

# Freaks and Geeks

How Freakonomics is ruining the dismal science.

NOAM SCHEIBER

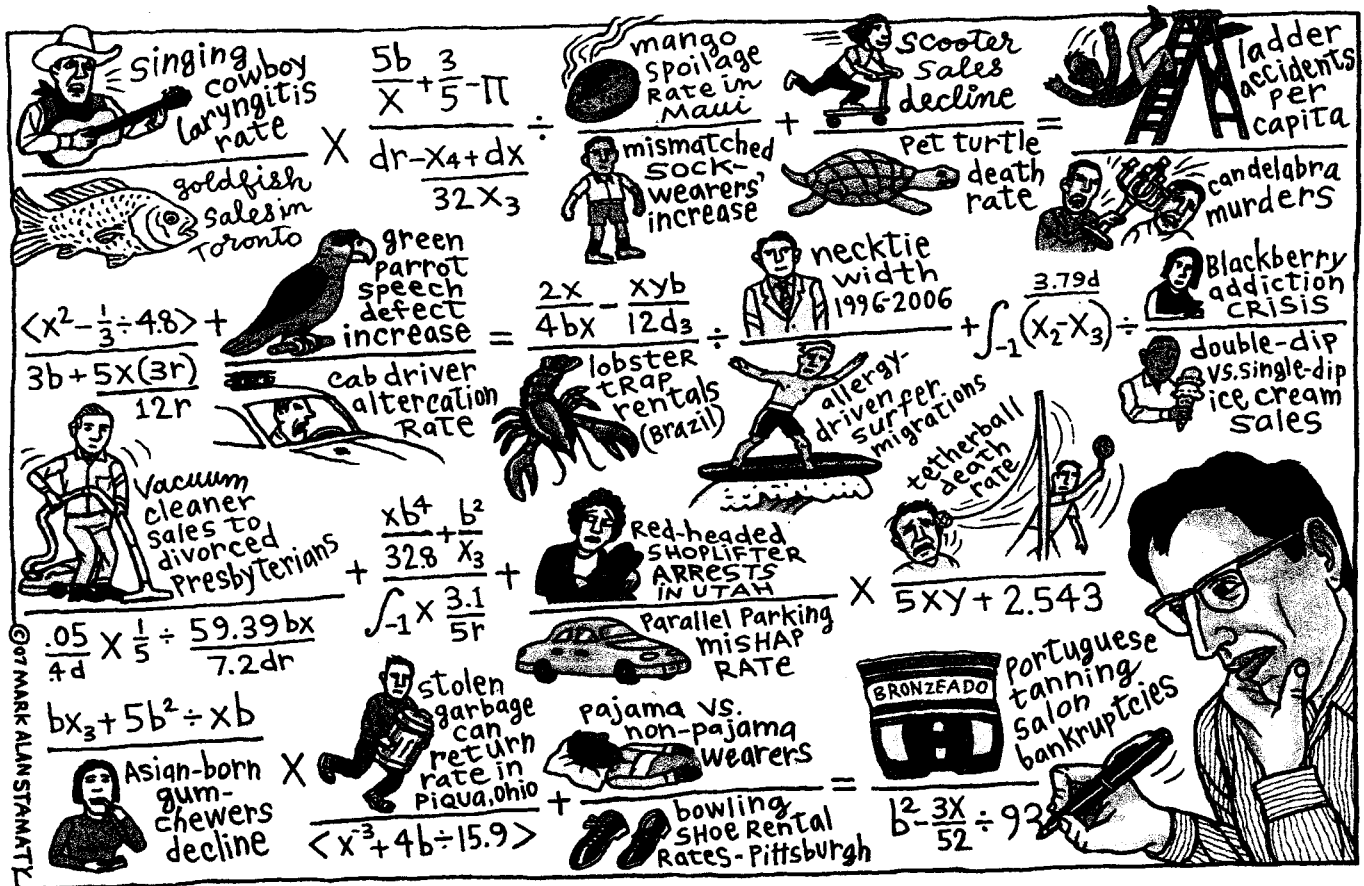
ONE OF THE FEW papers I actually read as a grad student was written by a pair of economists named Josh Angrist and Alan Krueger. In the early '90s, Angrist and Krueger set off to resolve a question that had been gnawing at economists for decades: Does going to school increase your future wages? Intuitively, it seemed obvious that it did. When you compared the salaries of, say, Ph.D.s with those of high-school dropouts, the grad-school set almost always did better. The question was whether education accounted for the difference. What if it was simply the case that smarter people spent more time in school and that their bigger salaries reflected intelligence, not education? One couldn't be sure. The only way to get to the bottom of it would be a ghastly social experiment, wherein you took a group of students and randomly sent half to the local vo-tech institute while forcing the other half to study feminist literary theory. Even an economist wouldn't be so audacious.

That's where Angrist and Krueger came in. In the paper, they pointed out that two features of the public school system allowed you to answer the question without all the uproar. First, most states force students to attend school until age 16. Second, for many decades, students started school the year they turned

six. The upshot was that, if I were born in January and you were born in December of the same year, and we both dropped out at 16, then the rules forced you to stay in school almost one year longer. (We'd start school the same year, but I'd turn 16 midway through tenth grade and you'd hit 16 midway through eleventh.) The additional schooling foisted upon one group by this arbitrary state of affairs produced a scaled-down version of our experiment, allowing Angrist and Krueger to conclude that education did, in fact, help people earn more money.

I probably first came across this paper in 1999. At the time, it struck me as a neat trick, certainly a welcome diversion from the stultifying technical literature economists must imbibe, but largely unremarkable beyond that. Three years later, having escaped the academic track, I was nonetheless spending a lot of time in Cambridge, Massachusetts, where I moonlighted as an economics poseur and consorted with the smartest young economists in the world. And one of the things I was shocked to learn was that Harvard grad students had invested the Angrist and Krueger paper with totemic significance.

It quickly became apparent that the way to win acclaim as a grad student was to devise a similarly ingenious method of tackling a problem. Several years after his paper on schooling, Angrist no-



MARK ALAN STAMATY

ticed that the Armed Forces Qualifying Test had been misgraded for a few years in the late '70s. This had opened the doors to thousands of subpar applicants and allowed Angrist to compare the lucky under-achievers with the people rejected once the glitch got corrected, thereby isolating the impact of military service on wages. The practical effect was to send the grad students scrambling to find other instances in which life-altering decisions had been handed down incorrectly. In 2000, a Harvard professor named Caroline Hoxby discovered that streams had often formed boundaries to nineteenth-century school districts, so that cities with more streams historically had more school districts, even if some districts had later merged. The discovery allowed Hoxby to show that competition between districts improved schools. It also prompted the Harvard students to wrack their brains for more ways in which arbitrary boundaries had placed similar people in different circumstances.

Every few weeks, when a student would stumble onto some new test-grading error or fatefully drawn boundary—what economists call “instruments”—word of the discovery would rocket through the department. The discoverer would become instantly, if momentarily, famous, like the holder of a winning card at a Bingo hall, and inspiring the same mix of reverence and jealousy. A typical conversation around the snack machine at the National Bureau of Economic Research, where many Harvard students had cubicles, went something like: *Hey, did you hear that so-and-so found this crazy example of excess tax refunds in western Manitoba in the early '60s?* At which point the other would reply, *Uh, no, wow, that's, uh, great*, and then scamper back to his desk to brainstorm for some similar quirk of public policy. At an age when most people brood that life is too random and arbitrary, these people's biggest complaint was that it wasn't random and arbitrary enough.

In retrospect, I have come to see this as the moment I realized economics had a cleverness problem. How was it that these students, who had arrived at the country's premier economics department intending to solve the world's most intractable problems—poverty, inequality, unemployment—had ended up facing off in what sometimes felt like an academic parlor game?

With the 2005 publication of *Freakonomics*, the breezy exposition of Chicago economist Steven D. Levitt's oeuvre, the rest of the world has come to see that economists are capable of spectacular feats of cleverness—and to a degree I

couldn't imagine back in my poseur days. In the search for what's known as “clean identification”—a situation in which it's easy to discern the causal forces in play—Levitt has turned to such offbeat contexts as Japanese sumo-wrestling and the seedy world of Chicago real estate. He has studied racial discrimination on “Weakest Link,” a once-popular game show, and reflected on the scourge of white-collar bagel-filching. This has, in turn, inspired a flurry of imitators, including papers on such topics as point-shaving in college basketball, underused gym memberships, and the parking tickets of U.N. diplomats.

Within the frequently tedious body of economics scholarship, these papers stand out as fantastically entertaining. Judging from the dizzying sales of *Freakonomics* and the thousands of lecture halls across the country now bursting with econ majors, they've also been wildly successful at ginning up interest in the discipline. But it does make you worry: What if, somewhere along the road from Angrist to Levitt to Levitt's growing list of imitators, all the cleverness has crowded out some of the truly deep questions we rely on economists to answer?

IT WOULD BE wrong to pretend that this cleverness problem only dates back a decade or two. Academic economics has sometimes valued ingenuity above usefulness since at least the Second World War. In 1996, days after winning the Nobel Prize in part for a theoretical model he'd pioneered in the '60s, Columbia economist William Vickrey confessed that the paper's practical contribution was next to nil. “At best, it's of minor significance in terms of human welfare,” he told *The New York Times*.

Still, for more than a generation, it was mostly theorists like Vickrey who had a monopoly on cleverness. The early post-war empiricists—that is, the people who dealt with real-world data—were, for the most part, earnest, stubborn men. They tackled the era's thorniest questions with a freakish determination. Harvard economist Zvi Griliches devoted decades to the problem of productivity growth, the chief determinant of rising living standards. His colleague Simon Kuznets spent half his career devising the measure of economic growth we still use today. If these men sometimes gave the impression that economics was a life-and-death proposition, well, that was no accident. *Life* to them was a life-and-death proposition. Griliches had survived the Holocaust. Kuznets had endured the German occupation of his native Ukraine.

By the '80s, however, the data-crunchers had come down with a crisis of confidence. In one famous episode, the eminent economist H. Gregg Lewis reviewed several studies on unions. What he found was alarming: Some papers reported that unions strongly increased wages; others reported exactly the opposite. The difference, in most cases, was simply the assumptions the authors had made.

Critiques like this tipped the discipline into a prolonged bout of soul-searching. The old approach had been sweeping in its ambition. But what good were ambitious goals if the best you could do was “on the one hand/on the other hand”—style equivocation or, worse, plain gibberish? “People didn't believe the estimates being produced,” recalls David Card, then a rising star at Princeton. “They felt the evidence in economics was not very credible.” Economists had long aspired to science. Suddenly they faced a harrowing thought: What if they were no better at pinning down truth than the average critical studies major?

Having glimpsed this nihilistic vision, many economists ran screaming in the opposite direction. They concluded that the path to knowledge lay in solid answers to modest questions. Henceforth, the emphasis would be on “clean identification,” on sorting out what caused what.

The early practitioners of this approach—Angrist, Krueger, Card—had well-earned reputations as crafty researchers. But, by and large, all three men used their creativity to chip away at important questions. It was only in the late '90s that the signs of overreach became apparent. To some professors at top departments, clean identification became a fetish. “Almost every student, myself included, had the terrible experience of getting up in front of the [professors] for whom identification is the Holy Grail, and getting cut to shreds when your identification strategy doesn't pass muster,” recalls a recent Harvard Ph.D. The problem is that there are only so many big questions that misgraded tests or arbitrary boundaries can shed light on. If you're wedded to these techniques, eventually they lead you in obscure directions. “People think about the question less than the method,” says Berkeley professor Raj Chetty, one of the most sought-after Harvard graduates in recent years (and a notable exception to this trend). “They're not thinking: ‘What important question should I answer?’ So you get weird papers, like sanitation facilities in Native American reservations.”

Many young economists began shunning big questions altogether. Jim Heckman, a Nobel Prize-winning labor

economist at the University of Chicago, illustrates the point with a story. A few years ago, he struck up a conversation with a promising assistant professor. Before long, Heckman began to gripe that economists lacked a comprehensive measure of all the obstacles a child might face in life—education, nutrition, family environment, and so on. There was no way to tell if childhood disadvantages were getting better or worse.

He encouraged the young economist to produce such a measure. "You're absolutely right. It's paramount, of first-order importance," the woman replied. But there was zero chance she'd pursue it. "It would take years to do," she explained apologetically. The woman had clearly made the right career decision: No one coming up for tenure could afford that kind of time while her colleagues published a stream of small-bore papers. On the other hand, says Heckman, "How long did it take Madame Curie to purify radium? Two or three years? If she was somehow told, like the grad students in our program, 'If you can't solve the problem in a week or a month, drop it,' a lot of big problems would have been dropped."

**H**ECKMAN'S ALLUSION to a certain pedagogical technique is almost surely a shot at his Chicago colleague, Steve Levitt. Levitt has been known to discourage students from laboring too long on a question for which the data are unlikely to produce a useful result. "I've always been someone who's thought it's better to answer a small question well than to fail to answer a big question," he says. This much should not be surprising. Levitt is the product of the same environment that birthed the clean-identification movement.

After graduating from Harvard and doing a brief stint in management consulting, Levitt earned a Ph.D. from MIT in 1994, completing the program in a mere three years. "In the early '90s, if you had a really great natural experiment, as we called it, you'd get yourself a job," he says. Levitt, it turned out, had many. While still a student, Levitt wondered whether money drives election results or if the better candidate simply raises more money. He ingeniously demonstrated the latter and published the results in a top journal. Another early paper found that a slight increase in the chance of arrest dramatically deterred auto theft. Levitt discerned this by studying cities that had approved the use of Lojack, a transmitter that leads police to stolen cars. In 2001, Levitt published what is probably his most controversial finding to date:

a paper highlighting the connection between the legalization of abortion in the '70s and the falling crime rates of the '90s. Levitt argued that unwanted children are most at risk of becoming criminals. Abortion, he concluded, lowered crime rates by reducing unwanted pregnancies.

Some of these papers made genuinely important contributions. The Lojack paper helped demonstrate that theft is a fundamentally rational phenomenon and can therefore be discouraged. This insight alone might have justified Levitt's John Bates Clark Medal, a prize awarded every two years to the most outstanding economist under 40. But, at times, Levitt gave the impression he was more interested in clever techniques than answers to questions. In a 1997 paper, for example, Levitt argued that hiring more police decreases crime, a proposition for which there was surprisingly little evidence. (The fact that municipalities expand police departments when crime rates rise tends to muddle the picture.) To prove it, Levitt needed to simulate an experiment in which the size of a police force was randomly increased. His solution was to exploit the fact that mayors often hire more police officers in the run-up to an election. The only hitch, as a grad student later pointed out, was that mayors up for reelection don't actually hire many police officers, at least not enough to show that they lower crime.

Whatever the flaws in this study, they were clearly the product of too much ambition, not too little. A few years later, however, Levitt debuted a new kind of paper: an investigation into offbeat phenomena from daily life. One of the earliest examples pondered the strategies soccer players employ when taking penalty kicks. Another paper studied corruption in sumo-wrestling tournaments as a window onto the power of incentives. Not long after, Levitt conducted an exhaustive inquiry into "Weakest Link," a game show in which contestants voted to remove a player after each round of trivia questions. Tallying the voting data revealed that contestants were discriminating against Latinos and the elderly but not blacks and women.

In some respects, the papers were just an extreme version of the Harvard approach—an attempt to shrink a question down to the point that you can answer it. Except that, in Levitt's case, the question hadn't just shrunk. It had traveled several hundred miles from its original starting point. Take the "Weakest Link" paper: While the game show did provide a pure setting for observing discrimination, there was no reason to

think we could extrapolate from "Weakest Link" contestants to, say, hiring and promotion decisions, where discrimination often intersects with economics. Set aside the fact that the logic governing "Weakest Link" votes may be completely different than the logic that reigns in the typical HR department. The bigger problem is that most hiring and promotion decisions don't take place in a Hollywood studio before a national TV audience.

Levitt is charmingly, almost painfully, self-effacing when you mention these reservations. "There's no question I have written some ridiculous papers," he says. By way of explanation, he draws an analogy to the fashion industry: haute couture versus prêt-à-porter. "Sometimes you write papers and they're less about the actual result, more about your vision of how you think the profession should be. And so I think some of my most ridiculous papers actually fall in the high-fashion category."

Predictably, not every economist buys this rationale. Not even every economist at Chicago, where Levitt's influence over grad students (whom he attracts), young faculty (whom he helps recruit), and the prestigious *Journal of Political Economy* (which he co-edits) have aroused the ire of an older generation of researcher.

Easily the most famous of the dissidents is Heckman. If Levitt is known for his novelty, the hallmark of Heckman's work on such issues as education and job training is its painstaking attention to detail. A few years ago, Heckman was rumored to be so upset over the direction of his department that he began looking to leave. Chicago had never been an ideal place to do empirical work. Nobel Prize-winning theorists like Gary Becker and Robert Lucas disliked dirtying their hands with data. Now the department was finally warming up to data-crunchers... and they were the kind Heckman deemed useless. "Chicago has been a little disappointing in that it hasn't been more firm in rejecting cute and clever," he laments.

Not that Heckman has been shocked by the development. It is sometimes said that healthy departments have a straightforward division of labor: The theorists generate predictions, and their colleagues test them with data. Alas, this process can be a drag for the people being scrutinized. It's much more fun to generate pie-in-the-sky predictions when you don't have some killjoy looking over your shoulder. Which is why, from the perspective of the theorists, the ideal colleague may be someone less fastidious. Someone who studies the off-beat and clever, not the discipline's central questions. Someone who, you might say, looks less like Jim Heckman and more like

Steve Levitt. "Rigorous theory and bullshit empirical work can co-exist," Heckman sighs. "It leaves the rigorous theorists to make up the numbers they want."

**E**ACH YEAR, the Harvard economics department invites a handful of eminent scholars to participate in its Political Economy Lecture Series. Standard operating procedure is for the speaker to discuss some new paper or avenue of research. Previous speakers have held forth on such eye-glazing subjects as "Rule Rationality versus Act Rationality" or "Equilibrium Contracts for the European Central Banker." In the spring of 2002, Harvard landed Levitt.

Levitt doesn't immediately strike you as a mesmerizing speaker. His voice is too high, except when it's trailing off at the end of a sentence. He leans heavily on the word "OK." He is lanky and concave-chested and often fails to make eye contact. But Levitt has a droll magnetism, a certain anti-charisma. Combined with his eclectic interests, it made the Harvard talk a hit.

Levitt framed his discussion as a how-to guide for practicing oddball economics. "He talked about his kick-ass creative papers, kind of a prelude to *Freakonomics*," recalls one person in attendance. "Here are the lessons you can draw to improve your own research, how you can do clever, appealing papers yourself." As he was wrapping up, Levitt reflected on the choices facing grad students: *If you think you can do as well in traditional topics as someone like Marty Feldstein—a giant of the profession—you should pursue that*, he said, at which point knowing laughter broke out. *But*, he continued, *if you don't feel like you're up to that, you might want to think about alternative topics*. The message resonated loudly. One student watched classmates

spend the next several weeks on high alert for some curiosity of daily life around which they could build a paper.

If Levitt were literally holed away in some ivory tower, emerging only every so often to distract economists from their thoughts about growth and inequality with another clever curiosity, he would undoubtedly be a force for good in the profession. It is not exactly earth-shattering to learn that sumo wrestlers respond to incentives like the rest of us or that soccer players behave strategically. But it's not completely trivial, either. The occasional confirmation of economic theory in contexts people can relate to may even act as a check against excess in the other direction—against useless esoterica. It would be far more disturbing if we *never* observed such behavior in these settings.

But if one Levitt is desirable, two may be pushing it, and more than that might be a problem. Which makes it somewhat alarming that, between his academic success, his popular acclaim, and his personal charm, Levitt may be shaping a generation of young economists.

One growth industry in recent years has been what you might call the lookie-here paper: a small-scale setting for observing some broad principle of economics. Many of these papers deal with sexy topics like corruption and, well, sex. One top journal recently published a paper deducing that Iraq had received billions in kickbacks from rogue buyers under the Oil-For-Food program. Another recent paper, this one in the *Journal of Political Economy*, demonstrated that johns in Mexico pay prostitutes more for unprotected sex than sex with a condom. Both of these findings may be of journalistic interest. But the fact that Saddam Hussein would try to profit from a scarce commodity, or that people might pay extra for services they value more, will

surprise no one who has opened an economics textbook lately.

A companion class of papers boast similar ambitions, but with one small difference: Instead of demonstrating all the offbeat places in which the standard economic model applies, they demonstrate where it breaks down. The first generation of so-called behavioral papers vastly enriched our understanding of how real human beings make decisions, as opposed to the constantly maximizing automatons of economic theory. Economists William Samuelson and Richard Zeckhauser famously observed that people have a huge bias in favor of the status quo, which sheds light on why, for example, they fail to enroll in 401(k).

But subsequent generations of behavioral work have increasingly focused on the picayune. One recent paper in a top journal, *The American Economic Review* (*AER*), documented how people pay hundreds of dollars to belong to gyms they visit infrequently. A forthcoming *AER* paper finds that catalog shoppers overindulge in cold-weather clothing when the temperature abruptly drops. These papers invariably predict that such irrationality will recur in more meaningful contexts—for example, in decisions about jobs or home purchases. Maybe. On the other hand, maybe the only reason people behave irrationally when buying snow boots is that *no one gives a damn about snow boots*. At least not the way they give a damn about their jobs and houses.

Unquestionably, much of what's driving this descent into "cute-o-nomics" is the burst of media attention now devoted to economics papers. "It's very clear that the incentives are to try to do work that gets some public attention. To some extent, that work gets accepted in top journals. You can't blame kids for doing it. I wouldn't blame Levitt," says David

### The Cute-O-Meter

Ranking the top six cute-o-nomics findings

	Chances you'd read about it in <i>Slate</i>	Surprising**	Irrelevance to ending world hunger	Overall cuteness score**
Mexican men pay prostitutes a premium for unprotected sex . . . . .	8	3	6	5.67
Nigerian diplomats run up lots of parking tickets in New York City . . . . .	10	1.5	6	5.83
People buy too much winter clothing from catalogs when it's cold outside . . . . .	6	2	10	6
Sumo wrestlers sometimes throw matches when the stakes are low . . . . .	7	4.5	7.5	6.3
Stock traders aren't sharp on Fridays because they're distracted by the weekend . . . . .	10	1	8.5	6.5
People pay too much for gym memberships they don't use . . . . .	9	1	10	6.67

\*A score of 1 is mathematically equivalent to the word "duh" \*\* Standard errors available upon request

Card, now of Berkeley. But people like Card sometimes worry about the consequences. "It is exactly like postmodernism in the humanities," he groans. "What is there to say about Beethoven anymore? ... Every moron can't understand technical orchestration, doesn't know the history of music. So you write about him having a gay affair with his nephew."

Sadly, Card isn't far off. A few years ago, Levitt supervised a Ph.D. student named Andrew Francis, who now teaches at Emory University. The paper Francis took with him into the job market was called "The Economics of Sexuality: The Effect of HIV/AIDS on Homosexual Behavior, Desire, and Identity in the United States." In it, Francis attempted to demonstrate that certain homosexual men become heterosexual when AIDS is widespread. Granted, there is a legitimate, if sometimes tawdry, literature examining the way sexual behavior responds to disease. But Francis wasn't talking about changes in behavior—less promiscuity, greater condom use, etc. He was talking about changes in *sexual orientation*.

This last may look like a different trend than the bull market in lookie-here papers—more economic imperialism than Harvard-style empiricism. It's not so much that Francis had dreamed up a clever context in which to get at a mainstream question. It's that he'd brought economics to bear on a distinctly non-economic phenomenon, one far more likely to be governed by the forces of biology and psychology than supply and demand. But, in a way, this style of research has also been enabled by the techniques practiced in Cambridge and refined by Levitt. The new imperialists have vast confidence in their ability to extract causality from data. This confidence has, in turn, sometimes led them into territory best left to other disciplines.

Perhaps the most infamous example is a paper written by a recent Harvard Ph.D. named Emily Oster. While still an undergraduate, Oster had become fascinated by the so-called "missing women" problem—the hypothesis, attributed to Amartya Sen, that gender discrimination in Asia has created a vast shortage of women. In some cases parents abort daughters, in some cases they commit infanticide, in some cases they simply don't care for their daughters as diligently as they should. Whatever the cause, Sen has suggested there could be as many as 100 million "missing women" in countries like China, India, and Pakistan.

Years later, while wrapping up her Ph.D., Oster stumbled onto a seemingly unrelated fact: a small medical literature

suggesting that women with hepatitis B were far more likely to give birth to boys. What followed was a series of sophisticated natural experiments, the upshot of which was to demonstrate that 100 million women hadn't gone missing after all. Instead, unusually high rates of hepatitis B had arranged it so that Asian mothers were producing far more boys than nature's track record would suggest.

It was a fabulously compelling result, one that partially absolved whole societies of lurid crimes against their children. It was also a vindication of the *Freakonomics* worldview. Levitt published Oster's paper in the *Journal of Political Economy*. He and his *Freakonomics* co-author, Stephen Dubner, took to the pages of *Slate* to breathlessly retell her "economics detective story." And then, just as suddenly, it all fell apart. A snot-nosed grad student from Berkeley pointed out that hepatitis B couldn't possibly explain the missing women problem. It turned out Asian women gave birth to daughters at the same rate as women everywhere else, at least during their first pregnancy. It was only during subsequent births that the ratios changed. Either a bunch of Asian women were running out to get hepatitis B in between their first and second pregnancies, or, as Sen feared, people were taking dramatic steps to avoid ending up with two girls.

**I**N A SENSE, the split between people like Jim Heckman and Steve Levitt is a split between the nerds and the clevers. The nerds complain that Levitt and his ilk are so far removed from using meat-and-potatoes economic theory they may as well be practicing journalism. "In some quarters of our profession, the level of discussion has sunk to the level of a *New Yorker* article," Heckman griped in a 2005 interview with the Minneapolis Fed. "The authors of these papers are usually unclear about the economic questions they address."

The clevers deny that their research turns its back on economic theory. They complain that the nerds insist upon a kind of sadomasochism that does little to advance the cause of knowledge but makes them feel like they're doing God's work. "The structural Gestapo out there acts as if it can't be useful if it's not hard or complicated. We should have a middle ground," Levitt told me.

All of which may just be garden-variety academic squabbling, except that Heckman does raise a valid point. One of the statements Levitt has become famous for since *Freakonomics* is his observation that "economics is a science with excellent tools for gaining answers but a se-

rious shortage of interesting questions." What is one to make of a discipline that heaps scorn on its own *raison d'être*?

This question turns out to be particularly poignant for someone like me. As a grad student, I was constantly torn between my own prurient interests and the belief that real economists tackle noble but boring subjects. My first shot at a thesis topic was a model of something I pompously dubbed "extended adolescence": basically, how it is that certain members of my generation get to piss away their twenties before settling on a career. (And, more to the point, why this slacker behavior should be subsidized.) When that didn't pan out, I drifted to a slightly more conventional but still narrow question about whether prison labor leads to higher wages once convicts rejoin society.

Finally, with time running out and no luck pinning down data, I settled on a genuinely important topic: how new technologies displace old technologies, using mass transit as a case study. I somehow convinced a man named Paul David to supervise the project. David, in addition to showering me with the patience and generosity of a parent, was arguably the world's foremost expert on technology adoption. Our collaboration taught me what it means to do serious, careful research.

It also happened to be one of the most tedious experiences of my life. So tedious, in fact, that previously tedious exercises—like reading scholarly works of history or *The New York Review of Books*—suddenly seemed like pleasant diversions. Increasingly, I felt oppressed by the prospect of spending years on some intractable problem. I chafed at the obligation to do "socially useful" work. I wanted the freedom to pursue whatever seemed interesting and to drop whatever didn't. Journalism, it turned out, was *precisely* what I wanted to practice. Maybe then I could tap my economics training in the highly seductive manner of a Malcolm Gladwell or a Michael Lewis, not wear it around like a hair shirt. How ironic, then, that academic economics was making room for us lightweighters at the very moment I decamped for journalism. How scary.

When I raise this issue with Levitt, he is almost apologetic: "There needs to be a core for work on the periphery to make any sense. I don't think we would want to have a whole profession with dilettantes like me out doing what I do." But, in nearly the same breath, he adds: "The simple fact is that it's hard to do good research. ... To the extent that you can do interesting research that teaches us something about the world, and entertains along the way, that's not so bad." ♦