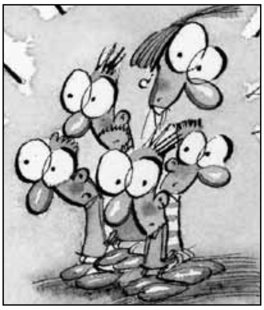
Who Gets What – and Why

BY ALVIN ROTH

A Alvin Roth, the author of this excerpt from *Who Gets What – and Why*,* won a Nobel in economics in 2012 for his work on the "the theory of stable allocations and the practice of market design." I know, I know: that's hardly an intro likely to induce you to dive right in. Most Nobels in economics, after all, are awarded for accomplishments that are too arcane for mere mortals to comprehend. And even the prize winners who do have something pressing to say to the public can rarely write their way out of that proverbial paper bag. ¶ But Roth and this book are spectacular exceptions. While he was really trained as a mathematician (his PhD is in a discipline called operations research), Roth's vision has never strayed far from the practical. And he's a natural-born writer to boot. ¶ Roth designs "matching markets," where price alone can't balance supply and demand – think of markets for everything from marriage to college admissions. Indeed, he's even saved lives by helping to design an ingenious way to match more donated kidneys to needy patients. ¶ The chapter excerpted here will give you a taste of his fine mind and formidable ability to make complicated ideas comprehensible. — *Peter Passell*



To understand the many ways in which markets fail, we must begin even before the beginning.

Part of making a market "thick" involves finding a time at which lots of people will participate at the same time. But gaming the system when the system is "first come, first served" can mean contriving to be earlier than your competitors.

That's why, for example, the recruitment of college freshmen to join fraternities and sororities is called rush. Back in the late 1800s, fraternities were mostly social clubs for college seniors. But in an effort to get a little ahead of their competitors in recruiting, some started "rushing" to recruit earlier and earlier. Fast forward to today, when it is first-semester students who are the targets of fraternity and sorority rush.

And that's not the only way the rush to transact sooner has entered the English language. It's also the reason that Oklahomans are called Sooners. The nickname was born on April 22, 1889, the beginning of the Oklahoma Land Rush, and truly entered the American vernacular four years later, on September 16, 1893, the height of the rush, known as the Cherokee Strip Land Run. In both cases, thousands of people – 50,000 in 1893 – lined up at the border of the former Indian territory and, at the sound of a cannon shot, raced off to stake out free land.

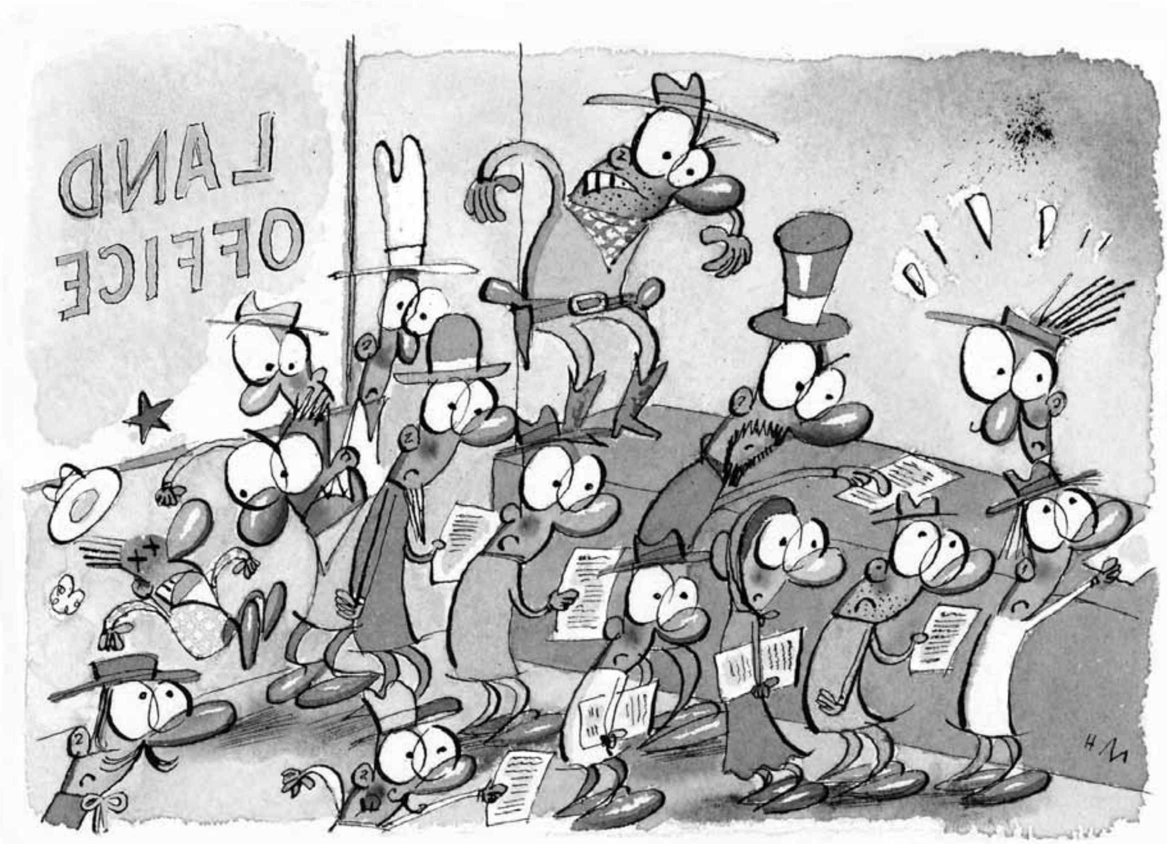
At least that was the plan. And most participants abided by the rules – not least because the U.S. Cavalry was patrolling the Strip with orders to shoot anyone found in the open territory or crossing the line before the signal sounded. To prove the cavalry's seriousness, when one unfortunate soul – perhaps confused by a pistol shot – took off early, they rode him down and shot him dead to the horror of thousands of onlookers.

When, finally, the cannon roared, those same thousands – on horseback, in wagons and even in carriages – surged forward in the most famous photographic image of the era. Fifteen miles away in what would, by afternoon, be the bustling municipality of Enid, America's newest city stood the only public building in the Strip, a land office/post office. About noon, the assistant postmaster, Pat Wilcox, took his binoculars and climbed up on the roof of the building. Looking south, he saw a lone rider, a 22-year-old cowboy named Walter Cook, appear on the crest of a low hill. Tearing toward him, and then rushing on past, Cook jubilantly staked his claim to a plot of land at the very center of the planned city.

Cook had played by the rules, waiting until he heard the signal to take off. But lots of other people, despite the cavalry's draconian efforts, had crossed the line earlier. These "claim jumpers" would come to be called Sooners, for their timing. And in the long tradition of turning pirates, bank robbers and other brazen criminals into lovable rogues, that would also become the nickname for all Oklahomans, and eventually for the University of Oklahoma's football team.

Gaming the system by entering the Oklahoma Territory to stake a claim before September 16 was illegal, but that didn't prevent it from happening. And claim jumping wasn't the only thing that didn't go according to plan in the course of that crazy day.

Take poor Walter Cook. His claim was quickly overrun by 300 false claimants to the



HAL MAYFORTH

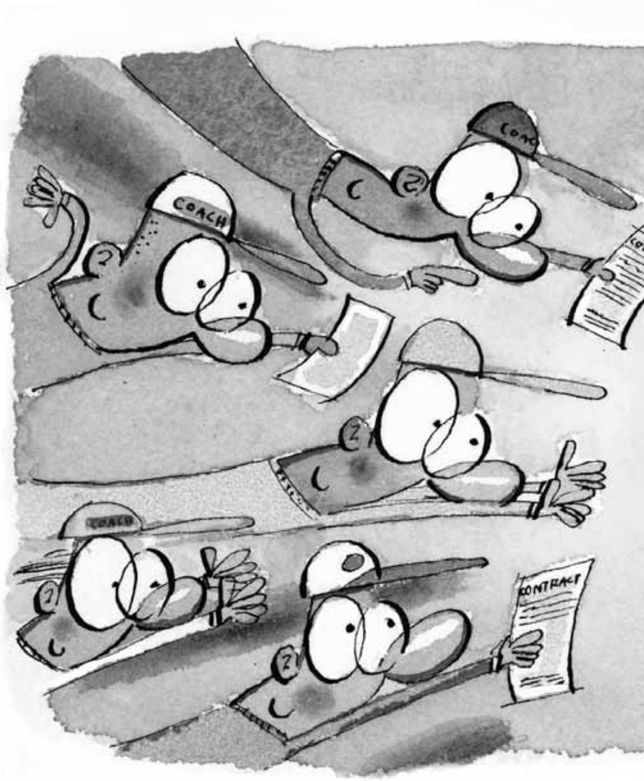
same plot of land, all taking advantage of the fact that the law wouldn't arrive for hours to validate anyone's claim. In the end, Cook got nothing except a lesson in the dangers of a poorly regulated, lawless market.

Cook might have had a chance if the land office had been open when he arrived and had processed his claim quickly. But instead, the line at the office quickly grew to hundreds of claimants, then thousands from throughout the Strip. Fights broke out; robberies occurred; at least one person died of a heart attack.

There were at least two ways in which the allocation of land failed to work well that day. First, the law-abiding citizens who followed the rules were often preceded by those who entered the territory sooner and marked their claims earlier. Second, the fact that those claims all had to be recorded at the land office in Enid on the same day led to congestion and confusion in which even some of those who had arrived in time to stake a claim, such as Walter Cook, couldn't get it recorded. The market wasn't fast enough to deal with all the claims made that day, and so it couldn't always sort out which claims came first.

Sometimes the problems of going too soon are subtler. Jumping the gun – in Oklahoma it was literally a cannon – can cause potentially thick markets to unravel. They become thin when too many participants try to transact before their competitors are fully awake and present in the market.

Let's look at those other Sooners, the ones who play football for the University of Oklahoma. We turn to college bowl games to see how "too early" can ruin a matching market's ability to make good matches. The matching of football teams to play in the big end-of-season bowl games suffered for many years because the teams that would play in those games were chosen too soon to make for good matchups.



HAL MAYFORTH

MAKE ME A MATCH, CATCH ME A CATCH

For those who love college football, there is no time of the year more exciting than bowl season, when the top teams from different conferences meet to determine which are better and, ultimately, to determine the national champion. But most college football fans have come to believe – and sometimes argue vociferously – that the system is broken. And unfortunately, they're right.

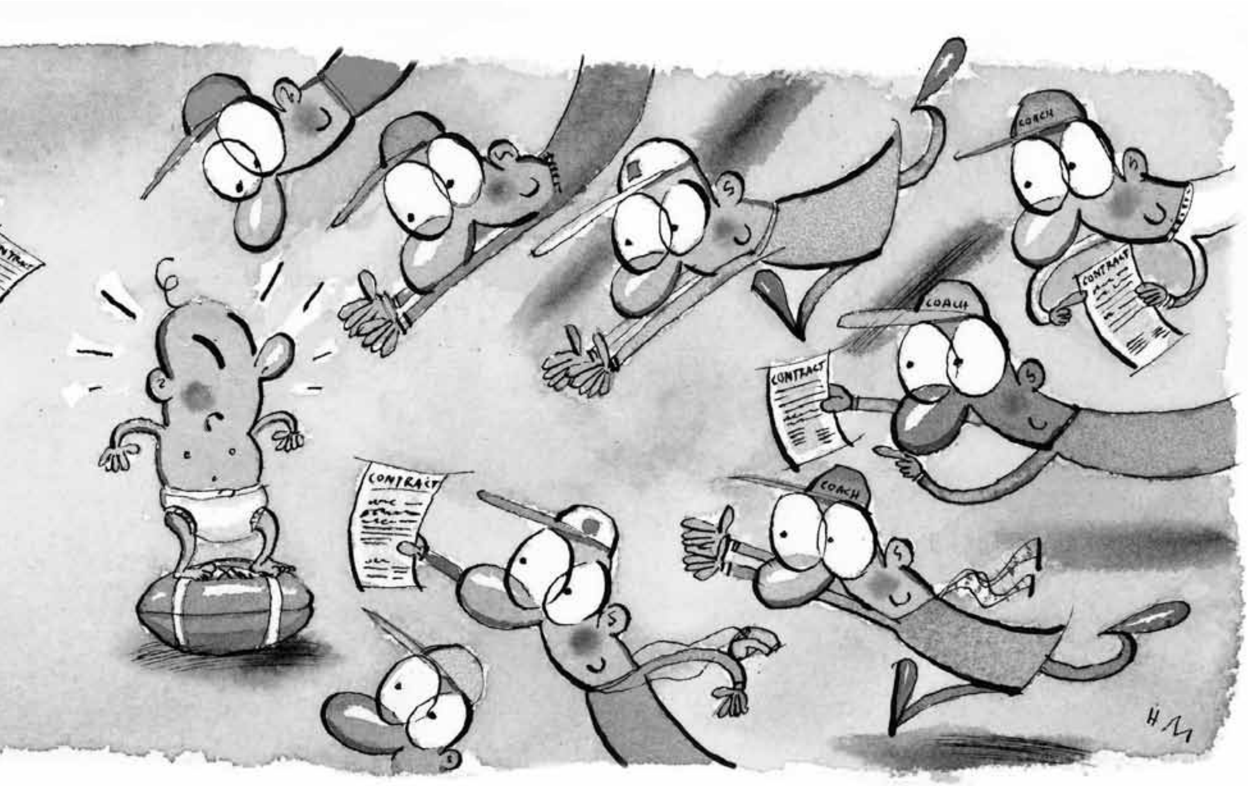
For a long time, teams and bowls succumbed to the temptation to do deals early. And while college football isn't the most important market in the world (except to its fans), the fact that new information is available every weekend about which teams have won or lost, and that those teams are then ranked according to polls of sportswriters and coaches, shows very clearly how important information can be ignored when the market moves before the results of the final games of the season have been played.

As the television audiences and advertising revenues became important, bowl committees began to recruit teams to play in the bowl games earlier and earlier – indeed, so early that the teams they recruited were sometimes, after an unexpected loss or two, no longer championship candidates by the time the game was played. That's one of the dangers associated with early transactions: they can come well before important information is available. And that can mean bad matches made and good ones missed.

Which teams get to play in which bowls is handled differently today than it used to be. Sports fans can disagree about how well the current system works, but no one disagrees about how badly it used to work before the market was redesigned.

The bowls are independent businesses that control a stadium and make contracts with television networks and corporate sponsors. Each would like to host a postseason game between the two teams that are ranked best in the country at season's end. For many years, the National Collegiate Athletic Association tried to make bowls and teams wait long enough to get good matches. But it consistently failed to do so, and after the 1990-91 season, it gave up trying.

That season there were 19 postseason bowls. The one that paid teams the most was the Rose Bowl, which was "closed." It had a long-term contract with the Big Ten and Pacific-10 football conferences, and each year the champions of those two conferences played each other in the Rose Bowl (and the two conferences shared the bowl revenues of their



champions). So the Rose Bowl wasn't involved in the unraveling we're examining here; it merely waited until the conference champions were determined.

But the other bowls had different arrangements. The Fiesta Bowl faced a unique challenge: as an "open" bowl, it needed to find two teams to play. The other top bowls were all "semi-closed"– that is, they each had a contract with one football conference, whose champion would be one of the teams that would play. Meanwhile, each of these bowls needed to find one additional team to provide quality competition. The available pool consisted of teams that were not in any football conferences (independents) or were in conferences that were not contractually tied to any bowl.

In 1990, the NCAA rule was that teams and bowls couldn't finalize bowl matchups until "pick-'em day," which that year was November 24. But some bowls and teams went ahead and made earlier arrangements. Notre Dame, an independent, had begun the season as the number-one ranked team and had recovered from an early loss to regain that position by early November. Meanwhile, Colorado had overcome an early season loss to become the number-four team in one of the rankings and number three in the other. When Colorado beat Oklahoma State and clinched the Big Eight conference championship, Colorado was ensured a berth in the Orange Bowl and rose to number two in the rankings.

The next day, 13 days before pick-'em day, an agreement was announced between the Orange Bowl and Notre Dame. This meant that the currently first- and second-ranked teams in the nation would meet in the Orange Bowl, and thus make that bowl the de facto national championship.

Virginia's acceptance of a bid from the Sugar Bowl to play the still-to-be-determined Southeastern Conference champion was announced the same day. And following the Orange Bowl agreement, the University of Miami agreed to play in the Cotton Bowl against the still-to-be-determined Southwest Conference champion. At this point, Notre Dame, Virginia and Miami all still had a few games to play. In college football, alas, a few games is forever.

The unraveling didn't stop until the conferences and bowls figured out new rules that removed the incentives to determine bowl matchups before the final rankings were known.

Sure enough, shortly after inking its agreement, Notre Dame lost a game and finished the regular season ranked number five. Meanwhile, Virginia, which had lost only one game before its agreement with the Sugar Bowl, lost two games and finished the regular season unranked in one poll (meaning it wasn't even in the top 25) and number 23 in the other. In the end, no bowl succeeded in getting the number one and number two teams (which turned out to be Colorado and Georgia Tech).

Thus, when the bowl games were over, there was no consensus national champion: Colorado was ranked first in one poll and Georgia Tech in the other. Since they hadn't played each other, the sportswriters and coaches who were surveyed for the national ranking felt entitled to their own opinions.

Faced with such a public failure to enforce pick-'em day, the NCAA abandoned the attempt for the 1991-92 season. The Football Bowl Association (FBA) responded with an attempt to enforce a pick-'em day of its own and voted to levy a fine of $250,000 on any member that violated this understanding. However, the FBA was no more successful than the NCAA, and not surprisingly the 1991-92 bowls also failed to produce a matchup of the top-two teams. Once again the postseason ended without producing a consensus national champion.

In retrospect, it's clear that several problematic market-design features prevented good bowl matches. Because the Rose Bowl dealt with only two conferences, these conference champions risked not being ranked close to each other and would seldom be the two highest-ranked teams nationally. But at least the Rose Bowl had a contract ensuring that the two teams that played each year would be the champs of their conferences.

The other major bowls enjoyed a substantial pool of conferences and teams from which to pick. But because of an unraveling of bids for their open slots, most were filled without the bowl committees knowing the end-of-season rankings of the teams invited to play. And because many bowls had one position reserved for a particular conference champion, this limited the matching flexibility of each of them and of the market as a whole.

It wasn't simple self-restraint that stopped colleges and bowls from going early, nor could a powerful organization like the NCAA stop them. In the end, the unraveling didn't stop until the conferences and bowls figured out new rules that removed the incentives to determine bowl matchups before the final rankings were known. They did this through a series of incremental, almost yearly reorganizations of the market, designed to make more teams available to be matched after the regular season – that is, to make the postseason market thicker.

One way to do this was to enlarge the football conferences, so that the champions of each conference would be the best of a larger group of teams. In 2011, the Pacific-10 conference became the Pacific-12 conference. The Big Ten kept its name but not its number: in 2011 it also expanded to 12 teams, and then expanded again to 14 teams for the 2014-15 college football season. In addition, coalitions of bowls formed to make the market thicker, and eventually the Rose Bowl joined with the other major bowls to create the Bowl Championship Series (BCS) in 1998.

Now the number-one and number-two teams in the country, determined according to the BCS ranking system, played a national championship game, and this game was rotated from year to year among the participating bowls.

The fact that teams and bowls were matched later in the thicker BCS market doesn't necessarily prove that the market was working better. It's often hard to get a quantitative measurement of how well a matching market is performing in some ultimate sense – for instance, how much social welfare it is producing, beyond how well it is serving the participants in the market itself. But if we think of football games as entertainment, then how many people decide to watch the games isn't a bad measure of how well the market is working.

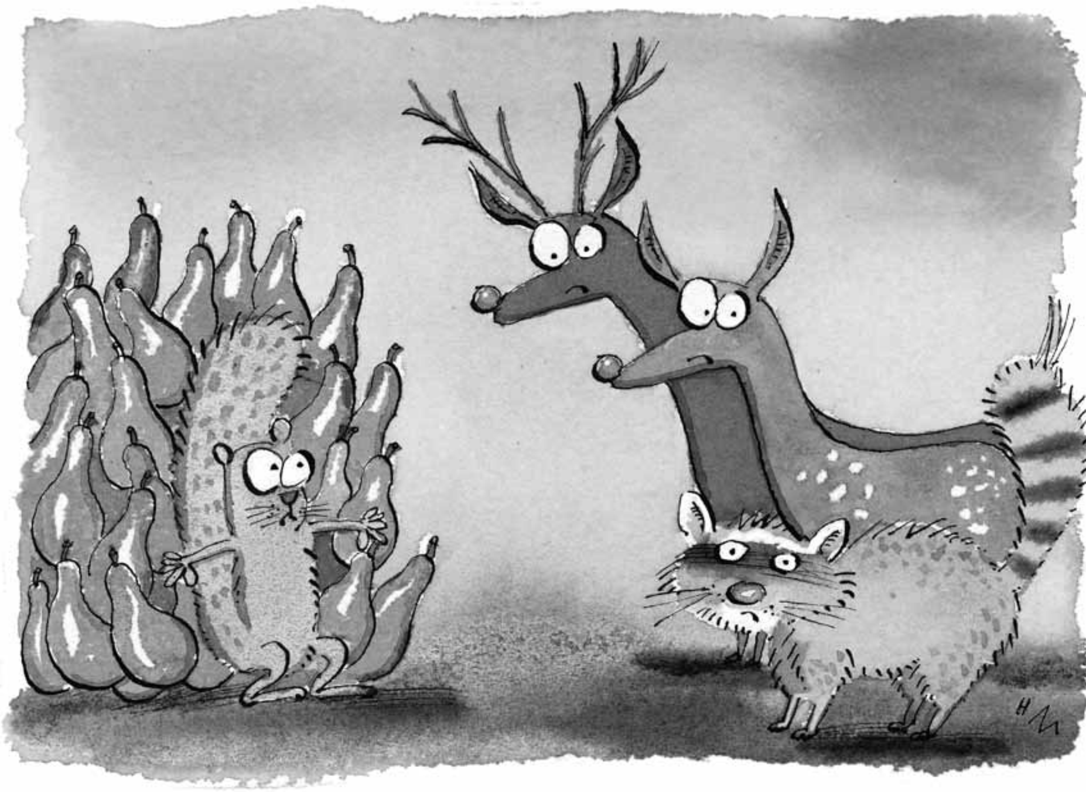
When Guillaume Fréchette, Utku Ünver and I looked at the Nielsen ratings for the televised bowl games over the years, we found that a game between the teams ranked first and second in the nation attracted so many more viewers that it was well worth it for the bowls to rotate such a game among themselves. This is why the BCS worked well when there was a consensus number-one and number-two at the end of the season and less well when there wasn't.

As I write this in 2014, plans are under way for a postseason playoff designed to more reliably produce a championship game that will attract many viewers. Four teams will be selected for the College Football Playoff, with the winners of the semifinals to meet for the championship game. The proposed new playoff model still has some of the old weaknesses of the BCS, but the uncertainty about which four teams to include should be less consequential in picking a national champ than uncertainty about which two teams to include.

The slow, incremental process by which the market for bowls evolved can be viewed as a kind of cultural evolution. Different bits of practice were reshaped over time, in ways that kept all the big players – the successful teams and conferences, the big bowls, the television networks – in business. Lots of interests had to be addressed to achieve any sort of coordination at all and to get some forward motion. Like football itself, that forward motion mostly came a yard at a time.

Incidentally, it's not just football teams that make early matches; often it's the players, too. For example, in 2012, Louisiana State University offered a football scholarship to Dylan Moses, a 14-year-old who had not yet begun the eighth grade and who wouldn't enter college for another five years. Who knows whether he'll be big enough, healthy enough and accomplished enough to play when he is finally old enough for college. But coaches worry that all the other teams are recruiting early, and if they don't do the same, they could miss out on a future star.

This "sooner" mentality isn't limited to the more high-profile college sports. When I meet varsity athletes at Stanford University, where I work, I ask them when they first met



HAL MAYFORTH

a Stanford Cardinals coach. So far, the earliest answer I've gotten was from a player on the women's basketball team, who met a Cardinals coach when she was in sixth grade. She hastened to add that she'd been a very tall sixth grader on a team with older players, and that the coach had been surprised to hear that she was only in the sixth grade. He was scouting eighth graders.

RUSH TO GLORY

Rushing to be sooner isn't just something in the history books or on the sports pages. If you know a recent college graduate who recently took a job with a big investment bank such as Goldman Sachs, there's a good chance that she'll get a call soon after beginning work. It will be from a big private equity firm such as Kohlberg Kravis Roberts, interested in signing her to a contract that would take effect after she's worked for Goldman for two years. And if you know someone who just graduated from law school and works for a big American law firm, he was most likely hired by that firm initially as a summer associate, about two years before he earned his degree.

Is this a good idea?

Remember the 1991 Orange Bowl. The same thing that happened then can happen to law firms that recruit years before their future employees earn their degrees. That top-notch first-year law student can go through a lot of changes over the next two years. At least the Orange Bowl selection committee knew how many teams (two) it would need for the game. Not so for law firms, which have to guess two years in advance how many lawyers they will need. Guess wrong, and they could be in a lot of trouble.

When a market's organization predictably leads to trouble, economists start asking whether it might be inefficient, meaning that a different organization might make everyone involved better off. We've already seen that going early can create bad matches, but it also could be that this approach benefits some people while hurting others. The market for new lawyers lets us see how unraveling can sometimes hurt everyone.

In particular, just about everyone could have been better off if that market had been less unraveled during the Great Recession, which reduced corporate demand for outside legal services. Hiring more than a year before the start of employment made it difficult for law firms to forecast their needs. Thus thousands of summer associates at large firms who'd accepted "permanent" offers shortly after their second-year summer associateships in August 2008 saw them rescinded or deferred before they started work in the autumn of 2009.

To maintain their reputations and their relationships, some of the firms paid those deferred employees a portion of their starting salaries and encouraged them to spend a year doing pro bono work – an outcome that was costly for both sides of the market.

If this two-year head start sounds bad, consider that in the late 1980s hiring was even earlier, with some students getting offers of summer associate positions right after they were accepted to a top law school, before they'd even taken their first class. Those firms undoubtedly would have liked to see how their prospective hires actually did in law school. But they worried that if they waited, other firms would snap up the best talent before them. So they told themselves that if Yale Law School wanted a student, that student also had a strong chance of becoming a good lawyer – just as at mid-season Notre Dame had a good chance of being number one when it played in the Orange Bowl.

If making offers very early makes it hard to identify good job candidates, you might think that some firms would take a little more time and make offers to candidates who had already received at least one offer from another firm. But the firms that made early offers prevented this by making their offers "exploding" – that is, take-it-or-leave-it offers of such short duration that they didn't leave enough time for another firm to jump in and compete for the same candidate, or for a candidate to get another offer for comparison.

Exploding offers are common in unraveled markets. These offers are both early and short-lived. So not only are firms making offers before they have as much information as they'd like about how candidates might perform in school, but the candidates themselves are confronted with accepting or rejecting an offer before they know what other offers might become available.

To put it another way, exploding offers make markets thin as well as early. So participants are deprived of information about both the quality of matches and what kind of matches the market might offer. In that situation, nobody has enough information to make an optimal decision.

More than the other sources of market failure that we'll explore, unraveling is a failure of self-control. Participants just can't stop themselves from transacting early, because if they resist the urge, they'll lose out to someone else. It's a little like what happened when my family planted a pear tree in our yard in Pittsburgh, right next to a wooded hillside. Each year, long before the pears were ripe, some squirrel would take them. I don't know whether squirrels like unripe pears or they just feared that if they waited any longer, the raccoons or the deer would get them.

Now, if a market is behaving badly and producing an inefficient outcome, it makes sense for participants to get together (if only for their own preservation) and design new rules to make the market work better. That's what happened in the 1980s. Student organizations, law schools and law firms supported a rule-making organization called the National Association for Law Placement (NALP),

It's hard to make rules constraining lawyers, because many lawyers earn their living by obeying the letter of the law while evading its intent.

which tried to bring some order to the lawless market for lawyers.

Because lawyers like precise rules, looking at these rules gives us a unique window on why unraveling is so hard to control. One rule was meant to give brand-new law students a chance to learn a little law before being confronted with an exploding offer from a law firm. This rule said that if an offer was made to a student who hadn't yet completed the first year of law school, that offer had to remain open until the end of the first semester, in December.

Unfortunately, it's hard to make rules constraining lawyers, because many lawyers earn their living by obeying the letter of the law while evading its intent. So this rule worked for a year or two, until some lawyer, in charge of hiring for his firm, had the bright idea of writing an offer letter that said, essentially, in keeping with NALP guidelines, this offer remains open until the end of the semester. But, the letter continued, the job didn't come with much of a salary. There was a handsome signing bonus, however, which would bring the salary up to the usual level. But that signing bonus would be paid only if the offer was accepted immediately.

Regulating the market for new lawyers soon became an arms race between the rule makers and the rule breakers. As of this writing, the most recent NALP rules say that exploding bonuses are against the rules, too.

RUSH TO JUDGMENT

At least lawyers and law firms make a show of obeying the rules while seeking ways around them. In the most prestigious part of the market for young lawyers, that of federal appellate judges hiring top students as law clerks, many judges openly flout the rules. Or perhaps a more "judicious" way to put it is that federal judges think they can make up their own rules.

Clerking for an appellate court judge is the classiest first job an ambitious young lawyer can have. For one thing, it's a ticket to the kind of career that makes people want to be lawyers. That's one reason the obituary of a retired senior partner of a big law firm often mentions his clerkship of decades before. (The first sentence might read: "Clancy Goldfinger, former managing partner of Catchum, Killum and Eatum, who graduated from Harvard Law School in 1951 and clerked for Judge X and Justice Y, passed away Tuesday.")

So there's a lot of competition among the best students at the elite law schools to clerk for one of the relatively few federal appellate judges. But at first glance, the clerkship market doesn't look like one that should experience unraveling, although it's easy to see why a law student would be tempted to accept an early offer from an appellate judge. Since there are so few judges and so many law students, however, every judge could get a very well-qualified clerk if only he or she would wait to see which law students did well.

But while there are only a small number of appellate judges, those judges realize there are an even smaller number of law students who will win the top awards at their schools or be elected to edit their law reviews. And those appellate judges are organized into circuit courts, not all of which are equally prestigious. Neither are all judges within a given circuit equally likely to have their clerks move up to the U.S. Supreme Court for a second, even more prestigious clerkship there.

So if all judges waited to recruit only third-year students as clerks, when it is clear who will be a law review editor or top student, only the most prestigious judges would be able to hire the best students from the handful of elite law schools. That's a very good motivation for slightly less prestigious judges to make offers before the students' third year.

It takes a brave student to turn down an offer from, say, a judge in the Ninth Circuit Court of Appeals (which covers all of California and more) in the hope that if she waits, she might get an offer from the even more prestigious D.C. Circuit. That could happen if she's lucky. But if she's only a little less lucky, she may have to settle for a much less attractive job than the one she's just been offered – and that she has to accept immediately or not at all.

Of course, the judge is gambling, too: a student who looks likely to win law school honors may fail to do so, and may turn into a clerk who won't live up to her early promise. If the market ran later, the matching of students and judges would be more predictable, with the top jobs reliably going to students who had earned the top honors.

Notice that the law students who get these early offers are hardly facing the prospect of unemployment. But that doesn't mean that they aren't facing difficult decisions. There will be positions for them even if they wait, but maybe not such good positions. They have to make quick, strategic decisions taking into account what the rest of the market is doing.

WEDDING BELLS' TOLL

Few of us will ever get an offer to clerk in a federal court of appeals. But once you understand this kind of strategic decision-making, you'll begin to see it all around you, from marrying to finding a parking spot. Quite a few of us may face such a dilemma when deciding whether to marry a current girlfriend or boyfriend, or to break up in the hope of finding a better match later. That's a different decision when the market is thick, such as when you're in college and there are lots of single people your age, than when the market is thin, such as when most of the people your age are already married. And some marriage markets are tougher than others.

Consider the teenage Bedouin bride in whose community polygamous marriages are common. Just such a woman lamented, "If you are 20 or older, you may be married as a second wife."

But even teenage brides don't face the earliest marriage decisions in the world. In some times and places, the marriage market has unraveled to the point that newborns are betrothed. In developing countries, it isn't unusual to find marriages arranged quite early, particularly for women, and particularly in places where women are in short supply because men compete for multiple wives. Some countries, such as India, have tried to stop this practice with minimum-age marriage laws. But those laws have proved difficult to enforce because private and informal matchmaking arrangements have emerged.

In searching for a striking example of unraveling, Xiaolin Xing and I considered places where child marriages occur, and even primitive societies in which unborn children may be betrothed. The most striking example we found involved the Arunta, an aboriginal people of Australia. Because the Arunta were polygamous, there was a shortage of women.

Marriages among the Arunta were frequently arranged by two men, one of whom had just fathered a baby boy and the other a baby girl. When two such men met to arrange a marriage, however, the union they were arranging wasn't between those two babies – it was much too late for that because the baby girl's marriage had already been arranged. Rather, the two fathers were agreeing that the baby boy would marry the first daughter of the baby girl. That is, they were agreeing that the infant girl would become the mother-in-law of the infant boy. This was a marriage arranged by the father of the infant boy on behalf of his son and the father of the infant girl on behalf of his granddaughter by his infant daughter.

The timing of transactions depends not just on what is available now, but what is likely to be available later.

In Arunta society, marriages could be transacted more than a generation in advance of when they would be consummated. You can understand how, as a responsible young father, you wouldn't feel safe letting your son's – or your granddaughter's – marriage arrangements lag behind their competitors'.

Notice that in many developed countries, ages at first marriage are increasing, not decreasing. As more women seek higher education and professional careers, they wait to get married. When I say it that way, I'm focusing on the choices made by women. But a woman can't simply choose a spouse, and neither is the choice of when to get married entirely an individual decision for either men or women.

Think back to the days when few women went to college. In 1947, for example, there were more than twice as many men as women in American colleges. A lot of people eventually married their high school sweethearts because high school provided a thick marriage market in which one could find a lot of single people of the opposite sex, and those opportunities wouldn't be so abundant later.

By 1980, many more men and women went to college, and in equal numbers, so there were opportunities to make a match there, too, and the pressure to marry early was reduced. Today, the growth of Internet dating sites also offers the possibility of a thicker marriage market for college graduates. Postponing marriage when there is still a thick market in the future isn't so risky, and more-mature brides and grooms might have a better chance of recognizing a good match.

So the timing of transactions depends not just on what is available now, but what is likely to be available later. When you're driving down a crowded street hoping to find on-street parking, you regularly face (for lower stakes) a decision like the one confronting a law school student with an exploding offer or someone thinking about marrying his high school sweetheart. You see a vacant spot while you're still some distance from your destination.

Should you take it? This spot probably will be taken before you can loop around and return, and you may have to settle for parking even farther away or in an expensive garage. Or should you risk waiting for a better choice: a spot right in front of your destination? That

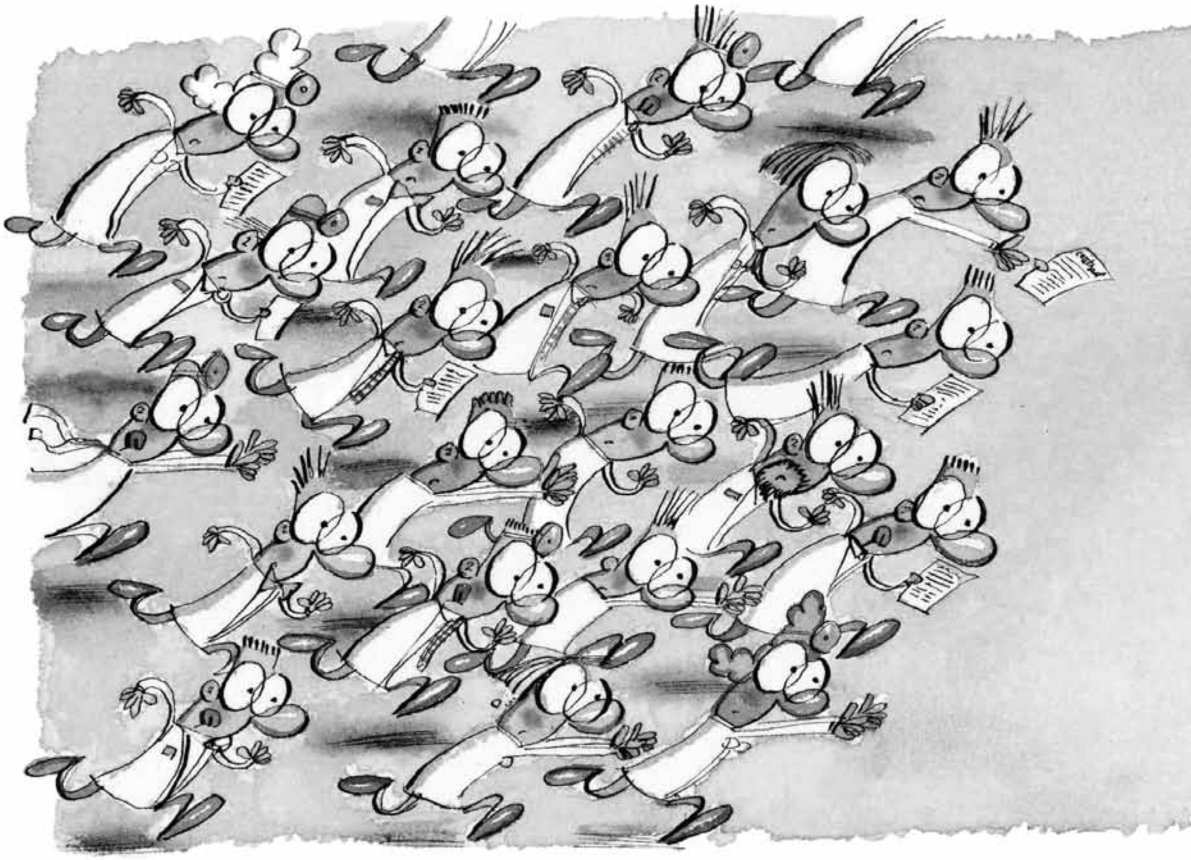
HAL MAYFORTH

would be a safer choice if you knew that lots of parking was available near your destination.

As you can see, unraveled markets aren't hard to find. We've just seen unraveling in matching markets ranging from sports to law to marriage, and in simple choices from whether to wait until a pear is ripe before picking it to how far from your destination to grab a parking spot.

But this kind of market failure wasn't widely remarked on when I first noticed unraveling in the 1980s, while studying the market for new doctors. Back in the 1940s, medical school students had to line up their first positions two years before they were due to graduate from medical school. Looking at it from the other perspective, hospitals had to hire their new interns and residents from a pool of medical students who hadn't even started on the clinical portions of their medical school education. Each side felt – correctly, given the circumstances – that if they didn't move quickly, the good positions and the top students would already be snapped up. It was a mess.

I first thought that unraveling was unusual, a kind of rare accident that happened in relatively special markets such as that for doctors. But as we've seen, lots of markets unravel. In fact, it's an even more widespread phenomenon than I've suggested thus far. For example,



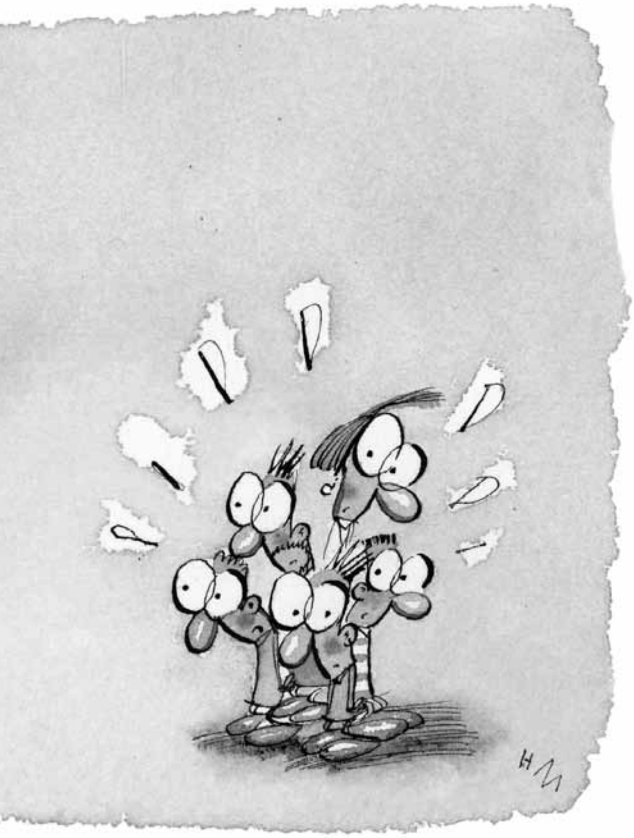
many selective colleges now fill more than half of their freshman slots through "binding early admission," a kind of exploding offer in which students apply early and commit to attend that school if accepted, without applying to other colleges.

Meanwhile, some private schools even enroll students at birth. At the Wetherby School in England, a school that Princes William and Harry attended, the spaces reserved by newborns fill up early each month, and the school advises women scheduling Cesarean sections to have them on the first of the month, if possible, to get a place before all the spots are gone.

In fact, unraveling is an ancient problem. In medieval England, it was sometimes a crime, called "forestalling," to trade before the official opening time of a market. It's not a crime today, but try telling that to the vendors who come to the farmers' market near my home and refuse to sell to me if I show up before the official opening time for fear of starting a race with the other sellers about who can set up first.

Those farmers are exerting self-control, maybe with a little help from the city, which licenses them to use the street only between certain hours. But unraveling can't always be limited by self-control. Even if a law firm can control itself and delay its own hiring until some reasonable time, if the firm's competitors hire earlier, it could be caught short. This fear of being left behind is part of what turned pioneers into Sooners.

HAL MAYFORTH



So unraveling is hard to control in a market that relies on self-control. Even if you have a lot of self-control, all you need is the suspicion that other participants might jump the gun and you will do so, too. It would be irrational not to. In many markets, what we see at first is a slow unraveling that suddenly tips over into a mad dash. It is often only then, when the participants see the profits from going a little early now getting swallowed up by the costs of racing with everyone else to go very early, that a consensus finally grows in support of reversing this unraveling. That's when new market designs might be considered.

Let me tell you about a simple market design solution that halted and reversed unraveling in a market in which unraveling was the only market failure that remained to be solved. The trick was to remove the need for self-control among those tempted to make early offers by handing over some control to the people on the receiving end of the offers.

FINDING THE GUTS TO WAIT

If you're under 50, you probably don't have to know what gastroenterologists do. Let's just say they are doctors who look after your digestive system. And after you turn 50, you're supposed to visit them so they can look for early signs of colon cancer.

To become a gastroenterologist, a doctor must participate in what's called a fellowship, which takes place following his first job, or residency, after graduating from medical school. The market for medical residencies was the first unraveled market I studied. Today that market is no longer unraveled, and new doctors are matched to residencies during their last year of medical school, in a market that is thick, uncongested and safe.

The medical residency that future gastroenterologists must complete for is in the field of internal medicine and takes three years. So gastroenterology fellows could, in theory, be hired after they have three years of medical experience. Unfortunately, the unraveling of the fellowship market led hiring to creep back earlier and earlier, until first-year residents sometimes found themselves being interviewed for jobs that they might begin two years in the future. Once again, this could be costly, both for fellowship directors hiring fellows while they were still inexperienced, and for young doctors having to choose a subspecialty before they'd had time to learn what they liked.

When my colleague Muriel Niederle and I studied the unraveling of this market, we observed that fellowship directors were increasingly hiring applicants who had done their residencies locally, because the only first-year residents it was safe to hire were those for whom those directors could get reliable recommendations from their own colleagues. This restriction in the candidate pool reduced the desirable diversity of fellows. What these directors didn't appreciate – until they saw our results – was that this local hiring was happening to everyone. Only then did they all realize that their own problem was in fact market-wide. As you might imagine, that generated a lot of interest in hiring later.

Muriel and I eventually helped them to plan a clearinghouse that operated later in the careers of medical residents, like the one that matched new doctors to residencies. But those same fellowship directors didn't trust each other to cooperate and wait for the clearinghouse; they all worried that the others would continue to hire via early exploding offers. If they waited to take part in the clearinghouse, they feared all the best candidates would already be hired.

This lack of trust threatened to keep everyone making early offers, just in case everyone else did – even when no one, or almost no one, wanted to. So we asked the four principal professional organizations of gastroenterologists if they couldn't simply forbid their members to hire before the clearinghouse opened for business. They told us they had no power to regulate the behavior of their members, the fellowship program directors.

We next asked those organizations if they could pass a resolution that would empower fellowship applicants who had accepted very early offers to change their minds if, later, at the time of the clearinghouse, they regretted their early decisions. That proposal caused some concern: administrators worried that the market would have many offers accepted and then rejected.

Using several kinds of evidence, we were able to convince them that this wouldn't happen, since the incentive to make an offer before you could tell which applicants were good ones would be eliminated if early offers and acceptances weren't binding. By freeing fellows to change their minds if they accepted early offers, the new approach deprived program directors of the incentive to make early offers and relieved them of the fear that others would do so. Thus they could safely wait and match to a great candidate later when the clearinghouse opened.

It doesn't take a lot of self-control to stop making early offers if they no longer get you what you want.

Part of our evidence came from the market for new PhD students in universities. Almost all American universities have agreed that students shouldn't have to accept offers before April 15 of each year. If students are pressed to accept an offer before that deadline, they can accept and then later decline in order to accept another offer before that date. This single rule has virtually eliminated all exploding offers for PhD candidates in the United States. Another part of our evidence was experimental: when we set up these rules in the laboratory and ran them in a simple artificial market, they eliminated exploding offers.

Still another part of our evidence was theoretical. Exploding offers won't occur when everyone has enough experience with the market to know what to expect. When that happens, economists say the market is "at equilibrium." In this case, equilibrium meant that everyone would expect that fellowship programs would be committed to hiring the young doctors to whom they made early offers but who subsequently performed below expectations. But they wouldn't actually get to employ those accepting early offers who exceeded expectations because those people would take better offers made later. Since the main point of early exploding offers is to "capture" better candidates than you could if you waited, program directors wouldn't make exploding offers if they no longer accomplished that goal. It doesn't take a lot of self-control to stop making early offers if they no longer get you what you want.

This worked for the gastroenterologists: they accepted the arguments and implemented the advice, then successfully organized a clearinghouse that now operates each year much closer to the time that gastroenterology fellows will actually start work. Exploding offers aren't a problem anymore, and so almost everyone succeeds in hiring and being hired in the clearinghouse, which operates later, along the lines of the successful market for residents. That clearinghouse provides a thick market that is worth waiting for, much as a thick marriage market in college makes it less pressing for people to marry their childhood sweethearts, or a lot of parking spaces near where you want to park makes it easy to pass up parking spots that you encounter when you're still far from your destination.

Our solution to the problem of hiring new gastroenterologists highlights one of the crucial facts about market design: successful designs depend greatly on the details of the market, including the culture and psychology of the participants. In the years that followed, we encountered a number of other markets facing problems that at first glance looked identical to the gastroenterology market. But in the end, some of these markets required very different solutions.

CULTURAL SHIFT

One excellent example of this is the market for orthopedic surgeons, which at first seemed to be a near clone of the market for gastroenterologists discussed above.

When I spoke with the orthopedic surgeons at Massachusetts General Hospital, it quickly became apparent that they had an unraveling problem: they were hiring new fellows up to three years ahead of time, when the fellows were still young surgical residents. The senior surgeons weren't too worried that they couldn't assess the dexterity of these residents while they were still so young. But they'd noticed that some of their new hires, when they eventually showed up to take their positions, had matured into operating room bullies who made the nurses and others reluctant to work with them.

This complicated scheduling and was bad for morale. If the senior surgeons could wait until after the young residents had time to grow into chief residents and assume more responsibility, they would be better able to assess what kind of colleagues the new surgeons would be, and not just how good they were with their tools. When Muriel and I began to look into the details, we found that the hiring of orthopedic surgeons looked almost exactly like that of gastroenterologists – early exploding offers, local hires, and all. So, naturally, we suggested to them that the solution that worked for the gastroenterology market might work for them as well. That is, if they could empower applicants to change their minds after accepting early offers, those early offers would cease and an orderly clearinghouse could be organized at a convenient later date.

But the orthopedic professional organizations – which include at least nine distinct subspecialties – quickly told us that they couldn't empower young surgeons to change their minds about agreements they'd made with senior surgeons. That would never happen, they said: senior surgeons were too powerful and imposing for younger surgeons to feel that they could really change their minds, no matter what anyone said.

Self-control is not a solution: you can control only yourself, and if others jump ahead of you, it might be in your self-interest to respond in kind.

They saw no obstacle, however, to imposing sanctions on fellowship directors who made early offers. One of the professional societies even told us they simply wouldn't let those doctors present papers at their annual meeting. So, by way of blunter methods, orthopedic surgeons were also prevailed upon to stop making exploding offers, which allowed some clearinghouses to be organized in orthopedic subspecialties.

Orthopedic surgeons needed a somewhat different market design than gastroenterologists to fix similar market failures because the two professions had distinctly different cultures. But in each case, they were able to find a way to prevent exploding offers.

The problems facing federal judges in the clerkship market are harder to solve because that market culture actually combines the difficulties faced by gastroenterologists and orthopedic surgeons. The organizations of judges – called "judicial conferences"– are like the gastroenterology organizations in that they have no way to prevent judges from making early exploding offers or punish those who do. Meanwhile, as with junior and senior orthopedic surgeons, law students aren't in a position to break promises to federal judges. These things make it difficult for judges to organize themselves in a way that lets them trust one another to obey the rules.

Markets unravel despite the collective benefit of having a thick market in which lots of people are present at the same time, with many opportunities to be considered and compared. Without a good market design, individual participants may still find it profitable to go a little early and engage in a kind of claim jumping. That's why self-control is not a solution: you can control only yourself, and if others jump ahead of you, it might be in your self-interest to respond in kind. These early movers become the equivalent of the Sooners in the Oklahoma Land Rush.

For both gastroenterologists and orthopedic surgeons, success had to do not just with setting a particular time at which the market should operate but also with having a well-designed market available at that time. Making the market operate within a narrow time frame, but without providing something, such as a clearinghouse, that brings order to the market at that time, usually isn't a good enough solution to the problem of unraveling. It can cause congestion, as when members of an unruly crowd all try to stake their claims at the same time – which can result in a different kind of market failure, when people feel pressed to make offers (and demand replies) too fast rather than too soon. A congested market may break down in a way that makes the participants risk the fate of poor Walter Cook, the man who was soon enough to stake a claim but not fast enough to register it – the biggest winner, and the biggest loser, in the Oklahoma Land Rush.