage, does not have the same effect in decision-making phase. Neither the coefficients on median household income nor turnout are statistically significant. Surprisingly, the share of the community that is Hispanic has a positive and statistically significantly coefficient; the coefficient on the share that is black is also positive, but not statistically significant. These results seem counter to a standard influence theory under which minority communities have less political voice. One possible explanation is that the EPA has paid more attention to sites in minority communities because of concerns about "environmental justice," which, for example, were formalized in a 1994 executive order. Thus, a traditional disadvantage may actually have sped cleanup.

Once again, the equations provide no evidence of legislative interference in this second stage. Although both point estimates suggests that decision-making is faster with more powerful legislators, neither coefficient is statistically different than zero.

A few of the remaining site characteristics appear to influence the length of this stage. Of the various facility types, decision-making is fastest at landfills. Although landfills may have a complex mix of contaminants, capping these sites (rather than removing or treating contaminants) has become a standard remedy. The EPA formalized this response as a "presumptive remedy" for these sites in September 1993. This result suggests that such standardization may considerably speed decision-making.[15] As reported in Table 4, landfills received 14% faster decision-making at sample means.

[15] However, a "presumptive remedy" variable — a dummy for landfills sites from 1993 on — did not enter the equations statistically significantly. It is possible that the standardization in remedies occurred much earlier than when formalized.

4.3 Completion of cleanup

The final column in Table 3 contains the parameters for the duration of cleanup (from the signing of the ROD to construction completion). The duration dependence parameters at the top of the table indicate positive duration dependence. As with the decision-making stage, the longer the site has been in the cleanup stage, the more likely it is to finish.

The funding variable is statistically significant and positive in the final stage.[16] As Table 4 reports, a one standard deviation increase in public funding would speed cleanup by 3.7%. Most spending in the later stages of the Superfund process is private funding as PRPs undertake cleanup at the sites under agreements with the EPA. Of sites that had commenced remedial action by January 1997, 75% had at least one PRP-financed remedial action or remedial design. Thus, the influence of public funding on the speed of this stage is somewhat surprising. It suggests that the resources that the EPA devotes to overseeing remedies have a substantial impact on their rate of progress.

Despite this evidence that EPA resource scarcities are important, the variables included to reflect health priorities do not enter with signs that suggest faster cleanup at more hazardous sites. Instead, the reverse appears to hold. Sites with higher groundwater and surface water Hazard Ranking System scores and higher population densities all spend longer in this final phase. This negative relationship may result from especially harmful sites receiving more extensive and complex remedies.

To address this hypothesis, the final stage was re-estimated with measures of the anticipated cost of chosen remedy included in the equation. In these estimates, presented in Table 5 in Appendix B, the remedy cost enters with a large and statistically significant coefficient. However, including the cost does not eliminate the negative coefficient estimates on the hazard and exposed population. Both remain negative and statistically significant. Thus, the

[16] By contrast, the point estimate of this coefficient is much smaller and not statistically significant when the model is estimated without unobserved heterogeneity. Models of multiple sequential transitions that lack unobserved heterogeneity may result in biased estimates for the coefficients on time-varying covariates because cross-site heterogeneity in previous stages renders the values of these covariates endogenous. The change in the funding coefficient thus reflects the presence of heterogeneity.

equations continue to provide no evidence of health-based priorities.

The coefficients suggest that private interests influence this stage of progress. Sites with a large publicly-traded PRP received 15% slower cleanup. Because the presence of these PRPs increases the likelihood that cleanup is privately financed, this result is consistent with PRPs exerting influence and with them using this influence to delay attention to their sites.[17]

In addition, communities appear to influence progress in the final stage. With higher voter turnout, sites move more rapidly through the Superfund process. The point estimate indicates that a one standard deviation increase in voter turnout reduces cleanup time by 2%. Thus, influential communities appear to get the EPA to prioritize their sites.

How powerful are communities relative to PRPs? It is difficult to come up with an exact comparison since the influence measures are not in comparable units. However, consider that the 95th percentile community in voter turnout is able to get cleanup that is 2.9 years faster than the 5th percentile community. By comparison, the presence of a large publicly-traded PRP reduces cleanup time by 1.7 years. Thus, a community's power can be at least as important as PRP influence.

Although influential communities appear to get faster cleanup, median household income may have the reverse effect of turnout. A one standard deviation increase in this variable may increase cleanup time by 1.5%. One possible explanation for this unexpected pattern is that these high income communities lobby for and receive more extensive cleanups; these more elaborate cleanups slow the cleanup time. Earlier research has not found a strong association between income and remedies (Hamilton and Viscusi, 1995; Gupta, Van Houtven, and Cropper, 1996). A direct test on the current data is more equivocal: when measures for the cost of the remedy are included (in Appendix B), the coefficient remains negative, but it is no longer statistically significant at the even the 10% level, providing some weak support for this explanation.

[17]When the cost of the selected remedy is included (Table 5 in Appendix B), the liable party coefficients change. The effect of a large publicly-traded PRP becomes nearly zero, but the orphan site coefficient is large and statistically significant, so the net effect of liable parties is still to delay progress significantly.

Unlike these direct private interests, however, there is no evidence that politicians themselves influence outcome. Neither measure of the power of the site's House members affects its cleanup speed.

As in previous stages, there is evidence that technical issues affect cleanup speeds. Although this variable has not been relevant in earlier stages, the size of the site matters to cleanup speed. Increasing site area by one standard deviation increases estimated cleanup time by 7%.

5 Welfare implications

The results can be used to provide an indication of the welfare effects of interest group politics. In particular, one might ask the cost of the delays imposed by PRPs. Evaluating the welfare consequences of these delays is difficult with the available data, but this section presents some rough calculations.

Congress is currently considering restricting or eliminating the liability funding of Superfund. How much would be gained if Superfund relied on taxes to finance cleanup rather than legal liability? The estimates suggest that the presence of PRPs in general delays progress by 1.8 years (based on 29% faster progress by orphan sites in the second stage). With a large publicly-traded PRP, the net effect of PRP presence is a delay of .8 years in the second stage and 1.7 years in the third stage, for a total of 2.5 years. Assuming the net benefits of cleanup arise entirely at completion and using a 5% social discount rate, a delay of 1.8 years reduces the net benefits of cleanup by 8.4% and a delay of 2.5 years reduces them by 11.5%.

To get a sense of the magnitude of these costs requires an estimate of the net benefits of completing cleanup. Recent research by Hamilton and Viscusi (1999b) on a sample of 150 Superfund sites finds that these sites yield a cost per cancer case avoided of $3 million in 1993 dollars. An average of 4.87 cancer cases was avoided per site at the sites they studied. If the value of a statistical life saved as $5 million (the mid-point of the range suggested by Viscusi (1993), based on labor-market hedonic studies), then the sites had an average net benefit of $9.7 million per site.[18] Using this value, the presence of PRPs would cost an average $.8 million per site and their presence with a large publicly-traded PRP would cost $1.1 million per site. Thus, the costs could be on the order of $1 billion for the entire NPL.

These values presume delay is uncorrelated with the net benefit of cleanup across site. Hamilton and Viscusi find great variation in the net benefits across sites. A very few sites in their study account for the positive mean net benefit; the median site would not pass a cost-benefit test. If sites with liable parties have negative net benefits, delays are actually welfare improving. However, research suggests that the EPA selects less costly remedies at sites with liable parties (Sigman, 1998), so net benefits at these sites may tend to be more favorable than average.

Thus, eliminating liability funding and moving to tax financing (as at orphan sites) might result in small gains from increased cleanup speeds. The policy implications of this result depend upon how these costs stack up against other costs and benefits of liability financing. On one hand, there are potential advantages of liability financing. Liability can provide incentives for precaution in managing potential contaminants (Kornhauser and Revesz, 1995).[19] Liability financing and resulting PRP participation may also help control clean-up costs (Stratmann, 1998). On the other hand, liability may involve high transactions costs from legal expenses. Dixon (1995) estimates that transactions costs will account for as much 30% of private spending on Superfund, so the delay costs would not loom large relative to these costs.

In summary, the empirical results suggest a few conclusions about bureaucratic priorities.

[18]This value may be inaccurate for several reasons. First, it includes only cancer reduction benefits. The full value of cleanup should also include other health effects and ecological benefits where they arise. Second, using valuation of workplace risks may be inappropriate for the risks reduced by Superfund. People may feel differently about environmental cancers than workplace deaths. In addition, as Hamilton and Viscusi point out, the cancers may require discounting because of their long latency. Finally, the values rely on risk estimates from the official site studies, which may be manipulated for political reasons or include other biases, such as the use of upper bound estimates of risks (Viscusi et al., 1997).

[19]However, a large component of Superfund liability is retroactive and thus cannot influence behavior. In addition, the Resource Conservation and Recovery Act (RCRA) regulates hazardous waste, so incentives for precaution through liability rules may be excessive from a social perspective. For a discussion of other advantages and disadvantages of liability financing, see Sigman (2000).

First, although the EPA is widely viewed as being a highly ideological agency, it does not appear to prioritize resources based on environmental goals. This evidence may weaken arguments for allowing agencies discretion in setting their agendas. Second, the EPA is sensitive to concentrated private interests in its priorities, providing support for capture theories. Such responsiveness to private interests may be helpful, if it encourages the agency to consider private costs and benefits in its calculations. However, it may also be harmful if it subverts broader social goals. Indeed, the empirical results presented here suggest that removing a set of concentrated interests by using a diffuse funding source (such as a broad-based tax) would significantly increase welfare. Finally, however important legislators' interventions with bureaucracies may be to election outcomes, this study finds no empirical evidence that they affect bureaucratic priorities.

A The Likelihood function

This appendix provides a technical overview of the construction of the likelihood function for the estimated model. The estimation strategy involves maximizing a joint likelihood function for the time that each transition j (j =1, 2 or 3) takes place. As before, the amount of time until transition j occurs at site i is t_{ij}. The hazard rate for transition j is conditional on observed variables, $x_{ij}(t)$, and an unobserved random variable, θ_i :

$$\lambda_j(t_{ij}|x_{ij}(t),\theta_i). \tag{4}$$

The observed covariates in x_{ij} may differ across the transitions and over time. However, a single unobserved θ_i (which might be associated with the complexity of the site) characterizes the site in all three durations.

The survivor function gives the probability that a site has not made the transition by time t_{ij}. Its relationship to the hazard rate is:

$$S_j(t_{ij}|x_{ij}(t),\theta_i) = \exp\left[-\int_0^{t_{ij}} \lambda_j(u|x_{ij}(u),\theta_i)du\right]. \tag{5}$$

The hazard and the survivor functions give rise to a conditional density function for the outcome t_{ij} as the time of transition j. This density function is:

$$f_j(t_{ij}|x_{ij}(t),\theta_i) = \lambda_j(t_{ij}|x_{ij}(t),\theta_i)S_j(t_{ij}|x_{ij}(t),\theta_i). \tag{6}$$

If a site has completed all its spells, the likelihood contribution for an observation is the product of this density function for all relevant spells. For censored observations, the densities for completed spells are multiplied by the survivor function the current transition. Thus, the likelihood contribution for site i is:

$$\prod_{j=1}^{3} f_j(t_{ij}|x_{ij}(t),\theta_i)^{\delta_{ij}} S_j(\overline{t_{ij}}|x_{ij}(t),\theta_i)^{\overline{\delta_{ij}}}. \tag{7}$$

If site i has made transition j, then $\delta_{ij} = 1$ and zero otherwise. If the site has not yet reached construction completion (the final transition), the exponent on the survivor function, $\overline{\delta_{ij}}$ equals one for the transition j that is one greater than the last completed transition. All other values of $\overline{\delta_{ij}}$ equal zero. In the survivor function, $\overline{t_{ij}}$ is the time that the site has been in the final state when censored on January 13, 1997.

Implementing the model requires a specific functional form assumption for the hazard rate. In the estimated equations, the hazard rate depends on a Box-Cox transformation of the time variable. This specification is the generalized functional form suggested by Heckman and Singer (1984) with first-order time dependence. For each duration, j, it has the form:

$$\lambda_j(t_{ij}|x_{ij}(t), \theta_i) = \exp(\frac{t_{ij}^{\sigma_j} - 1}{\sigma_j}\gamma_j + x_{ij}(t)\beta_j + c_j\theta_i) \tag{8}$$

where σ_j and γ_j characterize the duration dependence. For the special cases where $\sigma_j = 0$ and $\sigma_j = 1$, equation (8) implies Weibull and Gompertz distributions of failure times respectively. The function form for the relationship between λ_j and the covariates in x assures that the estimated hazard rate is always positive. The parameters $\sigma_j, \gamma_j, \beta_j$, and c_j may differ across transitions.

To use the likelihood contribution in equation (7) requires a specification for the distribution of the unobserved heterogeneity, θ. The model was estimated using two different approaches. First, Heckman and Singer's Nonparametric Maximum Likelihood procedure postulates a distribution with a small number of points of increase, and estimates the location and probability weight of these points directly (Heckman and Singer, 1984). However, the model failed to converge once even a small number of site characteristics (such as acreage or HRS scores) were included as covariates. Thus, the model was estimated using standard parametric distribution assumptions. The covariate estimates are not very sensitive to the choice of mixing distributions or indeed to the use of a mixing distribution at all, as discussed in footnote 12.

B Final stage with chosen remedies included

Table 5 presents estimates of the duration of the final stage that include a characterization of the remedy chosen for the site, namely the anticipated cost of remedy.[20] The cost variable derives from the Records of Decision for the sites (EPA, 1994). In addition to the new variable, there are two differences between the equations in Table 5 and those in Table 3 in the text. First, Table 5 uses the final stage only. A full multistage model with unobserved heterogeneity and remedy characteristics in the final stage may not be identified because the selected remedy could be sensitive to the duration of earlier stages. Second, the equations in Table 5 are censored earlier because anticipated cost data are available only for RODs signed by fiscal year 1993.

The equation (1) in Table 5 shows the equation for the final stage with no remedy characteristics included. This equation is estimated without heterogeneity and restricted to sites with anticipated cost data to allow a direct comparison of the coefficients with and without the remedy cost. The second equation includes the log of the anticipated cost of the selected remedy. Not surprisingly, this variable has a statistically significant negative effect: the more costly the remedy, the longer the cleanup.

Variables such as the PRP characteristics, community characteristics, and extent of hazard might enter the final stage through their influence the remedy selected rather the pace conditional on the remedy. However, comparing equation (1) with equation (2) in Table 5, few of the estimated coefficients change when the remedy cost is added to the explanatory variables.[21] In particular, even controlling for the chosen remedy, the variables that reflect the extent of hazard continue to have a negative effect of the speed of cleanup.

[20]A qualitative characteristic of the remedy — whether it involved treatment of contaminants — was included in previous estimates. Treatment remedies (e.g., incineration and social washing) modify or eliminate contaminants and are more extensive than other remedies. However, this variable did not enter with a statistically significant coefficient.

[21]The liable party variables are exceptions. See the discussion in footnote 17.

Table 5: **Maximum likelihood estimates for cleanup stage with remedy costs**

	(1)		(2)	
Duration dependence:				
Linear parameter (γ_j)	0.712	(.192)	0.956	(.213)
Box-Cox parameter (σ_j)	0.186	(.208)	0.166	(.167)
Log (Anticipated remedy cost)	–		-0.495	(.031)
Superfund budget authorization	0.432	(.389)	0.353	(.376)
Hazards:				
Hazard Ranking System scores:				
Groundwater scored	-0.206	(.474)	-0.109	(.498)
Value of groundwater HRS	-0.803	(.313)	-0.706	(.321)
Surface water scored	0.104	(.160)	0.095	(.155)
Value of surface water HRS	-1.568	(.426)	-0.976	(.422)
Air scored	0.590	(.663)	0.038	(.630)
Value of air HRS	-1.508	(1.19)	-0.066	(1.15)
Population density	-0.383	(.118)	-0.394	(.117)
Liability:				
Orphan site	0.246	(.203)	0.445	(.211)
Large publicly-traded PRP	-0.287	(.138)	0.034	(.146)
PRP financial stability (Z)	-0.154	(.102)	-0.232	(.106)
Community influence:				
Voter turnout in 1984	2.585	(.902)	1.348	(.900)
Median household income	-1.358	(.499)	-0.835	(.533)
Fraction black	-0.271	(.427)	-0.297	(.424)
Fraction Hispanic	1.113	(.424)	1.132	(.389)
Political oversight:				
House authorizing subcommittee	0.285	(.261)	0.099	(.260)
House seniority	-0.058	(.070)	-0.122	(.075)
Technical complexity:				
Log(Site acreage)	-0.235	(.041)	-0.198	(.046)
Facility type:				
Landfill	-0.197	(.232)	-0.360	(.233)
Surface impoundment	-0.544	(.222)	-0.683	(.239)
Chemical plant	-0.021	(.178)	-0.180	(.179)
Other manufacturing plant	-1.109	(.414)	-0.974	(.454)
Wood preserving site	-0.129	(.402)	-0.427	(.460)
Mine	-0.418	(.272)	-0.880	(.268)
Wellfield	-0.504	(.208)	-0.690	(.211)
Other facility (reference)	–		–	
Contaminants:				
Acids and bases	-0.132	(.147)	-0.193	(.148)
Dioxins	-0.248	(.271)	-0.288	(.297)
Metals	-0.042	(.130)	-0.216	(.133)
Radioactive materials	-0.447	(.830)	-1.503	(.935)
Other organic	0.250	(.156)	0.355	(.154)
Other inorganic	0.211	(.134)	0.261	(.135)
Intercept	-3.859	(.870)	-3.041	(.838)
Log-likelihood	-1091		-991	

Notes: Standard errors in parentheses.
Number entering spell: 800; Number completing spell: 347.

C Data and variable descriptions

Chronology: The durations are based on data from the January 1997 SCAP11, CERCLIS, and Construction Completion list. Throughout the analysis, site discovery dates are the maximum of the discovery date listed in SCAP or CERCLIS and December 11, 1980, when Congress enacted Superfund. This convention was adopted because some sites have discovery dates many years before the Superfund program began.

Orphan sites: Data on orphan sites are from the Resources for the Future National Priorities List Database (Probst, 1995).

PRP financial data: The EPA's Site Enforcement Tracking System (SETS) lists PRPs at each site. These lists are supplemented by any defendants at NPL sites in the EPA's Enforcement Docket data set. The PRP lists were matched based on the firm names to COMPUSTAT's Primary Industrial, Supplementary Industrial, and Tertiary (PST) file. Yuan Gao, Kerry Knight, Andres Lerner, and Geng Qu assisted with this match. Z scores were calculated based on the formula provided in footnote 5 and averaged for all PRPs at the site with financial data.

Site characteristics: The Superfund NPL data base provides data on contaminants and HRS scores. The site acreage was extracted from the verbal description of the sites. For about a hundred sites with no acreage data, the acreage is projected based on the site characteristics included in the equations. Facility types are from Probst (1995), supplemented with data from CERCLIS and verbal site descriptions where necessary.

Community characteristics: Median household income, minority composition, and population density derive from the 1990 Census of Population, matched to Superfund sites by census tract. About 10% of the sites had missing or invalid latitude-longitude data (which was used to identify the census tract). County-level averages were used for these sites.

Voter turnout: Votes cast for President in 1984 from the *1988 County and City Data Book* was divided by estimated 1984 eligible voters by county. The number of eligible voters in 1984 was interpolated from the 1980 and 1990 Censuses, assuming a constant rate of change during the decade.

Legislative data: Seniority derives from the *Roster of United States Congressional Officeholders*, supplemented with listings for the 104th Congress. The authorizing subcommittee is the Subcommittee on Transportation, Tourism, and Hazardous Materials. These data derive from the *Congressional Staffing Directory*.

References

[1] Acton, Jan Paul. *Understanding Superfund: A Progress Report.* Santa Monica, CA: RAND Institute for Civil Justice, 1989.

[2] Altman, Edward I. *Corporate Financial Distress.* New York: John Wiley and Sons, 1983.

[3] Barnett, Harold C. *Toxic Debts and the Superfund Dilemma,* Chapel Hill, NC: University of North Carolina Press, 1994.

[4] Been, Vicki with Francis Gupta, Coming to the nuisance or going to the barrios? A longitudinal analysis of environmental justice claims, *Ecology Law Quarterly,* 24 (1997), 1–56.

[5] Beider, Perry. *Analyzing the Duration of Cleanup at Sites on Superfund's National Priorities List.* Congressional Budget Office Memorandum, 1994.

[6] Dixon, Lloyd S. The transaction costs generated by Superfund's liability approach, in R.L. Revesz and R.B. Stewart, eds. *Analyzing Superfund: Economics, Science, and Law.* Washington, DC: Resources for the Future, 1995, pp. 171–85.

[7] Environmental Protection Agency. *ROD Annual Report: Records of Decision Signed in Fiscal Year 1992.* Washington, DC: The Agency, 1994.

[8] Environmental Protection Agency. *Remedial Program Managers (RPM) EPA National Site Data.* Washington, DC: The Agency, 1995.

[9] General Accouting Office. *Superfund: Times to Complete Assessment and Cleanup of Hazardous Waste Sites,* Washington DC: The Office, 1997.

[10] Gupta, Shreekant, George Van Houtven, and Maureen L. Cropper. Do benefits and costs matter in environmental regulations? An analysis of EPA decisions under Superfund, in R.L. Revesz and R.B. Stewart, eds. *Analyzing Superfund: Economics, Science, and Law.* Washington, DC: Resources for the Future, 1995, pp. 83–114.

[11] Gupta, Shreekant, George Van Houtven, and Maureen L. Cropper. Paying for permanence: An economic analysis of EPA's cleanup decisions at Superfund sites, *RAND Journal of Economics,* 27 (1996), 563–582.

[12] Hamilton, James T. Politics and social costs: Estimating the impact of collective action on hazardous waste facilities, *RAND Journal of Economics,* 24 (1993), 101–125.

[13] Hamilton, James T. and W. Kip Viscusi. The magnitude and policy implications of health risks from hazardous waste sites, in R.L. Revesz and R.B. Stewart, eds. *Analyzing Superfund: Economics, Science, and Law.* Washington, DC: Resources for the Future, 1995, pp. 55–82.

[14] Hamilton, James T. and W. Kip Viscusi. Environmental Equity at Superfund Sites in *Calculating Risks: The Spatial and Political Dimensions of Hazardous Waste Policy* Cambridge, MA: MIT Press, 1999.

[15] Hamilton, James T. and W. Kip Viscusi. How costly is "clean"? An analysis of the benefits and costs of Superfund site remediations, *Journal of Policy Analysis and Management,* 18 (1999), 2–27.

[16] Heckman, James J. and Burton Singer. Econometric duration analysis, *Journal of Econometrics,* 24 (1984), 63–132.

[17] Heckman, James J. and James R. Walker. The relationship between wages and income and the timing and spacing of births: Evidence from Swedish longitudinal data, *Econometrica,* 58 (1990), 1411–1441.

[18] Hird, John A. *Superfund: The Political Economy of Environmental Risk,* Baltimore, MD: Johns Hopkins University Press, 1994.

[19] Peltzman, Sam. Toward a more general theory of regulation, *Journal of Law and Economics,* 19 (1976), 211–240.

[20] Probst, Katherine. *The Resources for the Future National Priorities List Database* (Computer file). Washington, DC: Resources for the Future, 1995.

[21] Rausser, Gordon C., Leo K. Simon, and Jinhau Zhao, Information Assymetries, Uncertainties, and Cleanup Delays at Superfund Sites, *Journal of Environmental Economics and Management,* 35 (1998), 45–68.

[22] Reisch, Mark. *Superfund Reauthorization Issues in the 105th Congress.* Washington, DC: Congressional Research Service, 1998.

[23] Sigman, Hilary. Liability funding and Superfund clean-up remedies, *Journal of Environmental Economics and Management,* 35 (1998), 205–224.

[24] Sigman, Hilary. Hazardous waste and toxic substance policies, in P.R. Portney and R.N. Stavins, eds. *Public Policies for Environmental Protection, Second Edition* (Washington, DC: Resources for the Future, forthcoming in 2000).

[25] Stratmann, Thomas. The politics of Superfund, Mimeo, Montana State University, 1998.

[26] Stigler, George J. The economic theory of regulation, *The Bell Journal of Economics and Management Science,* 2 (1971), 3–21.

[27] Viscusi, W. Kip and James T. Hamilton. Are risk regulators rational?: Evidence from hazardous waste decisions, *American Economic Review,* 89 (1999), 1010–27.

[28] Viscusi, W. Kip, James T. Hamilton, P. C. Dockins, Conservative versus mean risk assessments: Implications for Superfund policies, *Journal of Environmental Economics and Management,* 34 (1997), 198–206.

[29] Weingast, Barry R. and Mark J. Moran. Bureaucratic discretion or congresssional control? Regulatory policymaking by the Federal Trade Commission *Journal of Political Economy,* 91 (1983), 765–800.