Proceedings of the 2023 Winter Simulation Conference
C. G. Corlu, S. R. Hunter, H. Lam, B. S. Onggo, J. Shortle, and B. Biller, eds.

PRACTICAL IMPACT AND ACADEMIA ARE NOT ANTONYMS

Shane G. Henderson

School of Operations Research and Information Engineering
233 Rhodes Hall, Cornell University
Ithaca, NY 14850, USA

ABSTRACT

This tutorial discusses principles and strategies for the interplay between applied work with organizations and an academic research agenda. I emphasize lessons I have learned through my own work and my own mistakes, with special focus on some high-stakes settings, including advising Cornell University's response to the COVID-19 pandemic and work with the emergency services, among other applications.

1 INTRODUCTION

The purpose of this article is to encourage more academics to tackle practical problems in concert with real organizations. My intended audience is the analysis methodology community in simulation, although most of this article is, I think, relevant for the operations research community. Drawing on my own experiences for context, I emphasize why tackling real problems is beneficial, how to get started, strategies for success and pitfalls to avoid. I also want to offer some lessons I have learned from working on some particularly high-stakes projects. Practical work is rewarding in and of itself, inspires important theoretical questions, and helps ensure that your research profile is relevant and interesting.

I do not mean to imply that *all* academics should adopt a research agenda of this form. Many academics whom I admire produce highly influential and important work without ever working directly with real organizations. Such individuals have a knack for determining important new directions and should keep up the good work! Still, I think many researchers struggle to identify important and engaging research directions, and I believe that such researchers might benefit from having practical work as a part of their research portfolio. If more researchers were willing to engage directly on practical problems in concert with real organizations, I believe that our collective research agenda could be significantly more exciting, interesting and impactful.

In this article I promote the following perspectives.

1. There is great professional satisfaction in working on practical problems.
2. Practical problems have helped me determine research directions that I perceive to be important and exciting, and are a constant source of research and teaching inspiration.
3. It is important to maintain a *portfolio* of projects, rather than a single project, partly due to the inherent risk in research where not all ideas work out, but also because one's contacts in organizations are not static - contacts can dissolve.
4. Whether you are considering taking on a project, in the midst of the project and trying to keep focused on the right questions or writing up your results at the conclusion of the project, Heilmeier's catechism (DARPA 2023) is valuable in focusing your thinking on questions that matter.
5. There will be questions that you cannot answer within your expertise; in such cases do not engage.
6. Problems that are of central importance in practice *and* that can be solved with exciting new methodological ideas on a timeframe that is relevant to practice are rare, perhaps exceptionally rare. It is still possible to have both of these qualities in one's work if one "separates the time scales."

7. I do not believe in *the* model when tackling real questions. Rather, I recommend using a collection of models that vary in detail and accuracy. A more complex model is not necessarily better, even if, in principle, it is a better match to reality, because it may rely on parameters you cannot estimate with sufficient accuracy to make quality decisions, in addition to the more obvious difficulties with verifying and maintaining such models.
8. Don't view yourself as a simulation person. Instead, be a problem solver who can bring a range of techniques to bear.
9. For particularly time-urgent situations, you must be responsive.
10. For high-stakes situations, be very careful with media you create, including files on your computer, email and Slack messages.
11. For high-stakes situations, be very careful what you say, and how you say it, in public.
12. Be lucky in opportunities and collaborators!

Most, and perhaps all, of these perspectives are not new. For example, the tutorials Sadowski (2007) and Hugan (2014) discuss closely related issues in working on simulation projects in industry. Also, Nelson (2016) provides perspectives that are important and helpful in technology transfer and that are relevant to the present topic. In an interview with Peter Horner (Horner 2015), Ed Kaplan encourages operations researchers to engage with important practical problems. The most fundamental principle of the Cornell Tech campus in New York city is that the work of its researchers must have a significant component of *engagement*, meaning a back and forth fertilization between practice and theory. Furthermore, much of the work I see coming out of MIT's OR Center has a practical inspiration and orientation, which was also the case at the Department of Engineering Science at the University of Auckland when I was a faculty member there from 1997–1999.

Section 2 discusses a number of projects from my own experience, highlighting aspects that are directly relevant to the present discussion. Section 3 returns to the central themes of this paper, with reference to the projects from Section 2, but organized to answer the key questions of *why* to engage with practical problems, *how* to identify and engage with practical problems, important practices *during* engagement with practical problems, considerations after a project has concluded and special comments for high-stakes situations. Section 4 contains final remarks.

2 SOME PROJECT EXPERIENCES

2.1 Early Missteps

One of my earliest projects was with Carter-Holt Harvey (CHH), when I was a lecturer (the equivalent of an Assistant Professor in the U.S. system) at the University of Auckland in New Zealand. The company CHH wanted to perform a risk assessment related to their costs for wood chip supply to their factory for making medium-density fibreboard. They were able to source wood chip from various suppliers, but each supplier's costs and potential supply was unknown. These quantities depended on a range of factors including whether a "windthrow" event would arise, where strong wind blows down a number of trees that are not otherwise ready to be milled.

After a site visit, I proceeded to build a stochastic linear programming model of the situation. I felt that the model needed optimization due to the presence of decisions after costs had become known. I stewed away in my office, trying to capture the most important aspects. In the end, the model I delivered was too complex for CHH and was not adopted.

I met with my primary contact at CHH at a later date. He, instead, used a rather simple spreadsheet-based simulation model that sampled costs and supplies from some distributions that were deemed plausible, then assumed all supply would be obtained from the cheapest supplier to the most expensive. The output of the model was a distribution on the cost of supply, for varying levels of desired supply. The model was simple to explain, captured the primary uncertainties, and was straightforward to develop and maintain.

The lessons I learned from this experience were 1) stick close to the sponsor and 2) start with the simplest model you can think of that still addresses the question at hand. I did not walk the sponsor through my initial model ideas. Had I done that, I would have learned early on that my ideas were too complicated and I wouldn't have spent pointless hours in uncertainty of what I was doing.

A second experience was also a bit of a blunder on my part, yet worked out well for me in the long run. The Police Communications Center (PCC) in Auckland, New Zealand handles the equivalent of North American 911 calls. Call takers receive the calls, determine the details of the call, and pass along those calls that require police response to dispatchers. The PCC asked that I help them schedule shifts for call takers that ensured very low call holding times at minimal cost. This is essentially a standard call-center staffing problem, and, with the help of a student, I performed a queueing analysis using estimates of call arrival rates, call handle times and so forth from data the PCC provided. The results were received in somewhat muted fashion, with the person in charge of the call center thinking our estimates of needed agent numbers were too high. I simply assumed the PCC was presently not sufficiently staffed and went on my merry way.

Later, I learned that one of the call center managers had noticed that average call handle times differed from one calltaker to another. They dug a little and found that some agents occasionally left a call open upon completion to give themselves a break between calls! The manager started posting the average handle times of different agents on a noticeboard and, sure enough, average handle times came down significantly for some agents, resulting in lower average handle times overall and improved on-time performance.

I learned 1) to pay attention to input data, especially any unexpected quirks, and 2) to not go into engagements assuming I already knew how to solve the problem, i.e., don't go into engagements "carrying your hammer!"

Still, this particular engagement worked out well for me. I had noted some over-dispersion, relative to the mean, in the hourly call counts, which led to a paper on this issue (Chen and Henderson 2001). There was one thing I got right in this engagement: I was also asked to advise on how many police units a single dispatcher could coordinate from his console. This question relates to the cognitive load a single dispatcher can handle, and was outside my expertise. I said I did not know, which I believe was the right answer in the circumstance.

2.2 America's Cup Defence

In 1997, Team New Zealand (TNZ), a yachting syndicate, was preparing to defend the prestigious America's Cup in early 2000 in Auckland, New Zealand. Andy Philpott, a sailing enthusiast and leading researcher in stochastic programming at the University of Auckland, was asked by TNZ to develop a yacht match-racing simulation to help with design decisions. A match race is a race between just two boats, consisting of a starting period when the boats jockey for position prior to the starting gun, and then several upwind-downwind legs. The fastest boat around the course doesn't always win, because a boat that sacrifices some downwind speed for upwind speed may get to the first mark (turning point) first, and then get to dictate terms in the remaining legs of the race. But how much downwind speed should be sacrificed? Quantifying this tradeoff for different boat designs was our primary challenge.

Masters student David Teirney, Andy and I built a match-race simulation (Philpott et al. 2003) that incorporated the physics of the boat, basic tactical elements such as when to tack, and a model of wind speed and direction as seen by the boats. We took as input the physical properties of two proposed designs and output the probability that the first design would beat the second design in a given average windspeed; see Figure 1, which appeared in Philpott et al. (2003). Two of the three curves, that each give the probability that one design beats another, are not monotone in average wind speed. We were confused by this observation and at first perceived it to be a bug. But after discussion with TNZ it became clear that as the wind speed gets sufficiently strong, the boats heel over beyond the heel angles for which they are primarily designed, and at such angles the odds in a race become closer to 50-50.

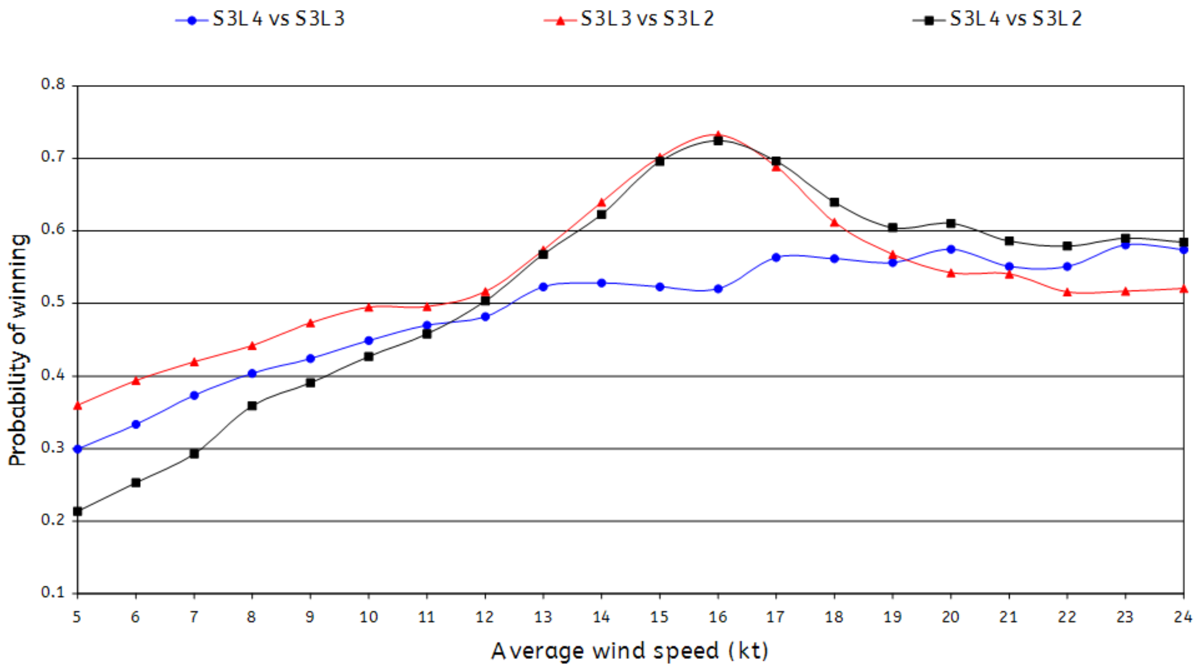

Figure 1: Probability one design beats another, for 3 design pairs.

Upon seeing plots like these giving the probability $p(\cdot)$ of a win as a function of average windspeed w, we thought the right next step was to estimate the probability density of the windspeed, $f(\cdot)$ say, and then select the design that maximized the overall expected win probability $\int p(w)f(w)\,dw$. Surprisingly, this approach was not appealing to the sailors. Rather than a design that performs well in expectation over the distribution of wind strengths, which could mean a specialist design tailored to the predominant conditions, they instead wanted a robust design that performed well across a wide range of wind conditions. Their rationale was that they wanted to be able to compete in all conditions, and would "win on their crew work," meaning they would perform maneouvers more efficiently and reliably than their competition.

Were they right? Were we? I don't know. But the important thing was that we had the conversation to decide on next steps, so went forward with a unified perspective. If we had instead worked on seeking to maximize the expectation then our findings would have been ignored and the work would be wasted. Contrast this experience with my effort with CHH, where I failed to keep in touch with the sponsor.

2.3 St. John's Ambulance

In 1997, my colleague Andrew Mason at the University of Auckland was asked by St John's Ambulance Service (henceforth, SJAS) to help design the shift schedules that would be used by their ambulance crews. Andrew kindly brought me into the project, and early discussions revealed that SJAS also sought help in determining the number of ambulance crews that should be scheduled in each hour of a typical week, which is required to determine the shift schedules. We explored the emergency services literature and implemented several queueing and optimization models that seemed promising, exploring their predictions using SJAS data to select the input parameters of the models. These models were useful in identifying an approximate sense of needs, but we were not confident in relying on their predictions because of the many modeling simplifications they required. Thus, we embarked on a simulation study of the ambulance system in Auckland, New Zealand. I did not know it at the time, but this area would become a primary focus in my research career that has continued to the present day.

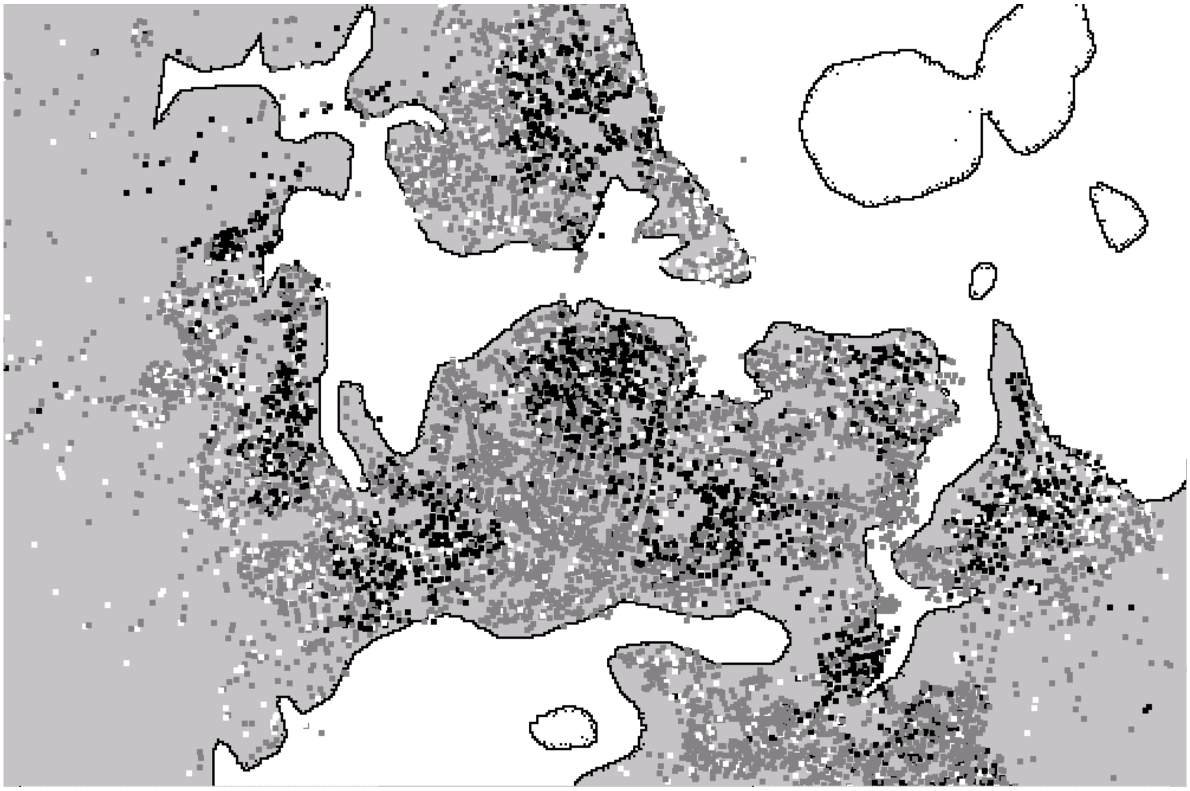

Figure 2: An example screenshot from the BARTSim package, where simulated call response times are categorized by shade; fast responses (darker colors) tend to cluster near ambulance bases and hospitals.

Over the next year or two we developed a software package (Henderson and Mason 2004) that could import ambulance data including data on calls, ambulance shifts, and travel times in a city, simulate the operations of the ambulance service and visualize both the raw data, the historical performance and the simulated results; see Figure 2 which appeared in Henderson and Mason (2004). This tool, BARTSim (Better Ambulance Rostering Technology Simulation), was subsequently licensed and later bought out entirely by The Optima Corporation, and to this day Optima's software operates in many cities worldwide. I benefited tremendously from Andrew's engineering insight, practical perspective, coding experience and hard work. Many times Andrew proposed ideas for how we should proceed and I was skeptical of the value of the proposed direction, but learned in retrospect how wrong I was. I learned the value of teaming up with amazing co-authors and mentors!

Challenges and eye-opening experiences included the following:

- Early in our work with SJAS we spent a shift riding along with an ambulance crew. While it was a quiet shift compared to normal, we still had some speedy rides on the way to high priority calls. We learned more about the crew experience while building goodwill with ambulance crews.
- There were many data challenges, some of which didn't emerge till we were able to visualize the data. For example, many calls seemed to be located in the sea, which Andrew eventually traced to a bug in the computer-aided dispatch software that SJAS was using that was introduced in a new software release. Even after correcting for that bug, there were a handful of call locations erroneously recorded in the Pacific ocean, some distance from New Zealand! Moreover, the data on times ambulance crews spent at the scene of calls was not reliable; investigation revealed that ambulance crews recorded timestamps manually for the various stages of a call (depart base, arrive at scene, complete at scene, reach hospital, clear at hospital, etc.) and crews sometimes forgot to

record the timestamps. In such cases, they would simply "push the button" to record timestamps after the call, so as to "catch up," leading to many recorded time intervals with durations of 1 second or so. Another issue we found in trying to locate ambulances was a large concentration of calls in West Auckland. We were unaware of a reason for such a large concentration. We learned, eventually, that in those early days SJAS was unable to locate calls that arose on highways, so all such calls were given the artificial location we had seen in West Auckland!

- The building of simple initial models, like queueing models, was very useful in helping us come to terms with the data. We also began with very simple models of travel times, because the recorded travel times seemed very noisy and difficult to understand. The modeling of travel times for emergency services would become an important component of my research agenda and that of others, e.g., Westgate et al. (2013), Westgate et al. (2016).
- We built the ability to animate the simulated ambulance operations. These animations were instrumental in detecting data errors, in gaining acceptance of our models by decision makers, and in improving our own understanding of ambulance operations and how they could be improved. For example, at a meeting with ambulance crews to demonstrate the model, an influential station manager stated that we hadn't appropriately captured the multiple priority levels of calls that SJAS handles. We subsequently updated the model to account for that fact and the station managers became more engaged with our modeling.
- Much of our early work with SJAS was in the form of consulting. If we had charged a typical hourly rate for all the coding we performed, we would have bankrupted the SJAS. Rather, we viewed much of our time on the project as some combination of a contribution to societal welfare and an investment in our future research careers. (At the time, we did not fully appreciate how entwined our future research careers would become with ambulance services).

The coding work was a very heavy load, even with Andrew assuming the lion's share. Access to something like a postdoctoral scholar or research programming team would have helped. I understand that Warren Powell's group at Princeton included such people (Powell 2003).

2.4 Beyond St. John's

The St. Johns work contains key lessons that opened my eyes to the challenges, excitement and reward of applied work, and set the stage for future research challenges. By being open to the opportunity of working with Andrew, I had opened a door to some exciting new questions that are sketched below.

Move-up The key finding from our analysis of SJAS operations was that, in contrast to a management belief that they simply needed to better align their resources to demand in space and time, there was no spare capacity in the system. The SJAS needed to increase resources or efficiency to meet their contractual obligations. This led us to explore the topic of "move-up," which is also known as system-status management (Wikipedia 2023). To ensure rapid response to incoming calls, the available ambulance fleet is spread over a city. When an ambulance is dispatched to a call, it leaves a hole in coverage. Move-up is a form of active fleet management, where one decides whether and how to fill such holes in coverage. Together with Ph.D. students Mateo Restrepo, Matt Maxwell and colleague Huseyin Topaloglu, I worked on the development of approximate dynamic programming (ADP) methods for the design of move-up policies (Restrepo 2008; Maxwell et al. 2010; Maxwell 2011). I believe that this was an early instance of the use of ADP in areas beyond "toy" applications such as video-games and the like, so the work had some influence in the research community. Our methods were never directly implemented in commercial software, but related methods were and are used. I learned that it is possible to significantly influence practice even in the absence of direct implementation of one's ideas in commercial software.

As part of this effort I had many conversations with ambulance providers about move-up, including learning of more "war stories." For example, in one major international city, an ambulance crew was asked to reposition themselves. They responded that they were moving, but the automatic vehicle location system (essentially, GPS) indicated that they hadn't moved. When a dispatcher contacted the crew to say that they had not yet seen movement, the crew responded that they were just finishing their coffee!

Ornge By now I had a track record in this industry, so when my friend Mike Carter at the University of Toronto was approached by Ornge, the provider of air-ambulance services in the province of Ontario, Canada, he referred them to me. We began with a team of Master of Engineering students building a simulation to determine how to deploy helicopters around the province. This early effort grew to multiple Master of Engineering projects, PhD students working in the summer to further develop software, and publications, e.g., MacDonald et al. (2011), Carnes et al. (2013), relating to both urgent and non-urgent aircraft deployment at Ornge. The work that was most innovative in this partnership was our (with colleague David Shmoys and several students) effort in solving dial-a-ride problems for non-urgent patient transfers. This innovative work would not have arisen had we not been open to questions beyond our initial expertise in emergency operations. In other words, if we had entered the relationship "carrying our hammer," we would have missed a golden opportunity.

Bounds and Expert Testimony One Friday in March of 2010 I was contacted by a law firm representing a major ambulance provider. The ambulance provider had lost a contract to provide ambulance services in a California county to a competitor during a tendering process. They believed the competitor could not achieve the contractual performance targets based on the ambulance hours the competitor had bid. The lawyers asked me to prove that this was the case.

This case involved a very large contract. To get a sense of how large, imagine that something like 50 ambulances were necessary to provide the required service (this is not the correct number). Each ambulance requires several crews to provide ambulance service 24/7, so a reasonable rule of thumb is that each ambulance costs on the order of 1 million dollars per year to operate, including crew costs (the dominant component), maintenance and depreciation. Thus, a 5 year contract is worth on the order of 250 million dollars.

When I asked when the lawyers needed a response they replied "Monday." This was a seriously daunting prospect with just a weekend to complete the work. With some thought on Friday, I suspected that I could devise an upper bound on performance given only the ambulance shifts that were proposed by the competitor, though I was not completely certain. I replied to the law firm that I would perform this computation, but made clear that the value of the bound could either provide the desired proof or be inconclusive. My PhD student, Matt Maxwell, and I had a very busy and sleep-deprived weekend, and explained our results to the law firm on Monday that established that the competitor could not meet the performance targets. Our report was delivered the next day. The report would not hold up to the most rigorous mathematical standards, but was certainly more than sufficient, beyond any reasonable practical doubt, to establish that the competitor bid was insufficient. Matt and I subsequently received payment for this consulting work, and I thought we were done.

Three years later the case went to court and I was called to testify. All our email records and computer files relating to the work were subpoenaed. The lawyers for the competitor were smart and thorough in deposition. I expected to be grilled in the courtroom based on watching TV shows and the like, but the lawyers treated me with respect and courtesy. I subsequently learned that if lawyers think a witness will be well-received by a jury, then they take care not to alienate the jury by grilling the witness. I believe this was the reason for "kid gloves" in my case.

The competitor subsequently won the case, i.e., "my side" lost. The lawyers for the competitor cleverly argued that in my analysis I had used travel times that were provided by the Optima Corporation, and their client did not have access to that information so couldn't be expected to have been able to repeat my analysis with Matt. This argument was legally clever, yet practically disingenuous. Certainly, the competitor needed some sense of travel times to enter their bid in the first place! Thus, my testimony was neutralized to some extent. I should emphasize that my testimony was just one part of the case brought by the ambulance organization, so one must not over-estimate the importance of my testimony and the counterarguments brought by the lawyers for the competitor.

In this experience I learned to be very careful about information trails in sensitive projects, that exciting questions can arise in practice, that deadlines can be brutal but you must be able to respond on the client's timeline, and that expert witness work can be stressful! Moreover, for expert witness work you must make yourself available when and where you are needed, so it can be very inconvenient.

Matt and I were inspired by this question of obtaining bounds. We wanted to publish the ideas we had developed over that weekend, but there was a residual mathematical gap between our analysis and a provable bound. Our understanding of ambulance operations meant that we did not expect this mathematical gap to make an appreciable difference in the practical application, but the "theorem" was not a theorem without closing the gap. With several co-authors we were later able to close the gap and to explore the properties of the bound in various settings, resulting in one of the papers about which I am most proud (Maxwell et al. 2014), despite it being far from my most cited work.

Community First Responders Community first responders (CFRs) are volunteers who are sometimes willing to respond to an emergency when notified. CFRs are notified by a smartphone app and can choose whether or not to respond. An important question relates to the desired density of CFRs across a city to ensure fast response times and thus medical outcomes. I have explored this question with co-authors in van den Berg et al. (2022). Another important question relates to CFR dispatch. To maximize survival rates for a call, one should alert *all* nearby CFRs, to maximize the chance of a fast response. But if multiple CFRs arrive on a scene then this can lead to frustration and potential disengagement for future calls. Thus, there is a question of how to dispatch CFRs in a time-staged manner so as to maximize the chance of survival while not creating too much disengagement due to multiple responses (Henderson et al. 2022).

My work in this sphere grew from an extended experience in a completely separate industry. Confidentiality agreements mean that I can't name that industry, but suffice to say that I would not have begun exploring this research area without that extended experience. The lesson I take from this vignette is that applied work in one sphere can seed work in another.

2.5 Citi Bike

I first learned about bike sharing applications from a talk given by Robert Hampshire at Cornell in 2010. At the time, I was inspired by the idea and wrote a simple test problem for use in the library SimOpt (D. J. Eckman and S. G. Henderson and R. Pasupathy and S. Shashaani 2020). I had alternative ideas to Robert's about how to formulate and solve the problem but shelved them due to other projects. Quite a bit later, Citi Bike was announced as a new initiative in New York City. My colleague at Cornell, David Shmoys, sensed a golden opportunity and after much effort managed to forge a connection out of nothing. Subsequently, he and his PhD student Eoin O'Mahony become the de facto "Analytics Group" for Citi Bike in its very early days.

At some point, David, Eoin and I devised a new approach to the problem. This early idea proved to be very important, permeating most of the decision problems faced by Citi Bike. First Eoin, and later PhD student Daniel Freund, built out the idea and worked to implement it for Citi Bike. We eventually called this central idea the "user dissatisfaction function" (UDF), which entails computation using time-dependent analysis of continuous-time Markov chains. Today almost all of Citi Bike's operational decisions are centered around UDFs.

This work had great impact at Citi Bike and led to Dantzig Dissertation Awards to both Eoin and Daniel, the INFORMS Wagner Prize, and a series of papers on the ideas, including, e.g., Freund et al. (2017), Freund et al. (2019), Freund et al. (2022). The main contribution does not include simulation modeling, though we did build simulation models and explored simulation-optimization methods for them (Jian and Henderson 2015; Jian et al. 2016).

What I learned from this experience is that 1) The inspiration for a great problem can come from anywhere, including departmental seminars; 2) It pays to have great colleagues, like David Shmoys, who can identify important problems and build links to companies like Citi Bike; 3) The use of a suite of models (our work included various combinatorial optimization formulations, linear programs, the use of discrete-convexity ideas, simulation and more) is almost always preferred to a single model; 4) Not every project is a simulation project; and 5) One must be very careful of the inputs to models — in this setting, the raw data on bike sharing is used to estimate demand for bike rides, but is censored since no-one initiates bike trips when no functional bikes are present in a station.

2.6 COVID-19 at Cornell

Unlike almost all universities in the world, Cornell's Ithaca campus reopened for (partial) in-person instruction in the fall semester of 2020, and has remained open since that time while keeping the campus population and those living in the surrounding community safe. There have been some hair-raising moments, such as when Omicron tore into our campus population 4 days after the variant was discovered in South Africa, but overall COVID-19 has been kept at bay. The story of how this was possible has many threads (Meredith et al. 2022), but mathematical modeling by the Cornell COVID Modeling Team was central to this success (Frazier et al. 2022). The team was led by my colleague Peter Frazier, with David Shmoys and me playing supporting roles in guiding the work of a cohort of PhD and undergraduate students. The work was simultaneously tremendously rewarding, eye-opening and exhausting, and extended over approximately 2 years. A multidisciplinary team from Cornell now has a research grant from the National Science Foundation to further develop the mathematical models and data-analysis techniques that we now know would have been of tremendous value in real time as the epidemic unfolded.

Like most of us, it all began for me in March of 2020. My wife and I were on sabbatical in New Zealand, and as the world began to reel from the spread of the virus, we returned to our home in Ithaca, New York. I began working on models for ventilator transshipment within New York state, which seemed at the time to be a critical issue. However, that work very quickly became obsolete as doctors learned to avoid the use of ventilators. I knew that Peter was working with some PhD students to explore the use of regular asymptomatic testing to keep the SARS-CoV-2 virus under control and it was clear he could use some help. So David and I volunteered.

Our first task was to build a simulation model of virus spread in the campus community, to gauge the frequency of testing needed to limit epidemic growth. The underlying model is a relatively straightforward stochastic multi-group compartment model (Frazier et al. 2022) that I won't describe here, but two issues are salient. First, much of our work centered on determining the input parameters to this model, which was complicated due to the fragmentary nature of data and evidence from news and scientific reports from around the world. We eventually developed an elaborate uncertainty quantification structure to capture this difficulty, and we repeatedly reused that structure as new variants emerged over the following two years. Second, our initial prediction was for approximately 1200 infections over the course of the fall 2020 semester should Cornell reopen for in-person instruction, but *what of the alternative*? I recall this being the pivotal question. Our analysis showed that if Cornell adopted the alternative of going online-only, then infections would have been far, far higher. This may seem counterintuitive, but is readily understood in retrospect. Many students would return to Cornell even in the online-only mode of instruction, because they had prepaid their (off-campus) leases and wanted to be near their friends. With online-only instruction, Cornell could not enforce behavioral compacts that include regular testing, so the virus could spread rapidly in the off-campus student population, meaning that there would be many more infections than in an in-person semester with regular required testing.

Figure 3 makes this point clear. Each point in this scatter plot represents the median number of infections under two scenarios for a given plausible sample from the probability distribution over the input parameters of the simulation model. The x axis depicts the median number of infections in an in-person instruction mode, while the the y axis depicts the corresponding number for online-only instruction. Virtually all points lie above the line $y = x$, and median infections for in-person instruction never greatly exceed those for online-only instruction. The safer decision is apparent, given this plot, and this was the courageous decision adopted by Cornell leadership in the summer of 2020. (Courageous in the sense that very few universities adopted this strategy and there were many severe critics of the plan.)

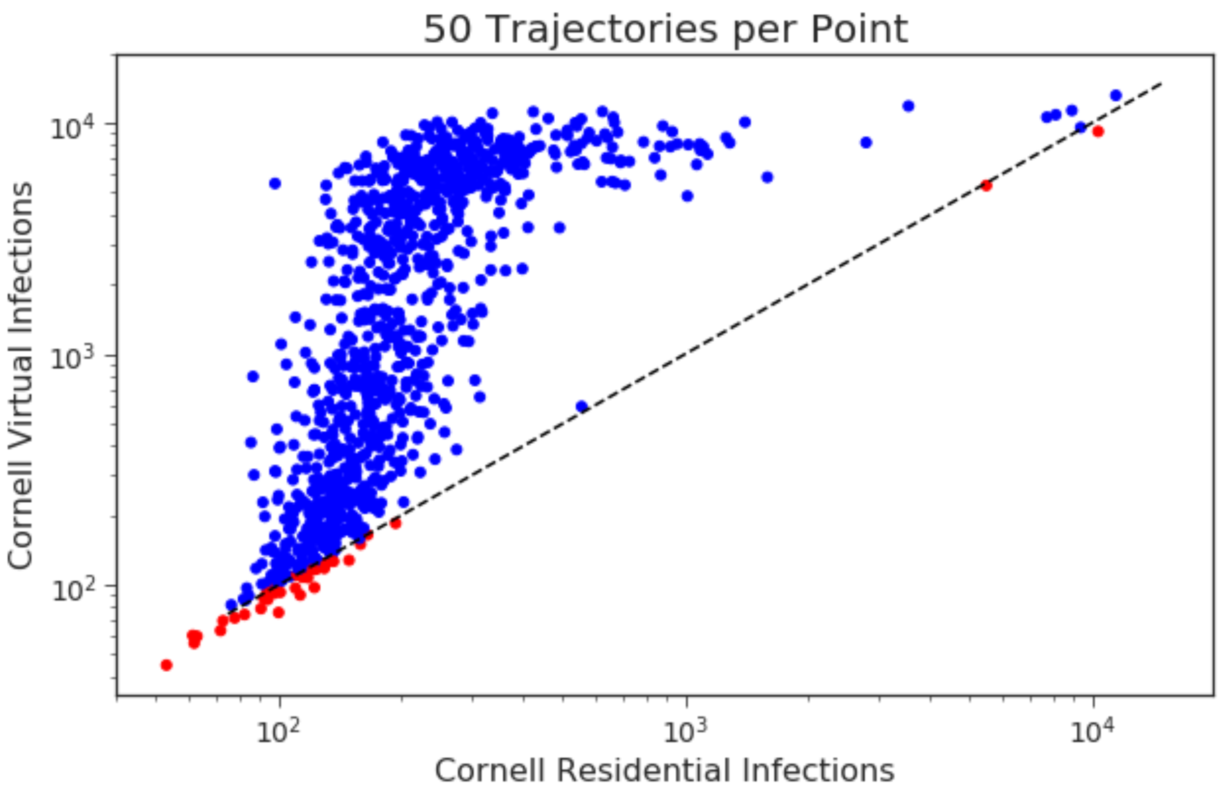

Figure 3: Scatter plot of median infections under an online-only mode of instruction (vertical axis) versus in-person instruction (horizontal axis) for the fall 2020 semester of Cornell's Ithaca campus. Each point reflects simulation results for a single choice of input parameters.

The modeling team was extremely busy over the summer of 2020, but its work was just beginning. We subsequently advised Cornell leadership on a dizzying number of policy questions, including differentiated testing frequencies for subpopulations on campus, quarantine and isolation capacity, the effect of vaccine and booster mandates, the effect of masking policies, the safety or lack thereof of fully dense classrooms post vaccination, the safety of running special events like summer camps and many more.

We have many "war stories" from this experience that enliven our research presentations and classroom discussions to this day. One of these stories is worth recapping here. In the late fall of 2021 we were building out a new class of models to determine Cornell policy for the upcoming spring 2022 semester, with Omicron dominating our thinking. Our students were pushing many updates to the github repository containing our code. This repository was public, partly to ensure that our work was seen to be evidence-driven and transparent, but it is fair to say that the public nature of the repository was not uppermost in our minds as we worked towards various meetings with the Cornell Provost. At some point, a Reddit thread emerged, where Reddit users were posting about the comments in our repository pushes and our code, attempting to determine the nature of our recommendations to the provost! While we had nothing to hide, it is important to not release preliminary findings that might change as code is debugged and parameter estimates are determined. We were fortunate that this episode did not give rise to major concerns about our work in the public eye. It helped that we quickly developed a *private* repository to contain work in progress, before publishing complete and debugged code to the public repository with carefully selected comments to go with the posting!

There were many lessons from the two years we spent on this project that are very much in line with those I've already expressed herein. Again, we saw the need to take great care with digital footprints.

Communication with key decision makers was critical, but so were the many town halls, public meetings and press interviews that Peter (almost entirely, Peter) presented. It was critical to properly represent our uncertainty in the input parameters to our models, and to identify decisions that were robust to those uncertainties. It was critical to be responsive to the needs of Cornell leadership; I remember many late nights and weekends working to impending deadlines for analysis and reports.

Our work in this sphere had the highest potential consequences. We played a major role in keeping the campus and surrounding community safe, even though there are still people who believe we endangered their lives and others who believe the threat was negligible. Cornell activities make up a very large part of the Ithaca economy, which would likely have taken a very heavy blow had Cornell gone online-only in the fall of 2020. Much of the Cornell community now has some appreciation for mathematical modeling and has a sense of what is "operations research." The students who worked in this area have had the best possible apprenticeship in the data analysis, modeling and report-writing work that forms a large part of the activities of scientists in the tech industry, as well as in preparation for academia, in providing context for research, war stories for teaching, etc.

I am very proud of this work.

3 BACK TO THE THEMES

3.1 Why Engage with Practical Projects?

There are many reasons.

- Practical projects are rewarding in and of themselves. It is gratifying to see one's ideas implemented and used, and to see the benefits that are attained, e.g., my work in ambulance services.
- Problem selection is the most important skill in research; if you aren't working on worthwhile questions, then no amount of effort will elevate your work to the forefront of research. Working on practical problems leads to a "flow" of quality research directions. Each question I tackle in an applied setting leads to further questions in that setting, but also interesting basic-research questions. For example, with the COVID work, we now have a team of 6 faculty members funded to work on central questions in mathematical epidemiology and beyond. Practical work on one application can also cross fertilize work in another, as with my work on community first responders.
- My intuition is stronger on research questions with context than on research questions absent of such context, which I think makes my work in those areas better.
- My teaching has greatly benefited from practical experiences. I have a wealth of war stories and examples that enrich my teaching. For example, when teaching confidence intervals for proportions, I discuss our America's Cup work on estimating the probability that one boat design beats another and the fact that we need approximately 10,000 simulated races to estimate this probability to approximately 2 decimal places of statistical accuracy.

3.2 How Can You Find and Initiate Projects?

I have been extraordinarily lucky in having friends who guided me into projects, including David Ryan, Andy Philpott and Andrew Mason at the University of Auckland, Mike Carter at the University of Toronto, and Peter Frazier and David Shmoys at Cornell University. My role, and yours, is to be open to such engagements and to possess expertise to add to the effort, which, presumably, relates to your methodological specialization. Once you gain a reputation for good work in a particular area, you will find that you get many referrals, as happened with Mike Carter and my work with Ornge.

Beyond mentors, how else can you get started? Often senior design projects and Master of Engineering projects are run with an industrial sponsor, which can lead to larger-scale activities, as with my work with Ornge. Advising such student teams can also help you build confidence, expertise, and a track record of success. An additional source of projects is through relationships with alumni; I have not been as successful with this avenue, though I often hear from my business school faculty friends of success in this pathway.

There are some general strategies and principles that may help in getting started:

- You should not think of yourself as (only) a simulation person. Instead, strive to be a problem solver who can draw on multiple tools.
- Do the simplest thing possible that still answers the questions posed to you; avoid my mistake with CHH. You can explore more complex methods once you have *an* answer for the client.
- In getting started on projects with a great deal of data, start with exploratory data analysis, i.e., plot the data every way possible! This was important in our work with the SJAS, where Andrew uncovered the call-location errors. In raising this issue, we provided value to the SJAS before doing any analysis.
- Use the Heilmeier catechism (DARPA 2023) to structure your thinking. If you can't answer those questions clearly, then you should revisit your high-level goals.
- If your client is already using sophisticated methods, then you may think that you can't add value, but check what they are doing. For example, one group with whom I worked included several physicists who were using simulated annealing to optimize a poorly structured objective function. It turned out that they were setting the temperature sequence to the constant 0! In doing so, they were simply performing local search, though they did not realize that!
- You need a portfolio of projects, because some projects will not bear fruit. For example, I tried to engage with the Ann Arbor police. Data sharing agreements were still being set up by lawyers when, two years after the initial overture, I left the University of Michigan.

3.3 Strategies and Pitfalls

You must stick close to the sponsor and listen to their preferences, as I learned in the work with CHH and with TNZ. You should try to get first-hand knowledge of the problem setting, as Andrew Mason and I did in ambulance planning, to improve your credibility with stakeholders and to ensure your understanding of the setting is correct.

Usually, companies work on a very different timeframe from that of academic research projects. Academics are used to working on a timescale associated with PhD theses, i.e., years, while companies are used to working on a timescale like weeks or even days! You cannot expect to perform high-quality academic work on such tight timescales. So what should you do? I solve problems in multiple ways. First, I do something for the company on their desired timescale, which must, by necessity, be relatively straightforward. In doing so, I try to extract a promise from my contacts at the company that they will continue to work with me on a (much) slower academic timescale, so that my team can develop new methods that we believe will better solve the problem. This way, the company gets what they need on their timescale, and I get what I need on mine. Think in terms of a *suite* of models, not just a *single* model.

Don't "carry a hammer," i.e., don't work on automatic pilot, assuming you know how to solve the problem. In many projects you will encounter interesting new features that may change your analysis and that may lead to future research directions, as I experienced with the Police Communications Centre.

In that work, the data had features that led to new research, in the realm of accounting for forecast errors in service systems, e.g., Steckley et al. (2005), Steckley et al. (2009). Our work on bike sharing brought out an additional issue, in that the data in that setting are censored, e.g., because trips cannot be initiated when no bikes are available. Such issues are in addition to the usual issues associated with data quality and errors therein. So pay attention to data.

Take great care with all forms of communication and media that you produce. All of that media can be subpoenaed, and great confusion and consternation can result if any of the media is misinterpreted.

There will be questions that you cannot answer within your expertise; in such cases do not engage. I mentioned such a situation with the Police Communications Centre in New Zealand. A second example comes from the question of the needed quarantine/isolation capacity at Cornell University during COVID-19. There, we provided an answer, but we took great pains to indicate how uncertain we were about the answer. This philosophy of communicating a range rather than a point forecast is a central theme in Silver (2012).

Avoid committing to providing an analysis to support a decision that has already been made. (See Sadowski 2007 for more on this point.) Indeed, I was asked to do exactly that in the California ambulance case. I was able to reframe the question to providing an analysis of the competitor's bid. The lawyers agreed that my report should simply report my findings as to whether I could rule out compliance with contractual requirements or not. Thus, I was free to report my findings and not a particular conclusion.

The standards for what constitutes an iron-clad observation are different in academic publications and in practice. In academia we are looking for the equivalent of a mathematical proof, or overwhelming empirical evidence. In industry, one often, or perhaps usually, only requires strong evidence for the benefits of a new policy or plan in order for the new idea to be tried. Of course, there are exceptions to this rule. Should academia change? No! Journals are right to seek rigorous evaluation of new ideas. Should industry change? No! Industry must be agile, to compete and innovate.

3.4 What Should You Do Afterwards?

After the initial effort on a project, on a timescale useful for the organization, is complete, you can then turn to other efforts, including writing papers on your work for journals like the INFORMS Journal on Applied Analytics, and exploring outgrowths of the practical work on academic time scales.

To that end, you should not view your practical experience as being confined to the project. Rather, reflect on what you saw and where you felt there was room for improvement in your methodology or in the application setting itself. When you identify interesting potential directions, you now have a strong line for potential research. It can be helpful at this stage to get "buy in" from your contacts from the practical project, to have them agree to (loosely) work with you over the next few years as you explore the research topic you identified. This approach has worked for me on multiple occasions, including with my ambulance work and the COVID modeling effort at Cornell.

Your contacts are valuable for ongoing collaboration and future opportunities. An ongoing challenge is that your contacts may be promoted or transferred within the organization, or they may even switch companies. But such moves usually entail an increase in responsibilities, and as such can be leveraged to scale your activities.

3.5 High-Stakes Settings

In the setting of high-stakes decisions, such as those I have encountered in COVID-19 modeling and in ambulance work, some of the themes above are amplified, but I am not sure that anything is fundamentally different. Agile models are important, since questions can evolve quickly, e.g., as new COVID variants emerge. Clarity in communication is even more important, e.g., through our outreach to the public during the COVID-19 work, and in my work as an expert witness. Finally, I want to reinforce the point that you must respond to your stakeholders on their timeframe, which can be very hard work and quite stressful. Just remember that you are the best person to help, and your best work can make a difference. Your research interest in some aspect of the work must wait. I maintain a list of interesting research questions at all times, and that list tends to lengthen considerably during high-stakes work! These themes are echoed, to some degree, in work of others in our general research community who tackle high-stakes problems, e.g., Bastani et al. (2022), Basso et al. (2023), Pitt et al. (2018), Kaplan and O'Keefe (1993)

4 FINAL REMARKS

My most cited work arises out of the interplay between applications and new methodology. Such work tends to capture the imagination of audiences. Indeed, when I give talks on such work, questions on the modeling come thick and fast, and I must carefully manage my time to ensure that I cover key messages.

This kind of work also helps with funding. Many funding agencies strive to justify the investments made by the public. Often, the research that comes out of the kinds of projects discussed herein fills that need well, which, happily, makes program directors open to future proposals.

Even if your work is not implemented directly, your efforts can often help an industry evolve. This was the case with my work on move-up, where the practice is now quite widespread. It is also true of some work I co-developed in radiation treatment planning for cancer therapy (Chu et al. 2005), where a workshop we organized to discuss our work and related ideas from others helped focus a lot more attention on the robustness of treatment plans to uncertainties in total delivered doses.

Let me close by urging you to **be lucky**, as I have been! I have been fortunate to have many wonderful mentors and research partners. Yet perhaps my biggest stroke of luck was the opportunity to complete a Ph.D. at Stanford with Peter Glynn. The theoretical foundation I gained under Peter's supervision has repeatedly allowed me to tackle practical problems with a well-appointed toolkit, and to have at least some insight into how I can translate some aspects of those experiences into basic research.

ACKNOWLEDGMENTS

I am grateful to project sponsors, mentors, students and other co-authors, whether mentioned or not. This work was partially supported by National Science Foundation Grants CMMI 2035086 and DMS 2230023.

REFERENCES

Basso, L. J., M. Goic, M. Olivares, D. Sauré, C. Thraves, A. Carranza, G. Y. Weintraub, J. Covarrubia, C. Escobedo, N. Jara, A. Moreno, D. Arancibia, M. Fuenzalida, J. P. Uribe, F. Zúñiga, M. Zúñiga, M. O'Ryan, E. Santelices, J. P. Torres, M. Badal, M. Bozanic, S. Cancino-Espinoza, E. Lara, and I. Neira. 2023. "Analytics Saves Lives During the COVID-19 Crisis in Chile". *INFORMS Journal on Applied Analytics* 53(1):9–31.

Bastani, H., K. Drakopoulos, V. Gupta, J. Vlachogiannis, C. Hadjichristodoulou, P. Lagiou, G. Magiorkinis, D. Paraskevis, and S. Tsiodras. 2022. "Interpretable Operations Research for High-Stakes Decisions: Designing the GReek COVID-19 Testing System". *INFORMS Journal on Applied Analytics* 52(5):398–411.

Carnes, T. A., S. G. Henderson, D. B. Shmoys, M. Ahghari, and R. D. MacDonald. 2013. "Mathematical Programming Guides Air-Ambulance Routing at Ornge". *Interfaces* 43:232–239.

Chen, B. P. K., and S. G. Henderson. 2001. "Two Issues in Setting Call Center Staffing Levels". *Annals of Operations Research* 108:175–192.

Chu, M., Y. Zinchenko, S. G. Henderson, and M. B. Sharpe. 2005. "Robust Optimization for Intensity Modulated Radiation Therapy Treatment Planning Under Uncertainty". *Physics in Medicine and Biology* 50:5463–5477.

DARPA 2023. "The Heilmeier Catechism". https://www.darpa.mil/work-with-us/heilmeier-catechism. [Online; Accessed March 20, 2023].

D. J. Eckman and S. G. Henderson and R. Pasupathy and S. Shashaani 2020. "Simulation Optimization Library". http://www.simopt.org. [Online; Accessed March 20, 2023].

Frazier, P. I., J. M. Cashore, N. Duan, S. G. Henderson, A. Janmohamed, B. Liu, D. B. Shmoys, J. Wan, and Y. Zhang. 2022. "Modeling for COVID-19 College Reopening Decisions: Cornell, a Case Study". *Proceedings of the National Academy of Sciences* 119(2):e2112532119.

Freund, D., S. G. Henderson, E. O'Mahony, and D. B. Shmoys. 2019. "Analytics and Bikes: Riding Tandem With Motivate To Improve Mobility". *INFORMS Journal on Applied Analytics* 49(5):310–323.

Freund, D., S. G. Henderson, and D. B. Shmoys. 2017. "Minimizing Multimodular Functions and Allocating Capacity in Bike-Sharing Systems". In *Integer Programming and Combinatorial Optimization Proceedings*, edited by F. Eisenbrand and J. Koenemann, Volume 10328 of *Lecture Notes in Computer Science*, 186–198: Springer.

Freund, D., S. G. Henderson, and D. B. Shmoys. 2022. "Minimizing Multimodular Functions and Allocating Capacity in Bike-Sharing Systems". *Operations Research* 70(5):2715–2731.

Henderson, S. G., and A. J. Mason. 2004. "Ambulance Service Planning: Simulation and Data Visualization". In *Operations Research and Health Care: A Handbook of Methods and Applications*, edited by M. L. Brandeau, F. Sainfort, and W. P. Pierskalla, 77–102. Boston: Kluwer Academic.

Henderson, S. G., P. L. van den Berg, C. J. Jagtenberg, and H. Li. 2022. "How Should Volunteers Be Dispatched to Out-Of-Hospital Cardiac Arrest Cases?". *Queueing Systems* 100(3-4):437–439.

Horner, P. 2015. "Meet the 'Member in Chief': Interview With Ed Kaplan, the Next President of INFORMS Who Plans to Lead by Example". *OR/MS Today* 42(6):36–41.

Hugan, J. C. 2014. “A Practical Look at Simulation Project Management”. In *Proceedings of the 2014 Winter Simulation Conference*, edited by A. Tolk, D. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, 98–102. Piscataway NJ: Institute of Electrical and Electronics Engineers.

Jian, N., D. Freund, H. Wiberg, and S. G. Henderson. 2016. “Simulation Optimization for a Large-Scale Bike-Sharing System”. In *Proceedings of the 2016 Winter Simulation Conference*, edited by T. M. K. Roeder, P. I. Frazier, R. Szechtman, and E. Zhou, 602–613. Piscataway NJ: Institute of Electrical and Electronics Engineers.

Jian, N., and S. G. Henderson. 2015. “An Introduction to Simulation Optimization”. In *Proceedings of the 2015 Winter Simulation Conference*, edited by L. Yilmaz, W. K. V. Chan, T. M. K. Roeder, C. Macal, and M. Rosetti, 1780–1794. Piscataway NJ: Institute of Electrical and Electronics Engineers.

Kaplan, E. H., and E. O’Keefe. 1993. “Let the Needles Do the Talking! Evaluating the New Haven Needle Exchange”. *Interfaces* 23(1):7–26.

MacDonald, R. D., M. Ahghari, T. A. Carnes, S. G. Henderson, and D. B. Shmoys. 2011. “Use of a Novel Application to Optimize Aircraft Utilization for Non-urgent Patient Transfers”. *Air Medical Journal* 30(5):255.

Maxwell, M. S. 2011. *Approximate Dynamic Programming Policies and Performance Bounds for Ambulance Redeployment*. Ph. D. thesis, School of Operations Research and Information Engineering, Cornell University, Ithaca, NY.

Maxwell, M. S., E. C. Ni, C. Tong, S. R. Hunter, S. G. Henderson, and H. Topaloglu. 2014. “A Bound on the Performance of an Optimal Ambulance Redeployment Policy”. *Operations Research* 62(5):1014–1027.

Maxwell, M. S., M. Restrepo, S. G. Henderson, and H. Topaloglu. 2010. “Approximate Dynamic Programming for Ambulance Redeployment”. *INFORMS Journal on Computing* 22:266–281.

Meredith, G. R., D. G. Diel, P. I. Frazier, S. G. Henderson, G. A. Koretzky, J. Wan, and L. D. Warnick. 2022. “Routine Surveillance and Vaccination on a University Campus During the Spread of the SARS-CoV-2 Omicron Variant”. *JAMA Network Open* 5(5):e2212906–e2212906.

Nelson, B. L. 2016. “Technology Transfer of Simulation Analysis Methodology: One Man’s Opinion”. In *Proceedings of the 2016 Winter Simulation Conference*, edited by T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, 248–254. Piscataway NJ: Institute of Electrical and Electronics Engineers.

Philpott, A. B., S. G. Henderson, and D. Teirney. 2003. “A Simulation Model for Predicting Yacht Match Race Outcomes”. *Operations Research* 52:1–16.

Pitt, A. L., K. Humphreys, and M. L. Brandeau. 2018. “Modeling Health Benefits and Harms of Public Policy Responses to the US Opioid Epidemic”. *American Journal of Public Health* 108(10):1394–1400.

Warren B. Powell 2003. “Personal Communication”.

Restrepo, M. 2008. *Computational Methods for Static Allocation and Real-Time Redeployment of Ambulances*. Ph. D. thesis, Cornell University, Ithaca NY, USA.

Sadowski, D. A. 2007. “Tips for Successful Practice of Simulation”. In *Proceedings of the 2007 Winter Simulation Conference*, edited by S. G. Henderson, B. Biller, M. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, 107–112. Piscataway NJ: Institute of Electrical and Electronics Engineers.

Silver, N. 2012. *The Signal and the Noise: Why So Many Predictions Fail-but Some Don’t*. New York, NY: Penguin.

Steckley, S. G., S. G. Henderson, and V. Mehrotra. 2005. “Performance Measures for Service Systems With a Random Arrival Rate”. In *Proceedings of the 2005 Winter Simulation Conference*, edited by M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 566–575. Piscataway, NJ: Institute of Electrical and Electronics Engineers.

Steckley, S. G., S. G. Henderson, and V. Mehrotra. 2009. “Forecast Errors in Service Systems”. *Probability in the Engineering and Informational Sciences* 23(2):305–332.

van den Berg, P. L., S. G. Henderson, C. J. Jagtenberg, and H. Li. 2022. “Modeling the Impact of Community First Responder Systems”. http://ssrn.com/abstract=3825060, accessed March 20, 2023.

Westgate, B. S., D. B. Woodard, D. S. Matteson, and S. G. Henderson. 2013. “Travel Time Estimation for Ambulances Using Bayesian Data Augmentation”. *The Annals of Applied Statistics* 7(2):1139–1161.

Westgate, B. S., D. B. Woodard, D. S. Matteson, and S. G. Henderson. 2016. “Large-Network Travel Time Distribution Estimation for Ambulances”. *European Journal of Operational Research* 252(1):322–333.

Wikipedia 2023. “Public Utility Model”. https://en.wikipedia.org/wiki/Public_utility_model, accessed March 20, 2023.

AUTHOR BIOGRAPHY

SHANE G. HENDERSON holds the Charles W. Lake, Jr. Chair in Productivity in the School of Operations Research and Information Engineering at Cornell University. His research interests include a range of topics in discrete-event simulation and simulation optimization, and he has worked for some time with emergency services, bike-sharing and other transportation applications. He worked intensively on modeling efforts to support decision making by Cornell University leadership during the COVID-19 pandemic. He is a co-creator of SimOpt, a testbed of simulation optimization problems and solvers. He is an INFORMS Fellow. His web page is http://people.orie.cornell.edu/shane and his email address is sgh9@cornell.edu.