Illinois Class V Automotive Well Program:
Proposed Statistical Methodology

February 5, 2009

Prepared for:

Andrew Jankowski
Illinois EPA

Prepared by:

The Cadmus Group, Inc.
Michael Crow
A. Richard Bolstein

1. Introduction

To comply with USEPA's Class V Well regulations regarding the management of motor vehicle waste disposal wells (64 FR 68546, Dec 7, 1999; the Rule), Illinois EPA (the State) is planning to implement an Environmental Results Program (ERP) with a grant from EPA's National Center for Environmental Innovation.

In two targeted counties, the State used available records to identify the potential universe of automotive and truck maintenance or repair facilities, and sent those facilities self-certification packets enabling them to determine and report whether they have waste disposal wells subject to the Rule. The State will follow up with those who have declared that they have wells (4 of the facilities that responded) to ensure that they come into compliance with the Rule (either by closing the well with State supervision or by applying for a State permit--it is anticipated that most or all facilities will choose well closure, since the rules for permitting are very strict). The State will conduct site visits at a representative sample of facilities from the original universe to

verify the accuracy of self-reporting and characterize the pool of non-respondents. (These site visits are referred to as "inspections" below, consistent with standard ERP terminology, but note that they are not compliance inspections. They will be conducted by staff from the Office of Pollution Prevention who are not directly responsible for enforcing the Rule. Facility participation is strictly voluntary.) This paper provides an approach for conducting this follow-up survey. The appendix identifies the facilities that are to be inspected.

2. Sampling Plan for Follow-up Inspections

The universe of facilities identified as potentially subject to the Rule in the targeted counties can be subdivided into five categories, based on how they responded to the self-certification process.

1. **Non-respondents**. These are facilities that did not return self-certification forms. Of the 675 facilities in the targeted counties, 526 did not respond. (This includes facilities whose self-certification packets were returned by the post office as undeliverable.)

2. **Not applicable**. Forty-seven respondents to the mailing reported that they were not automotive maintenance or repair facilities and therefore are not subject to the Rule. In follow-up calls to these respondents, the State determined that two of the 47 are in fact automotive facilities. Two additional respondents could not be reached during the follow-up phone calls.

3. **Has a well**. Four automotive respondents reported having a well.

4. **Does not have a well**. Ninety-one automotive respondents reported they do not have a well.

5. **Uncertain response**. Seven automotive respondents provided responses that were unclear or raised questions about whether they have a well.

A different approach will be taken for each category. Among facilities that claimed not to be automotive (category 2), the State will inspect those whose claim appears to be false and those whose claim could not be verified. The State will inspect each facility that provided an uncertain response about well status (category 5). The four facilities that acknowledged having wells (category 3) will be handled through the State's normal administrative processes.

The State will inspect a random sample of non-respondents (category 1). It also will inspect a random sample of facilities that reported they do not have wells (category 4). The state will conduct on-site inspections at randomly selected facilities within each category. The facilities are spread between two counties. The State is not interested in differences between the counties; therefore, the samples will not be stratified by county.

Considering budgetary and other constraints, the State will be able to inspect approximately 60 facilities to verify the accuracy of self reporting and to characterize the non-respondents. Eleven inspections will be directed at the following facilities: facilities whose statement that they are not subject to the Rule appears to be incorrect (two facilities from category 2) or could not be verified (two more facilities from category 2), and facilities that provided uncertain answers

about the presence of wells (all seven facilities in category 5). The remaining 49 inspections will be conducted among randomly selected non-respondents (category 1) and facilities that reported they do not have wells (category 4).

Twenty-two facilities will be selected from category 4. A simple random sample of facilities will be drawn without replacement. Excel's random number generator will be used to draw the sample.

Twenty-seven non-respondents (category 1) will be inspected. A substantial number of facilities in this category are likely not automotive maintenance and repair facilities. To minimize the chances that inspectors will go to non-automotive repair facilities, a two-step process will be implemented. The list of non-respondents will be placed in random order using Excel's random number generator. Working down the randomly ordered list, State staff will first call facilities to verify automotive status. If State staff determine that a facility does not perform automotive maintenance or repair, they will move on to the next facility on the list. If they determine that the facility is an automotive repair shop, or if they are unable to make a determination, State staff will visit and inspect the facility. State staff will continue to work down the randomly ordered list of facilities, without skipping over any, until 27 facilities have been inspected.

The State will make a reasonable effort to make contact by phone (e.g., calling several times, at different times of day) before determining that a category 1 facility's automotive status is unknown and that a site visit is therefore necessary. (On the phone with facilities, State staff could start by asking if the self-certification packet was received, and follow up by verifying the automotive status of the facility. State staff should also verify each facility's name and address. But as discussed below, State staff will not attempt to schedule a visit over the phone.) A log of calls and their results should be kept and the information should eventually be entered into the project database. Corrections to facility names and addresses may be useful at a later date, for example, to identify duplicates and reduce the effective size of the universe of facilities.

While the phone calls to facilities in category 1 need to be made in a particular order, the site visits do not. For the sake of convenience the State may wish to conduct all phone calls first, and then group category 1 site visits geographically. Or they may wish to intersperse them, as convenient, with planned visits to facilities from other categories.

State staff will make unannounced visits to facilities slated for inspection, rather than scheduling inspections in advance. The reason is that inspectors are less likely to be turned away if they show up in person than if they phone in advance. The State will visit each facility only once, and not schedule return visits.

Failure to gain entry to a facility could bias the results of the study and reduce its precision; therefore, the State should make all reasonable efforts to gain entry. (Bias would be introduced if, for example, facilities that refuse entry tend to be less compliant with the Rule than facilities that allow entry.) The State should characterize and track the circumstances when entry is denied to allow for qualitative assessment of potential bias.

If inspectors discover evidence of an active motor vehicle waste disposal well, they will share that information with program staff to be dealt with appropriately. They will advise those

responsible for such wells of the steps they need to take and where to go for more information. The intent of the project is to work cooperatively with facility owners/operators who act in good faith to bring their facilities into compliance.

If inspectors discover evidence of other violations, they will handle that at their discretion, or as they normally would during a pollution prevention assistance visit.

See the appendix of this document for a complete list of facilities to be inspected in each category. Facilities in categories 1 and 4 are ordered randomly.

3. Margin of Error

Some measurements that may be made in this project include:

- The percentage of facilities in each category, and on the whole, that are automotive

- The percentage of automotive facilities in each category, and on a whole, that have a motor vehicle waste disposal well

- The percentage of facilities in each category, and on the whole, that responded accurately to questions about automotive status and well status.

Statistical inferences will be made where random sampling was performed: specifically, in category 1 and category 4. The margin of error of a statistical inference depends on several factors. For example, when estimating the number of facilities in category 4 who have wells, the margin of error depends on the size of the category (the number of facilities in category 4), the observed proportion (the percentage of inspected facilities that had wells), the desired confidence level (discussed below), and the sample size (the number of inspected facilities in that group). The confidence level and sample size are set by the State; its decisions will determine the margin of error of estimates based on the samples.

The margin of error is represented by a confidence interval, which gives a range of values that is believed to contain the actual population proportion or mean with the confidence level prescribed by the user. There are several different acceptable ways of constructing confidence intervals. The standard Wald confidence interval is symmetric about the point estimate. The Score interval is not symmetrical except for a point estimate of 50 percent, but it is often shorter (i.e., more precise). It is particularly useful for small sample sizes and for estimating single proportions, especially when they are outside the 30-70 percent range.[1] We expect that a relatively small proportion of facilities in category 1 and category 4 will have wells, given that the overall incidence of wells among auto repair shops is, in our experience, generally no higher than 10-20 percent. Therefore, the estimates presented below are Score intervals. The size of the Score confidence interval is twice the margin of error when the observed proportion is 50 percent (e.g.,

---

[1] Agresti, Alan and Brent A. Couli, "Approximate is Better than 'Exact' for Interval Estimation of Binomial Proportions," *The American Statistician*, No. 2, May 1998.

expressed as 50 percent +/- the margin of error). At other proportions the Score confidence interval is narrower and not symmetrical.

The confidence level is the percentage of all possible random samples of size *n* whose corresponding confidence intervals contain the actual (unknown) population proportion or mean. A 90-percent confidence level means that approximately 90 percent of all random samples of n facilities from the population will produce a confidence interval that includes the true proportion. The State is selecting just one random sample from each category of facilities. The corresponding confidence interval may or may not contain the true proportion, but the fact that 90 percent of all possible samples would contain the true proportion is expressed by saying we have 90 percent confidence that the particular interval contains the true proportion. (We do not know for sure that this is the case, however.) Confidence levels of 95 percent and 90 percent are standard. We recommend using a 90 percent confidence level because the sample sizes are relatively small.

With regard to sample size, as a default option we recommend splitting the site visits between category 1 and category 4 in a manner that achieves the same margin of error in each. That means inspecting 27 shops from category 1 (non-respondents) and 22 shops from category 4 (facilities that report they do not have a well). The margin of error would be approximately 14.7 percent in each case. In other words, if the proportion of facilities with wells in either category were 50 percent, the 90 percent confidence interval would be 35.3 percent to 64.7 percent. (If the proportion is not 50 percent, the interval will be narrower and will not be symmetric.)

The State may decide to perform more site visits and gain precision in one category, at the cost of losing site visits and losing precision in the other category (assuming the State stays within its planned site visit budget). For example, if the State is denied entry at a disproportionate number of facilities in one category, it may want to perform additional inspections in that category to compensate for the loss of precision caused by the denials. As an outside limit, we recommend that the sample size in each category be no smaller than 15.

The following table shows the margin of error of the estimates for several different sample sizes.

	Margin of Error (Half-width of 90 Percent Confidence Interval)	
Sample Size	**Sample of Non-respondents (Category 1) total = 526**	**Sample of Facilities that Reported No Wells (Category 4) total = 91**
15	0.193	0.181
20	0.170	0.155
22	0.162	**0.147**
25	0.153	0.136
27	**0.147**	0.129
30	0.140	0.120

| 35 | 0.130 | 0.107 |

Note: the recommended sample size for non-respondents is 27. The recommended sample size for facilities that reported no wells is 22. The margin of error in both samples would be 0.147.

Please note:

- Since the State will make an effort to contact non-respondents by phone and inspect only those that cannot be ruled out as non-automotive, the sample size for non-respondents called and/or inspected is expected to be larger than 27, even as the sample size of shops visited is limited to 27. With a larger sample size, the margin of error will fall.

- The margins of error listed above assume that a sample is used to make inferences about all facilities in a category. If conclusions are to be drawn about a subset of facilities in a category (e.g., the percentage of *automotive* non-respondents that have wells), the margin of error will be different than described above.

- An estimate that applies to facilities in multiple categories (e.g., the percentage of automotive facilities in all categories that had wells, or the percentage of facilities in all categories that answered truthfully about automotive status) will also have an associated margin of error. The calculation of that margin of error will be based, in part, on category-specific margins of error.

- The key measures the State is examining in each sample are yes/no (or dichotomous) questions: e.g., whether a facility is subject to the Rule, and whether a well is present in the facility. If continuous measures are developed as part of the analysis of the data (e.g., based on measurements made by inspectors rather than yes/no questions), the margin of error for those variables will be different than reported here.

Appendix: Facilities to be Inspected According to Proposed Sampling Plan

Category 1 (Non-respondents): 27 site visits

Facilities in this category have been randomly ordered. Facilities are to be contacted by phone, in the order presented. If the State is able to determine via phone contact that the facility is non-automotive, no site visit is necessary. If the State determines that the facility is automotive, or is unable to make a determination, a site visit is necessary. The State will make as many calls as necessary (in order down the list, skipping none) until 27 facilities have been selected for inspection. The result of calls should be entered into the "Inspections" form in the project database, regardless of whether or not a facility receives a site visit.

If the State decides to inspect additional facilities in this category beyond the 27 proposed in this sampling plan, State staff should continue down the list in order, making additional calls until enough additional facilities have been designated for inspection.

Rank in Random Order	Facility ID	Rank in Random Order	Facility ID	Rank in Random Order	Facility ID	Rank in Random Order	Facility ID
1	IL000169	26	IL000325	51	IL000266	76	IL000529
2	IL000568	27	IL000349	52	IL000389	77	IL000695
3	IL000160	28	IL000648	53	IL000436	78	IL000647
4	IL000336	29	IL000355	54	IL000079	79	IL000230
5	IL000390	30	IL000639	55	IL000534	80	IL000055
6	IL000084	31	IL000599	56	IL000137	81	IL000348
7	IL000403	32	IL000133	57	IL000164	82	IL000543
8	IL000488	33	IL000542	58	IL000636	83	IL000459
9	IL000153	34	IL000504	59	IL000287	84	IL000660
10	IL000720	35	IL000650	60	IL000501	85	IL000159
11	IL000362	36	IL000478	61	IL000419	86	IL000539
12	IL000126	37	IL000329	62	IL000652	87	IL000608
13	IL000158	38	IL000530	63	IL000704	88	IL000333
14	IL000629	39	IL000208	64	IL000197	89	IL000548
15	IL000453	40	IL000260	65	IL000587	90	IL000538
16	IL000056	41	IL000125	66	IL000128	91	IL000580
17	IL000494	42	IL000614	67	IL000565	92	IL000049
18	IL000123	43	IL000463	68	IL000442	93	IL000103
19	IL000030	44	IL000371	69	IL000244	94	IL000280
20	IL000465	45	IL000721	70	IL000577	95	IL000550
21	IL000531	46	IL000553	71	IL000511	96	IL000036
22	IL000221	47	IL000022	72	IL000640	97	IL000521
23	IL000661	48	IL000129	73	IL000384	98	IL000594
24	IL000117	49	IL000405	74	IL000671	99	IL000592
25	IL000115	50	IL000157	75	IL000395	100	IL000519

Rank in Random Order	Facility ID	Rank in Random Order	Facility ID	Rank in Random Order	Facility ID	Rank in Random Order	Facility ID
101	IL000288	141	IL000512	181	IL000602	221	IL000262
102	IL000502	142	IL000068	182	IL000378	222	IL000653
103	IL000324	143	IL000541	183	IL000209	223	IL000431
104	IL000673	144	IL000454	184	IL000513	224	IL000603
105	IL000658	145	IL000700	185	IL000070	225	IL000505
106	IL000065	146	IL000439	186	IL000054	226	IL000455
107	IL000183	147	IL000353	187	IL000422	227	IL000083
108	IL000490	148	IL000309	188	IL000669	228	IL000340
109	IL000434	149	IL000657	189	IL000495	229	IL000662
110	IL000552	150	IL000516	190	IL000308	230	IL000257
111	IL000613	151	IL000622	191	IL000593	231	IL000330
112	IL000475	152	IL000643	192	IL000229	232	IL000651
113	IL000522	153	IL000383	193	IL000754	233	IL000427
114	IL000566	154	IL000232	194	IL000113	234	IL000536
115	IL000089	155	IL000433	195	IL000131	235	IL000642
116	IL000601	156	IL000514	196	IL000443	236	IL000551
117	IL000156	157	IL000264	197	IL000206	237	IL000310
118	IL000588	158	IL000546	198	IL000426	238	IL000143
119	IL000233	159	IL000005	199	IL000483	239	IL000269
120	IL000616	160	IL000167	200	IL000503	240	IL000604
121	IL000540	161	IL000510	201	IL000567	241	IL000345
122	IL000425	162	IL000621	202	IL000124	242	IL000381
123	IL000154	163	IL000059	203	IL000163	243	IL000088
124	IL000606	164	IL000368	204	IL000702	244	IL000184
125	IL000718	165	IL000717	205	IL000066	245	IL000591
126	IL000412	166	IL000537	206	IL000341	246	IL000699
127	IL000649	167	IL000668	207	IL000365	247	IL000161
128	IL000148	168	IL000486	208	IL000707	248	IL000667
129	IL000421	169	IL000526	209	IL000212	249	IL000334
130	IL000093	170	IL000581	210	IL000046	250	IL000044
131	IL000414	171	IL000654	211	IL000107	251	IL000177
132	IL000165	172	IL000663	212	IL000596	252	IL000358
133	IL000058	173	IL000585	213	IL000610	253	IL000279
134	IL000415	174	IL000456	214	IL000042	254	IL000337
135	IL000470	175	IL000315	215	IL000202	255	IL000713
136	IL000258	176	IL000041	216	IL000380	256	IL000429
137	IL000213	177	IL000569	217	IL000549	257	IL000075
138	IL000226	178	IL000528	218	IL000575	258	IL000080
139	IL000326	179	IL000243	219	IL000694	259	IL000205
140	IL000645	180	IL000709	220	IL000724	260	IL000322

Rank in Random Order	Facility ID	Rank in Random Order	Facility ID	Rank in Random Order	Facility ID	Rank in Random Order	Facility ID
261	IL000562	301	IL000705	341	IL000576	381	IL000457
262	IL000632	302	IL000618	342	IL000374	382	IL000135
263	IL000609	303	IL000138	343	IL000626	383	IL000489
264	IL000711	304	IL000477	344	IL000025	384	IL000716
265	IL000468	305	IL000339	345	IL000496	385	IL000619
266	IL000404	306	IL000644	346	IL000664	386	IL000301
267	IL000245	307	IL000525	347	IL000369	387	IL000196
268	IL000211	308	IL000151	348	IL000533	388	IL000361
269	IL000087	309	IL000377	349	IL000719	389	IL000335
270	IL000449	310	IL000634	350	IL000047	390	IL000746
271	IL000485	311	IL000370	351	IL000586	391	IL000571
272	IL000458	312	IL000034	352	IL000392	392	IL000190
273	IL000423	313	IL000524	353	IL000267	393	IL000627
274	IL000120	314	IL000570	354	IL000012	394	IL000194
275	IL000583	315	IL000142	355	IL000360	395	IL000466
276	IL000515	316	IL000074	356	IL000316	396	IL000300
277	IL000696	317	IL000574	357	IL000637	397	IL000556
278	IL000547	318	IL000590	358	IL000597	398	IL000141
279	IL000015	319	IL000021	359	IL000136	399	IL000250
280	IL000564	320	IL000560	360	IL000461	400	IL000121
281	IL000659	321	IL000493	361	IL000607	401	IL000420
282	IL000239	322	IL000026	362	IL000555	402	IL000101
283	IL000545	323	IL000215	363	IL000479	403	IL000520
284	IL000428	324	IL000655	364	IL000584	404	IL000445
285	IL000270	325	IL000544	365	IL000302	405	IL000517
286	IL000248	326	IL000462	366	IL000094	406	IL000633
287	IL000082	327	IL000509	367	IL000641	407	IL000723
288	IL000451	328	IL000098	368	IL000247	408	IL000444
289	IL000706	329	IL000491	369	IL000109	409	IL000296
290	IL000623	330	IL000628	370	IL000366	410	IL000180
291	IL000096	331	IL000350	371	IL000498	411	IL000535
292	IL000612	332	IL000364	372	IL000407	412	IL000323
293	IL000067	333	IL000497	373	IL000573	413	IL000398
294	IL000500	334	IL000283	374	IL000561	414	IL000193
295	IL000278	335	IL000726	375	IL000100	415	IL000469
296	IL000487	336	IL000703	376	IL000281	416	IL000060
297	IL000011	337	IL000432	377	IL000572	417	IL000166
298	IL000307	338	IL000532	378	IL000595	418	IL000473
299	IL000589	339	IL000037	379	IL000140	419	IL000708
300	IL000554	340	IL000476	380	IL000615	420	IL000691

Rank in Random Order	Facility ID	Rank in Random Order	Facility ID	Rank in Random Order	Facility ID		
421	IL000582	461	IL000387	501	IL000523		
422	IL000057	462	IL000198	502	IL000715		
423	IL000031	463	IL000413	503	IL000578		
424	IL000406	464	IL000185	504	IL000294		
425	IL000630	465	IL000508	505	IL000701		
426	IL000271	466	IL000265	506	IL000354		
427	IL000482	467	IL000665	507	IL000097		
428	IL000480	468	IL000670	508	IL000518		
429	IL000295	469	IL000182	509	IL000409		
430	IL000240	470	IL000617	510	IL000396		
431	IL000666	471	IL000069	511	IL000448		
432	IL000274	472	IL000027	512	IL000386		
433	IL000471	473	IL000217	513	IL000611		
434	IL000298	474	IL000199	514	IL000017		
435	IL000725	475	IL000134	515	IL000020		
436	IL000099	476	IL000352	516	IL000446		
437	IL000024	477	IL000014	517	IL000437		
438	IL000338	478	IL000356	518	IL000631		
439	IL000474	479	IL000656	519	IL000710		
440	IL000342	480	IL000357	520	IL000507		
441	IL000484	481	IL000219	521	IL000440		
442	IL000638	482	IL000078	522	IL000303		
443	IL000393	483	IL000210	523	IL000424		
444	IL000472	484	IL000481	524	IL000174		
445	IL000297	485	IL000447	525	IL000698		
446	IL000410	486	IL000231	526	IL000646		
447	IL000438	487	IL000506				
448	IL000222	488	IL000359				
449	IL000320	489	IL000697				
450	IL000733	490	IL000145				
451	IL000418	491	IL000062				
452	IL000061	492	IL000579				
453	IL000220	493	IL000452				
454	IL000600	494	IL000311				
455	IL000321	495	IL000050				
456	IL000375	496	IL000467				
457	IL000712	497	IL000460				
458	IL000203	498	IL000672				
459	IL000256	499	IL000450				
460	IL000293	500	IL000499				

Category 2 (Not applicable): 4 site visits

Facility ID	Reason for Inspection
IL000071	Facility certified as nonautomotive, but phone call revealed that it is automotive
IL000290	Facility certified as nonautomotive, but phone call revealed that it is automotive
IL000168	Based on phone call, State unable to determine whether facility is automotive
IL000314	Based on phone call, State unable to determine whether facility is automotive

Category 3 (Has a well): 0 inspections

Facilities in this category will be handled through the State's normal administrative process.

Category 4 (Does not have a well): 22 inspections

A random sample of 22 facilities has been selected for inspection out of the 91 facilities in the category.

Rank in Random Order	Facility ID	Rank in Random Order	Facility ID
1	IL000317	12	IL000191
2	IL000035	13	IL000040
3	IL000328	14	IL000152
4	IL000214	15	IL000343
5	IL000327	16	IL000118
6	IL000275	17	IL000170
7	IL000092	18	IL000139
8	IL000372	19	IL000254
9	IL000218	20	IL000304
10	IL000408	21	IL000108
11	IL000176	22	IL000144

If the State decides to inspect additional facilities in category 4, those facilities should be taken in order from the following continuation of the randomized list:

Rank in Random Order	Facility ID	Rank in Random Order	Facility ID	Rank in Random Order	Facility ID	Rank in Random Order	Facility ID
23	IL000246	41	IL000112	59	IL000028	77	IL000155
24	IL000312	42	IL000043	60	IL000268	78	IL000282
25	IL000313	43	IL000289	61	IL000259	79	IL000086
26	IL000727	44	IL000111	62	IL000740	80	IL000347
27	IL000417	45	IL000285	63	IL000029	81	IL000273
28	IL000391	46	IL000085	64	IL000344	82	IL000263
29	IL000253	47	IL000192	65	IL000179	83	IL000201
30	IL000172	48	IL000291	66	IL000063	84	IL000073
31	IL000722	49	IL000095	67	IL000276	85	IL000624
32	IL000224	50	IL000013	68	IL000033	86	IL000204
33	IL000076	51	IL000127	69	IL000306	87	IL000729
34	IL000399	52	IL000373	70	IL000162	88	IL000228
35	IL000004	53	IL000053	71	IL000081	89	IL000045
36	IL000237	54	IL000388	72	IL000018	90	IL000367
37	IL000252	55	IL000284	73	IL000249	91	IL000351
38	IL000178	56	IL000114	74	IL000286		
39	IL000236	57	IL000003	75	IL000319		
40	IL000130	58	IL000119	76	IL000150		

Category 5 (Uncertain response): 7 inspections

All seven facilities in this category are to be inspected, as their self-certification forms did not provide enough information to determine whether they have a well.

Facility ID
IL000007
IL000016
IL000023
IL000189
IL000251
IL000394
IL000731