

CONFIDENTIAL: SUBJECT TO PROTECTIVE ORDER

EXPERT REPORT OF PHILIP B. STARK, PH.D.

8 May 2006

REDACTED

ACLU v. Gonzales

Civ. Action No. 98-5591 (E.D. Pa.)

EXPERT REPORT OF PHILIP B. STARK, PH.D.  
8 May 2006

I. INTRODUCTION

1. I am Professor of Statistics at the University of California, Berkeley, and the author of many research articles. I have lectured at universities and professional societies around the world. I have research and consulting experience relating to the Internet. Appendix III gives my qualifications in more detail, and lists recent testimony. Appendix IV is my CV.

2. This report studies the prevalence of sexually explicit websites on the world-wide web, the frequency with which searches return sexually explicit websites, and the efficacy of content filters in blocking those websites. Among other things, the report gives estimates of:

- The percentage of sexually explicit websites<sup>1</sup> among all websites cataloged by the MSN and Google search engines.
- The percentage of websites based in the United States among all sexually explicit websites cataloged by the MSN and Google search engines.
- The percentage of AOL, MSN and Yahoo! queries that retrieve at least one sexually explicit website.
- The effectiveness of several filters at blocking sexually explicit websites.

---

<sup>1</sup> In this document I generally use the terms “website” and “URL” (Uniform Resource Locator) interchangeably. “Sexually explicit” is defined in paragraph 8.

## II. DATA

3. The estimates in this report are based on counting websites and search queries.

When a search engine receives a query, the search engine retrieves results from a catalog of websites, called an index. Each search engine has its own index, and the major search providers add and remove websites from their indexes regularly.

They also record the queries that they receive.

4. The government obtained 50,000 random websites from the Google index and 1 million random websites from the MSN index.<sup>2</sup> I specified the method and the random numbers used to select the websites.<sup>3</sup> The government also obtained a week of search queries from AOL, MSN and Yahoo!.<sup>4</sup> The government asked Google for queries, but Google refused. Google, Yahoo!, MSN and AOL have the

---

2 The MSN sample was drawn in November 2005 and the Google sample was drawn in March 2006. Yahoo! also provided 1 million websites from its index, but those data were not reliable enough to use: two domains—[www.cracks.me.uk](http://www.cracks.me.uk) and the anesthesiology department at the University of Washington—comprised about 5% of the Yahoo! sample. At the time, [www.cracks.me.uk](http://www.cracks.me.uk) had a sexually explicit banner ad. (A domain is the “root” of a web address. For example, [www.stat.berkeley.edu](http://www.stat.berkeley.edu) is a domain; [www.stat.berkeley.edu/users/stark/index.html](http://www.stat.berkeley.edu/users/stark/index.html) is one of the pages in that domain.)

3 The search providers produced the URLs that corresponded to random numbers I drew.

4 The queries were from 2005: AOL from 22–28 July, MSN from 17–23 July, and Yahoo! from 18–24 August. The AOL and MSN queries had weights, which appear to be the number of times each query was run each day. In round numbers, AOL provided [REDACTED] queries with total weight [REDACTED]; MSN provided [REDACTED] queries with total weight [REDACTED]; and Yahoo! provided [REDACTED] queries. Problems in decompressing the AOL data resulted in the exclusion of a set of queries with total weight 8.4 million. The estimates in this report are thus based on a sample from nearly [REDACTED] searches. Data on the queries that were excluded could not change the estimates materially: 8.4 million is about [REDACTED].

bulk of the search market, although AOL does not have its own search engine.<sup>5</sup>

Thus, the government had websites from Google and MSN, and queries from AOL, MSN and Yahoo!. I drew random samples of the queries.<sup>6</sup>

5. These random samples of websites and queries were given to CRA International, which, I understand, performed the following steps.<sup>7</sup> CRA International viewed random samples of 11,100 of the Google websites and 39,999 of the MSN websites and categorized their content. Sexually explicit adult entertainment websites belong to category 5f. Websites with no nudity or sexual content belong to category 1a. Websites with sexual content belong to other categories if the context is educational, medical or artistic. CRA International ran random samples of queries through search engines<sup>8</sup> and categorized the first ten websites each search retrieved.<sup>9</sup> (See paragraph 14.) CRA International sent me lists of the category 1a websites they had found from indexes and search results. I drew random samples from those lists and sent the samples back. CRA International then tested filters on the random samples of category 1a websites and

---

5 It is my understanding that Google processes searches for AOL, and that differences in the results of Google and AOL searches are small.

6 I took random samples without replacement from each search provider's queries. The weights provided by AOL and MSN were factored into estimates. See Appendix II.

7 More detail will be found in the 8 May 2006 Expert Report of Paul Mewett.

8 CRA International ran [REDACTED] AOL queries through the Google search engine, [REDACTED] MSN queries through the MSN search engine, and [REDACTED] Yahoo! queries through the Yahoo! search engine. I provided lists of queries in random order for CRA International to study.

9 If fewer than ten websites were retrieved by a query, all were used.

on all the category 5f websites.

6. Some queries are far more popular than average. Wordtracker markets lists of the most popular search terms.<sup>10</sup> CRA International processed 685 of the top Wordtracker queries<sup>11</sup> the same way they processed the other queries.<sup>12</sup>

7. Finally, CRA International determined the country of origin of the category 5f websites they had found.

8. CRA International sent me a database that classified websites by content. The database showed which websites were used to test each filter, and whether the websites were blocked. It also gave the country of origin of the category 5f websites. These data are the basis of all the estimates in this report. I checked the internal consistency of the data. Beyond that, I did not verify the work CRA International performed. Throughout this report, “sexually explicit” means material CRA International would put in category 5f, and “clean” means material CRA International would put in category 1a. A website is “domestic” if CRA International would identify its host country to be the United States.

---

<sup>10</sup> I understand that Wordtracker collects queries from Dogpile.com and MetaCrawler, “meta search” services that send users’ queries to a variety of search engines—including MSN, Yahoo!, Google, and Ask.com—and collate the results.

<sup>11</sup> These were queries from 12 November 2005 through 20 February 2006.

<sup>12</sup> That processing included testing filters on a random sample of the category 1a search results. I drew the random sample.

### III. METHODS AND RESULTS

9. Random samples can be used to make statistical inferences about the populations from which they are drawn. This report uses samples to make inferences about websites cataloged by Google or MSN and about the set of queries from which the sample was drawn.<sup>13</sup> Appendix I gives detailed results and some uncertainty estimates. Appendix II summarizes the statistical methods used.

#### IIIA. SEARCH ENGINE INDEXES

10. I estimate that 1.1 percent of the websites cataloged by Google and 1.1 percent of the websites cataloged by MSN are sexually explicit. The numbers are the same to one decimal place. I estimate that 44.2 percent of the sexually explicit websites in the Google index are domestic and that 56.6 percent of the sexually explicit websites in the MSN index are domestic.

11. Content filters fail to block some sexually explicit websites. I estimate that 8.8 percent to 60.2 percent of the sexually explicit websites in the Google and MSN indexes are not blocked by the filters CRA International tested—different filters block different percentages. I estimate that 31.6 percent to 57.1 percent of the sexually explicit websites that filters do not block are domestic. Tables 2 and 4 in

<sup>13</sup> All the estimates are specific to the times the data were collected. Some variation with time is to be expected.

Appendix I give detailed results.

12. Content filters block some clean websites. Of the clean websites cataloged by Google or MSN, I estimate that 0.4 percent to 23.6 percent are blocked by filters. Table 2 in Appendix I gives detailed results.

13. Generally, if a filter blocks more of the sexually explicit websites, it will block more of the clean websites. To take an extreme example, a parent could block all sexually explicit websites by turning off the computer. The filter that blocked all but 8.8 percent of the sexually explicit websites in the Google and MSN indexes also blocked over 22 percent of the clean websites. Tables 2, 5 and 7 in Appendix I show the tradeoff for all the filters that were tested.

### IIIB. QUERIES

14. The estimates presented so far concern websites cataloged by Google and MSN. Next, I will give similar estimates for search results. Figure 1 shows an example of Google search results.<sup>14</sup> The query, “jenna jameson,”<sup>15</sup> is in the box next to the gray button labeled “Search.” Below that, the text in the shaded blue bar says “Results 1–10 of about 12,400,000 for jenna jameson.” The page gives the first ten results Google retrieved from its index in response to the query. There

---

<sup>14</sup> The search was run at 2:05pm PDT on 3 May 2006.

<sup>15</sup> A recent Wordtracker report lists “jenna jameson” as the 20<sup>th</sup> most popular query.

are blocks of text on the left side of the page, below the shaded blue bar. Each block starts with a blue underlined phrase in a larger font. The underlined phrase is a search result, that is, a link to a website in the Google index. The rest of the block contains a description, the URL<sup>16</sup> (web address) of the search result, the size of the page when Google last checked it, a link to a copy of the page Google stored in its own computers (if Google cached the page), and a link to a list of related websites. For instance, the first search result in Figure 1 is a page titled “Jenna Jameson (I),” which has the URL <http://www.imdb.com/name/nm0001398/>. The blocks of text and links on the right side of the page are paid advertisements.<sup>17</sup> The estimates that will be discussed next concern the first ten search results retrieved by each query. Typically, the first page of search results contains about ten websites.

---

16 The Google search result does not show the protocol portion of the URL (<http://>).

17 “Sponsored links” and other advertisements were not studied.



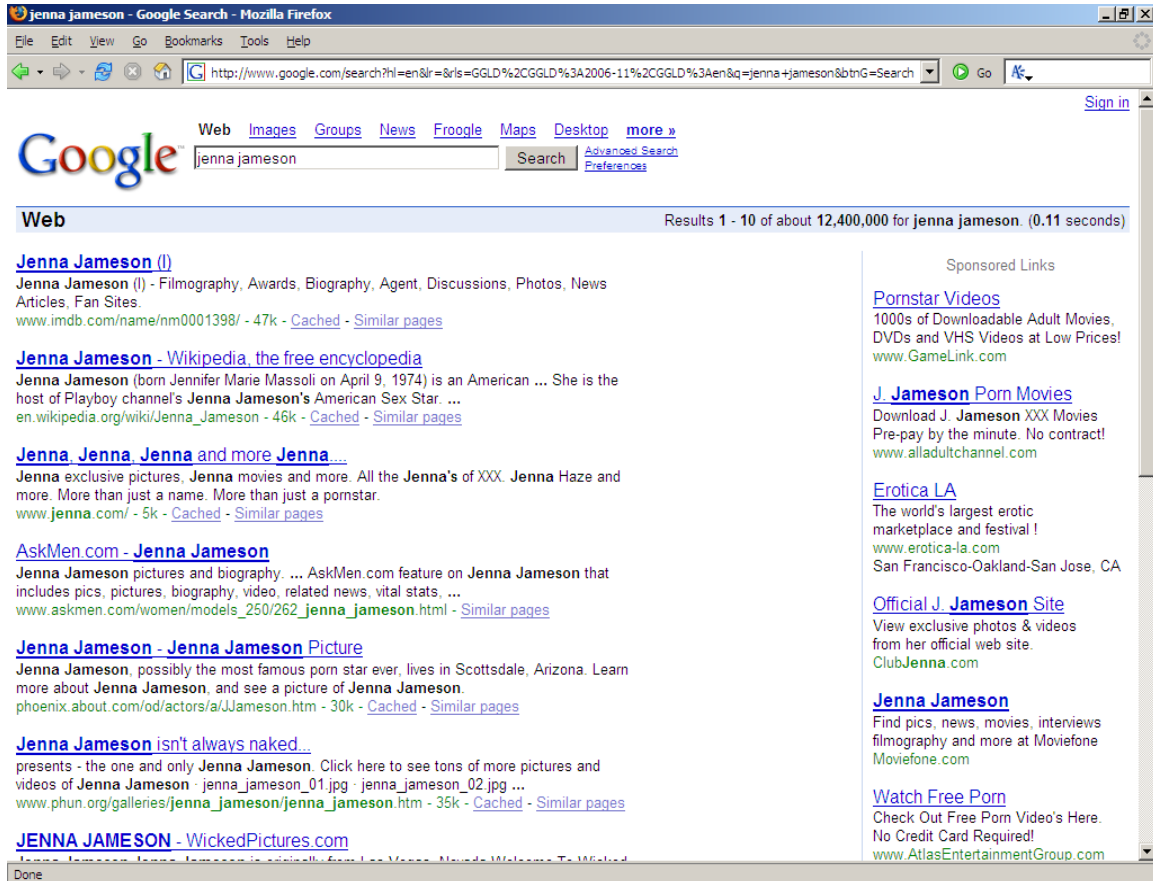


Figure 1: Google search results for “jenna jameson” on 3 May 2006. See paragraph 14.

15. I now report estimates for queries. I estimate that 6 percent of the AOL, MSN, and Yahoo! queries retrieve at least one sexually explicit website and that 5.7 percent retrieve at least one domestic sexually explicit website. I estimate that 1.7 percent of the AOL, MSN and Yahoo! search results are sexually explicit. This is somewhat higher than the corresponding estimate for the Google and MSN indexes, 1.1 percent.<sup>18</sup> Sexually explicit websites seem to be disproportionately popular. Of the sexually explicit websites returned by AOL, MSN and Yahoo!

<sup>18</sup> See paragraph 10.

queries, I estimate that 87 percent are domestic. This, too, is higher than for the Google and MSN indexes: domestic sexually explicit websites seem especially popular.

16. Of the sexually explicit websites returned by AOL, MSN and Yahoo! queries, I estimate that 6.2 percent to 43.4 percent are not blocked by filters. Of those, I estimate that 33.8 percent to 91.9 percent are domestic. Of the clean websites retrieved by the AOL, MSN and Yahoo! searches, I estimate that 0 percent<sup>19</sup> to 20.7 percent are blocked by filters. These are percentages of websites retrieved by queries.

17. Now I focus on queries that retrieve at least one sexually explicit website. Of the AOL, MSN and Yahoo! queries that retrieve sexually explicit websites, I estimate that 15.6 percent to 65.6 percent retrieve at least one sexually explicit website that would not be blocked by filters. Table 5 in Appendix I gives more detail.

18. The queries discussed so far were drawn at random from AOL, MSN and Yahoo! queries. As mentioned in paragraph 6, a relatively small number of queries are extremely popular. Among the 685 queries Wordtracker determined to be most

<sup>19</sup> Filters that did not block any clean websites failed to block 20.4 percent to 43.4 percent of the sexually explicit websites.

popular, 37.3 percent retrieved at least one sexually explicit website and 37.2 percent retrieved at least one domestic sexually explicit website. Among the Wordtracker search results, 13.9 percent were sexually explicit websites, of which 87.4 percent were domestic. The percentage of sexually explicit results for popular queries is over six times larger than for the AOL, MSN and Yahoo! queries; the same is true for the percentage of queries that return sexually explicit results. The most popular searches retrieve far more than their share of sexually explicit results.

19. I now report the performance of filters on results from Wordtracker searches. Of the sexually explicit Wordtracker search results, 1.4 percent to 12.8 percent were not blocked by filters. Of those, 67.3 percent to 91.5 percent were domestic. Filters blocked 2.9 percent to 32.8 percent of the clean Wordtracker results. Of the Wordtracker queries that returned at least one sexually explicit website, 4.6 percent to 34.2 percent retrieved at least one sexually explicit website that was not blocked. Table 7 in Appendix I gives detailed results.

#### IV. BIASES AND UNCERTAINTIES

20. There are many downward biases in the estimates of prevalence of sexually explicit websites and of the rates at which filters fail to block sexually explicit websites or block clean websites. For example, I counted queries that did not retrieve any working websites in the denominator of estimates of the prevalence of

sexually explicit material. And I understand that category 5f is very restrictive: there must be sexually explicit content that is clearly adult entertainment, and that content must be visible without clicking anything—not even the “play” button of a video. I also understand that category 1a is quite restrictive: there can be no nudity or sexual content whatsoever, not even in a medical or educational context. See Report of Paul Mewett, 8 May 2006.

21. Estimates from random samples are subject to uncertainty from bias and sampling error (the luck of the draw). Confidence limits measure the uncertainty due to sampling error. Conservative confidence limits<sup>20</sup> can be calculated for many of the estimates given here; approximate confidence limits can be calculated for all the estimates. See Appendices I and II.

## V. SUMMARY

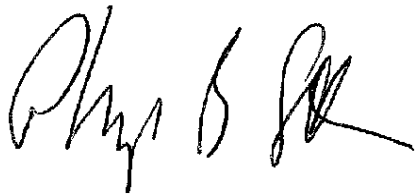
22. This study reports on the Google and MSN indexes, on AOL, MSN and Yahoo! queries, and on the most popular Wordtracker queries. About 1 percent of the websites in the Google and MSN indexes are sexually explicit. About 6 percent of queries retrieve a sexually explicit website. Nearly 40 percent of the most popular queries retrieve a sexually explicit website. Close to 90 percent of the sexually explicit websites retrieved by queries are domestic. Filters that block

---

<sup>20</sup> “Conservative” means the confidence level is greater than the claimed 95% level.

more of the sexually explicit websites also block more of the clean websites. The most restrictive filter blocks about 94 percent of the sexually explicit search results, but also blocks about 13 percent of the clean results. Of the sexually explicit websites that get through the filters, 30 percent to 90 percent are domestic.

23. The number of sexually explicit websites is huge. Search results often include sexually explicit material. A lot of sexually explicit material is not blocked by filters. Of that, a substantial percentage is domestic.



Dated 8 May 2006.

---

Philip B. Stark

## APPENDIX I: DETAILED ESTIMATES AND UNCERTAINTIES

I.1. As noted in paragraph 21, it is possible to compute conservative confidence limits for many of the estimated quantities, and it is possible to compute approximate confidence intervals for all of them. This appendix presents some conservative confidence limits. It also gives estimates of the performance of individual filters.

I.2 Table 1 shows lower 95% confidence limits for the percentage of sexually explicit websites and the percentage of domestic sexually explicit websites in the Google and MSN indexes. Table 1 also gives lower confidence limits for the percentage of the AOL, MSN and Yahoo! queries that retrieve at least one sexually explicit website, and for the percentage that retrieve at least one domestic sexually explicit website. Appendix II explains how I calculated confidence limits.

	Google websites	MSN websites	AOL, MSN and Yahoo! queries
Sexually explicit	1.0%	1.0%	2.5%
Domestic sexually explicit	0.4%	0.5%	2.2%

Table 1: Conservative 95% lower confidence limits. The second and third columns are lower confidence limits for the percentage of sexually explicit websites among all websites in the Google and MSN indexes, and for the percentage of domestic sexually explicit websites. The fourth column gives lower confidence limits for the percentage of the AOL, MSN and Yahoo! queries that return at least one sexually explicit website, and for the percentage of AOL, MSN and Yahoo! queries that return at least one domestic sexually explicit website.

I.3. “Underblocking” is the percentage of sexually explicit websites that filters do not block, and “overblocking” is the percentage of clean websites that filters block.

Table 2 gives estimates and Table 3 gives lower confidence limits, for the Google and MSN indexes. Table 4 gives estimates of the percentage of domestic sexually explicit websites among the sexually explicit websites in the Google and MSN indexes that filters do not block. Table 5 gives estimates of overblocking and underblocking for websites retrieved by AOL, MSN and Yahoo! queries. Table 6 gives lower confidence limits for the percentage of AOL, MSN and Yahoo! queries that return at least one sexually explicit website that would not be blocked by a filter. Table 7 shows underblocking and overblocking for Wordtracker queries.

Filter <sup>21</sup>	Underblocking		Overblocking	
	Google	MSN	Google	MSN
1	8.9%	8.8%	22.6%	23.6%
2a	16.8%	18.9%	19.6%	10.3%
2b	17.7%	20.5%	21.9%	18.9%
3a	38.3%	45.3%	2.8%	3.0%
3b	28.3%	46.6%	1.4%	0.7%
4	31.0%	33.7%	1.4%	0.9%
5a	12.7%	16.7%	3.6%	4.1%
5b	12.4%	19.1%	4.0%	3.7%
6	16.1%	26.2%	12.4%	13.2%
7	44.0%	46.3%	3.3%	2.2%
8a	60.2%	55.0%	1.4%	0.7%
8b	58.4%	54.3%	0.9%	0.4%

Table 2: Estimated underblocking and overblocking of websites in the Google and MSN indexes. Among sexually explicit websites, the percentage that are not blocked by a filter is the rate of underblocking. Among clean websites, the percentage that are blocked by a filter is the rate of overblocking. The filter settings and testing protocol will be explained in the 8 May 2006 Expert Report of Paul Mewett.

---

21 The filters are as follows; settings will be described more fully in the 8 May 2006 Expert Report of Paul Mewett. 1: AOL Mature Teen. 2a: MSN Pornography. 2b: MSN Teen. 3a: ContentProtect Default setting. 3b: ContentProtect Custom setting. 4: CyberPatrol Custom setting. 5a: CyberSitter Default setting. 5b: CyberSitter Custom setting. 6: McAfee Young Teen. 7: Net Nanny Level 2. 8a: Norton Default setting. 8b: Norton Custom setting.

Filter	Underblocking		Overblocking	
	Google	MSN	Google	MSN
1	5.5%	6.7%	18.4%	21.0%
2a	12.0%	15.9%	15.8%	8.5%
2b	12.8%	17.3%	17.8%	16.6%
3a	31.3%	41.2%	1.5%	2.1%
3b	22.2%	42.5%	0.6%	0.4%
4	24.6%	29.9%	0.6%	0.5%
5a	8.6%	13.8%	2.1%	3.1%
5b	8.4%	16.1%	2.4%	2.7%
6	11.4%	22.7%	9.3%	11.3%
7	36.8%	42.0%	1.9%	1.5%
8a	52.9%	50.8%	0.6%	0.4%
8b	51.1%	50.2%	0.4%	0.2%

Table 3: 95% lower confidence limits for the entries in Table 2. For illustration, at 95% confidence, filter 2b fails to block at least 12.8% of the sexually explicit websites in the Google index. Similarly, at 95% confidence, filter 2b blocks at least 16.6% of the clean websites in the MSN index.

Filter	Estimated Domestic Underblocking	
	Google	MSN
1	40.0%	39.4%
2a	31.6%	42.3%
2b	40.0%	37.7%
3a	39.0%	45.8%
3b	40.6%	47.1%
4	48.6%	43.7%
5a	50.0%	32.3%
5b	57.1%	35.7%
6	44.4%	37.1%
7	41.7%	47.8%
8a	35.3%	49.0%
8b	36.4%	49.5%

Table 4: Of the sexually explicit websites in the Google and MSN indexes that filters do not block, the estimated percentage that are domestic websites.



Filter	Underblocking for results	Overblocking for results	Domestic Underblocking	Underblocking for queries
1	6.2%	12.5%	57.0%	15.6%
2a	21.4%	4.4%	86.1%	32.3%
2b	20.1%	5.8%	91.9%	28.1%
3a	18.4%	6.4%	70.1%	46.2%
3b	20.4%	0.0%	62.1%	42.2%
4	34.6%	0.4%	91.9%	65.6%
5a	11.2%	4.6%	33.8%	23.2%
5b	10.0%	5.3%	44.1%	20.1%
6	14.2%	20.7%	80.7%	30.9%
7	28.1%	3.7%	76.6%	36.6%
8a	42.1%	0.8%	82.9%	51.6%
8b	43.4%	0.0%	83.4%	56.1%

Table 5: Estimated underblocking and overblocking of the results of AOL, MSN and Yahoo! searches. “Underblocking for results” is the fraction of sexually explicit search results that are not blocked. “Overblocking for results” is the percentage of clean search results that are blocked. “Domestic underblocking” is the percentage of domestic websites among the sexually explicit websites the filters do not block. “Underblocking for queries” is, among the queries that retrieve any sexually explicit websites, the percentage that retrieve at least one sexually explicit website that is not blocked. Table 6 gives lower 95% confidence limits for underblocking for queries.

Filter	Underblocking for queries
1	5.3%
2a	20.9%
2b	18.8%
3a	10.0%
3b	25.4%
4	24.4%
5a	11.2%
5b	8.1%
6	10.4%
7	20.8%
8a	49.3%
8b	54.3%

Table 6: Lower 95% confidence limits for the rightmost column in Table 5. Appendix II explains how these limits were calculated.

Filter	Underblocking for results	Overblocking for results	Domestic Underblocking	Underblocking for queries
1	1.4%	19.5%	70.8%	4.6%
2a	2.7%	13.3%	91.5%	8.1%
2b	2.6%	13.8%	88.7%	8.2%
3a	7.7%	12.5%	84.3%	23.4%
3b	8.3%	7.8%	85.1%	25.6%
4	4.1%	9.2%	85.2%	10.8%
5a	1.6%	20.0%	67.3%	5.7%
5b	3.0%	18.1%	84.4%	9.6%
6	3.1%	32.8%	72.7%	10.0%
7	12.8%	9.1%	84.1%	34.2%
8a	10.2%	4.6%	79.7%	25.6%
8b	10.5%	2.9%	79.7%	26.2%

Table 7: Underblocking and estimated overblocking for the results of Wordtracker queries. “Underblocking for results” is the percentage of sexually explicit search results that are not blocked. “Overblocking for results” is the percentage of clean search results the filter blocks. “Domestic underblocking” is the percentage of domestic websites among the sexually explicit websites the filters do not block. “Underblocking for queries” is, among the queries that retrieve any sexually explicit websites, the percentage that retrieve at least one sexually explicit website that is not blocked. Overblocking was estimated from a random sample of clean search results. Underblocking was determined from all the sexually explicit search results.

## APPENDIX II: TECHNICAL NOTES

II.1. I used sample percentages to estimate the percentage of sexually explicit websites in the Google and MSN indexes, the percentage of sexually explicit websites that originate in the United States, and overblocking and underblocking (Table 2). Sample percentages are unbiased in this problem. Because the samples are minute fractions of the populations, the number of items in the sample with a given property has essentially a binomial distribution. (The exact distribution is hypergeometric, which is less dispersed than the binomial; the binomial approximation results in conservative confidence limits.) I used the binomial distribution to find lower confidence limits for the percentage of sexually explicit websites in the Google and MSN indexes, for the percentage of domestic sexually explicit websites, and for underblocking and overblocking of websites in the Google and MSN indexes (Tables 1 and 3). The confidence limits are conditional on the number of working URLs in each sample.

II.2. I estimated the percentage of AOL, MSN and Yahoo! queries that had various properties using a weighted sample percentage. The sampled lines from the AOL and MSN files were weighted using the frequencies AOL and MSN supplied. To get an overall estimate for the pool of queries from which the sample was drawn, I combined weighted sample percentages for AOL, MSN and Yahoo! using the fraction of queries each vendor supplied. In this problem, the weighted sample

percentage is a ratio estimator, which may have some bias, although I expect the bias to be small.

II.3. The exact probability distribution of the weighted sample percentage is not computable in this problem. To calculate uncertainties for population percentages of queries, I gave all queries from a given search provider equal weight. The lower confidence limits are therefore biased downward, because queries that retrieve sexually explicit materials have higher than average weight. (The Wordtracker, AOL and MSN query data confirm this.) As a result, the confidence limits are conservative: the confidence level is greater than claimed. Because each sample is a small fraction of the set of queries from which it is drawn, the number of queries with a given property in the sample has essentially a binomial distribution. I used the independence of the samples to find a conservative overall confidence limit.

### APPENDIX III: QUALIFICATIONS AND EXPERIENCE

I have been on the Statistics faculty at the University of California, Berkeley, since 1988. I have been a Miller Research Professor, a Dodson Fellow, a Presidential Chair Fellow, and a Mellon/Library Faculty Fellow. I received a Bachelor's degree from Princeton University in 1980 and a Ph.D. from the University of California, San Diego, in 1986. I was a Presidential Young Investigator and a National Science Foundation Postdoctoral Fellow in Mathematical Sciences. I have been on the editorial board of several journals. I have written over 65 articles and technical reports. I have given roughly 130 invited lectures at scientific conferences and universities in 16 countries. I have testified to the U.S. House of Representatives Subcommittee on the Census and the California Senate Natural Resources Committee. I have consulted for the U.S. Department of Justice, the Federal Trade Commission, the U.S. Department of Agriculture, the U.S. Census Bureau, the U.S. Attorney's Office of the Northern District of California, the U.S. Department of Veterans Affairs, the Los Angeles County Superior Court, the National Solar Observatory, public utilities, major corporations, and numerous law firms. I have been an expert witness or non-testifying expert in cases involving antitrust, consumer class actions, employment discrimination, equal protection, fairness in lending, federal legislation, insurance, intellectual property, product liability, trade secrets, truth in advertising, wage and hour disputes, and other matters. Some of

my consulting and research relates to the Internet, including characterizing and predicting online consumer behavior and developing search algorithms. I created a web-based statistics course using HTML, JavaScript and Java, the most widely used web languages. I have been on the advisory boards of a web marketing firm and two online publishers.

In the last four years, I have been deposed or given testimony in three cases:

Richison et al. vs. American Cemwood Corporation (San Joaquin County Superior Court, Case No. 005532), Pacific Gas and Electric Co. vs. City and County of San Francisco (U.S. District Court, Northern District of California, Case No. C99-2071 VRW), and Star Scientific, Inc. vs. R.J. Reynolds Tobacco Company (U.S. District Court, Maryland District, Northern Division, Case Nos. MJG-01 1504 and MJG-02 2504). I charge \$600 per hour for consulting and \$750 per hour for deposition and testimony. My compensation does not depend on the outcome of the case. My CV is attached as Appendix IV.

## APPENDIX IV

# Curriculum Vitae

Philip Bradford Stark

### Biographical Information

- **Born:** 7 October 1960, Houston, Texas.
- **Citizenship:** U.S.A.

### Interests

- **Theory:** Inverse problems, multiplicity, nonparametrics, optimization, restricted parameters
- **Applications:** Astrophysics, cosmology, geophysics, legislation and litigation, hearing, educational technology, web computing, information retrieval, marketing

### Education

- B.A. 1980, Princeton University, Princeton, New Jersey
- Ph.D. 1986, University of California, San Diego, La Jolla, California

### Awards and Fellowships

- Mellon Library/Faculty Fellow for Undergraduate Research (2006–2007)
- Presidential Chair Fellow, University of California, Berkeley (2003–2004)
- Fellow, Institute of Physics (elected 1999)
- Miller Research Professor, Miller Institute for Basic Research in Science (1999)
- Dobson Fellow, University of California, Berkeley (1998, 1999)
- Presidential Young Investigator (1989–1995)
- National Science Foundation Postdoctoral Fellowship in Mathematical Sciences (1987–1989)
- University Fellowship, University of Texas at Austin (1982–1983)

### Societies and Affiliations

- American Geophysical Union
- Bernoulli Society for Mathematical Statistics and Probability
- Center for Astrostatistics (Penn State)
- Center for Data Analysis Technology and Applications (DATA)
- Global Oscillation Network Group (GONG)
- Fellow and Chartered Physicist, Institute of Physics

- Institute of Mathematical Statistics
- National Partnership for Advanced Computational Infrastructure (NPACI)
- Royal Astronomical Society
- Solar and Heliospheric Observatory Solar Oscillations Investigation (SOHO-SOI)
- Space Sciences Laboratory, University of California, Berkeley
- Theoretical Astrophysics Center, University of California, Berkeley

## **Employment**

- 7/98–present Professor, Department of Statistics, University of California, Berkeley
- 7/01–6/03 Faculty Assistant in Educational Technology (to Vice Provost for Undergraduate Education), University of California, Berkeley
- 7/94–6/98 Associate Professor, Department of Statistics, University of California, Berkeley
- 7/88–6/94 Assistant Professor, Department of Statistics, University of California, Berkeley
- 7/87–6/90 National Science Foundation Postdoctoral Fellow in Mathematical Sciences
- 1/87–6/87 Postgraduate Research, Department of Statistics, University of California, Berkeley
- 8/86–12/86 Postgraduate Research, Institute for Geophysics and Planetary Physics, University of California, San Diego

## **Visiting Positions**

- 6/96 Visiting Associate Professor, School of Mathematical Sciences, Tel Aviv University, Tel Aviv, Israel

## **Former Students and Postdocs**

- Imola K. Fodor, Lawrence Livermore National Laboratory
- Christopher R. Genovese, Carnegie Mellon University
- Niklaus W. Hengartner, Los Alamos National Laboratory



- R. Jay Pulliam, University of Texas
- Chad M. Schafer, Carnegie Mellon University

## Mentors

- Robert L. Parker, Institute for Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California, San Diego (PhD dissertation advisor)
- George E. Backus, Institute for Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California, San Diego (postdoctoral advisor)
- David L. Donoho, Department of Statistics, Stanford University (postdoctoral advisor)

## Publications

### Refereed Publications

1. Stark, P.B. and C. Frohlich, 1985. The depths of the deepest deep Earthquakes, *J. Geophys. Res.*, *90*, 1859–1869.
2. Stark, P.B., R.L. Parker, G. Masters and J.A. Orcutt, 1986. Strict bounds on seismic velocity in the spherical Earth, *J. Geophys. Res.*, *91*, 13,892–13,902.
3. Stark, P.B. 1986. *Travel–Time Inversion: Regularization and Inference*, Ph.D. Thesis, Scripps Institution of Oceanography, University of California, San Diego, 106pp.
4. Stark, P.B. and R.L. Parker, 1987. Smooth profiles from tau (p) and X(p) data, *Geophys. J. R. astr. Soc.*, *89*, 2713–2719.
5. Stark, P.B. and R.L. Parker, 1987. Velocity bounds from statistical estimates of tau(p) and X(p), *J. Geophys. Res.*, *92*, 2713–2719.
6. Stark, P.B., 1987. Rigorous velocity bounds from soft tau (p) and X(p) data, *Geophys. J. R. astr. Soc.*, *89*, 987–996.
7. Orcutt, J.A., R.L. Parker, P.B. Stark and J.D. Garmany, 1988. Comment concerning "A method of obtaining a velocity-depth envelope from wide-angle seismic data" by R. Mithal and J.B. Diebold. *Geophys. J.*, *95*, 209–212.
8. Stark, P.B. and R.L. Parker, 1988. Correction to "Velocity bounds from statistical estimates of tau(p) and X(p)." *J. Geophys. Res.*, *93*, 13,821–13,822.
9. Donoho, D.L. and P.B. Stark, 1989. Uncertainty principles and signal recovery. *SIAM J. Appl. Math.*, *49*, 906–931.
10. Stark, P.B., 1992. Affine minimax confidence intervals for a bounded Normal mean, *Statistics and Probability Letters*, *13*, 39–44.
11. Stark, P.B., 1992. Minimax confidence intervals in geomagnetism, *Geophys. J.*

- Intl.*, 108, 329–338.
12. Stark, P.B., 1992. Inference in infinite-dimensional inverse problems: Discretization and duality, *J. Geophys. Res.*, 97, 14,055–14,082.
  13. Donoho, D.L. and P.B. Stark, 1993. A note on rearrangements, spectral concentration, and the zero-order prolate spheroidal wavefunction. *IEEE-IT*, 39, 257–260.
  14. Pulliam, R.J. and P.B. Stark, 1993. Bumps on the core-mantle boundary: Are they facts or artifacts?, *J. Geophys. Res.*, 98, 1943–1956.
  15. Stark, P.B. and N.W. Hengartner, 1993. Reproducing Earth's kernel: Uncertainty of the shape of the core-mantle boundary from PKP and PcP travel-times, *J. Geophys. Res.*, 98, 1957–1972.
  16. Stark, P.B., 1993. Uncertainty of the COBE quadrupole detection, *Ap. J. Lett.*, 408, L73–L76.
  17. Stark, P.B. and D.I. Nikolayev, 1993. Toward tubular tomography, *J. Geophys. Res.*, 98, 8095–8106.
  18. Constable, C.G., R.L. Parker and P.B. Stark, 1993. Geomagnetic field models incorporating frozen-flux constraints, *Geophys. J. Intl.*, 113, 419–433.
  19. Gough, D.O. and P.B. Stark, 1993. Are the 1986–1988 changes in solar free-oscillation frequency splitting significant?, *Ap. J.*, 415, 376–382.
  20. Stark, P.B., M.M. Herron and A. Matteson, 1993. Empirically minimax affine mineralogy estimates from Fourier-transform infrared spectroscopy data using a decimated wavelet basis, *Applied Spectroscopy*, 47, 1820–1829.
  21. Pulliam, R.J. and P.B. Stark, 1994. Confidence regions for mantle heterogeneity, *J. Geophys. Res.*, 99, 6931–6943.
  22. Genovese, C.R., P.B. Stark and M.J. Thompson, 1995. Uncertainties for Two-Dimensional Models of Solar Rotation from Helioseismic Eigenfrequency Splitting, *Ap. J.*, 443, 843–854.
  23. Stark, P.B. and R.L. Parker, 1995. Bounded-variable least-squares: an algorithm and applications, *Comput. Stat.*, 10, 129–141.
  24. Hengartner, N.W. and P.B. Stark, 1995. Finite-sample confidence envelopes for shape-restricted densities, *Ann. Stat.*, 23, 525–550.
  25. Stark, P.B., 1995. Reply to Comment by Morelli and Dziewonski, *J. Geophys. Res.*, 100, 15,399–15,402.
  26. Gough, D.O., T. Sekii, and P.B. Stark, 1996. Inferring spatial variation of solar properties from helioseismic data, *Ap. J.*, 459, 779–791.
  27. Benjamini, Y. and Stark, P.B., 1996. Non-equivariant simultaneous confidence intervals less likely to contain zero, *J. Am. Stat. Assoc.*, 91, 329–337.
  28. Hill, F., P.B. Stark, R.T. Stebbins, E.R. Anderson, H.M. Antia, T.M. Brown, T.L. Duvall, Jr., D.A. Haber, J.W. Harvey, D.H. Hathaway, R. Howe, R. Hubbard, H.P.

- Jones, J.R. Kennedy, S.G. Korzennik, A.G. Kosovichev, J.W. Leibacher, K.G. Libbrecht, J.A. Pintar, E.J. Rhodes, Jr., J. Schou, M.J. Thompson, S. Tomczyk, C.G. Toner, R. Toussaint, and W.E. Williams, 1996. The solar acoustic spectrum and eigenmode parameters, *Science*, 272, 1292–1295.
29. Thompson, M.J., J. Toomre, E.R. Anderson, H.M. Antia, G. Berthomieu, D. Burtonclay, S.M. Chitre, J. Christensen-Dalsgaard, T. Corbard, M. DeRosa, C.R. Genovese, D.O. Gough, D.A. Haber, J.W. Harvey, F. Hill, R. Howe, S.G. Korzennik, A.G. Kosovichev, J.W. Leibacher, F.P. Pijpers, J. Provost, E.J. Rhodes, Jr., J. Schou, T. Sekii, P.B. Stark, and P.R. Wilson, 1996. Differential rotation and dynamics of the solar interior, *Science*, 272, 1300–1305.
  30. Stark, P.B., 1996. A few considerations for ascribing statistical significance to earthquake predictions, *Geophys. Res. Lett.*, 23, 1399–1402.
  31. Evans, S.N., and P.B. Stark, 1996. Shrinkage estimators, Skorokhod's problem, and stochastic integration by parts, *Ann. Stat.*, 24, 809–815.
  32. Genovese, C.R. and P.B. Stark, 1996. Data Reduction and Statistical Consistency in Linear Inverse Problems, *Phys. Earth Planet. Inter.*, 98, 143–162.
  33. Stark, P.B., 1997. Earthquake prediction: the null hypothesis, *Geophys. J. Intl.*, 131, 495–499.
  34. Benjamini, Y., Y. Hochberg, and P.B. Stark, 1998. Confidence Intervals with more Power to determine the Sign: Two Ends constrain the Means, *J. Amer. Stat. Assoc.*, 93, 309–317.
  35. Tenorio, L., P.B. Stark, and C.H. Lineweaver, 1999. Bigger uncertainties and the Big Bang, *Inverse Problems*, 15, 329–341.
  36. Stark, P.B., 1999. Geophysics, Statistics in, in *Encyclopedia of Statistical Sciences, Update Volume 3*, S. Kotz, C.B. Read, and D.L. Banks, eds., John Wiley and Sons, NY. Invited.
  37. Komm, R., Y. Gu, P.B. Stark, and I. Fodor, 1999. Multitaper Spectral Analysis and Wavelet Denoising Applied to Helioseismic Data, *Astrophysical J.*, 519, 407–421.
  38. Freedman, D.A., and P.B. Stark, 1999. The swine flu vaccine and Guillain-Barré syndrome: a case study in relative risk and specific causation, *Evaluation Review*, 23, 619–647.
  39. Fodor, I. and P.B. Stark, 2000. Multitaper Spectrum Estimation for Time Series with Gaps, *IEEE Trans. Signal Processing*, 48, 3472–3483.
  40. Freedman, D.A., P.B. Stark, and K.W. Wachter, 2001. A probability model for census adjustment, *Mathematical Population Studies*, 9, 165–180.
  41. Freedman, D.A. and P.B. Stark, 2001. The swine flu vaccine and Guillain-Barré syndrome, *Law and Contemporary Problems*, 64, 49–62.
  42. Evans, S.N. and P.B. Stark, 2002. Inverse Problems as Statistics, *Inverse Problems*, 18, R1–R43. Invited.

43. Schafer, C.M. and P.B. Stark, 2004. Using what we know: inference with physical constraints, *Proceedings of the Conference on Statistical Problems in Particle Physics, Astrophysics and Cosmology PHYSTAT2003*, L. Lyons, R. Mount and R. Reitmeyer, eds., Stanford Linear Accelerator Center, Menlo Park, CA, 25–34.
44. Evans, S.N., B. Hansen, and P.B. Stark, 2005. Minimax Expected Measure Confidence Sets for Restricted Location Parameters. *Bernoulli*, *11*, 571–590.
45. Divenyi, P., P.B. Stark, and K. Haupt, 2005. Decline of Speech Understanding and Auditory Thresholds in the Elderly, *J. Acoustical Soc. Am.*, *118*, 1089–1100.

### **Technical Reports and Unrefereed Publications**

46. Stark, P.B., 1988. Strict bounds and applications. in *Some Topics on Inverse Problems*, P.C. Sabatier, ed., World Scientific, Singapore.
47. Donoho, D.L. and Stark, P.B., 1988. Rearrangements and Smoothing, Tech. Rept. 148, Dept. Stat., Univ. Calif. Berkeley.
48. Donoho, D.L. and P.B. Stark, 1989. Recovery of a Sparse Signal When the Low Frequency Information is Missing, Tech. Rept. 179, Dept. Statistics, Univ. Calif. Berkeley.
49. Stark, P.B., 1990. Rigorous computer solutions to infinite-dimensional inverse problems. in *Inverse Methods in Action*, P.C. Sabatier, ed., Springer-Verlag. 462–467.
50. Hengartner, N.W. and P.B. Stark, 1992. Conservative finite-sample confidence envelopes for monotone and unimodal densities, Tech. Rept. 341, Dept. Statistics, Univ. Calif. Berkeley.
51. Hengartner, N.W. and P.B. Stark, 1992. Confidence bounds on the probability density of aftershocks, Tech. Rept. 352, Dept. Statistics, Univ. Calif. Berkeley.
52. Stark, P.B., 1992. The Cosmic Microwave Background and Earth's Core-Mantle Boundary: A Tale of Two CMB's, Tech. Rept. 371, Dept. Statistics, Univ. Calif. Berkeley.
53. Genovese, C. and P.B. Stark, 1993.  $l_1$  spectral estimation: Algorithms and tests of super-resolution, in *GONG 1992: Seismic Investigations of the Sun and Stars, Proc. Astr. Soc. Pac. Conf. Ser.*, *42*, T. Brown, ed., 453–456.
54. Gough, D.O. and P.B. Stark, 1993. The significance of changes in solar free-oscillation splitting from 1986–1990, in *GONG 1992: Seismic Investigations of the Sun and Stars, Proc. Astr. Soc. Pac. Conf. Ser.*, *42*, T. Brown, ed., 221–224.
55. Stark, P.B., 1994. Simultaneous Confidence Intervals for Linear Estimates of Linear Functionals, Tech. Rept. 417, Dept. Statistics, Univ. Calif. Berkeley.
56. Sekii, T., C.R. Genovese, D.O. Gough, and P.B. Stark, 1995. Observational constraints on the internal solar angular velocity, in *Fourth SOHO Workshop: Helioseismology*, J.T. Hoeksema, V. Domingo, B. Fleck and B. Battrick, eds., ESA Publications Division SP-376, Noordwijk, *2*, 279–283.
57. Stark, P.B., 1997. Data Sampling Rate Reduction for the OERSTED Geomagnetic

Satellite.

58. Fodor, I.K., J. G. Berryman, and P. B. Stark, 1997. Comparison of Autoregressive and Multitaper Spectral Analysis for Long Time Series, *Stanford Exploration Project*, **95**, 331–355.
59. Komm, R.W., Y. Gu, F. Hill, P.B. Stark, and I.K. Fodor, 1998. Multitaper Spectral Analysis and Wavelet Denoising Applied to Helioseismic Data, *Proc. Tenth Cambridge Workshop on Cool Stars, Stellar Systems and the Sun*, ASP Conference Series, **154**, CDR 783–790.
60. Komm, R.W., E. Anderson, F. Hill, R. Howe, A.G. Kosovichev, P.H. Scherrer, J. Schou, I. Fodor, and P. Stark, 1998. Comparison of SOHO-SOI/MDI and GONG Spectra, *Proceedings of the SOHO 6/GONG 98 Workshop*, 'Structure and Dynamics of the Interior of the Sun and Sun-like Stars,' Boston, USA, 1–4 June 1998, ESA SP-418, pp. 253–256.
61. Komm, R.W., E. Anderson, F. Hill, R. Howe, I. Fodor, and P. Stark, 1998. Multitaper analysis applied to a 3-month time series, *Proceedings of the SOHO 6/GONG 98 Workshop*, 'Structure and Dynamics of the Interior of the Sun and Sun-like Stars,' Boston, USA, 1–4 June 1998, ESA SP-418, pp. 257–260.
62. Stark, P.B., 1999. Letter to the Editor of USA Today regarding Sampling to Adjust the 2000 Census, 19 January.
63. Stark, P.B., 1999. The 1990 and 2000 Census Adjustment Plans, Tech. Rept. 550, Dept. Statistics, Univ. Calif. Berkeley (revised May 2000)
64. Fodor, I.K. and P.B. Stark, 1999. Multitaper Spectrum Estimates for Time Series with Missing Values, *Computing Science and Statistics, 31: Models, Predictions, and Computing*. K. Berk and M. Pourahmadi, eds., pp. 383–387.
65. Stark, P.B., 2000. Inverse Problems as Statistics, in *Surveys on Solution Methods for Inverse Problems*, Colton, D., H.W. Engl, A.K. Louis, J.R. Mclaughlin and W. Rundell, eds., Springer-Verlag, New York, pp. 253–275. Invited.
66. Stark, P.B., 2001. Review of *Who Counts?* by Margo J. Anderson and Stephen E. Fienberg, *Journal of Economic Literature*, **XXXIX**, pp. 593–595. Invited.
67. Stark, P.B., 2002. Capture-recapture. *Encyclopedia of Social Science Research Methods*, Sage Publications, Thousand Oaks, CA. Invited.
68. Stark, P.B., 2002. Census Adjustment. *Encyclopedia of Social Science Research Methods*, Sage Publications, Thousand Oaks, CA. Invited.
69. Freedman, D.A. and P.B. Stark, 2003. What is the Chance of an Earthquake? in *Earthquake Science and Seismic Risk Reduction*, F. Mulargia and R.J. Geller, eds., NATO Science Series IV: Earth and Environmental Sciences, v. 32, Kluwer, Dordrecht, The Netherlands, pp. 201–213. Invited.
70. Stark, P.B., 2004. Estimating power spectra of galactic structure: can Statistics help?, in *Penetrating Bars Through Masks of Cosmic Dust: The Hubble Tuning Fork Strikes a New Note*, D.L. Block, I. Puerari, K.C. Freeman, R. Gross and

- E.K. Block, eds., Springer, The Netherlands, pp. 613–617. Invited.
71. Freedman, D.A. and P.B. Stark, 2005. Ecological correlation and the ecological fallacy, *Encyclopedia of Law and Society*, to appear. Invited.
  72. Luen, B. and P.B. Stark, 2006. Testing earthquake predictions. Festschrift for David A. Freedman, Institute for Mathematical Statistics, to appear. Invited.

### Online Documents (see [www.stat.berkeley.edu/~stark](http://www.stat.berkeley.edu/~stark))

1. **SticiGui**<sup>®</sup>: Statistics Tools for Internet and Classroom Instruction with a Graphical User Interface. [www.stat.berkeley.edu/~stark/SticiGui](http://www.stat.berkeley.edu/~stark/SticiGui)
2. Testimony before U.S. House of Representatives Subcommittee on the Census, 5 May 1998.
3. Response to 25 Questions from Representative C. Maloney, Ranking Minority Member, U.S. House of Representatives Subcommittee on the Census, 13 May 1998.

### Software

- Stark, P.B., and R.L. Parker, 1994. BVLS (Bounded-Variable Least Squares), STATLIB (Carnegie-Mellon University ftp server)
- unofficially published software: see [www.stat.berkeley.edu/~stark](http://www.stat.berkeley.edu/~stark)

### Invited Presentations

(The text of some recent seminars is available online; see [www.stat.berkeley.edu/~stark](http://www.stat.berkeley.edu/~stark))

2006

- **Measuring Resolution in Nonlinear and Constrained Inverse Problems**, *Workshop on Statistical Inverse Problems*, Institute for Mathematical Stochastics, Göttingen, Germany, 23–25 March.

2005

- **Resolution in Nonlinear and Constrained Inverse Problems**, *Workshop on Computational and Mathematical Geoscience*, Colorado School of Mines, Golden CO, 15–17 June.

2004

- **Quantifying uncertainty in inverse problems**, *Summer school: Mathematical Geophysics and Uncertainty in Earth Models*, Colorado School of Mines, Golden CO, 14–25 June.
- **Estimating power spectra of galaxy structure: can Statistics help?**, *Penetrating bars through masks of cosmic dust: the Hubble tuning fork strikes a new note*, Pilanesberg National Park, South Africa, 7-12 June.

2003

- **Quantifying uncertainty in inverse problems**, *Institute for Pure and Applied Mathematics (IPAM) Conference on Statistical Methods for Inverse Problems*, 5–6 November, IPAM, Los Angeles, CA
- **Guest**, *The Fred Ebert Show* program on probability and statistics, 27 October, KIRO 710, Seattle, WA
- **Using what we know: inference with physical constraints**, *PhyStat 2003: Statistical Problems in Particle Physics, Astrophysics and Cosmology*, 8-10 September, Stanford Linear Accelerator Center, Stanford, CA

2002

- **Strategic Planning and Implementation I: The Challenge of Adapting Organizations and Creating Partnerships to Target New Markets**, *University Teaching as E-business?*, Center for Studies in Higher Education, 26-27 October, Berkeley, CA.
- **Inverse Problems and Data Errors**, *New Developments in Astrophysical Fluid Dynamics*, Chateau de Mons, 25–29 June, Caussens, France.
- **Data Reduction and Inverse Problems in Helioseismology**, *Workshop Statistics of inverse problems*, Institut Henri Poincaré, 28–29 May, Paris, France.
- **Why Statistics is worth the Stigma**, *Letters and Sciences Faculty Forum*, 23 April, University of California, Berkeley
- **Inverse Problems in Helioseismology**, Second MaPhySto Workshop on Inverse Problems: Inverse problems from a Statistical Perspective, 28–31 March, Aalborg, Denmark

2000

- **What are the Chances?** *NATO Advanced Research Workshop: State of scientific knowledge regarding earthquake occurrence and implications for public policy*, Le Dune, Piscinas - Arbus, Sardinia, Italy, 15–19 October.
- **Why Unadjusted Census Results should be Used for Reapportionment and Funding within the State of California**. *13th Annual Demographic Workshop*, U.S. Bureau of the Census, California State Census Data Center, and the Population Research Laboratory of the University of Southern California, Los Angeles, CA, 15 May.
- Invited Discussant, *Workshop of the National Academy of Sciences Panel to Review the 2000 Census*, Washington, D.C., 2–3 February.

1999

- Invited Discussant, *Panel Discussion on the role of sampling in the US Census*, San Francisco Bay Area Chapter of the American Statistical Association, 20 December.
- Lecturer, *Mathematical Geophysics Summer School*, Stanford University,

Stanford, CA, 2–20 August.

- **Less Asymptotic Tomography.** *9th SOHO Workshop: Helioseismic Diagnostics of Solar Convection and Activity*, Stanford University, Stanford, CA, 12–15 July.
- Panelist, *Reinventing Undergraduate Education: Technology Enhanced Learning in the Sciences, Math, and Engineering*, University of California, Berkeley, CA, 23 April.
- **Error in Numerical Models Fitted to Data.** *DSRC/DARPA Study on Numerical Simulation of Physical Systems: The State of the Art, and Opportunities for Further Advances, Kick-Off Meeting*, Arlington, VA, 19–20 January.
- **Sampling to Adjust the U.S. Census.** *Miller Institute for Basic Research in Science*, University of California, Berkeley, CA, 12 January.

1998

- **A Statistician's Perspective on Census Adjustment,** *Berkeley Breakfast Club*, Berkeley, CA, 5 December.
- **SticiGui<sup>®</sup>: Melts in your Browser, not in your Brain,** *Joint Berkeley-Stanford Statistics Colloquium*, Department of Statistics, Stanford University, Stanford, CA, 27 October.
- **SticiGui: Statistics Tools for Internet and Classroom Instruction with a Graphical User Interface,** *1998 Joint Statistical Meetings of the American Statistical Association, International Biometric Society, and Institute of Mathematical Statistics*, Orlando, FL, 12 August.
- **Presidential Panel on Statistics in Public Policy,** *1998 Joint Statistical Meetings of the American Statistical Association, International Biometric Society, and Institute of Mathematical Statistics*, Orlando, FL, 10 August.
- Guest, *KQED-FM Forum program on the 2000 Census*, San Francisco, CA, 17 July.
- **Misfit Measures and Statistical Inconsistency in Linear Inverse Problems.** *AMS/IMS/SIAM Joint Summer Research Conferences in the Mathematical Sciences, Mathematical Methods in Inverse Problems for Partial Differential Equations*, Mt. Holyoke, MA, 4–9 July.
- **Uncertainties for functions from incomplete, erroneous data.** *NSF/DOE Workshop on Uncertainty in Modeling*, National Science Foundation, Arlington, VA, 11–12 June.
- **Sampling to adjust the 1990 Census for Undercount.** U.S. House of Representatives Subcommittee on the Census, May.
- **Sounding the Sun: Helioseismology.** *1998 American Association for the Advancement of Science (AAAS) Annual Meeting and Science Innovation Exposition*, Philadelphia, PA., February.



1997

- **Does God play dice with the Earth, and if so, are they loaded?** *Fourth SIAM Conference on Mathematical and Computational Methods in the Geosciences*, Albuquerque, NM.
- **Solving Problems for a Large Statistics Lecture Course using a Website** *UC Berkeley Academic Senate Workshop on Classroom Technology*, Berkeley, CA.
- **Deficiencies of the simple theories**, *Local Helioseismology Workshop*, University of Cambridge, Cambridge, England.

1996

- **CMB's**, *Royal Astronomical Society Ordinary Meeting*, London, England.
- **The Null Hypothesis**, *Royal Astronomical Society and Joint Associations for Geophysics discussion meeting on Assessment of Schemes for Earthquake Prediction*, London, England.
- **On the consistency of multiple inference in inverse problems using  $I_p$  confidence sets**, *International Conference on Multiple Comparisons*, Tel Aviv, Israel.

1995

- **Confidence Intervals in Inverse Problems**, *Conference in Honor of George Backus*, Institute for Geophysics and Planetary Physics, La Jolla, CA
- **The Need for Wave-Equation Travel-Time Tomography**, *Institute for Mathematics and Its Applications, Conference on Tomography*, Minneapolis, MN
- **Inference, Prior Information, and Misfit Measures**, *Interdisciplinary Inversion Conference on Methodology, Computation and Integrated Applications*, University of Aarhus, Aarhus, Denmark
- **Optimization and Inference in Travel-Time Seismology**, *National Research Council Board on Mathematical Sciences Symposium on Mathematical Sciences in Seismology*, Washington, DC
- **Prior Information and Confidence Intervals in Inverse Problems**, *International Union of Geodesy and Geophysics Meeting*, Boulder, CO
- **Something AGAINST Nothing: A Confidence Game**, *Joint Statistical Meetings of the American Statistical Association, International Biometric Society, and Institute of Mathematical Statistics*, Orlando, FL
- **Uncertainties in Travel-Time Seismology**, *SIAM/GAMM Symposium on Inverse Problems: Geophysical Applications*, Fish Camp, CA

1994

- **Toward Tubular Tomography**, *27th General Assembly of the Int. Assoc. of Seismology and Phys. of the Earth's Inter. (IASPEI)*, Wellington, New Zealand
- **Alternative Data Analysis Techniques**, *Global Oscillation Network Group annual meeting*, Los Angeles, CA (presented by C. Genovese due to illness).

- **Mathematical Aspects of Integral Equation Inversion**, *Global Oscillation Network Group workshop*, Sydney, Australia.

1993

- **Conservative Finite-Sample Confidence Envelopes for Monotone and Unimodal Densities**, *Mathematisches Forschungsinstitut Oberwolfach meeting on Curves, Images and Massive Computation*, Oberwolfach, Germany
- Invited Discussant, *Joint IMS/ASA/ENAR Meeting*, Philadelphia, PA
- **Uncertainty of the Quadrupole Component of the Cosmic Microwave Background**, *Israel Statistical Association Annual Meeting*, Tel Aviv
- **Brute-Force Minimax Estimation in Geochemistry**, *Joint Statistical Meetings of the American Statistical Association, International Biometric Society, and Institute of Mathematical Statistics*, San Francisco, CA

1992

- **Conservative Numerical Uncertainty Estimates in Inverse Problems**, *SIAM 40th Anniversary Meeting*, Los Angeles, CA

1991

- **Minimax Estimation in Geomagnetism**, *European Geophysical Society Annual Meeting*, Wiesbaden, Germany
- **Minimax Estimation in Geophysical Inverse Problems: Applications to Seismic Tomography and Geomagnetism**, Schmitt Institute for Physics of the Earth, Academy of Sciences of the USSR, Moscow
- **Imagining Earth's Interior: Controversies in Seismology and Geomagnetism**, *Mathematical Sciences Research Institute Workshop on Statistical Methods in Imaging*, Berkeley, CA

1990

- **Discretization and its Discontents: New Methods in Inverse Theory**, *Institute for Theoretical Physics program "Helioseismology—Probing the Interior of a Star,"* National Science Foundation Institute for Theoretical Physics, University of California, Santa Barbara
- **Inference in Infinite-Dimensional Inverse Problems**, Schmitt Institute for Physics of the Earth, Academy of Sciences of the USSR, Moscow
- **Inference in Infinite-Dimensions: Discretization and Duality**, *Israel Statistical Association Annual Meeting*, Jerusalem
- **Superresolution: What, When and How?**, *Institute for Theoretical Physics program "Helioseismology—Probing the Interior of a Star,"* National Science Foundation Institute for Theoretical Physics, University of California, Santa Barbara

1989

- **Sparsity-Constrained Deconvolution**, *International Union of Radio Science Meeting*, Boulder, CO
- Invited Discussant. *Statistics, Earth and Space Sciences Meeting of the Bernoulli Society*, Leuven, Belgium
- **Rigorous Computer Solutions to Infinite-Dimensional Inverse Problems**, *rcp 264 problemes inverses*, Montpellier, France

1988

- **Duality and Discretization Error**, *Conference on Mathematical Geophysics*, Blanes, Spain

1987

- **Spectral extrapolation with positivity**, *International Union of Radio Science Meeting*, Boulder, CO

1986

- **Travel-Time Constraints on Core Structure**, *Special Session on Geophysics of the Core and Core-Mantle Boundary*, American Geophysical Union Spring Meeting, Baltimore, MD
- **Smooth Models from  $\tau(p)$  and  $X(p)$  Data**, *Scripps Industrial Associates Short Course on Inverse Theory*, Scripps Institution of Oceanography, La Jolla, CA

### Other invited seminars

- California State University, Chico (Mathematics 1993)
- Colorado School of Mines (Dept. of Mathematical and Computer Sciences, 1997)
- Copenhagen University (Niels Bohr Institute for Astronomy, Physics, and Geophysics 1996)
- Hebrew University of Jerusalem (Statistics 1993)
- National Solar Observatory (1997)
- Naval Postgraduate School (Operations Research, 2001)
- Schlumberger-Doll Research (1988, 1990, 1991, 1992)
- Southern Methodist University (Statistical Sciences, 1998)
- Stanford University (Center for Space Physics and Astrophysics 1992; Mathematics, 1997; Geology and Geophysics, 1993, 1997; Statistics 1988, 1993, 1995)
- The Technion (Statistics 1987)
- Tel-Aviv University (Geology and Geophysics 1988, 1991; Statistics 1991)
- University of British Columbia (Geophysics and Astronomy 1996)
- University of California, Berkeley (Astronomy 1996; Center for Pure and Applied Mathematics 1988; Geology and Geophysics 1988; Materials Science and Mineral

- Engineering 1988; Physics, 2001; Seismographic Stations, 1991, 1992, 1996; Statistics 1987, 1988(2),1989(2), 1990, 1991, 1992, 1994, 1996(2), 1997)
- University of California, Davis (Statistics 1995; Mathematics 2000)
  - University of California, Los Angeles (Mathematics 1992; Statistics 2000)
  - University of California, Riverside (Earth Sciences 1996; Statistics 1996)
  - University of California, San Diego (Institute for Geophysics and Planetary Physics 1985, 1986, 1987, 1988(2), 1990, 1998, 2005; Mathematics 1994)
  - University of Cambridge (Institute for Astronomy 1992, 1997)
  - University of Chicago (Statistics 1990)
  - University of Edinburgh (Earth Sciences, 1998)
  - University of Texas at Austin (Geological Sciences 1988; Mathematics 1990, 1991; Institute for Geophysics 1990)
  - Yale University (Geology and Geophysics 1988; Statistics 1988)

## **Service**

### **Professional Societies and Government Agencies**

2006

Consultant, U.S. Department of Justice, Civil Division

2005

Consultant, U.S. Department of Justice, Civil Division  
 Consultant, Department of Veterans Affairs Medical Center  
 Consultant, Habeas Corpus Resource Center

2004

Reviewer, National Science Foundation  
 Consultant, U.S. Department of Justice, Civil Division  
 Consultant, U.S. Attorney's Office  
 Consultant, Department of Veterans Affairs Medical Center

2003

Reviewer, National Science Foundation  
 Referee, National Sciences and Engineering Research Council of Canada  
 Consultant, Department of Veterans Affairs Medical Center

2002

Consultant, U.S. Department of Agriculture  
 Consultant, U.S. Department of Justice, Civil Division

2001

Consultant, U.S. Department of Justice, Civil Division

- Co-organizer, Institute for Mathematics and Its Applications 2001–2002 Program  
*Mathematics in the Geosciences* and workshop on Inverse Problems and the  
Quantification of Uncertainty
- 2000
- Discussant, National Academy of Science Committee on National Statistics  
workshop on dual-system estimation for the 2000 Census  
Consultant, U.S. Department of Justice, Civil Division
- 1998
- Witness, U.S. House of Representatives Subcommittee on the Census.  
Panelist, National Science Foundation
- 1997
- Session organizer, International Statistical Institute and Bernoulli Society Meeting,  
Istanbul, Turkey
- 1996–
- Global Oscillation Network Group (GONG) Data Users Committee (Chair, 1996–  
1998)  
Reviewer for United States Geological Survey
- 1996–1999
- Consultant, National Security Agency
- 1995
- Institute of Mathematical Statistics Program Chair, Joint Statistical Meetings of the  
American Statistical Association, International Biometric Society, and Institute of  
Mathematical Statistics, Orlando, FL
- 1994–1996
- Consultant to Federal Trade Commission
- 1993
- Session organizer and chair, IMS/ASA/ENAR meeting, Philadelphia, PA  
Session organizer and chair, Joint Statistical Meetings of the American Statistical  
Association, International Biometric Society, and Institute of Mathematical  
Statistics, San Francisco, CA
- 1992
- Faculty sponsor, Department of Energy TRAC program
- 1990–1994
- Bernoulli Society Committee on Statistics in the Physical Sciences
- 1991–present
- Reviewer for National Aeronautics and Space Administration (Space Physics  
Division)

1991

Local organizer and session chair, Mathematical Sciences Research Institute  
Workshop on Statistical Methods in Imaging, Berkeley, CA

1989

Session organizer and chair, Bernoulli Society Satellite Meeting, Leuven, Belgium

1989–present

Reviewer for National Science Foundation (Atmospheric Sciences, Infrastructure,  
International Programs, Mathematical Sciences, Solar-Terrestrial Program,  
Statistics and Probability)

### **Private Industry**

2000–2001

Technical Advisory Board, Cogit.com  
National Advisory Board, eTextbooksOnline.com

2000–2002

Technical Advisory Board, Atomic Dog Publishing

### **Editorial Service**

1998–1999

Editor, Statistical Science

1997–2000

Editorial Board, Inverse Problems

1994–1998

Associate Editor, Journal of Geophysical Research

### **Referee Service**

- American Association for the Advancement of Science
- Annales Geophysicae
- Annals of the Institute of Statistical Mathematics
- Annals of Statistics
- Arabian Journal for Science and Engineering
- Bulletin of the Seismological Society of America
- Cambridge University Press
- Geophysical Journal International
- Geophysical Research Letters
- Geophysics
- Geophysical & Astrophysical Fluid Dynamics
- IEEE Journal on Acoustics, Speech and Signal Processing
- IEEE Journal on Information Theory
- Inverse Problems

- Journal of the American Statistical Association
- Journal of Computational Physics
- Journal of Economic Literature
- Journal of Geophysical Research
- Jurimetrics
- Physics of the Earth and Planetary Interiors
- Science
- SIAM Review
- Tectonophysics
- Chapman-Hall
- HarperCollins
- Simon and Schuster
- Springer-Verlag

## **University Service**

2004-2005

Chair, Educational Technology Committee  
 e-Berkeley Steering Committee  
 e-Berkeley Committee of Chairs  
 e-Berkeley Implementation Task Force  
 CourseWeb Steering Committee  
 Faculty Athletic Fellow

2003–2004

Chair, Educational Technology Committee  
 e-Berkeley Steering Committee  
 e-Berkeley Implementation Task Force  
 Student Systems Policy Committee  
 CourseWeb Steering Committee

2002–2003

Faculty Assistant in Educational Technology (to Vice Provost for Undergraduate Education)  
 Chair, Educational Technology Committee  
 Provost's Academic Council  
 e-Berkeley Steering Committee  
 e-Berkeley Implementation Task Force  
 Campus Committee on Classroom Policy and Management (CCCPM)  
 Student Systems Policy Committee  
 e-Berkeley Symposium Program Committee  
 Faculty Search Committee, Graduate School of Education  
 CourseWeb Steering Committee

2001–2002

Faculty Assistant in Educational Technology (to Vice Provost for Undergraduate Education)  
Chair, Educational Technology Committee  
Provost's Academic Council  
e-Berkeley Steering Committee  
e-Berkeley Implementation Task Force  
Campus Committee on Classroom Policy and Management (CCCPM)  
Academic Senate Committee on Academic Planning and Resource Allocation (CAPRA)  
CITRIS II Program Committee  
TeleBEARS and Bear Facts Committees (Student Systems Policy Committee as of 3/2002)  
e-Berkeley Portal Working Group

2000–2001

Academic Senate Committee on Academic Planning and Resource Allocation (CAPRA)  
Space Allocation and Capital Improvements Committee (SACI)  
CAPRA Subcommittee on Expanded Enrollment  
CAPRA Subcommittee on changes to Academic Coordinator title  
*Ad hoc* hiring/tenure committee

1999–2000

Academic Senate Library Committee (LIBR)  
Academic Senate Committee on Academic Planning and Resource Allocation (CAPRA), Physical Planning Subcommittee, *ex officio* representative from Library Committee  
Space Allocation and Capital Improvements Committee (SACI)  
Academic Effects Study Committee, Molecular Engineering Building  
*Ad hoc* tenure/promotion committee  
SACI subcommittee to audit space in Barrows Hall

1998–1999

Space Allocation and Capital Improvements Committee (SACI)  
Electronic Dissertations Project  
Planning Space for the Physical Sciences Libraries

1997–1998

*Ad hoc* tenure/promotion committee

1996

Review of College of Science, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia



1994–1999

University review committee for Department of Agricultural and Resource Economics, University of California, Berkeley

1993–1995

Physical Sciences Division committee for Graduate Affirmative Action and Retention

Physical Sciences Division committee for Science and Mathematics Academic Re-Training (SMART)

## Grants

1. PI, NASA Grant NAG 5-883, “Constructing Core Fields Consistent with Geomagnetic Data and Geophysical Constraints,” 1987–1990.
2. Project Director and PI, NSF Grant DMS-8810192, “Inference in Curved-Ray Tomography: Solid Earth Structure,” 1989–1992.
3. PI, NSF Grant INT-9205103, “Long and Medium-Term Research: Inference in Seismological Investigations of Subducting Lithosphere,” 1992–1994.
4. PI, NSF Grant DMS-930006P, “Estimating the Sun's Internal Angular Velocity from Free-Oscillation Frequency Splittings,” 1993–1994.
5. PI, NSF Presidential Young Investigator Award DMS-8957573, 1989–1995.
6. Co-I, NASA Grant NAG5-2438, “The Analysis of Cobe DMR Sky Maps,” 1993–1994. PI: J. Silk
7. PI, NASA Grant NAGW-2515, “New Methods for Inversion and Analysis of Solar Free-Oscillation Data,” 1991–1995.
8. PI, NSF Grant DMS-9404276, “New Methods for Inference From COBE Data,” 1994–1997.
9. PI, NSF Grant AST-9504410, “Function Estimation and Inference in Helioseismology,” 1995–1998.
10. PI, LLNL/IGPP Grant 97-AP028, “Helioseismology with Solar Luminosity Constraints,” 1996–1997.
11. Co-I, NASA Grant NAG5-3941, “Development of data analysis, compression and visualization tools for large data sets in astrophysics and cosmology,” 1997–1998. PI: J. Silk
12. PI, NASA Grant NRA-96-09-OSS-034SOHO, “Modern Statistical Methods for Helioseismic Spectrum Estimation,” 1997–1998.
13. PI, NASA Grant NAG 5-3919, “Data Sampling Rate Reduction for the Oersted Satellite,” 1997–1998.
14. PI, UC Berkeley Classroom Technologies Grant, “Statistics *Statim*,” 1997–1998.
15. Co-I, NSF Grant DMS-9872979, “KDI: Computational Challenges in Cosmology,” 1998–2000. PI: A. Jaffe.

16. Co-I, NSF Grant IIS-98-17353, “Re-Inventing Scholarly Information Dissemination and Use,” 4/1/1999 – 3/31/2004. PI: R. Wilensky and D. Forsythe.
17. PI, Hewlett Packard Company Grant 89293, “Applied Mobile Technology Solutions in Learning Environments,” 3/19/2003–8/31/2004.
18. PI, Hewlett Packard Company Grant 14928, “Applied Mobile Technology Solutions in Learning Environments—2004 Extension,” 4/1/2004–6/30/2005.

## **Consulting**

- Bramson, Plutzik, Mahler & Birkhaeuser LLP, Walnut Creek, CA: consumer class action litigation
- Brinks, Hofer, Gilson & Lione, Chicago, IL: intellectual property litigation (client: R.J. Reynolds Tobacco Co.)
- Cisco Systems: predicting email spool fill
- City of Santa Rosa, CA: water treatment monitoring
- Cogit.com, San Francisco, CA: technical advisory board; targeted web advertising
- Contra Costa County Public Defender, Richmond, CA: equal protection
- Crosby, Heafey, Roach, & May, Oakland, CA: insurance litigation (client: Farmers Insurance)
- Dept. of Veterans Affairs Medical Center, Martinez, CA: speech and non-speech hearing segregation in aging
- East Bay Municipal Utilities District: water treatment monitoring
- EEG Systems Laboratory, San Francisco, CA: inverse problems for electrical activity of the brain
- eTextbooksOnline.com, New York, NY: National Advisory board
- Federal Trade Commission, San Francisco, CA: sampling in litigation
- Fuller-Austin Joint Defense Group: Modeling in litigation
- Habeas Corpus Resource Center, San Francisco, CA: bias in jury selection
- Howard, Rice, Nemerovski, Canady, Falk, & Rabkin, San Francisco, CA: sampling in litigation; inference from retail sales data (clients K-Mart Corp., R.J. Reynolds Tobacco Co.)
- Kaiser Permanente Northern California, Redwood City, CA: clinical trials in oncology
- KLA Instruments Corporation, San Jose, CA: calibration of algorithms to detect IC mask flaws
- Kramer, Levin, Naftalis, & Frankel, New York, NY: sampling in litigation
- Law Offices of Gorman & Miller, San Jose, CA: trade secret litigation
- Law Offices of Ilson W. New, San Francisco, CA: natural resource legislation

- Law Offices of Ramirez, Tollner, Stebbins, Bahrick, & Sasseen, San Jose, CA: trade secret litigation
- Law Offices of Welebir & McCune, Woodside, CA: product liability litigation
- Law offices of Wells, Pinckney & McHugh, Austin, TX: employment discrimination arbitration
- Law Offices of Wolkin & Timpane, San Francisco, CA: insurance litigation
- Law Offices of Scott K. Zimmerman, Brentwood, CA: product liability litigation
- Life Chiropractic College West, Hayward, CA: experimental design
- Los Angeles Superior Court, Central District: sampling in employment litigation
- Mayer, Brown, Rowe & Maw, LLP, Chicago, IL: intellectual property litigation
- Morrison & Foerster, San Francisco, CA: product liability class action litigation
- National Security Agency: adaptive filtering, combining expert opinions, digital communications, information retrieval, estimation
- National Solar Observatory, Tucson, AZ: spectrum estimation
- Pacific Gas & Electric Co., San Francisco, CA: statistics and causal inference in litigation
- Paul Hastings, Washington, DC: intellectual property litigation
- Schlumberger-Doll Research, Ridgefield, CT: inverse problems, signal processing
- Shearman & Sterling, Washington, DC: survival analysis
- Skadden, Arps, Slate, Meagher & Flom, LLP, San Francisco, CA: case-control studies in litigation
- Spriggs & Hollingsworth, Washington, DC: environmental litigation
- St. Paul Fire and Marine Insurance Company, Baltimore, MD: projecting tort liability
- U.S. Attorney's Office, Northern District of California: ethnic bias in grand jury selection
- U.S. Department of Agriculture, Washington, D.C.: fairness in lending
- U.S. Department of Justice, Civil Division, Federal Programs Branch, Washington, D.C.: sampling the internet and web-browsing behavior; USDA import restrictions on cattle and beef
- U.S. House of Representatives, Washington, D.C.: sampling to adjust the U.S. Census
- Willoughby, Stuart & Bening, San Jose, CA: insurance litigation.
- Zimmerman Reed, Scottsdale, AZ: consumer class action litigation.

Last modified 7 May 2006. A current version of this document in Adobe Acrobat (.pdf) format is available at [www.stat.berkeley.edu/~stark/bio.pdf](http://www.stat.berkeley.edu/~stark/bio.pdf); for an html version, see [www.stat.berkeley.edu/~stark/bio.htm](http://www.stat.berkeley.edu/~stark/bio.htm)