

Interdisciplinary Sojourns

by

Jeffrey S. Rosenthal¹

(May, 2014. Upon receiving the SSC 2013 Gold Medal.)

Abstract. I was very honoured to receive the SSC 2013 Gold Medal, which caused me to reflect on different aspects of my research career. In this paper, I provide a personal reflection on those projects of mine which could be regarded as “interdisciplinary”. I present informal summaries of my experiences with these projects, and consider possible lessons to be learned.

1 Introduction

By training, I am a mathematician. My undergraduate and graduate studies were both primarily in mathematics, and I’ve always been thrilled by rigorous mathematical arguments involving formal ϵ - δ constructions and the like. It’s true that I’ve always wanted to somehow “apply” my mathematical thinking – but as an undergraduate that just meant taking lots of physics courses, and as a graduate student that eventually (after much uncertainty) meant studying mathematical probability theory since it had some slight connections to “the real world”. In typical mathematician fashion, these efforts comprised the full extent of my early “application” efforts.

My probability PhD supervisor eventually steered me towards Markov chain Monte Carlo (MCMC) computer algorithms, which are very widely used by researchers in many fields. But even then, my papers remained firmly on the theoretical side, proving theorems about abstract convergence properties without “getting my hands dirty” with application details. Some applied researchers did start citing my papers, but usually just with passing mentions about “related theoretical results” rather than because my work had any direct influence on their investigations. I was still thoroughly an abstract and theoretical researcher.

Very gradually, over the twenty or so years since my PhD, I have engaged more and more in truly interdisciplinary research, i.e. in actually working with people who are in other fields outside of my usual mathematics / statistics / computer science comfort zone. The transition hasn’t been easy, and is still far from complete. But it has led to various brief visits – sojourns – to different subjects and different sorts of collaborations.

¹Department of Statistics, University of Toronto, Toronto, Ontario, Canada M5S 3G3. Supported in part by NSERC of Canada. E-mail: jeff@math.toronto.edu. Web: <http://probability.ca/jeff/>

When I was asked to write a paper about my research career (on the occasion of my receiving the SSC 2013 Gold Medal, for which I am very honoured and grateful), I realised that I have already written a number of summaries of my core research area of probability theory applied to MCMC (e.g. Rosenthal, 2014b; Craiu and Rosenthal, 2014; Roberts and Rosenthal, 2001, 2004; for a rather different reflection on my academic career see Rosenthal, 2014c). So, instead, in this paper, I will present brief summaries and personal recollections and reflections about my various interdisciplinary projects, and what lessons I have learned from them – in the hopes that my thoughts might be of some help to other statistical researchers who are considering moving in more interdisciplinary directions.

2 First Forays

I was initially fairly slow to move away from my core research areas of mathematics and probability theory with applications to MCMC. As a result, my first interdisciplinary projects still involved a lot of probability theory. Nevertheless, they did provide the foundations for me to move further afield later on.

2.1 Economics: Mixed Nash Equilibria

My first truly interdisciplinary collaboration was fairly painless. A senior economics professor at my university, whom I also knew personally, approached me for help with a probability problem. He and his colleague were studying models for community voting in which the *median* opinion prevails. Intuitively, in such cases only those with *extreme* opinions will bother to vote, since moderate opinions probably won't affect the median very much anyway. The economists were studying this concept formally, using the game theory concept of “Nash equilibria” strategies in which no single player can improve their payoff by changing their action. Unfortunately for them, in some of their models no pure Nash equilibrium existed. Instead, they had to consider “mixed” Nash equilibria, in which players acted in different ways with different – you guessed it – *probabilities*. So they decided to contact me.

This was quite an ideal first interdisciplinary collaboration for me: the economists were respected and knowledgeable researchers in their field, they already had a clearly defined question and some progress towards its solution, the part they were missing was specifically about probability theory, they truly wanted me to collaborate with them, and they were helpful and appreciative as I worked towards a solution. (I would later learn that such favourable conditions are not always present!) With their help, I was able to derive some

theorems guaranteeing certain properties of their mixed Nash equilibria, which bolstered their case that moderate voters would tend not to bother to vote. They incorporated my work into their paper, and it led to a nice publication in one of the very top academic economics journals (Osborne et al., 2000). A success all around.

2.2 Finance: Generators for Credit Ratings

My second truly interdisciplinary collaboration was only a little less painless. A professor of finance from my university's business school, whom I did not know, was studying credit ratings. Agencies like Moody's and Standard & Poor's give investment opportunities like bonds a "credit rating" which is one of 8 ratings: AAA, AA, A, BBB, BB, B, CCC, or Default. These ratings change over time, e.g. a bond might be downgraded from AA to A. Certain finance researchers were studying the pattern of these credit rating changes by considering them as an 8-state Markov chain. However, their models were all in discrete time, allowing changes only when new ratings were published every six months or so. The professor of finance wanted to convert these results to continuous-time Markov processes, i.e. to instantaneous *rates* of change. This led to the question of whether discrete-time Markov chain transition probabilities could always be embedded into an underlying continuous-time Markov process. That's when the professor decided to contact me.

It turned out that the answer to his question was no. Many discrete-time transitions could *not* be embedded in a continuous-time process. Sure, you could use logarithms to come up with a matrix whose exponentiation would equal the original discrete-time transitions, but it wouldn't be a true generator for a continuous-time process since its entries would have the wrong signs. On the other hand, it was easy enough to find a true generator matrix whose exponentiation would *approximately* equal the discrete-time transitions, i.e. the problem could be *approximately* solved without difficulty. The finance professor was pleased with all of these results, and we started to move towards publication.

There were still some wrinkles. One was that I couldn't figure out what other related work had been done previously regarding the existence and approximation of generators. Web searches and enquiries with colleagues didn't yield much, so I decided to *post to an online discussion forum* of all things. This led to a lot of useless answers, but also to one very helpful reply from an older math professor at UBC. I started e-mailing with the UBC professor, which yielded further clarifications and refinements. Before I knew it, we had written a three-author paper which was published in a well-known finance journal (Israel et al., 2000). Another success!

2.3 Political Economics: Computing Capitalism

Shortly thereafter, I was referred by a math colleague to another finance professor, this one retired and of Marxist bent. He wanted to use simple models of capitalist enterprises to establish certain flaws, or at least strict requirements, in the capitalist system.

This project proved challenging primarily because it was hard to pin the retired professor down on precisely what models he wanted to consider, and precisely what simulations he required, and precisely what results he was after. However, after many back-and-forth e-mails, and lots of face-to-face meetings, I was able to prove some simple limit theorems about his models, and also to run numerous computer simulations to illustrate his points. Eventually he was happy enough to write the project up as a formal paper, including an Appendix with my detailed mathematical results. After some further revisions, the paper was published in a respected economics journal (Gordon and Rosenthal, 2003).

So, this collaboration also had a happy ending. However, a later follow-up project got bogged down by the retired professor's increasing demands and failing health, and never got finished, leaving some simulations of mine (and of a PhD student I had recruited to the project) unused and wasted. You win some, you lose some.

3 Criminology: Modeling Young Offenders

My longest-running interdisciplinary collaboration began when a mathematician-friend's wife's psychologist-colleague heard that "hidden Markov models" could be useful in studying the criminal trajectories of young offenders. He didn't know what hidden Markov models were, so he asked the mathematician-friend's wife to recommend someone who knew about "Markov" things, and the mathematician-friend mentioned that I worked on "Markov chains", and introductions were made.

It thus came to pass that four psychology researchers, all collecting and studying young criminal offender data, came to my office to "audition" me to collaborate with them and further analyse their data. I had to confess to them that my MCMC algorithms were rather different from the hidden Markov models they had heard of, but it turned out they didn't "really" need hidden Markov models after all so that was fine. Of greater concern to me was that due to my unusual background, I was a statistics professor who hadn't ever really done any major large-scale data analysis at all – but I decided *not* to confess that bit. Instead, I confidently declared that the key to any statistical analysis is getting good data, so I would be delighted to work with them. Apparently I, like John Lennon in 1969, "passed the audition", and the next thing I knew they were sharing huge Excel files with me.

At this point, a few problems arose. For one thing, their data turned out not to be as clean as I'd anticipated. Their Excel files had over one hundred columns, some of which were hard to interpret, many of which had missing data, some of which were inaccurate, and all of which were tricky to load correctly into a computer program. Sorting out the data turned out to be a major challenge.

In addition, it soon became clear that the psychologists did not have a precise question they wanted answered. They wanted to better understand the relationship between young people's criminal activities, and their later criminal activities in adulthood, with an eye towards improving prevention techniques. But how to model this as a statistical question was unclear. After much discussion, we agreed (or I insisted) that we could focus on the issue of how best to *predict* adult criminal behaviour from adolescent criminal behaviour.

Furthermore, we wanted to adjust the offense rates by taking into account the dates when offenders were in secure custody and thus unable to re-offend. Now, the secure-custody data was available, but it was very difficult to parse, and also it couldn't be matched up precisely with commission of new offenses since we had only the *conviction* dates, not the actual *offense* dates. In short, the data was messy and the challenges were immense – not such a great “first serious data analysis” project for me after all.

In desperation, I enlisted a junior colleague with more applied leanings. He agreed to help out, and contributed greatly to the project over the ensuing years. But one aspect of his collaboration was distinctly *unhelpful*: he departed within a year for another university, thus necessitating long-distance collaboration through e-mails and phone calls and visits. Despite this additional challenge, we finally got a handle on the data and made some progress.

We then made the error of getting *too ambitious*. Previous studies of offender data had used very simple latent-class models and just estimated a few parameters. We thought this was too limiting, so we set out to use *all* of the adolescent data (offense dates, offense types, time in custody, etc.) to predict the *entire* adult offense trajectory (using Cox proportional hazards models), and to carefully compare different model approaches using cross-validation. We worked hard and eventually got some results, but the reality was that the modest data that we had (just a few hundred offenders) was not up to the ambitious task we had set ourselves. The resulting paper eventually got published (Day et al., 2005), but it wasn't very influential and I wasn't very happy with it. This made me realise that I still had a lot to learn about how to handle interdisciplinary projects when the precise questions and goals weren't so clear.

On the positive side, the collaboration continued over the years, and led to a number of other publications on related matters (e.g. Ward et al., 2010; Day et al., 2012). We also decided to further consider latent-class models (which we had initially rejected), and

to develop automated software to apply cross-validation techniques to them. For this we enlisted the help of another young statistician who had worked on related software, and he developed an R package (“*crimCV*”) for our purposes and became lead author on a paper about that (Nielsen et al., 2014).

(As an aside, this collaboration also represented my first major encounter with the old “order of authors” issue. In my theoretical work, I generally stick with the mathematical tradition of listing all authors *alphabetically*, rather than by perceived order of importance. This has not always served me well: my surname comes fairly late in the alphabet, and my closest collaborator Gareth Roberts beats me alphabetically (just barely!), so my name usually appears towards the end. In one very recent example (Craiu et al., 2014), I was listed sixth out of six alphabetically, even though by order of importance I would likely have been first! Nevertheless, I hadn’t thought much about this issue until the psychologists started carefully considering who had contributed more than whom on each paper. I was not overly fond of such discussions, but I always remained flexible about where my name appears, and overall I have found this issue to be more amusing than problematic.)

In the past few months, I have worked with that same junior colleague (who is not so junior any more!) on studying the time lags between offense dates and conviction dates, about which the psychologists recently acquired new data. We obtained some preliminary results, but it is not clear how to interpret them due to problems of right-censoring (i.e., we don’t know about offenses for which convictions haven’t yet been recorded). So, that may or may not lead to another publication. In any case, this collaboration has now gone on for over a decade and will probably continue, one way or another, in the years ahead.

4 Solitary Voyages

The above interdisciplinary projects were instigated primarily by my co-authors: experts in their own subjects who enlisted my assistance with probability aspects. That has continued to be the main model of my interdisciplinary work, as it probably should be. I have, however, occasionally ventured into other areas on my own initiative.

4.1 Educational Studies: Active-Learning Strategies

Early in my teaching career, I experimented with alternative teaching arrangements, such as having students work together in small-groups during class time, write short essays about their impressions of mathematical ideas, critique each others’ work, etc. (Such practices are now quite common, but back in the early 1990s I may have been slightly ahead of

my time.) Encouraged by an educational researcher friend of mine, and by an educational practices discussion group I had been attending regularly, I decided to conduct a few simple “experiments” on my students (essentially just before-and-after anonymous surveys), and to write up the results in a short informal paper which emphasised the potential positive benefits of “active” student learning. After some revisions and struggles, I managed to get my article published in a respected education journal (Rosenthal, 1995). This publication was rather minor, but it was my first-ever non-mathematical publication.

4.2 Philosophy: Sleeping Beauty’s Frustrating Odds

On a trip to Australia, I had dinner with some old friends who were philosophy professors. During the main course, one of them suddenly turned to me and said, “Any Sleeping Beauty insights?” I had no idea what he was talking about, and was brave enough to admit it. He and his friend then explained to me the infamous *Sleeping Beauty Problem* which had apparently occupied many philosophers of science in recent years. The problem is frustratingly simple: depending on whether a fair coin flip is heads or tails, a woman (Sleeping Beauty) is woken either once or twice, without knowing which (due to induced amnesia); what probably should she assign, upon waking, to the coin being heads?

My reaction at the dinner was to say that the probability of heads is $1/2$ regardless of who is woken when. However, upon my return to Canada I realised the question was more subtle – and worst of all it could not be solved by a direct application of the usual conditional probability formulas! I then thought of a “clever” way to adjust the problem, after which the answer clearly became $1/3$. So, I wrote up my note and submitted it. A couple of philosophy journals rejected it since it turned out they had already published a number of other articles about this strange little problem, including some others which also argued for the answer $1/3$. But eventually I got my paper published (Rosenthal, 2009), thus marking the first and only “philosophical” publication of my career.

4.3 Political Science: Those Evil Conservatives

I’ve always followed Canadian electoral politics fairly closely on a personal level, generally from a centre-left perspective. I was quite opposed to the right-wing orientation of Canada’s still-current Conservative federal government, but I contented myself during the period 2006–2011 with the fact that they were a minority government and thus had some constraints against overly extreme measures. In May 2011, a new federal election was held, and I was worried that the Conservative government might achieve their long-sought majority status and thus remove those constraints.

In the weeks before the 2011 election, loads of newspaper articles and political commentators and pollsters confidently declared that the Conservatives would fall short of a majority, so I started to relax. Then, the day before the election, I happened upon a blog post which argued that a conservative majority was “a little more likely than not” (Cosh, 2011). This blog post seemed to be making sensible arguments, carefully going through the various polling numbers and scenarios. My gosh, I thought – perhaps everybody else is wrong, and this blog post is right.

Inspired (or perhaps *worried*) by this thought, I started looking over the recent opinion poll breakdowns even more closely than before, and running some simple computer models and simulations. Sure enough, I saw that one could make the case that a Conservative majority was more likely than not. And, indeed, the next day the Conservatives won their majority, taking 166 of 308 seats. How could all the pundits have gotten their predictions so wrong?

I decided to write up my notes and observations more formally, together with lots of quotations about how “sure” everyone had been that there would be no majority. I submitted the resulting article for publication 9 days after the election. It was published in a statistics journal, in fact in this very journal (Rosenthal, 2011). However, it is actually more of a political science paper – and certainly quite far removed from my usual research topics.

5 Branching Out from the Academy

As nervous as I was about working with academics outside my usual areas, I was even less comfortable with the world “out there” beyond the university setting. After all, universities are fairly insular in some ways, and have very different attitudes about research and thinking than do private businesses and the media and so on. Still, gradually over the years, I have managed to bridge this gap to some extent as well.

5.1 Consulting for an Online Casino

My first work for a non-academic was rewarding but also scary. I received a call from my department’s statistical consulting service, about a lawyer in a hurry to hire a probability expert. Although I was in the middle of a busy day, I squeezed in an appointment that afternoon. When the client arrived, he knocked loudly at my office door, and marched in aggressively. He explained that he worked for an online casino, and he seemed to look and act the part. (Yes, he wore a tough-looking black suit, and even drove a black BMW.) As further intimidation, he explained that his legal background consisted largely of tracking

down people who owed money, to see if and how they should be sued – and he even managed to use *my* name in his example (gulp).

I listened nervously as he explained his situation. His company offered an online version of the gambling game “keno”. The lawyer felt from personal observation that too many of the keno balls chosen by the computer were physically close to the number 17 on the screen.

I calmly replied – with more confidence than I actually felt – that I was indeed a probability expert, and that I could indeed investigate this issue for him. He seemed satisfied with this, and asked me what my “rate” was. I figured he must be referring to how much money I wanted to charge for my services, but I had never dealt with such issues before! I knew that consulting rates are often fairly high, so with trepidation I awkwardly suggested a figure that sounded awfully large to me. He immediately converted my figure to an hourly rate in his head – not bad for a lawyer – and quickly snapped, “Okay, I can do that.” I was a bit stunned and just stared at him, so he continued, “I’m not here to nickel-and-dime you.” I was amazed that I had just managed to negotiate a pay rate which was higher than I ever imagined getting paid for anything. (Since then I have done a fair bit of consulting, sometimes for even higher rates, but back then it was all new to me.)

We then had to arrange a formal contract (where I was *very* careful to make sure they couldn’t sue me even if I screwed up), and transfer data for me to analyse, and so on. After that, the work proceeded fairly smoothly. I did statistical analysis of one million simulated draws from their keno program, and found (as I suspected) that the lawyer’s fears were unfounded and there was no statistical significant deviation from uniformity. I also did some other work for them, the most dramatic of which was analysing the rules of one of their complicated online poker games and proving, as the lawyer suspected, that it was actually possible for a clever customer to win on average (which meant that their Las Vegas contacts, who had designed the game, had screwed up). Later on they had me do some additional work, studying a particular customer who was consistently winning at one gambling game in clear defiance of the odds (I determined that it must be some sort of “inside job” or computer security problem; they reacted by banning that customer from gambling on their site).

Overall I found the work fairly interesting, not to mention lucrative (though due to my ethical qualms about online gambling, I donated part of my fee to an addiction research foundation). I never did totally relax around that lawyer, though in the end I didn’t work with him too closely – I worked more with a technical computer guy in their company whom I quite liked. One practical lesson I learned was that my reports shouldn’t be too technical, or at least should begin with a non-technical “executive summary”: my first report to them contained lots of equations and mathematical arguments, and the lawyer responded by dismissively saying that he couldn’t understand it. More importantly, I learned to my

surprise that I could succeed with consulting work in a non-academic setting, thus giving me more confidence to take on other such work in the future.

5.2 Detecting Lottery Fraud

At the urging of my journalist in-laws, I ended up writing a book about probabilities in everyday life, for general readers (Rosenthal, 2005a). To everyone’s surprise – especially mine – the book was quite successful, making some bestseller lists and getting translated into nine other languages. This in turn led to my giving lots of media interviews and public lectures, and meeting all sorts of different people, and gaining a somewhat higher public profile (Rosenthal, 2005b). So, this was a very interesting development in my life, though not really “interdisciplinary work” *per se*.

Unexpectedly, my higher public profile in turn led to new work opportunities, as different sorts of people sought me out. The most dramatic such opportunity started out rather mundane: yet another call from the media with yet another question about lotteries (the bane of every public probabilist). However, this question was more interesting than most. Specifically, an investigative television program had learned that one Ontario customer was defrauded out of his large lottery prize by an unscrupulous retail seller, who managed to keep the customer’s winning ticket and cash it in herself. The television people suspected there were other similar cases out there, and they wanted me to analyse some retailer win data to see if a case could be made.

Regarding this as “yet another media enquiry”, I did a fairly quick and simplistic analysis, and wrote up a report for them. They videotaped a brief interview of me discussing my findings, and I figured that was that. They later obtained some new information, and recorded a fresh interview with me. Finally the program went to air. To my surprise, it created a *huge* media sensation as lottery customers across Canada worried that they too had been defrauded out of lottery winnings. (Ironically, I myself had nothing to worry about, since I never buy lottery tickets!) Various provincial governments and police forces then conducted investigations, eventually leading to legislative debate, the firing of two CEOs, several criminal charges, jail time, and payouts to defrauded customers totalling over twenty million dollars. I was later hired by some provincial governments looking to better oversee their own retailer lottery wins – so the story was not only dramatic front-page material, but it also ultimately led to some paid work for me. In any case, it had so many twists and turns that I later wrote up all the details, including links to loads of media reports, in a separate article (Rosenthal, 2014a).

5.3 Optimising Cell Phone Plans

A year or so later, I was contacted by an executive at a marketing company. He had seen me in the media and read my book and was apparently a fan, and he wanted to hire me to do some analysis for him. By this point I was reasonably confident about working as a consultant, so I happily agreed to meet him for lunch. He enthusiastically explained his work, and we agreed to terms, and he said he would send me a contract and retainer the next day. Then I heard nothing for several days, so I assumed the deal was off. However, it turned out he was just busy and disorganised, and he did eventually send me the materials for me to get to work. (He also purchased 20 copies of my book to share with his clients, so I liked him immediately.)

I ended up working on several different projects for the executive, but the main one involved analysis of cell phone customer plans. (This was again ironic, since up to that point I had never owned a cell phone!) The executive, together with his clients at a major cell phone company, explained that customers have a huge array of choices in terms of what cell phone billing plans they select: how many free minutes, what charge for texting, whether incoming calls are free, whether long-distance calls are included, whether weekend calls are charged, etc. Depending on the customer's usage patterns, some plan choices could be much cheaper than others, but it was hard for customers to choose. The clients wanted to help the customers make better choices and thus reduce their bills – in an effort to dissuade them from deciding to change cell phone providers.

This sounded straightforward enough, and I immediately started to imagine big computer “for” loops which would cycle through all the various cell phone plan choices, compute the customer's cost under each plan, and find the minimum. I even started to contemplate further statistical analysis about how the customer's usage varies from month to month, so I could take such monthly variances into account when optimising. This would be exciting!

Unfortunately, complications soon arose. For one thing, the cell phone plans were very complicated, and poorly explained, and frequently changed, and some customers were “grandfathered in” to keep using old plans even after they were discontinued, and so on, leading to a myriad of different and confusing options. So, those “simple for loops” turned out not to be so simple after all. But worse than that, the internal billing information at the cell phone company was run by a different division, and we couldn't access it! Instead, we had to proceed solely using pdf files of the actual bills (that would actually be mailed to the customer). So, my biggest challenge turned out to be parsing poorly-formatted pdf files to extract all of the customer's cell phone usage, and figure out how many minutes were long-distance and weekends and incoming calls and so on, so that I could then compute what

their cost would have been under the various plans and options available to them. The task got more and more complicated, and sometimes the computations from the pdf bills didn't quite match up with the total they were charged.

Nevertheless, I carried on and worked very hard on what was becoming more of a computer programming assignment than statistical analysis. I wrote and tested complicated C programs which finally managed to parse the bills fairly well. And I got paid well for each hour spent. However, as I proceeded, I worried more and more that not only were these pdf bills awkward for me, but they would be awkward for the customer too: To use our optimiser, a customer would first have to download their bill from the cell phone company's web site, then upload it to our separate optimiser web page, then receive a recommendation about how to modify their cell plan options, and then return to the cell phone web site to actually make the changes. The clients explained to me that this awkward situation was unavoidable at first – though if our product was a success then they might later be able to convince their billing department to integrate the two sites together.

Even after my computer program was ready to go, there were numerous additional delays as the cell phone company designers fussed over the precise “look and feel” that the optimiser web page would have. Eventually the product was launched and announced to select customers, and even received a few write-ups in industry blog postings. Unfortunately, the excitement was short-lived: that awkwardness of use that I had worried about earlier did indeed severely limit the number of customers who bothered to use the optimiser product. Ultimately, rather than maintain and update the optimiser, the cell phone company decided to scrap it. I still got paid, but all my hard work ended up on the cutting room floor.

6 You Can't Win Them All

Naturally, even academic collaborations do not always conclude successfully. In fact, in some cases they fail spectacularly. Two examples come to mind.

6.1 Sociology: A Markovian Roller-Coaster

In an effort to broaden my horizons, I had started attending a weekly informal interdisciplinary lunchtime seminar at one of the “colleges” within my university. The participants were mostly humanities and social science researchers; I was the token scientist who was invited to join after the publicity from my book. One week the speaker was a sociologist, and we chatted over lunch before his talk. When he heard I was from the statistics department, he asked me if by any chance I knew anyone who knew anything about Markov chains.

I confess that my first reaction was to feel *insulted* that he hadn't already heard of me. After all, Markov chains were my main research area, and I had received a fair bit of recent publicity from my book and media work, and I was a regular participant in the seminar, and he was supposedly searching for a Markov chain expert – how could he not know who I was? Fortunately, I managed to put such thoughts out of my mind, and ask him what he wanted to know.

It turned out that he was studying gambling addiction, and had data from questionnaires assessing the gambler's attitudes and opinions. He wanted to model attitude changes using simple finite state space Markov chains, and needed assistance with that. This seemed like a perfect project for me, since it involved an established researcher, an important topic, a gambling connection, and even Markov chain analysis! So, I happily agreed to get involved, and over the next few days he sent me various data and documents to get started.

Unfortunately, the project fell apart almost as soon as it ramped up. For one thing, it seemed that he didn't really have longitudinal data over time, rather he was going to pretend that the second-highest responses represented a previous state – which didn't make sense to me. Also, he boldly declared that he doesn't like writing articles, only books, which seemed to indicate that any payoff from this project would only come well down the road if at all. But most bizarrely, when I mentioned that a paper from my young offender analysis (discussed above) had a small connection to one aspect of his project, he read that paper and proceeded to *tell me all the ways in which he didn't like it*. In short, it quickly started to seem like working on this project would be “difficult” in many ways.

So, I e-mailed the sociology professor to tell him that I would rather not continue with the project. What happened next was the strangest of all. He replied with a long impassioned plea for me to reconsider, saying that he was “disappointed” and had “really looked forward” to working with me. He apologised for his harsh critique of my previous paper. He acknowledged my concerns about his Markov model. And he allowed that papers might result from our collaboration after all. He closed by asking me to “re-open our discussion of collaboration”, and to accept his apology for his previous e-mail.

I was touched by his plea, and – despite my misgivings about the project, plus being very busy with other work – I quickly replied to say that I would consider his request further and would reply again soon. Then, to my shock, I received another e-mail from him *later that same day*. He wrote that he had just been asked to revise one of his textbooks, so he would have no time to work on this latest project after all. In short, he had pleaded so hard for me not to quit, just so that he could fire me 9 hours later. Thus ended my sociology career.

6.2 Psychiatry: Wasting Time with fMRI's

One evening at an extended family dinner gathering, I had a friendly chat with a psychiatry professor at Columbia University who was actually my second cousin. He later e-mailed me asking for some statistical advice about regressions with correlated coefficients, and I managed (with help from a colleague) to provide a useful response.

Then, two years later, I heard from him again. This time, he was seeking help analysing fMRI measurements. He and his colleagues were measuring blood flows in different regions of the brains of clinically depressed patients, and of healthy control patients, while performing different tasks. They used these blood flow measurements to generate network graphs based on the connections between the different brain regions. They wanted to analyse these network graphs statistically, to better understand the difference between depressed and healthy patients, with the goal of better treating clinical depression symptoms.

The psychiatry professor asked me to recommend an expert on network flows for him to contact. But no perfect person came to my mind, and meanwhile I was feeling somewhat bold. I had done a bit of research on networks in the context of World Wide Web searches (Borodin et al., 2001, 2005), and I was looking for new projects and topics and collaborators. Furthermore, this seemed like the perfect context: a respected researcher at a leading university with interesting data and fundamental questions about the nature of depression – with the possibility of improving treatments! So, I boldly suggested that I join the project myself, and he agreed.

Next, he sent me lots of large files of patient fMRI data, together with some preliminary visual analysis indicating that the depressed patients tended to have much denser networks than the controls. I busily and excitedly set to work writing C programs to read and parse the data, and examine various connections and networks and blood flow values, and compare depressed to control patients, and so on.

After many days in front of the computer, I started to get results. However, they weren't what I expected! Instead of seeing much denser networks in the depressed patients, I saw very little difference, with the depressed patients' networks actually being slightly *less* dense. What was going on here? I spent more precious hours and days checking my code and verifying my results, but they didn't change.

Then came the awkward clarifications with the psychiatry professor. I told him that my results weren't agreeing with his preliminary analysis. He replied that fMRI data needs a lot of pre-processing, and his earlier preliminary analysis had been done on less-processed versions of the data and should not be relied on. He agreed that the effect he had originally observed might have disappeared once the data was processed correctly – as I had discovered.

He told me to wait while he consulted further with his colleagues.

So, I waited. I heard nothing further for two months, after which I e-mailed him again. He then explained that he had ended up using a “different (and better) method for de-noising the data”, and the resulting statistical analysis “is pretty straightforward, so I don’t know that it would be very interesting” for me to remain involved. In short, I was being declared redundant. I replied that I would still be happy to continue to work on the analysis if he wanted, and he said he would keep me posted. I never heard from him again.

7 New Frontiers

In recent years, I have engaged in a number of new interdisciplinary projects, in various different areas.

7.1 Climate Change in Pre-Industrial Iceland

One summer, I arranged to have lunch with the economics professor who had been the third author on my earlier Nash equilibrium project discussed above. I expected a mere social chat, but he launched straight into a proposal for a new collaboration: the effect of climate on the historic population of Iceland!

He explained that he had visited Iceland the previous year, and had gone to a museum where he saw a large graph of the yearly population of Iceland going all the way back to the Middle Ages. He was impressed with Iceland’s national record-keeping, and wanted to somehow make use of this data. Meanwhile, he knew that historic temperatures, while not recorded directly, can be estimated from such proxies as ice cores, tree rings, ocean sediments, etc. His idea was to combine these two data sources together, to study the effect that temperature changes had on the population of Iceland before the rise of modern technology. His hypothesis was that the colder Iceland got, the more deaths and fewer births would result. He even suggested that our study could have some implications for the consequences of modern climate change – though that aspect was more subtle since the modern climate change threat comes from excessive *warming* rather than cooling.

This was all very far removed from anything I had ever worked on before, but it seemed like an interesting project so I agreed to join in (and I soon enlisted two MSc project students to assist). The next step was to start tracking down some data. We managed to find the source of that Middle Ages graph in a research paper (written in Icelandic!), but apparently it was just a loose estimate based on numerous general assumptions, so it wasn’t sufficiently accurate on a year-to-year basis to be useful to us: the graph which had inspired our project

was actually useless! Nevertheless, we soldiered on, and learned that Iceland did record careful annual population counts from 1735 onwards, giving us about 150 years' worth of data before the modern industrial age. Good enough to press forward.

The temperature estimates were more subtle. Actual temperatures were only recorded in Iceland from the late 1800's onwards, so we indeed needed to use proxy measurements, and ice cores seemed most promising. We couldn't find data for any ice cores in Iceland, but we did find five in nearby Greenland – good enough. We averaged them year-by-year, used the later recorded temperatures to calibrate them, and voila – we had the annual temperature estimates we required.

Then came the moment of truth: were the population count annual changes actually affected by temperature changes? I had maintained a healthy skepticism throughout the project, thinking that we probably wouldn't find any effect – especially given the small changes and large uncertainties in our annual temperature estimates. So I was amazed and delighted when clear effects *did* emerge: numerous regressions confirmed that temperature changes in previous years did indeed have a significant effect on Iceland population counts! This effect remained even as we modified such parameters as the effect time lag, what year ranges we included, how we calibrated the temperature estimates, what other covariates were or were not controlled for, etc. We also managed to estimate secondary effects like the extent to which populations “adapted” to increasing cold when it persisted for a number of years in a row. The results seemed more and more reliable – backed up by pages and pages of robustness regression tests – and my skepticism slowly melted away. We had done it!

We eventually got our results published in a prestigious economics conference proceedings (Turner et al., 2012a), and also in a statistics journal (Turner et al., 2012b). They didn't have as much impact as I might have anticipated, perhaps because the connection to modern climate change challenges was too tenuous. But I was still pleased that I had participated in a research project related to climate change in pre-industrial Iceland – something I would never have anticipated.

7.2 Legal Studies: Judging the Judges

Several years ago, I was approached by a law professor who was also the son-in-law of a recently retired statistics colleague of mine. He wanted to investigate the extent to which judges get their junior law clerks to help them write their official legal judgments – a closely guarded secret. He wondered if quantitative text analysis (i.e., “stylometry”) on the published judgments could shed light on this question, similar to previous analyses of authorship of Shakespeare's plays and of the Federalist Papers. This was rather different

from any question I had ever worked on, but it sounded interesting, so I bravely offered to get involved.

We quickly agreed that one useful measure is the judgments' frequency of use of various *function words* (i.e., words like *all*, *have*, *not*, *than*, etc.), which tend to reflect writing style rather than content. After further discussions, we agreed that judges who rely more on law clerks should have higher *variability* in these frequencies, since their judgments are actually being written by many different clerks rather than always by the same judge. So, I wrote some computer programs to automatically download the U.S. Supreme Court legal decisions from the web, and then compute these usage frequency variabilities for each judge.

Our results were fairly promising, in that some of the judges who were rumoured to rely more heavily on their law clerks did indeed have larger variability scores. On the other hand, there was at least one judge who was generally believed *not* to rely on his clerks but still had a high score – meaning that either this general belief or our methodology was faulty (we couldn't determine which). Anyway, I was happy enough with our results, so we wrote up a paper (with the law professor writing the legal bits, and me writing the technical bits), and submitted it to *JASA*. To my surprise, they rejected it quickly and firmly, saying it might be more appropriate for an educational journal like *Chance*. That annoyed me at the time, but made its later acceptance in the *Annals of Applied Statistics* (Rosenthal and Yoon, 2011a) all the more sweet. A follow-up paper was published in the *Cornell Law Review* (Rosenthal and Yoon, 2011b), and was mentioned in the *New York Times* (Liptak, 2010). Later related papers followed about the Canadian Supreme Court (Bodwin et al., 2013a), and about one specific judge who kindly told us which of his judgments his clerks had written (Bodwin et al., 2013b). All in all, quite a satisfying and successful collaboration.

7.3 Educational Policy: How GPAs Discourage Math

I later collaborated with the same law professor on a different topic, this time related to educational policy. Specifically, the law professor was worried that university students often avoid certain subjects which tend to give low grades (e.g. mathematics), even if they wish to – and perhaps should – study them. This raises the question of how raw GPA values can be appropriately “adjusted” to take into account the difficulty of courses taken. The problem, of course, is that it is hard to determine how “difficult” a course is: even if the grades are low, this could simply be because weaker students tend to take it. So, what to do?

The law professor argued that these different effects could be untangled by examining cases of *overlap*. For example, suppose a math student and an English student both take a sociology class, and the math student gets a higher grade. This may provide some evidence

that the math student is stronger overall, and thus their math grades should be adjusted upwards relative to the English student's English grades. Such a comparison is rather problematic and involves a host of uncontrolled variables, but hopefully those effects mostly average out when thousands of students are considered.

We were fortunate that my university agreed to share with us ten years' worth of undergraduate grades – a total of 1.66 million different grades. We also received my university's law school admission and first-year-grades data, which we could regard as outcome measures (at least for those of our students who applied to our law school). In the end we had a huge amount of data, which was very complicated to process, requiring another big C program.

My initial effort involved using each overlapping grades pair to “tweak” the GPAs, but those results were inconclusive. The law professor then contacted a political science professor friend of his, who did a more systematic Bayesian analysis using MCMC (I should have thought of that myself!), and got more promising results. This eventually provided appropriate adjustments to GPAs which somewhat more accurately predicted future success (in first year law school, at least). This project ended up being quite a long slog, but it did lead to a paper which is currently under revision and will hopefully be published in due course (Bailey et al., 2014).

7.4 Speech Pathology: Overlapping Tongues

Another recent interdisciplinary collaboration has also been fruitful. It began several years ago when I met an enthusiastic young speech pathology professor at a lecture. We started chatting, and upon hearing that I was a statistician she said there were many statistical issues that arise in her work and she would like to collaborate. I was pleased and handed her my business card, but I heard nothing so I forgot all about it. Then, a full *two years later*, she e-mailed me about a project.

She explained that she was studying subjects' tongue positions when producing different sounds, with an eye towards helping ALS patients regain their lost speech abilities. She had tongue position data for many subjects, producing many different sounds, obtained by placing electronic sensors on subjects' tongues (!), and wanted me to help analyse it. Fortunately, I had studied a bit of linguistics as an undergraduate, and remembered a bit about sound production (e.g. for 'th' the tongue touches the teeth, for 'y' the tongue touches the hard [middle] palate, for 'k' the tongue touches the soft [back] palate, etc.). But the speech pathology professor wanted to refine this picture further, and in particular determine the extent to which different subjects do or do not articulate the same sounds in the same positions, and do or do not articulate different sounds in different positions.

This ended up being a fun, if intense, project. (The speech pathology professor was passionate by nature, plus she was coming up for tenure the following year, giving our collaboration a slightly frantic feeling.) It wasn't clear how to proceed, so I had a lot of freedom to try different approaches. I ended up deciding to form *density estimates* for tongue locations for different subjects, and then estimate the *overlaps* (i.e., integrals of the minima) of different pairs of densities for different sounds or different subjects. This required some computer programming but wasn't too difficult, and after many refinements it led to conclusions that the speech pathology professor was happy with. Our results were published in a leading acoustics journal which, best of all, had those same famous initials *JASA* (Yunusova et al., 2012).

(As an aside, this paper provided another interesting order-of-authors issue. In speech pathology, alphabetical order is never used. But in addition to the usual order-of-importance tradition, there is also a tradition of listing the most senior collaborator last as the “lab director”. In this case, the speech pathology professor would be first author, and she asked me if I preferred to be second author or last author. I replied second author, which pleased her since she could then honour a different collaborator – who was involved with earlier efforts to collect the data, but less involved with the current project – as last author. Tricky!)

After a brief respite, the speech pathology professor contacted me again about a different project. This time, she was studying how ALS patients move from a “non-bulbar” phase where their speech is normal, to a “bulbar” phase where their speech is seriously impaired. She wanted to find methods for early detection of the change, using other related speech measurements such as how quickly the patients say certain phrases, how often they pause while speaking, how much air flows through their nose while they speak, and so on. This project ended up involving various regressions and correlations and variable selections (some of which we did using overlap measures as in our previous project, for old time's sake). It never produced very clean results, largely due to insufficiently large data sets. But we managed to complete a preliminary analysis for a conference proceedings (Yunusova et al., 2014), and will hopefully continue with a more complete analysis – perhaps with additional data – in the near future. As with many collaborations, it is never really completely finished.

8 Interdisciplinary Reflections

Although my main research area continues to be mathematical probability theory applied to MCMC algorithms, I am pleased that over the years I have done more and more interdisciplinary work. So what have I learned from all of this?

One observation is that it is not easy to “prepare” for interdisciplinary collaborations,

since the topics and required skills vary so much from one project to the next. However, one constant seems to be that a variety of computer skills – including the ability and inclination to deal with different file formats and processing utilities and programming languages and so on – often come in handy. (Other knowledge and skills can become useful in unexpected ways, e.g. my linguistics studies for the speech pathology project, but that is hard to anticipate.) Beyond that, it is very important to listen carefully to your collaborators and other researchers in the discipline, and make a sincere effort to understand what they are interested in. I have also found that a strange combination of humility (“What does that mean?”) and over-confidence (“Sure, I can figure that out!”) is required to join the projects you want while still acquiring the information you need.

Another observation is that I have not developed any systematic way of *finding* new interdisciplinary contacts. I have never worked in a statistical consulting service, nor organised an interdisciplinary forum, nor otherwise sought out interdisciplinary connections in any organised fashion (though perhaps I should). Instead, my interdisciplinary work has all arisen either from my own special interests (active-learning strategies, Canadian politics), or personal contacts (professor friend, colleague’s son-in-law, friend’s wife’s colleague), or chance meetings at talks (sociologist, speech pathologist), or being contacted based on my reputation (lottery fraud investigators, cell phone marketer). I have been quite fortunate in this way, but hopefully similar opportunities can arise for any active statistical researcher: if you build your statistical career, then interdisciplinary contacts will come.

Of course, it is also important to be *open* to such opportunities when they do arise. This means that if you are at a party and get into an interesting chat with a researcher in an interesting subject, then you might want to turn the conversation over to ways that statistics might help. If you go to a talk and are intrigued, then try to find out more and see if perhaps you can get involved. If you have a particular interest in an important topic, consider taking your interest further into the research realm. And if someone leaves you a phone message asking for statistical advice, be sure to return their call. Such contacts won’t always lead to a collaborative research project, but every once in a while they will.

On the flip side, it is also important (and I have sometimes failed at this) to be *selective* about which projects you will or will not pursue. I have found that my interdisciplinary projects have been most successful when they involve collaborators who are leading researchers in their field, with clearly-defined questions about an interesting and relevant topic, a strong desire for and appreciation of statistical assistance, and the patience and kindness and communication skills to help me learn what I need to join in. Now, not every collaboration will meet this standard, and lots of flexibility is required too. But it is okay to walk away from a project if it doesn’t seem promising, or if the collaborators are not

providing the assistance and context you require.

My interdisciplinary work has sometimes made me feel like a fraud. Indeed, almost by definition, working in other areas means working on topics in which you are not an expert. In some cases, I didn't know even the most basic and simple facts about a subject on which I was co-authoring a research paper. Once again, it seems that a combination of humility and over-confidence is required – that fine balance between standing by your own statistical contributions, while pleading ignorance about certain other aspects of the project. It usually works out okay, but at times it has made me feel insecure – and glad that I have tenure.

Related to this, I have found – sometimes to my disappointment – that none of my interdisciplinary work has led to a “second career” in which I became a true expert on another topic. That is, my interdisciplinary efforts have typically resulted in just a single joint paper, or perhaps two or three papers, before drying up – because there is little more I can contribute to the topic, or the research is becoming less successful and satisfying, or my collaborator moves in a different direction, and so on. So, despite all of my interdisciplinary projects, there isn't really any subject outside of my core (i.e. math / statistics / computer science) in which I can claim to be any kind of “expert”.

Despite these caveats, I have found interdisciplinary work to be very rewarding and stimulating. It has allowed me to learn about new areas, meet new researchers, think in new ways, and publish in new journals. Hopefully it does the same for you, too.

References

M.A. Bailey, J.S. Rosenthal, and A.H. Yoon (2014), “Grades and Incentives: Assessing Competing GPA Measures and Post-Graduate Outcomes”. Preprint. Available at: <http://probability.ca/jeff/research.html>

K. Bodwin, J.S. Rosenthal, and A.H. Yoon (2013a), “Opinion Writing and Authorship on the Supreme Court of Canada”. *University of Toronto Law Journal* **63**, 159–192.

K. Bodwin, J.S. Rosenthal, and A.H. Yoon (2013b), “A Statistical Approach to Judicial Authorship: A Case Study of Judge Easterbrook”. *Advances and Applications in Statistics* **37**(2), 123–148.

A. Borodin, G.O. Roberts, J.S. Rosenthal, and P. Tsaparas (2001), “Finding Authorities and Hubs From Link Structures on the World Wide Web”. *Proceedings of the Tenth International World Wide Web Conference*. (Chosen as *Selected Area Highlight* for Hypermedia.)

A. Borodin, G.O. Roberts, J.S. Rosenthal, and P. Tsaparas (2005), “Link Analysis Ranking: Algorithms, Theory, and Experiments”. *ACM Transactions on Internet Technology* **5**(1), 231–297.

- C. Cosh (2011), Blog post of May 1, 2011. Available at:
<http://www2.macleans.ca/2011/05/01/election-forecast-dance-on-a-razors-edge/>
- R.V. Craiu, L. Gray, K. Latuszynski, N. Madras, G.O. Roberts, and J.S. Rosenthal (2014), “Stability of Adversarial Markov Chains, with an Application to Adaptive MCMC Algorithms”. Submitted.
- R.V. Craiu and J.S. Rosenthal (2014), “Bayesian Computation via Markov chain Monte Carlo”. *Annual Review of Statistics and Its Application* **1**, 179–201.
- D.M. Day, I. Bevc, T. Duchesne, J.S. Rosenthal, L. Rossman, and F. Theodor (2007), “Comparison of Adult Offense Prediction Methods based on Juvenile Offense Trajectories using Cross-Validation”. *Advances and Applications in Statistics* **7**, 1–46.
- D.M. Day, J.D. Nielsen, A.K. Ward, Y. Sun, J.S. Rosenthal, T. Duchesne, I. Bevc, and L. Rossman (2012), “Long-term Follow-up of the Criminal Activity of Adjudicated Youth in Ontario”. *Canadian Journal of Criminology and Criminal Justice* **54(4)**, 378–413.
- M.J. Gordon and J.S. Rosenthal (2003), “Capitalism’s Growth Imperative”. *Cambridge Journal of Economics* **27**, 25–48.
- R.B. Israel, J.S. Rosenthal, and J.Z. Wei (2000), “Finding generators for Markov chains via empirical transition matrices, with applications to credit ratings”. *Mathematical Finance* **11**, 245–265.
- A. Liptak (2010), “Justices Are Long on Words but Short on Guidance”. *New York Times*, Nov. 17, 2010. Available at:
<http://www.nytimes.com/2010/11/18/us/18rulings.html>
- J.D. Nielsen, J.S. Rosenthal, Y. Sun, D.M. Day, I. Bevc, and T. Duchesne (2014), “Group-based Criminal Trajectory Analysis using Cross-Validation Criteria”. *Communications in Statistics – Theory and Methods*, to appear.
- M.J. Osborne, J.S. Rosenthal, and M.A. Turner (2000), “Meetings with costly participation”. *American Economic Review* **90**, 927–943.
- G.O. Roberts and J.S. Rosenthal (2001), “Optimal scaling for various Metropolis-Hastings algorithms”. *Statistical Science* **16**, 351–367.
- G.O. Roberts and J.S. Rosenthal (2004), “General state space Markov chains and MCMC algorithms”. *Probability Surveys* **1**, 20–71.
- J.S. Rosenthal (1995), “Active-Learning Strategies in Advanced Mathematics Classes”. *Studies in Higher Education*, Vol. **20** No. **2**, 223–228.
- J.S. Rosenthal (2005a), “Struck by Lightning: The Curious World of Probabilities” (book for the general public). HarperCollins Canada.
- J.S. Rosenthal (2005b), Web page about the book Rosenthal (2005a). <http://probability.ca/sbl/>

J.S. Rosenthal (2009), “A mathematical analysis of the Sleeping Beauty problem”. *The Mathematical Intelligencer* **31(3)**, 32–37.

J.S. Rosenthal (2011), “Was the Conservative majority predictable?” *Canadian Journal of Statistics* **39(4)**, 721–733.

J.S. Rosenthal (2014a), “Statistics and the Ontario Lottery Retailer Scandal”. *CHANCE Magazine* **27(1)** (February 2014). Available at: <http://probability.ca/lotteryscandal/>

J.S. Rosenthal (2014b), “Optimising and Adapting the Metropolis Algorithm”. Chapter 6 of the SSC volume “Statistics in Action: A Canadian Outlook” (J.F. Lawless, ed.), pp. 93–108. Chapman & Hall / CRC.

J.S. Rosenthal (2014c), “Lessons From a Twisted Career Path”. Chapter 11 of the COPSS award winners volume “Past, Present, and Future of Statistical Science” (X. Lin, C. Genest, D.L. Banks, G. Molenberghs, D.W. Scott, and J.-L. Wang, eds.), 117–128, 2014. CRC Press.

J.S. Rosenthal and A.H. Yoon (2011a), “Detecting Multiple Authorship of United States Supreme Court Legal Decisions Using Function Words.” *Annals of Applied Statistics* **5(1)**, 283–308.

J.S. Rosenthal and A.H. Yoon (2011b), “Judicial Ghostwriting: Authorship on the U.S. Supreme Court”. *Cornell Law Review* **96**, 1307–1343.

M.A. Turner, J.S. Rosenthal, J. Chen, and C. Hao (2012a), “Adaptation to Climate Change in Preindustrial Iceland”. *American Economic Review: Papers & Proceedings* **102(3)**, 1–8.

M.A. Turner, J.S. Rosenthal, J. Chen, and C. Hao (2012b), “Adaption to Climate Change: Evidence from 18th and 19th Century Iceland”. *Advances and Applications in Statistics* **28(2)**, 77–108.

A.K. Ward, D.M. Day, I. Bevc, Y. Sun, J.S. Rosenthal, and T. Duchesne (2010), “Criminal Trajectories and Risk Factors in a Canadian Sample of Offenders”. *Criminal Justice and Behavior* **37(11)**, 1278–1300.

Y. Yunusova, J.S. Rosenthal, K. Rudy, M. Baljko, and J. Daskalogiannakis (2012), “Place of articulation for lingual consonants defined using electromagnetic articulography”. *Journal of the Acoustical Society of America* **132(2)**, 1027–1038.

Y. Yunusova, J.S. Rosenthal, J.R. Green, S.Shellikeri, J. Wang, and L. Zinman (2014), “Detection of Bulbar ALS using a comprehensive speech assessment battery”. Proceedings of MAVEDA 2013 – the 8th International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications, 2014.