"True Believers" or Numerical Terrorism at the Nuclear Power Plant

Walter Krämer
Gerhard Arminger

CESIFO WORKING PAPER NO. 3180
CATEGORY 12: EMPIRICAL AND THEORETICAL METHODS
SEPTEMBER 2010

CESifo *Working Paper No. 3180*

"True Believers" or Numerical Terrorism at the Nuclear Power Plant

Abstract

For decades, there has been a heated debate about whether or not nuclear power plants contribute to childhood cancer in their respective neighbourhoods, with statisticians testifying on both sides. The present paper points to some flaws in the pro-arguments, taking a recent study prepared for the political party "Bündnis 90 / Grüne" as a specimen. Typical mistakes include an understatement of the size of tests of significance, disregard of important covariates and extreme reliance on very few selected data points.

Walter Krämer
Faculty of Statistics
Technical University Dortmund
44221 Dortmund
Germany
walterk@statistik.uni-dortmund.de

Gerhard Arminger
Schumpeter School of Business and Economics
Bergische Universität Wuppertal
42119 Wuppertal
Germany
arminger@statistik.uni-wuppertal.de

Paper prepared for the special issue "Methodological artefacts, data manipulation and fraud in Economics and Social science" of Jahrbücher für Nationalökonomie und Statistik, version Sept. 6. 2010.
The title is taken from Dewdney (1996) and refers to misleading media coverage of nuclear power generating plants in the U.S.
We are grateful to Marc Drolet and Stefano Grasso for helpful criticism and comments.

1. Introduction and summary

In the fall of 2009, the German political party Bündnis 90/Grüne (2009) produced a temporary stir in the German media by claiming final proof that nuclear power plants induce childhood leukemia. "AKW erhöhen das Leukämierisiko (nuclear power plants increase risk of leukemia)" was the heading of a press release. While not even the meta-analysis by Greiser (2009), which formed the basis of this press release, has any such claim in it (since Greiser is well aware of the difference between correlation and causation), the press release strongly contributed to the fiercely held belief by many Germans that nuclear power is bad for you.

The present paper shows that presumably not even the correlation claimed by Greiser (2009) does exist. We use his study to exemplify various mistakes that are often made when statistical analyses are guided by strong a priori beliefs which are so typical in the leukemia vs. nuclear power debate. The first and most prominent source of error is an understatement of the true size of tests of significance which results from the well known publication bias. We provide a brief survey of this literature and show that there is ample reason to believe that this bias also prevails in the leukemia debate. In technical terms, the true significance level of such tests is much larger than the nominal one reported in the respective papers.

Other mistakes include the disregard of important covariables and the heavy reliance on outliers which, other removing them, reverse the patterns observed before. Then there is the well known phenomenon called HARKing (“Hypothesizing After the Results are Known”), where tests of significance are taking place only after some abnormal data has been observed. This seems to apply in particular to the leukemia debate, where many studies were undertaken only after the media had aroused attention to abnormal incidence or mortality close to nuclear installations of various types. Taken together, these deficiencies seem to invalidate any “proof” that nuclear power correlates with childhood leukemia, let alone that is responsible for it. While it might still be true that some such relationship exists, it certainly cannot be derived from the evidence that is available so far. Even if one does not subscribe to the well known Taubes (1995) – thesis that epidemiological evidence of any sort should only be taken seriously if there is at least a twofold increase in the risk observed, one needs much more and in

particular much more convincing data before sounding the kind of alarm that is so popular among true believers in science and in the media alike. [3]

2. Empirical studies of cancer incidence around nuclear power plants

There is an enormous literature in statistics, epidemiology and public health on childhood cancer, in particular childhood leukemia, in the vicinity of nuclear installations of all sorts. It dates back to a 1982 British television documentary entitled "Windscale: the Nuclear Laundry", which reported an abnormal incidence of leukemia in young people living in the village of Seascale close to the nuclear site of Sellafield, and has subsequently spawned an enormous interest in similar clusters elsewhere. Among studies which did find such clusters, or at least "abnormal" rates of incidence or mortality, are Heasman et al. (1987), Ewings et al. (1989), Clarke et al. (1991), Körblein and Hoffmann (1999) or Hoffmann et al. (1996, 2007), just to name a few. Alexander (1999) and Laurier and Bard (1998) provide convenient summaries of the earlier literature, and Baker and Hood (2007) and later Greiser (2009) collect many of these studies for meta-analyses which led to similar results.

Also – partially – included in these meta analyses were studies which could not find any excess incidence or excess mortality. Because they are so rarely cited, we here present their main conclusions:

> "No excess cases were found in small towns around the plant" (Sofer et al 1991, p. 191).
>
> "Our study gives no evidence for an increased risk of childhood leukaemia … in the vicinity of nuclear installations" (Michaelis et al. 1992, p. 262).
>
> "No increase of Leukaemia and lymphoma mortality in the vicinity of nuclear power stations in Japan" (Iwasaki. et al. 1995).
>
> "We see no statistically significant clustering of the observed cases about the four nuclear power plants in Sweden" (Waller et al. 1995, p. 14).

[3] "With epidemiology you can tell a little thing from a big thing. What's very hard to do is to tell a little thing from nothing at all." This is a quotation attributed by Taubes (1995, p. 164) to the director of analytical epidemiology of the American Cancer Society.

> "There was no evidence of a generally increased risk of childhood leukaemia … around nuclear sites in Scotland" (Sharp et al. 1996, p. 823).
>
> "Over the entire zone, children do not have an increased risk of malignant haematology disease" (Bouges et al. 1999, p. 205).
>
> "Our study shows no evidence of a generally increased risk of childhood leukaemia within 20km of the 29 nuclear sites under study" (White Koning et al. 2004).
>
> "There is no indication of any effect on the incidence of childhood cancer" (COMARE 2006, p. 115).
>
> "It is concluded that there is no evidence that acute leukaemia in children aged under five has a higher incidence close to NPSs in Britain" (Bithell et al. 2008, p. 196).
>
> "Neither for the whole study region nor for the individual NPP areas was a statistically significant average observed" (Kaatsch et al. 2008b, p. 727).
>
> "Our results do not indicate an increase in childhood leukemia and other cancers in the vicinity of Finnish NPPs" (Heinävara et al. 2009).

In the next section we argue that such studies, i.e. studies which report no effect at all, or no "significant" effect, have much lower chances of being undertaken in the first place and later getting published in the second. Or how often does one stumble on a journal article like "Pet ownership and childhood acute leukemia" (Swensen et al. 2001), which, after protracted investigations, finds that "no relationship was found between exposure to an ill pet and childhood leukemia" (p. 301)? This certainly does not happen very often, with the net result that meta-analyses such as Greiser (2009) are much more likely to summarize positive than negative results and are therefore much more likely than the nominal α-error claims to find effects where none exist.

3. Publication bias and errors of the third kind

A significance level of 5% for a statistical test means that, even without any effect being present, the test will claim one in roughly 5 out of 100 trials. This is the well known error of the first kind, which among the uninitiated often leads to an error of the third

kind: to assume that a significant test implies that the alternative is true. “The sin comes in believing a causal hypothesis is true because your study came up with a positive result” (Sander Greenland form UCLA, as quoted in Taubes, 1995, p. 169).

This error of the third kind, or some variant such as “the null hypothesis is wrong with 95 % probability” occurs even among professional statisticians. Haller and Krauss (2002) asked 30 statistics instructors, 44 statistics students and 39 practicing researchers from six psychology departments in Germany about the meaning of a significant two-sample t-test (significance level = 1%). The test was supposed to detect a possible treatment effect based on a control group and a treatment group. The subjects were asked to comment upon the following six statements (all of which are false). They were told in advance that several or perhaps none of the statements were correct.

> 1) You have absolutely disproved the null hypothesis (that is, that there is no difference between the population means). ❍ true / false ❍
>
> 2) You have found the probability of the null hypothesis being true. ❍ true / false ❍
>
> 3) You have absolutely proved your experimental hypothesis (that there is a difference between the population means). ❍ true / false ❍
>
> 4) You can deduce the probability of the experimental hypothesis being true. ❍ true / false ❍
>
> 5) You know, if you decide to reject the null hypothesis, the probability that you are making the wrong decision. ❍ true / false ❍
>
> 6) You have a reliable experimental finding in the sense that if, hypothetically, the experiment were repeated a great number of times, you would obtain a significant result on 99% of occasions. ❍ true / false ❍

All of the statistics students, 90% of the practicing psychologists and 80% of the methodology instructors marked at least one of the above faulty statements as correct. And what is more, even lots of statistics textbooks do. Examples from the American market include Guilford (1942, and later editions), which was probably the most widely read textbook in the 1940s and 50s, Miller & Buckhout (1973, statistical appendix by Brown, p. 523) or Nunally (1975, pp. 194 – 196). On the German market, there is Wyss (1991, p. 547) or Schuchard-Fischer et al. (1982), who on p. 83 of their best-selling textbook explicitly advise their readers that a rejection of the null at 5% implies a probability of 95% that the alternative is correct. For details, see Gigerenzer (2002, chap. 13), Krämer and Gigerenzer (2005), or Krämer (2008, chapter 8).

Another mistake, unrelated to but often occurring in tandem with the one above, is to report some nominal significance level α when in reality the reported test statistic is the most significant one among n trials, each conducted at the level α. The true significance level is then simply the probability that the maximum of n test statistics is larger than some critical value and increases rapidly with n. Table 1 gives some examples for independent trials and various nominal and true significance levels of the test.

Table 1: True significance level when rejection is based on the most unfavourable of n independent trials

number of trials	Nominal significance level		
	1%	5%	10%
2	1,9 %	9,8	19,0 %
3	3,0 %	14,3	27,1 %
4	3,9 %	18,5 %	34,4 %
5	4,9 %	22,6%	41,0 %
10	9,6 %	40, 1 %	65,1 %

Krämer and Runde (1992) have used this trick to establish what they call the "Krämer-Runde-seven-modulo 1 effect." This means in words, that on days of the month Nr. 1, 8, 15, 22, and 29 the German stock price index DAX performs significantly better than average (t=3.161). Or in technical terms, the null hypothesis that stocks perform the same on these days as on others could be rejected, given the available data, at a level of 5%. What Krämer and Runde also did, and also reported, were additional tests of many other hypotheses: There is no six-modulo-2-effect, there is no six-modulo-3-effect, there is no seven-modulo- 2-effect, eight-modulo-3-effect, and so on, ad nauseam. Given a particular data set and one hundred such hypotheses, all of them true, one is still bound to find about five "significant" effects, i.e. rejections of the null. And it is well known (see e.g. McCloskey 1983 or Ziliak and McCloskey 2008) that many other authors procede along similar lines, without reporting the unsuccessful trials, see also Krämer (2010, chapter 15). And although an increasing number of authors seem to be aware of this (see e.g. Fertig and Tamm 2010), only few take recourse to the impressive

toolbox of multiple testing procedures which have been developed to control for this effect.

In economics, this habit of reporting only the most unfavorable (to the null hypothesis) results is sometimes referred to as "data mining" (Lovell, 1983)[4]. It is of course strictly illegal and rightly frowned upon. Not illegal, but equally misleading, is the related phenomenon known as “publication bias”: 100 authors, each testing at 5%, are searching for effects, but there are none. Five studies still observe significant results. All studies are submitted for publication. Which have higher chances for acceptance?

One does not have to think hard (see section 2). Let us assume that 4 of the 5 studies with positive results and 36 of the 95 studies with negative results find their way into some scientific journal. This means that the true significance level of the tests is not 5% but 10%, and this happens even when no individual investigator engages in data mining. Denton (1985) calls this "collective data mining" and provides a rule of thumb to adjust for it in some selected applications.

It is common knowledge that such „collective data mining” is happening in almost every field where formal tests of significance are employed. “There is some evidence that in fields where statistical tests of significance are commonly used, research which yields nonsignificant results is not published” (Sterling 1959, p. 30). “Such research being unknown to other investigators may be repeated independently until eventually by chance a significant result occurs.” Taken to the limit, this argument implies that a “significant“ effect will be found eventually almost surely, no matter what.

In psychology, this bias is also known as the file drawer problem: negative results remain stuck in the file drawer. In medicine, Stern and Simes (1997) report that among 748 studies approved by the Royal Prince Alfred Hospital Ethics committee between 1979 and 1988, about 85% were eventually published if they reported significant results at levels 5% or less. Among studies which did not report significant results, this percentage of published papers was only 50 %. See also Beck-Bernholdt and Dubben (2004).

In economics, it is above all McCloskey who has repeatedly, although with little effect, drawn attention to this phenomenon, and the implications that this form of statistical

[4] Not to be confused with the serious business of the same name that is a modern subject of computer science

nonsense has for the field as such: “The progress of economic science has been seriously damaged. You can’t believe anything that comes out of [it]. Not a word. It is all nonsense, which future generations of economists are going to have to do all over again. Most of what appears in the best journals of economics is unscientific rubbish. I find this unspeakably sad. All my friends, my dear, dear friends in economics, have been wasting their time....They are vigorous, difficult, demanding activities, like hard chess problems. But they are worthless as science“ (2002, p. 44).

This is rather harsh judgement, and a bit beside the point. For instance, the large area of specification testing, where there is no particular alternative, and therefore no “effect” to be established, has certainly improved empirical economic work a lot. But whenever significance tests are meant, not to test the validity of some model (which in case of rejection is to be substituted by a better one), but to establish a particular and prearranged alternative, pitfalls abound.

4. Data mining in radiation epidemiology

At the time of this writing, there are 439 commercial nuclear power reactors operating worldwide. Some sites have more than one reactor (in Germany, Biblis is an example), so the number of different sites is only 210. In addition, there are 368 operational research reactors, 10 reprocessing plants, 14 uranium refineries, and several dozen uranium mining and milling facilities and atomic weapon factories each (the exact number of the latter being, for obvious reasons, hard to validate). Adding the well above 300 nuclear sites which had been in operation sometime but have by now been decommissioned or shut down, there are well above 1000 geographical locations worldwide available for testing[5]. Greiser (2009) singles out 80 of these.[6]

The respective data are mostly from previous studies, which, like the Seascale studies in the UK, have in turn often been undertaken subsequent to the occurrence of leukemia clusters. This HARKing (Hypothezing After the Results are Known) reinforces the data

[5] The numbers are from Wikipedia and the websites of Atomforum (http://www.kernenergie.de/kernenergie/Themen/Kernkraftwerke/Kernkraftwerke_weltweit/index.php) and the International Atomic Energy Agency (http://nucleus.iaea.org/RRDB/RR/ReactorSearch.aspx?rf=1 eew).

[6] In fact, the number of sites on which his tables are based is even smaller than he claims: 69 rather than 75 in his table 4, for instance.

mining effect. In Germany for instance, testing on a massive scale started only after an abnormal cluster of leukemia cases was observed close to the Krümmel power generation plant.

Another important degree of freedom is the time period under consideration. The literature abounds with examples where excess mortality or morbidity was found in certain periods, but not in others (Heasman 1987, Möhner and Stabenow 1993, Kaatsch et al. 2008). For instance, the studies form Canada quoted by Greiser (2009), reporting excess incidence of childhood leukemia around Canadian nuclear power plants, cover only years up to 1986. It is rather safe to assume (and confirmed by private information from Canadian authorities) that no excess incidence was observed thereafter.

Then one has to choose a distance from the potential source of radiation. Conventional choices are 6.5 km (Evrard et al. 2006)[7], 15 km (Kaletsch et al. 1997, Möhner et al. 1993), 20 km (Laurier et al. 2008), 25 km or 50 km (COMARE 2005, 2006) or complete counties, like in most studies from Canada and the U.S.. Again, there is an abundance of examples where excess incidence or mortality was observed for some distances, but not for others. It is also not true that incidence necessarily increases with proximity to power plants. Laurier et al. (2006, table 1) for instance report 5.2 expected and 5 observed cases within a 5 km distance from 19 French nuclear power plants, as compared to 69.3 expected and 71 observed cases when the distance is increased to 20 km. Similar results are also given in Bithell et al. (2008, p. 195), who find "that there is no association between childhood cancer and proximity to NPs in the UK."

Then there is the type of cancer (myeloid leukaemia – ML, acute lymphoblastic leukaemia – ALL, acute non-lymphoblastic leukemia, Non-Hodgkin lymphoma, other cancers), which likewise might lead to an excess for one type and a deficit for another. Kaatsch et al. (2008b, p. 530) for instance find an excess of leukemia, but a deficit of other childhood cancers close to nuclear power plants in Germany. And sometimes there is an excess of ML but not of ALL, or vice versa, so any investigator has a large number of choices where to investigate. In addition, the age group of the children is also important. Laurier et al. (2006, p. 402) and Evrard et al. (2006, table 2), among many others, report an excess of leukemia for some age groups, and a deficit for others.

[7] not 40 km, as claimed by Greiser (2009). An area of 40 square km and an area of 40km x 40 km are not the same.

It is obvious that by judiciously adjusting these parameters it is trivial to establish "significant" effects of any sort. A prime example is Körblein and Hoffmann (1999, p. 18), who, being dissatisfied with negative results from another epidemiological study, got what they wanted using the same data set: "A reanalysis of the data … reveals a statistically significant increase in childhood cancers … when the evaluation is restricted to commercial power reactors, the vicinities closest to the plants and children of the youngest age group."

Greiser (2009) uses all data available to him from previous studies, plus data from various U.S. cancer registries. The following table, compiled from his table 4, p. 20-21, gives the number of leukemia cases for the age group 0-4. As this is also the age group where radiation induced susceptibility to leukaemia is supposed to be highest, we focus on this data set in what follows.

Table 2: Observed vs. expected leukemia cases for age group 0-4, version I

Country	Number of sites	Expected cases	Observed cases
Canada	2	47.7	58
France	19	108	114
Germany	15	524. 8	593
U.K.	9	43.8	50
U.S.	24	1244. 4	1312
total	69	1968.7	2127

The data for the UK cover only myeloid leukemia, which comprises about 20% of all leukemia cases, and are therefore rather small. The data for Germany, from Kaatsch et al. (2008a), who report an excess incidence of 13%, are not explicitly given by Greiser (2009), and are taken form the initial study. Also, the number of sites -75 – which Greiser quotes is not correct. Still, according to table 2, the expected value of leukemia cases, if incidence around nuclear power plants were equal to the national average, is 1969, as compared to an actual number of 2127, so there certainly appears to be some reason for concern.

The particular statistical procedure which was employed by Greiser to show that this excess is "significant" shall not concern us here. Rather, the point we want to make is that any "significance", no matter how it was obtained, is bound to disappear once some obvious deficiencies have been accounted for. For instance, what if some plants outside the scope of Greiser (2009) had also been included? According to Sofer (1991), Waller (1995) or Heinävara et. al (2009), cancer incidence around nuclear power plants in Sweden, Israel and Finland is no higher than elsewhere and sometimes well below. Also, no excess incidence has so far been reported for nuclear sites in Japan, Spain and Switzerland. Given the enormous media interest in occurrences of this kind, one can certainly be sure that any leukemia cluster close to a nuclear facility in these counties would have made headlines there as well.[8] Therefore, the absence of such headlines provides evidence that no such clusters have occurred.

5. Disregard of confounding factors

As mentioned before, childhood leukemia often comes in clusters. Contrary to what most true believers claim, there is no consensus on the underlying causes. Extremely high doses of radiation might theoretically be responsible, but have never been observed or even been approximated in routine practice close to nuclear power plants. In fact, if there is any agreement at all among partisans in this debate, then this concerns the impossibility of routine doses of industrial radiation to cause cancer in the first place: "Based on the findings of radiation research such a connection seems implausible, because the radiation emitted by an NPP in normal operation is at least 1000 times lower than 'background radiation', i.e. the 1.5mSv of natural radiation to which the average German is exposed in a year" (Kaatsch et al. 2008b, p. 729).

According to Ries et al. (1999, figure 6 and table 1.5), and confirmed by many others, risk factors which are really important in practice are race and sex. For instance, childhood cancer incidence in the U.S. is 30% higher for boys as compared to girls and almost double for whites as compared to blacks. For leukemia only, the highest incidence rates are observed among hispanics (48.5 per million as compared to 41.6 per

[8] In fact, there was a preliminary examination in Switzerland following the KiKK-excitement, which produced no effect and was therefore neglected by the media, see Reichmuth (2010). The final results will be available in 2011.

million for whites and 25,8 per million for blacks). By far the lowest rates for any type of childhood cancer are observed for American Indians.

Also, leukemia incidence correlates strongly with income – the higher the income of the parents, the larger the risk of leukemia for kids (Borugian et al. 2005, COMARE 2006 and many others). The true underlying cause is still subject to debate; current hypotheses include an increased susceptibility of wealthy children to non-specific infectious agents (COMARE 2006, p. 12; wealthy children are brought up in "cleaner" environments and develop less antibodies) or a higher incidence of parental cosanguinity. In Scotland, for instance, the incidence of childhood leukemia between the richest and the poorest subpopulations differs by as much as 50%.

Other risk factors which have been identified so far are population density (more cases per 1000 children in densely populated as compared to sparsely populated areas: "it can be seen that the incidence of … tumours increases as population density increases at both county district and word level" (COMARE 2006, p.26) and population mixing (Kinlen 1995, Kinlen and Doll 2004, COMARE 2005, p.8)). Like population density, this might likewise lead to an increased exposure of susceptible individuals to infections and local epidemics which in turn could later promote the onset of cancers of many types.

It would be surprising if these established covariates did not also affect the numbers in table 2. For instance, the plant that contributes most to the surplus of 158 leukemia cases reported in the table is San Onofre Nuclear Generating Station in Southern California. It is located in the northwestern corner of San Diego County, south of the city of San Clemente, and started operations in 1967. Its initial unit is no longer in service, but two additional units, built in the early eighties, have licences to operate until 2022. According to Greiser (2009, p. 21, table 4) there were 281 cases of childhood leukemia close to San Onofre (which in this case means: in San Diego County) in the 2001-2006 time period, compared to only 177 expected cases, an excess of 104. Therefore, this single data point contributes almost all of the excess cases in table 1.

Now, looking closer at the San Onofre site (see figure 1), it appears that the power plant is almost 300 km away from the south-eastern border of San Diego County, where it is supposed to be responsible for cancer. Attributing cancer cases there to radiation in San Onofre is like attributing cancer in Hanover to the Krümmel nuclear power plant on the

river Elbe one hundred miles to the north. This is mistake Nr. 1: The geographical area for investigating leukemia cases connected to San Onofre is much too large. Even if one focused on more densely populated areas, this argument would still apply, since the metropolitan area of San Diego, where the bulk of the population of San Diego County lives, is still more than 100 kilometers away. This means that even if there were an impact of San Onofre on childhood leukemia, it could hardly be detected with the Greiser (2009) data set.

Figure 1: San Diego County and San Onofre nuclear power generating station

Even more important is mistake Nr. 2: The neglect of virtually all confounding factors which have so far been established in the literature. For instance, San Diego County is rather wealthy. According to Forbes Magazine, San Diego is the 4th wealthiest city in the U.S., and household income in San Diego County overall is 20 % above the national average, see table 3. In addition, San Diego County has an above-average population of Hispanics and very few blacks (in the city of San Clemente, which is closest to San Onofre, blacks compose less than 1% of the population). In fact, among children under the age of 18, the largest proportion in the meantime is hispanic (which is also the ethnic group where leukemia incidence among children is highest). Also, both

population density and population mixing are more pronounced in San Diego County than elsewhere in the U.S.. San Diego is the largest concentration of naval facilities in the world, with a constant moving in and out of families, which is even further accentuated by a large university and many more military facilities such as training camps, airbases, Marine corps recruit depots and coast guard stations. All of these variables correlate strongly with childhood leukemia.

Summing up, among factors which are known to correlate positively with childhood leukemia, almost every one is larger in San Diego County than elsewhere in the United States. Not surprisingly, therefore, taking account of these covariables and using data from the early days of operation of the plant, Enstrom (1983) found that childhood leukemia is no more prevalent around San Onofre than elsewhere.

Table 3: San Diego County vs. National Average (Census 2002)

Variable	San Diego County	National Average
mean household income	$47.067	$41.994
percentage blacks	5.7%	12.3%
percentage white	66.5%	75.1%
percentage Hispanic or latino	26.7%	12.5%
number of white children <5	110 739	
number of black children <5	13 276	
number of Hispanic children <5	80 261	

However, removing San Onofre from the Greiser (2009) data set, and adding some studies he has overlooked (for instance Bithell et al. (2008) and Kaatsch et al. (2008b)) the initial surplus of leukemia cases among children aged 0-4 turns into a deficit (table 4): Other than in table 2, the data for the UK now comprises all sorts of acute leukemia as specified by International Classification of Childhood Cancer Groups 11 and 12; therefore, incidence is lager. The data from Germany was collected by almost the same research group which had supplied the German data for table 2 (Kaatsch et al. 2008a), but covers a longer time span. Therefore, the data base for table 4 is both more comprehensive and less prone to omitted variable bias (due to the deletion of San Onofre) than table 2.

Table 4: Observed vs. expected leukemia cases for age group 0-4, version II

Country	Number of sites	Expected cases	Observed cases
Canada	2	47,7	58
France	19	108	114
Germany	15	623,7	619
U.K.	13	374,9	360
U.S.	23	1067,9	1031
Together	72	2222,2	2182

Will there ever be a study claiming that nuclear power protects against leukemia? With some proper data mining, and a convenient choice of statistical model, this salutary side effect can almost certainly be made highly "significant".

References

Alexander, F.E. (1999): "Clusters and clustering of childhood cancer: A review", *European Journal of Epidemiology* 15, 847 – 852.

Baker, P. J. and Hoel, D. G. (2007): "Meta-analysis of standardised incidence and mortality rates of childhood leukaemia in proximity to nuclear facilities", *European Journal of Cancer Care* 16, 355 – 363.

Beck-Bornholdt, H.-P. and Dubben, H.-H. (2004): *Unausgewogene Berichterstattung in der medizinischen Wissenschaft - publication bias-*, Hamburg (Institut für Allgemeinmedizin des Universitätsklinikums Hamburg-Eppendorf).

Bithell, J. F., Keegan, T. J., Kroll, M. E., Murphy, M. F. and Vincent, T. J. (2008): "Childhood leukaemia near British nuclear installations: Methodological issues and recent results", *Radiation Protection Dosimetry* 132, 191 – 197.

Borugian, M. J., Spinelli, J., Mezei, G., Wilkins, R., Abanto, Z., and Mc Bride, M. L.: (2005): "Childhood leukaemia and socioeconomic status in Canada", *Epidemiology* 16, 526-531.

Bündnis 90 / Grüne (2009): "AKWs erhöhen das Leukämierisiko", Pressrelease, 07.09.2009.

Committee on medical aspects of radiation in the environment (COMARE) (2005): "The incidence of childhood cancer around nuclear installations in Great Britain", *10th report*, London.

Committee on medical aspects of radiation in the environment (COMARE) (2006):"The incidence of childhood cancer around nuclear installations in Great Britain", *11th report*, London.

Clarke, E.A., McLaughlin, J., Anderson, T. W. (1991): "Childhood leukaemia around Canadian nuclear facilities. Phase II. Final report", Atomic Energy Control Board, Ottawa.

Denton, F.T. (1985): „Data mining as an industry", *The Review of Economics and Statistics* 67, 124 – 127.

Dewdney, A.K. (1996): *200 % of Nothing: An Eye Opening Tour through the Twists and Turns of Math Abuse and Innumeracy*, New York (Wiley).

Enstrom, J.E. (1983): "Cancer mortality patterns around the San Onofre nuclear power plant, 1960-1978", *American Journal of Public Health* 73, 83 – 92.

Evrard, A. S., Hémon, D., Morin, A., Laurier, D., Tirmarche, M., Backe, J. C., Chartier M. and Clavel J. (2006): "Childhood leukaemia around French nuclear installations using geographic zoning based on gaseous discharge dose estimates", *British Journal of Cancer* 94, 1342 – 1347.

Ewings, P.D., Bowie, C., Phillips, M.J., Johnson, S.A.N. (1989): "Incidence of leukemia in young people in the vicinity of Hinkley Point nuclear power station, 1959-86", *British Medical Journal* 299, 289 – 293.

Fertig, M. and Tamm, M. (2010); "Always poor or never poor and nothing in between? duration of child poverty in Germany", *German Economic Review* 11, 150-168.

Gigerenzer, G. (2002): *Calculated risks: How to know when numbers deceive you*. New York (Simon & Schuster). [British edition: Reckoning with risk. London (Penguin).]

Greiser, E. (2009): „Leukämie-Erkrankungen bei Kindern und Jugendliche in der Umgebung von Kernkraftwerken in fünf Ländern", Report prepared for the political party Bündnis 90 / Grüne, see http://www.gruene-bundestag.de/cms/archiv/dokbin/302/302113.studie_leukaemierisiko.pdf

Haller, H. and S. Kraus (2002): "Misinterpretation of significance: A problem students share with their teachers?" *Methods of Psychological Research Online* 7, 1 – 20.

Heasman, M.A., Urquhart, J.D., Black, R.J. and Kemp, I.W. (1987): "Leukemia in young persons in Scotland: a study of its geographical distribution and relationship to nuclear installations", *Health Bulletin* 45, 147 – 151.

Heinävara, S., Toikkanen, S., Pasanen, K., Verkasalo, P. K., Päivi, K and Auvinen, A. (2010): „Cancer incidence in the vicinity of Finnish nuclear power plants: an emphasis on childhood leukemia", *Cancer Causes and Control* 21, 587 – 595.

Hoffmann, W., Kuni, H. and Ziggel, H. (1996): "Leukämiesterblichkeit in der Nähe von japanischen Atomkraftwerken doch erhöht", *Strahlentelex* 238, 2 – 5.

Hoffmann, W., Terschueren, C., Richardson, D.B. (2007): "Childhood leukemia in the vicinity of the Geesthacht nuclear establishments near Hamburg, Germany", *Enviromental Health Perspectives* 115, 947 – 952.

Iwasaki, T., Nishizawa, K. and Murata, M. (1995): "Leukaemia and lymphoma mortality in the vicinity of nuclear power stations in Japan", *Journal of Radiological Protection* 15, 271 – 288.

Kaatsch, P., Spix, C., Schulze-Rath, R., Schmiedel, S. and Blettner. M. (2008a): „Leukaemia in young children living in the vicinity of German nuclear power plants“, *International Journal of Cancer* 721-726.

Kaatsch, P., Spix, C., Jung, I. and Blettner, M. (2008b): “Childhood leukemia in the vicinity of nuclear power plants in Germany”, *Deutsches Ärzteblatt International* 105, 725 – 732.

Kinlen, L.J. (1995): “Epidemiological evidence for an infective basis in childhood leukemia”, British Journal of Cancer 71, 1-5.

Kinlen, L.J. and Doll, R. (2004): “Population mixing and childhood leukemia: Fallon and other US clusters,” *British Journal of Cancer* 91, 1-3.

Körblein, A. and Hoffmann, W. (1999): “Childhood cancer in the vicinity of German nuclear power plants“, Medicine & Global Survival 6, 18-23.

Krämer. W. (2008): *Denkste – Trugschlüsse aus der Welt des Zufalls und der Zahlen*. 8th paperback edition, München (Piper).

Krämer, W. (2010): *So lügt man mit Statistik*. 10th paperback edition, München (Piper).

Krämer, W. and Runde, R. (1992): “The holiday effect: yet another capital market anomaly?“ in Schach, S. and Trenkler, G. (Hrsg.): *Data analysis and statistical inference: Festschrift in honour of Friedhelm Eicker*, Bergisch-Gladbach (Eul-Verlag), 453 – 462.

Krämer, W. and Gigerenzer, G. (2005): „How to confuse with statistics. The use and misuse of conditional probabilities,” *Statistical Science* 20, 223-230.

Laurier, D. and Bard, L.: “Epidemiologic studies of leukemia among Persons under 25 years of age living near nuclear sites“, *Epidemiologic Reviews* 21, 188 – 206.

Lovell, M.C. (1983): “Data mining”, *Review of Economics and Statistics* 65, 1 – 12.

McCloskey, D. (1983): “The rhetoric of economics*”*, *Journal of Economic Literature* 21, 481-517.

McCloskey, D. (2002): *The Secret Sins of Economics*, New York (Wiley).

Möhner, M. und Stabenow, R. (1993): “Childhood malignancies around nuclear installations in the former GDR” *Medizinische Forschung* 6, 59 – 67.

Michaelis, J., Keller, B., Haaf, G. and Kaatsch, P. (1992): “Incidence of childhood malignancies in the vicinity of West German nuclear power plants“, *Cancer Causes Control* 3, 255 – 263.

Miller, G. A. and R. Buckhout (1973): *Psychology: The science of mental life*. New York (Harper & Row).

Nunally, J. C. (1975): *Introduction to statistics for psychology and education*. New York (McGraw-Hill).

Reichmuth, A. (2010): “Angstmacherei mit Atomkraft”, *Die Weltwoche* Nr. 16, 22. April.

Ries, L. A. G., Smith, M. A., Gurney, J. G., Linet, M., Tamra, T., Young, J. L. and Bunin, G. R. (1999): “Cancer Incidence and Survival among Children and

Adolescents: United States SEER Program 1975-1995", National Cancer Institute, Bethesda, MD.

Schuchard-Fischer, C., K. Backhaus, H. Hummel, W. Lohrberg, W. Plinke and W. Schreiner (1982): *Multivariate Analysemethoden – Eine anwendungsorientierte Einführung*. 2nd edition, Berlin (Springer).

Sharp, L., Black, R. J., Harkness, E. F. and McKinney, P. A. (1996): "Incidence of childhood leukaemia and non-Hodgkin's lymphoma in the vicinity of nuclear sites in Scotland", *Occupational and Environmental Medicine* 53, 823 – 831.

Sterling, T. R. (1959): "Publication decisions and their possible effects on inferences drawn from tests of significance – or vice versa", *Journal of the American Statistical Association* 54, 30-34.

Stern J.M., Simes, R.J. (1997): "Publication bias: evidence of delayed publication in a cohort study of clinical research projects", *British Medical Journal* 315, 640 – 645.

Sofer, T., Goldsmith, J. R., Nusselder, I. and Katz, L. (1991): "Geographic and temporal trends of childhood leukaemia in relation to the nuclear power plant in the Negev", *Public Health Review* 19, 191 – 198.

Swensen, A. P., Ross, J.A., Shu, X.O., Reaman, G. H., Steinbuch, M. and Robison, L. L. (2001): "Pet ownership and childhood acute leukemia", *Cancer Causes and Control* 12, 301-303.

Taubes, G. (1995): "Epidemiology faces its limits", *Science* 269, 164-169.

Waller, L. A., Turnbull, B.W., Gustafsson, G., Hjalmars, U. and Andersson, B. (1995): "Detection and assessment of clusters of disease: an application to nuclear power plant facilities and childhood leukaemia in Sweden", *Statistics in Medicine* 14, 3 – 16.

White-Koning, M. L., Hémon, D., Laurier, D., Tirmarche, M. Jougla, E. Goubin, A. and Clavel J. (2004): "Incidence of childhood leukaemia in the vicinity of nuclear sites in France, 1990-1998", *British Journal of Cancer* 91, 916 – 922.

W. Wyss (1991): *Marktforschung von A – Z*. Lucerne (Demascope).

Ziliak, S. and McCloskey, D. (2008): *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice and Lives*, Ann Arbor (University of Michigan Press).

CESifo Working Paper Series

for full list see www.cesifo-group.org/wp
(address: Poschingerstr. 5, 81679 Munich, Germany, office@cesifo.de)

3116 M. Hashem Pesaran, Predictability of Asset Returns and the Efficient Market Hypothesis, July 2010

3117 Dorothee Crayen, Christa Hainz and Christiane Ströh de Martínez, Remittances, Banking Status and the Usage of Insurance Schemes, July 2010

3118 Eric O'N. Fisher, Heckscher-Ohlin Theory when Countries have Different Technologies, July 2010

3119 Huw Dixon and Hervé Le Bihan, Generalized Taylor and Generalized Calvo Price and Wage-Setting: Micro Evidence with Macro Implications, July 2010

3120 Laszlo Goerke and Markus Pannenberg, 'Take it or Go to Court' – The Impact of Sec. 1a of the German Protection against Dismissal Act on Severance Payments -, July 2010

3121 Robert S. Chirinko and Daniel J. Wilson, Can Lower Tax Rates be Bought? Business Rent-Seeking and Tax Competition among U.S. States, July 2010

3122 Douglas Gollin and Christian Zimmermann, Global Climate Change and the Resurgence of Tropical Disease: An Economic Approach, July 2010

3123 Francesco Daveri and Maria Laura Parisi, Experience, Innovation and Productivity – Empirical Evidence from Italy's Slowdown, July 2010

3124 Carlo V. Fiorio and Massimo Florio, A Fair Price for Energy? Ownership versus Market Opening in the EU15, July 2010

3125 Frederick van der Ploeg, Natural Resources: Curse or Blessing?, July 2010

3126 Kaisa Kotakorpi and Panu Poutvaara, Pay for Politicians and Candidate Selection: An Empirical Analysis, July 2010

3127 Jun-ichi Itaya, Makoto Okamura and Chikara Yamaguchi, Partial Tax Coordination in a Repeated Game Setting, July 2010

3128 Volker Meier and Helmut Rainer, On the Optimality of Joint Taxation for Non-Cooperative Couples, July 2010

3129 Ryan Oprea, Keith Henwood and Daniel Friedman, Separating the Hawks from the Doves: Evidence from Continuous Time Laboratory Games, July 2010

3130 Mari Rege and Ingeborg F. Solli, The Impact of Paternity Leave on Long-term Father Involvement, July 2010

3131 Olaf Posch, Risk Premia in General Equilibrium, July 2010

3132 John Komlos and Marek Brabec, The Trend of BMI Values by Centiles of US Adults, Birth Cohorts 1882-1986, July 2010

3133 Emin Karagözoğlu and Arno Riedl, Information, Uncertainty, and Subjective Entitlements in Bargaining, July 2010

3134 John Boyd, Gianni De Nicolò and Elena Loukoianova, Banking Crises and Crisis Dating: Theory and Evidence, July 2010

3135 Michael R. Baye, Dan Kovenock and Casper G. de Vries, The Herodotus Paradox, July 2010

3136 Martin Kolmar and Hendrik Rommeswinkel, Group Contests with Complementarities in Efforts, July 2010

3137 Carolina Manzano and Xavier Vives, Public and Private Learning from Prices, Strategic Substitutability and Complementarity, and Equilibrium Multiplicity, July 2010

3138 Axel Löffler, Gunther Schnabl and Franziska Schobert, Inflation Targeting by Debtor Central Banks in Emerging Market Economies, July 2010

3139 Yu-Fu Chen and Michael Funke, Global Warming and Extreme Events: Rethinking the Timing and Intensity of Environmental Policy, July 2010

3140 Lawrence M. Kahn, Labor Market Policy: A Comparative View on the Costs and Benefits of Labor Market Flexibility, July 2010

3141 Ben J. Heijdra, Jochen O. Mierau and Laurie S.M. Reijnders, The Tragedy of Annuitization, July 2010

3142 Erkki Koskela, Outsourcing Cost and Tax Progression under Nash Wage Bargaining with Flexible Outsourcing, July 2010

3143 Daniel Osberghaus and Christiane Reif, Total Costs and Budgetary Effects of Adaptation to Climate Change: An Assessment for the European Union, August 2010

3144 Philip E. Graves, Benefit-Cost Analysis of Environmental Projects: A Plethora of Systematic Biases, August 2010

3145 Sabrina Di Addario and Daniela Vuri, Entrepreneurship and Market Size. The Case of Young College Graduates in Italy, August 2010

3146 Shoshana Amyra Grossbard and Alfredo Marvão Pereira, Will Women Save more than Men? A Theoretical Model of Savings and Marriage, August 2010

3147 Jarko Fidrmuc, Time-Varying Exchange Rate Basket in China from 2005 to 2009, August 2010

3148 Ilja Neustadt and Peter Zweifel, Is the Welfare State Sustainable? Experimental Evidence on Citizens' Preferences for Redistribution, August 2010

3149 Marcus Dittrich and Andreas Knabe, Wage and Employment Effects of Non-Binding Minimum Wages, August 2010

3150 Shutao Cao, Enchuan Shao and Pedro Silos, Fixed-Term and Permanent Employment Contracts: Theory and Evidence, August 2010

3151 Ludger Woessmann, Cross-Country Evidence on Teacher Performance Pay, August 2010

3152 Lorenzo C. G. Pozzi, Casper G. de Vries and Jorn Zenhorst, World Equity Premium Based Risk Aversion Estimates, August 2010

3153 Volker Grossmann, Thomas M. Steger and Timo Trimborn, Dynamically Optimal R&D Subsidization, August 2010

3154 Alexander Haupt, Tim Krieger and Thomas Lange, A Note on Brain Gain and Brain Drain: Permanent Migration and Education Policy, August 2010

3155 António Afonso and Christophe Rault, Long-run Determinants of Sovereign Yields, August 2010

3156 Franziska Tausch, Jan Potters and Arno Riedl, Preferences for Redistribution and Pensions. What can we Learn from Experiments?, August 2010

3157 Martin Kolmar and Andreas Wagener, Inefficient Group Organization as Optimal Adaption to Dominant Environments, August 2010

3158 Kai Carstensen, Klaus Wohlrabe and Christina Ziegler, Predictive Ability of Business Cycle Indicators under Test: A Case Study for the Euro Area Industrial Production, August 2010

3159 Horst Rottmann and Timo Wollmershäuser, A Micro Data Approach to the Identification of Credit Crunches, August 2010

3160 Philip E. Graves, Appropriate Fiscal Policy over the Business Cycle: Proper Stimulus Policies Can Work, August 2010

3161 Michael Binder and Marcel Bluhm, On the Conditional Effects of IMF Program Participation on Output Growth, August 2010

3162 Michael Binder, Qianying Chen, and Xuan Zhang, On the Effects of Monetary Policy Shocks on Exchange Rates, August 2010

3163 Felix J. Bierbrauer, On the Optimality of Optimal Income Taxation, August 2010

3164 Nikolaus Wolf, Europe's Great Depression – Coordination Failure after the First World War, September 2010

3165 Dan Kovenock and Brian Roberson, Conflicts with Multiple Battlefields, September 2010

3166 Jean-Pierre Ponssard and Catherine Thomas, Capacity Investment under Demand Uncertainty. An Empirical Study of the US Cement Industry, 1994-2006, September 2010

3167 Jørgen Juel Andersen, Jon H. Fiva and Gisle James Natvik, Voting when the Stakes are High, September 2010

3168 Michael Hoel, Is there a Green Paradox?, September 2010

3169 Scott Alan Carson, Nineteenth Century US African-American and White Female Statures: Insight from US Prison Records, September 2010

3170 Gil S. Epstein, Yosef Mealem and Shmuel Nitzan, Political Culture and Discrimination in Contests, September 2010

3171 Sara Fisher Ellison, Jeffrey Greenbaum and Wallace P. Mullin, Diversity, Social Goods Provision, and Performance in the Firm, September 2010

3172 Silvia Dominguez-Martinez, Randolph Sloof and Ferdinand von Siemens, Monitoring your Friends, not your Foes: Strategic Ignorance and the Delegation of Real Authority, September 2010

3173 Marcus Dittrich and Beate Schirwitz, Union Membership and Employment Dynamics: A Note, September 2010

3174 Francesco Daveri, Paolo Manasse and Danila Serra, The Twin Effects of Globalization – Evidence from a Sample of Indian Manufacturing Firms, September 2010

3175 Florian Blöchl, Fabian J. Theis, Fernando Vega-Redondo and Eric O'N. Fisher, Which Sectors of a Modern Economy are most Central?, September 2010

3176 Dag Morten Dalen, Marilena Locatelli and Steinar Strøm, Longitudinal Analysis of Generic Substitution, September 2010

3177 Armin Falk, Stephan Meier and Christian Zehnder, Did we Overestimate the Role of Social Preferences? The Case of Self-Selected Student Samples, September 2010

3178 Christian Fahrholz and Cezary Wójcik, The Bail-Out! Positive Political Economics of Greek-type Crises in the EMU, September 2010

3179 Klaus Abberger and Wolfgang Nierhaus, The Ifo Business Cycle Clock: Circular Correlation with the Real GDP, September 2010

3180 Walter Krämer and Gerhard Arminger, "True Believers" or Numerical Terrorism at the Nuclear Power Plant, September 2010