

BOAT SPEED

Through statistical modeling, Boston College economics professor Chris Maxwell can tell you which college crews are likely to win. But his numbers reveal something more: the spring season is a search for speed.

CALCULUS

By Tom Matlack
Photos Devin Swett

A

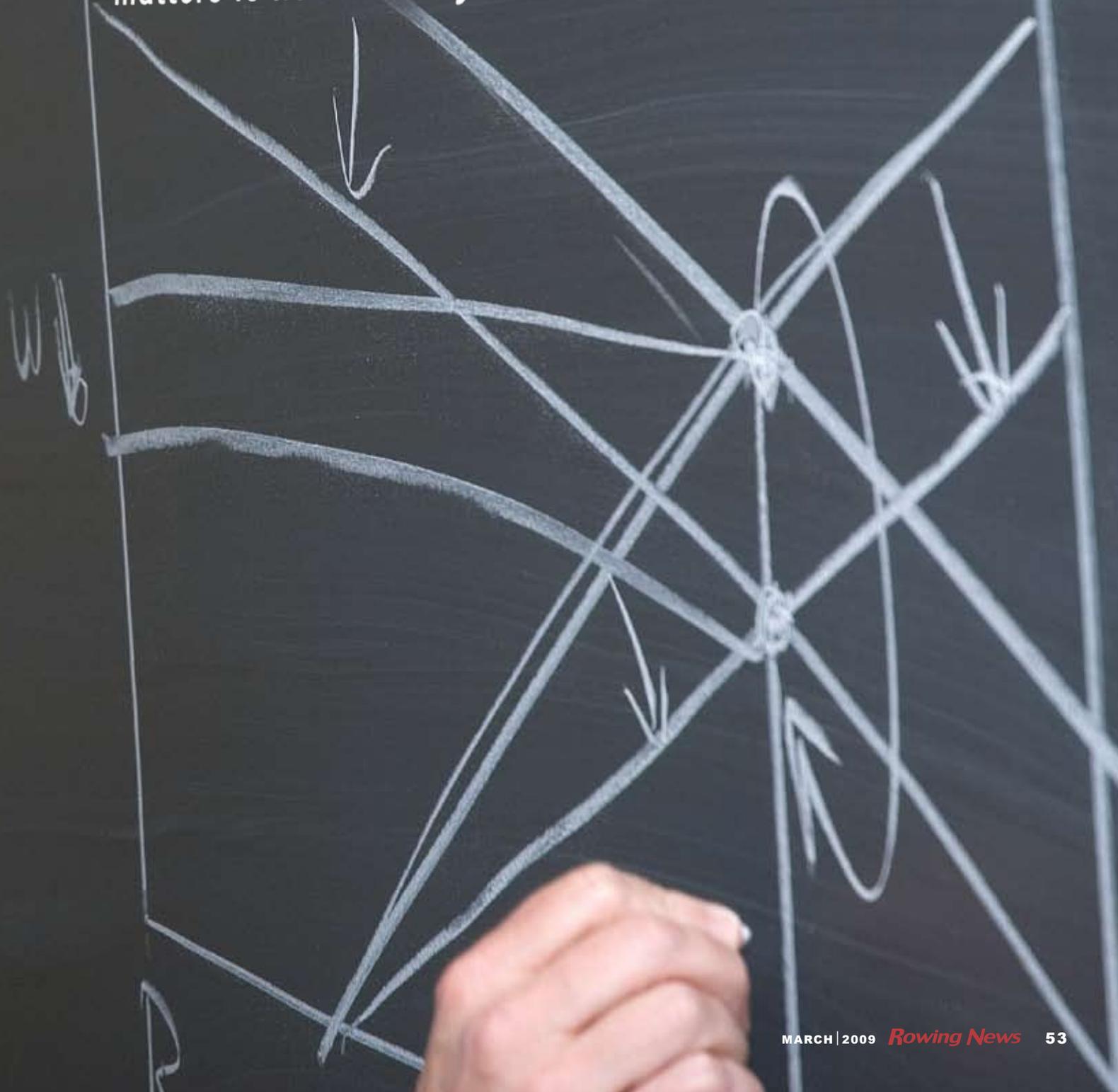
lumni Stadium in Chestnut Hill, Mass. has been the site of more than one miraculous play. During the last Boston College home game of the 2008 season, in a must-win situation against ACC rival Maryland, Billy Flutie, the backup quarterback and nephew of Heisman Trophy winner Doug Flutie, threw a nine-yard touchdown on a fake field goal to win the game 28-21. Flutie's trick play, like his uncle's epic last-second bomb against Miami, is just the kind of sports phenomenon that could never be predicted ahead of time. Or could it?

On a Thursday morning just before the Maryland game, Chris Maxwell ambles into a classroom across a parking lot from Alumni Stadium and Conte Forum where the hockey and basketball teams play their games (the buildings are connected so luxury boxes on the common wall have a view of football, basketball, and the national champion hockey teams in action). Maxwell has the soft voice and gentle eyes of a man infinitely confident in his chosen profession but he becomes more introverted when the topic moves away from statistical modeling. He's wearing a pink shirt and khaki pants. His salt and pepper gray hair gives him an air of distinction. As he plugs his laptop into an overhead screen 20 students wander in, mostly senior boys. A couple of them wear Red Sox caps on backwards. There's exhaustion in the room after a late night running mathematical scenarios in preparation for Maxwell's Sports Economics class. Almost every student has a big cup of coffee in his or her hands along with their computer printouts.

There is no theoretical instruction in this class. Maxwell has divided the students into six teams. For each class, he asks his groups to model a specific sports question, in this case the causes of changes in NFL ticket prices by city from 1996-2005. Then he models the same question. Grades are determined by how well the teams do against each other as well as the master himself. Maxwell has confided in me ahead of time that five out of the six groups did a good job, though they didn't beat him. One group, however, is in trouble. He will make the poor performers go first so they don't know how badly they have done until later in the session.



"The coaches are trying to do something different. I'm modeling speed. Whether you won or lost doesn't matter to me. What matters is how much you won or lost by."

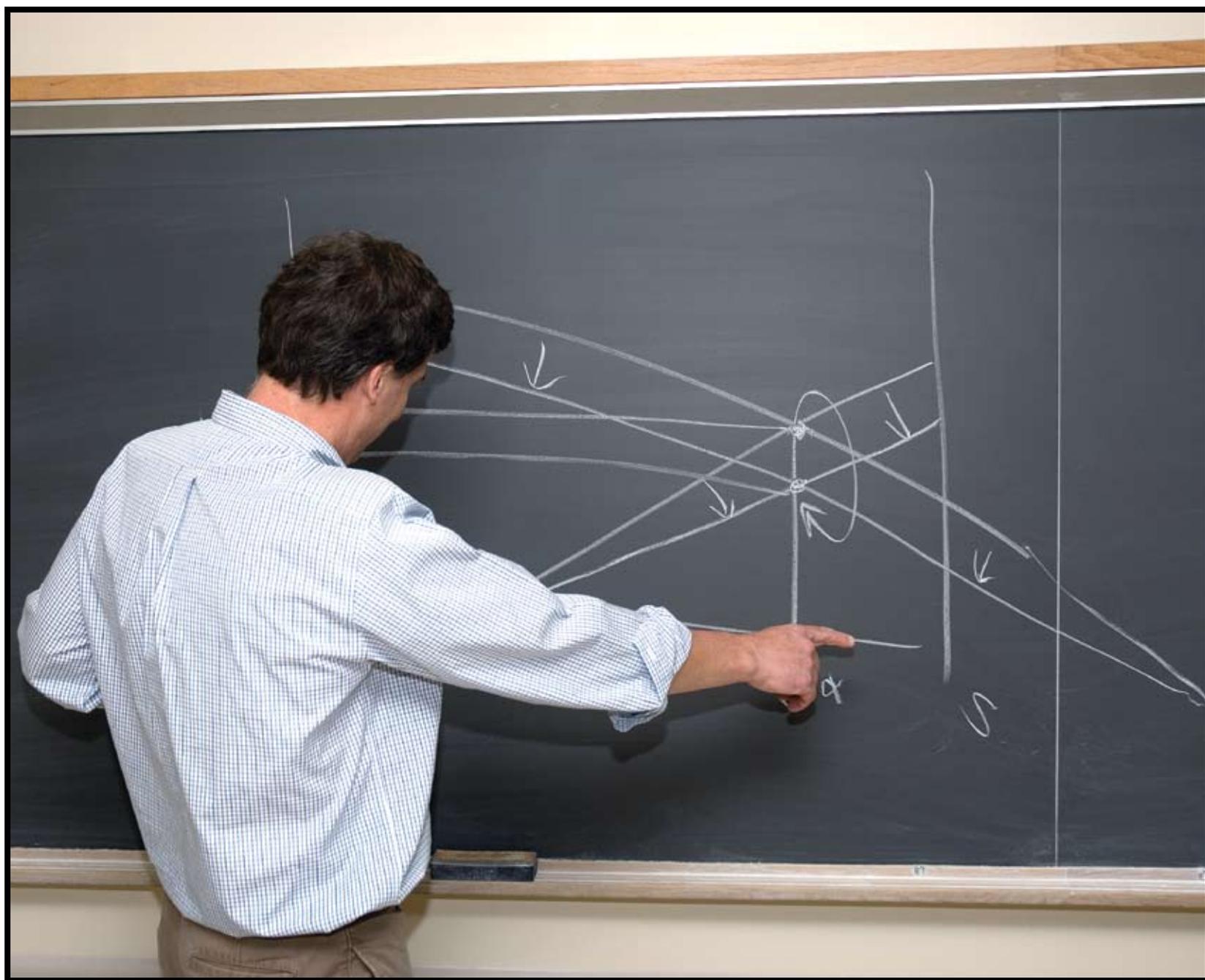


Once class has started, Maxwell hands the students a summary sheet with seven columns: one for each team and the seventh titled "cMax." Down the side of the sheet is a listing of the potential underlying causes of movements in ticket prices: wins in the previous season, offensive and defensive ranking, per capita income, the price of a hot dog, among many others. This is what Chris Maxwell does better than anyone else. He looks at professional and college sports through the lens of modern mathematical modeling. Maxwell has spent the first few classes improving upon the theories of baseball guru Bill James, in many ways the grandfather of "Money Ball" taken up by

Red Sox GM Theo Epstein and others with such devotion that baseball has become a different sport. Maxwell's class is intended to do nothing less than produce the next wave of Money-Ball-stat-jocks that now dominate baseball, football, and even was an underlying reason for the Celtics recent championship (individual defensive statistical analytics, like the shooting percentage of individual opposing players when guarded by a particular Celtic, are posted before the game and discussed at length afterwards in the Celtics' locker room).

Halfway through class a student raises his hand and asks, "Professor Maxwell, why

don't you use your models to take down Vegas?" Maxwell lets out a quiet laugh like the quantum physicist who has just been asked why rocks fall to the earth. "I did download 20 years of NCAA basketball data last weekend just for fun," he admits, "but if I could beat Vegas I wouldn't be here" he says with a sparkle in his eye. Truth is, Maxwell probably could take down Vegas, but he would never admit that to his students, or anyone else for that matter. His interest is academic not financial. Besides, he has an obsession that has nothing to do with gambling or even football or basketball. He spends 10 hours a week poring over data of a different kind in mad pursuit of one thing: boat speed.



2008 NCAA cMax Ratings

Division I, Eight, Grand Final

Pos.	School	sec.
1	Yale University	-
2	Brown University	0.2
3	University of California at Berkeley	0.7
4	Stanford University	1.2
5	Radcliffe University	2.1
6	University of Virginia	3.0

Division I, Eight, Petite Final

7	University of Washington
8	Michigan State University
9	Washington State University
10	University of Southern California
11	University of Michigan
12	Princeton University

2008 NCAA Actual Results

Division I, Eight, Grand Final

Pos.	School	Times
1	Yale University	6:34.05
2	Stanford University	6:34.95
3	Brown University	6:35.25
4	University of California at Berkeley	6:36.21
5	University of Washington	6:38.25
6	Michigan State University	6:40.26

Division I, Eight, Petite Final

7	Harvard University	6:35.30
8	University of Virginia	6:38.52
9	Washington State University	6:42.40
10	University of Southern California	6:43.32
11	University of Michigan	6:43.91
12	Princeton University	6:45.04



Stanford	2.2	Stanford	6.5	MichSU	7.2	9WSU	8.1
MichSU	4.7	MichSU	7.3	WSU	8.5	10USC	9.1
WSU	5.1	USC	8.6	USC	9.8	Michigan	10.1
USC	5.3	Michigan	9.8	Michigan	10.1	Yale	10.2
Michigan	5.4	OhioSU	10.1	Yale	10.2	Princeton	10.3
OhioSU	5.9	Princeton	10.3	Princeton	10.3	Yale	10.4



UCD	13.3	Columbia	14.6	UCD	15.3
UCF	14.6	Iowa	14.8	Iowa	15.3
SacSt	14.8	Syracuse	14.9	Syracuse	15.3
Iowa	15.6	UCF	15.4	UCF	15.3
Miam	15.7	SacSt	15.8	SacSt	15.3
Syrac	15.7	Iowa	16.0	Iowa	15.3
Indian	17.0	Syracuse	16.0	Syracuse	15.3

2008 IRA cMax Ratings

Division I, Eight, Grand Final

pos.	School	Rating
1	University of Washington	2.5
2	University of Wisconsin	2.9
3	Brown University	3.8
4	University of California at Berkeley	5.2
5	Stanford University	6.5
6	Stanford University	6.5

Division I, Eight, Petite Final

7	Harvard University	7.6
8	Yale University	7.6
9	Northeastern University	8.4
10	U.S. Naval Academy	9.6
11	Columbia University	12.2
12	Cornell University	12.2

2008 IRA Actual Results

Division I, Eight, Grand Final

pos.	School	Times
1	University of Wisconsin	5:31.173
2	University of Washington	5:32.894
3	University of California at Berkeley	5:39.021
4	Northeastern University	5:42.399
5	Stanford University	5:43.324
6	Columbia University	5:44.565

Division I, Eight, Petite Final

7	Cornell University	5:38.888
8	Brown University	5:41.764
9	Yale University	5:42.848
10	Princeton University	5:44.423
11	Harvard University	5:45.851
12	U.S. Naval Academy	5:53.597

UCD	15.3	UCD	15.3	UCD	15.3
Iowa	14.6	Iowa	14.6	Iowa	14.6
UCF	14.8	UCF	14.8	UCF	14.8
SacSt	15.6	SacSt	15.6	SacSt	15.6
Iowa	15.7	Iowa	15.7	Iowa	15.7
Syracuse	15.7	Syracuse	15.7	Syracuse	15.7
Indian	17.0	Indian	17.0	Indian	17.0
MichSU	4.7	MichSU	4.7	MichSU	4.7
WSU	5.2	WSU	5.2	WSU	5.2
USC	5.3	USC	5.3	USC	5.3
Tenn	5.5	Tenn	5.5	Tenn	5.5
Stanford	5.9	Stanford	5.9	Stanford	5.9
Washington	6.6	Washington	6.6	Washington	6.6
NEastern	7.9	NEastern	7.9	NEastern	7.9
OhioSU	8.6	OhioSU	8.6	OhioSU	8.6
ND	8.7	ND	8.7	ND	8.7
Michigan	7.7	Michigan	7.7	Michigan	7.7
Clemson	7.9	Clemson	7.9	Clemson	7.9
Bucknell	8.6	Bucknell	8.6	Bucknell	8.6
Dartmouth	8.7	Dartmouth	8.7	Dartmouth	8.7
OregonSt	9.5	OregonSt	9.5	OregonSt	9.5
Louisville	10.2	Louisville	10.2	Louisville	10.2
Penn	11.0	Penn	11.0	Penn	11.0
Wisconsin	12.0	Wisconsin	12.0	Wisconsin	12.0
UCLA	12.1	UCLA	12.1	UCLA	12.1
Gonzaga	12.2	Gonzaga	12.2	Gonzaga	12.2
UCD	13.4	UCD	13.4	UCD	13.4
Minn	14.5	Minn	14.5	Minn	14.5
UCDavis	14.9	UCDavis	14.9	UCDavis	14.9
Columbia	15.4	Columbia	15.4	Columbia	15.4
UCF	15.8	UCF	15.8	UCF	15.8
Syracuse	16.6	Syracuse	16.6	Syracuse	16.6
Texas	19.1	Texas	19.1	Texas	19.1
USD	25.5	USD	25.5	USD	25.5
SacSt	26.0	SacSt	26.0	SacSt	26.0
UCD	25.5	UCD	25.5	UCD	25.5
UCF	26.0	UCF	26.0	UCF	26.0



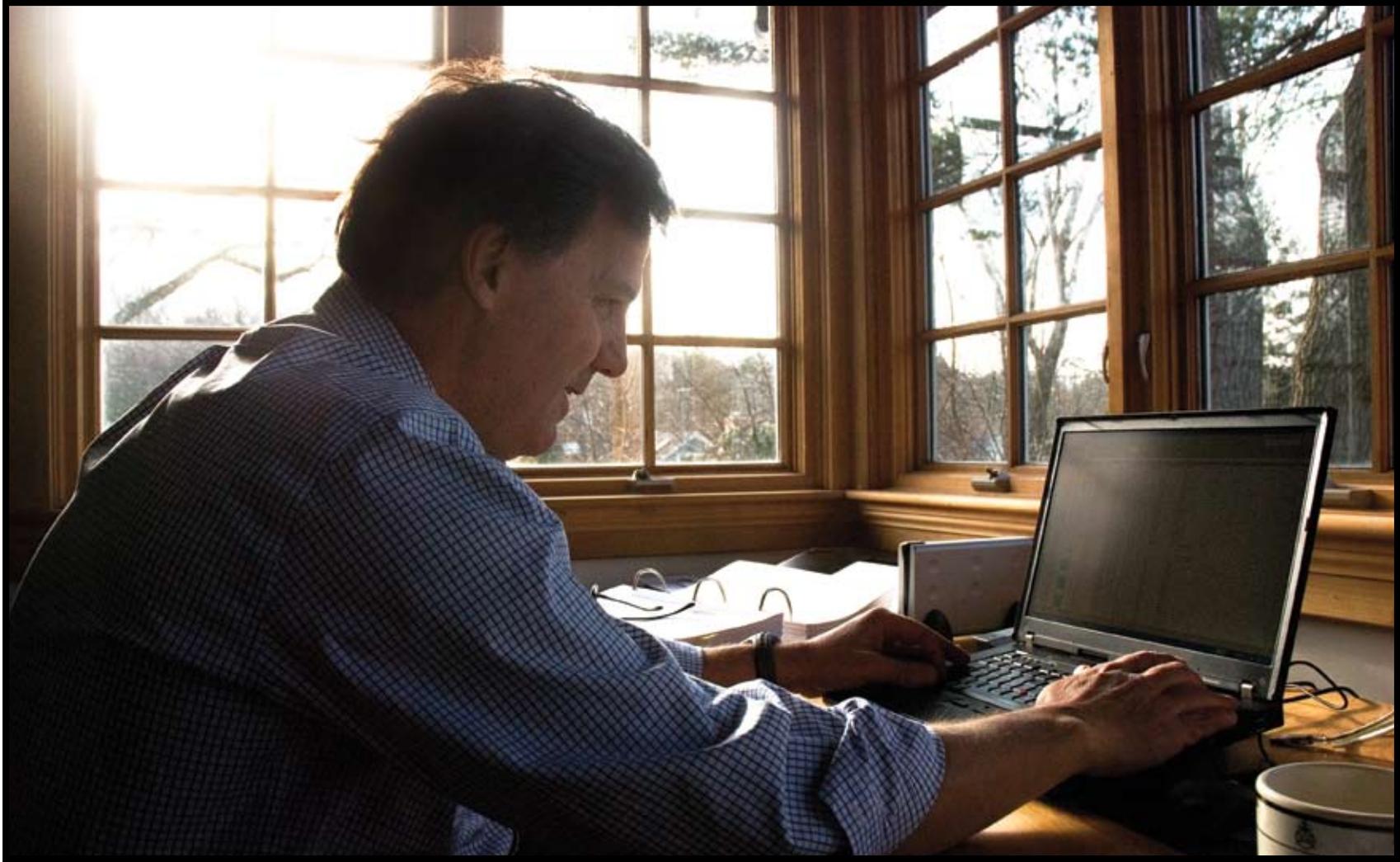
It all started with third grade girls' soccer 15 years ago," Maxwell tells me over a steaming cup of coffee at the Panera Bakery in Brookline's Coolidge Corner. He had been a microeconomics professor at Boston College in the 1980s and left for a 20-year stint as an economic consultant working on models for things like predatory pricing. His wife had become coordinator of the soccer league for one of their three daughters. "They had played five games and needed to schedule the rest of the season,"

Maxwell continues. He realized he could use his economic models to solve his wife's problem. "I rated the teams in terms of how many goals they had scored against common opponents, set the schedule for the rest of the season, and then went to watch how well I had done. The good thing about girls' soccer, at least in Newton, is that they make sure every kid plays the same amount in every game so there is a fairly consistent product on the field. I'd show up at games that were 2-0 in the first half and say, 'Don't worry, it'll be tied by the end of the game and it would be!' That was the start of cMax."

Next up was girls' basketball. "With three daughters, I coached 18 teams in total and

did all the scheduling for seven years. We'd play three games before Christmas, I'd run my model, and then match up the strong against the strong teams and the weaker teams against the weaker teams." It wasn't until his eldest daughter starting rowing at Andover that Maxwell turned his mathematical mind to the relatively obscure sport. During his daughter's sophomore year, Andover had a strong boat including a couple of junior national team members.

Maxwell knew nothing about the sport, but logged onto the boards on row2k. He gives a lot of credit for his cMax ratings to the quality of the data he found on the



rowing website. “A lot depends on how good the underlying data is. I already knew it was easy to find good data on lacrosse, on golf, on soccer, but without row2k’s collecting the data in one place my ratings never would have gotten off the ground.” The first model rated scholastic girls’ rowing in New England. When he shared it in the row2k community he was told, “Ah, you’re an idiot! You know nothing.” But in fact Maxwell was quite surprised by how well even his first crude models performed. “Going into the exercise, I didn’t think my model would do very well, but the fact is it did far better than I expected.” He was hooked. Models often produce nonsense results for long periods of time, but Maxwell’s model was surprisingly good from the very start.

At the time, Maxwell happened to work with Malcolm Baker, an Olympic oarsman in Atlanta, and Tom Murray, a member of the U.S. national team for more than a decade.

He still didn’t know much about the sport so he pestered his colleagues about how boats were put together and what made them fast. Maxwell attempted his first comprehensive collegiate cMax model in 2002, when his daughter was a junior at Andover. He printed out all the data in six-inch binders and key stroked it in by hand, the same way he still does. He sent the results to three friends. To Maxwell, it was just a way to learn about collegiate rowing before his daughter got there (she eventually went to Harvard). But to Murray, what Maxwell was doing was much more than that. Murray eventually contacted row2k about the cMax model.

Maxwell resisted going public with his model until 2004, when he finally agreed to let row2k publish the ratings. There was only one problem: for the first time, Maxwell’s results made absolutely no sense. He had developed the methodology of waiting until after the San Diego Crew Classic to post

his first rating. One of the most difficult parts of the model is attempting to compare East and West Coast teams and the Classic always gave Maxwell enough data points to get started. But he couldn’t figure out why, just on the cusp of going public, the rankings were trash in 2004. In the end, he decided to call off publishing for the whole year to figure out what had gone wrong. “The model is very sensitive to data errors,” Maxwell admits. “I finally discovered that I had entered a Texas crew wrong by a minute. Because they raced East and West Coast crews that data point was very important.”

The mantra for Maxwell’s model is “the rowing season is a search for speed.” His rating appears every Wednesday and gives the relative speed of every crew in the country for the next weekend. “The coaches are trying to do something different. I’m modeling speed. Whether you won or lost doesn’t matter to me. What matters is

how much you won or lost by. If you win by 0.1 second, the coaches would call it a great win, but I would consider it a basically a tie.” Maxwell points out that teams like the first Patriots’ championship squad, who won many games by three points or less but still won when it counted, are not captured by his model. Last year, Wisconsin beat Navy but not by as big a margin as he had predicted, for instance, so his model failed to predict that they would win the national championship. “I love being right,” he says, “but I love even more being proven wrong.”

Maxwell has developed similar models for football, basketball, ice hockey, lacrosse, golf, cross-country running, and squash. Of these, golf and cross-country are the most similar because every course is different and every race includes “pair-wise” matchups, even though they are not team sports. One of the particular challenges for rowing is the extent to which teams change things during the season—changing line-ups and combinations—in the pursuit of speed. Teams are also most focused on the final championship races, making early data

“You win or lose on the water not based on any rating. But I do a pretty good job of getting it right.”

points less significant. “I worry a lot about whether a boat is challenged,” Maxwell points out. If a boat is challenged that model is calibrated to pay more attention to that observation than if it’s a horizon job. Still, Maxwell points out that he had Western Washington winning the Division 2 championships by a big spread last year. And they did.

In the end, the NCAA championship is the barometer of the model’s strength; the regatta pits East and West Coast teams against each other, many of whom have not met face to face and share few common competitors. Last season, the model picked up on a subtle change in boat speed at the league championship, which knocked the California women out of the top spot in the

ratings just before the NCAAs. California remained at the top of the coaches’ poll going into the race. This time cMax was right and the coaches were wrong. California lost. “I get a lot wrong too,” Maxwell hastens to point out. “I had Brown beating Yale too, which didn’t happen.”

As we finish our coffee, Maxwell pulls out a sheet of paper with summary results for cMax’s 2008 performance. There were 896 races and 3,158 observations last year (in multi-boat races each pair of boats counts as one observation). cMax predicted the correct winner 84 percent of the time, meaning the rating from the prior Wednesday correctly predicted the following weekend’s results. If you took the final cMax model and ran it “retrodictively” back through all the data after the season was over, it would pick the correct winner 89 percent of the time.

“I’m just trying to get it right,” Maxwell says with a shy smile as we shake hands on Harvard Street. “You win or lose on the water not based on any rating. But I do a pretty good job of getting it right.” □

