THE WALL STREET JOURNAL.
WSJ.com

THE NUMBERS GUY | NOVEMBER 11, 2009

Statistical Time Travel Helps to Answer What-Ifs

Researchers Devise Systems to Explore How Supreme Court Justices and Baseball Players Compare With Their Predecessors

By CARL BIALIK

The debates often take place on barstools and can devolve into a bleary tedium: If Babe Ruth were alive today, would he still be among the greatest to have ever played baseball?

Such what-ifs are no longer limited to happy hour and disrupted by last call. In recent years, statisticians have created time machines to answer a wide range of historical hypotheticals, from how today's Supreme Court would have voted on Roe v. Wade to what sort of scientific papers Einstein might write today.

These inquiries derive from statistical techniques that aren't all that new. But their full force has been unleashed through leaps in number-crunching computing power. They have helped better explain how the past and its crucial players, from baseball sluggers to judges, would stack up in the present. Even if they are far from perfect as a result of some tricky assumptions, they are a lot better than guesswork.

"The famous statistician George Box once wrote that 'all models are wrong, but some are useful,' " Kevin Quinn, a professor of law at the University of California, Berkeley, who has studied changing attitudes of Supreme Court justices, said in an email. "I think that is a useful way to approach what we're doing."

Prof. Quinn and Andrew Martin of Washington University's law school used the time-machine techniques to track judges' ideological evolution and compare them to each other. The researchers were trying to identify where on the political spectrum, for each case, the judges and possible verdicts would lie. Then they would use that information to estimate the probability of how each judge would vote in each case. In that way, they used judges' past votes to construct a numerical model of probabilities that they could then apply to cases the judges never heard.

But how would the researchers compare two justices -- say Hugo Black and William Rehnquist -- who never served together? One answer was Justice Potter Stewart, who served with both and thus becomes a good reference point. By placing Justice Stewart on an ideological spectrum, they could compare him to both Justice Black, who was more liberal, and Justice Rehnquist, who was more conservative.

Such overlaps to bridge the gap between far-flung eras is at the core of these statistical time machines. They make use of a kind of bucket-brigade technique called Markov chains that allows the past to be

compared to subsequent events that ultimately reach the present.

The results of the Supreme Court study are striking reminders of how unpredictable justices' ideology can be over time, which many presidents have discovered. For example, William O. Douglas began his career as a liberal, moderated for several years and then became more liberal. Overall, their model had a good success rate, and was able to accurately classify three out of four votes.

In a related paper, Georgetown University political scientist Michael A. Bailey tried to characterize the political leanings of the other branches of government beyond the judiciary. He plotted justices alongside presidents and members of Congress, using Supreme Court cases as points of comparison.

The challenge was finding overlaps between the distinct branches of government. Prof. Bailey found them in congressional votes and public statements by presidents about court cases.

Such expressions of sentiment provided data to bridge the gulf between governmental branches, and also between politicians separated by decades. The study also helped devise an ideological scoring system that could track subtle changes in governing philosophy. For instance, Prof. Bailey's numbers show President Ronald Reagan becoming more conservative during his term, while President Bill Clinton also moved to the right. And every president was more partisan than the other federal branches.

These analyses are far from final words on political history. The greater the interval of time between two eras being compared, the more statistical errors introduced. And some judges don't fit neatly on a spectrum, such as someone who is economically conservative and socially liberal. Prof. Bailey's analysis also excludes cases on foreign policy and other issues, which leaves out many important functions of government.

Still, keeping in mind such caveats, these time machines can reveal how successful presidents have been at nominating Supreme Court justices who reflect their world views. Despite his difficulties with the Harriet Miers nomination, for example, George W. Bush has done well in this regard -- if his nominees don't change much. The analyses also provide a measure of how closely court ideology corresponds to that of the House of Representatives.

The analyses can also provide strong clues to how today's justices would have voted on abortion. The current court would have issued a 5-4 decision in Roe v. Wade that would have allowed states to ban abortion, the research shows -- though Prof. Bailey warns that such results should be taken with "a huge grain of salt and perhaps a tongue in cheek."

Such time machines don't have to be applied only to people. David Blei, a computer scientist at Princeton University, has been studying a vast trove of scientific journal articles. His system identifies topics from scratch and assigns topic scores -- say, 80% neuroscience and 20% philosophy, or 40% biology and 60% chemistry. Any papers that have the same topic scores could then be grouped together, even if they are decades apart and keywords or concepts didn't yet exist. (Think of quarks or H1N1.)

Here the critical bridge -- the necessary overlap to relate past decades to the present -- were keywords that were associated with others before they faded. In other words, they performed the same bridging function Justice Potter Stewart did. Such techniques connected an 1880 paper on orangutan brains with a 1976 paper on monkey brains.

That technique helps dig up research that was ahead of its time. For instance, these very time machines, including Dr. Blei's, make use of so-called Bayesian statistics, which were developed decades before

there was sufficient computing power to use them fully. "Part of science is uncovering long-forgotten discoveries," Dr. Blei says.

For sports, the time machines can help settle the debate between nostalgic fans and those who insist the games have never been played better. A paper a decade ago attempted to bridge the gap between baseball players over the past century. Co-authors Shane Reese and Scott Berry found that performance was improving, but not uniformly. For instance, baseball players' improvement in hitting home runs was much more dramatic than their batting-average gains -- perhaps because of steroids, which wasn't controlled for.

Researchers for Baseball Prospectus, a sort of think tank for the diamond, also quantified how the sport got more difficult over time. The analyses found that Babe Ruth would continue to shine in today's game.

Mr. Ruth would have had a lower batting average, but he would have hit 199 more homers, making him the greatest home-run hitter ever. End of debate -- nearly.

Write to Carl Bialik at numbersguy@wsj.com