Outliers

To “learn from our mistakes” is a message drilled into most of us by our parents. This message is of particular relevance when building a regression model. Our “mistakes,” of course, are the residuals in our current model – the errors we would have made had we used the model to make predictions for the individuals in our sample.

We’ve already discussed one modeling “trick” based on an examination of the residuals: Investigate the observations with the most extreme residuals, searching for systematic differences between those individuals for whom the current model most overpredicts, and those for whom it most underpredicts, with the goal of finding new explanatory variables that play a role in the relationship being studied.

If some individuals lie “outside” the range of typical observations, they can have a distorting influence on the estimates of the regression coefficients – and can provide opportunities for constructive refinement of our model.

Consider a study of the relationship between the length of time an administrative employee has worked for a municipal government, and the employee’s current salary.

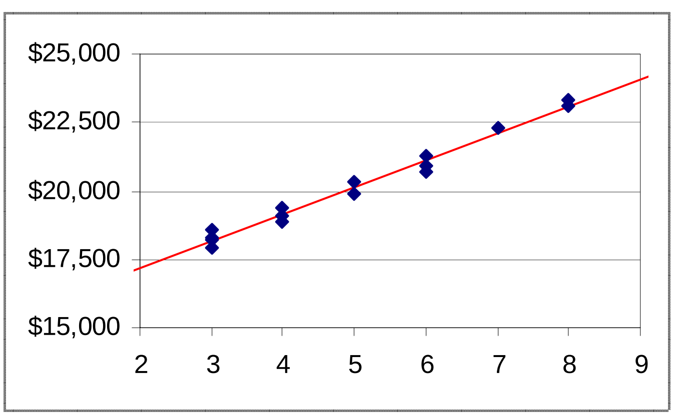
Salary	Experience
$19,100	4
$18,600	3
$22,300	7
$18,300	3
$19,900	5
$23,100	8
$20,700	6
$18,900	4
$18,200	3
$20,300	5
$20,900	6
$23,300	8
$21,300	6
$19,400	4
$17,900	3

Regressing Salary onto Experience yields:

Regression: Salary

	constant	Experience
coefficient	15226.21	984.09
std error of coef	200.87	38.01
significance	0.0000%	0.0000%
standard error of regression		252.10
adjusted coef of determination		97.95%

The associated scatterplot, with the regression line superimposed, looks like this:

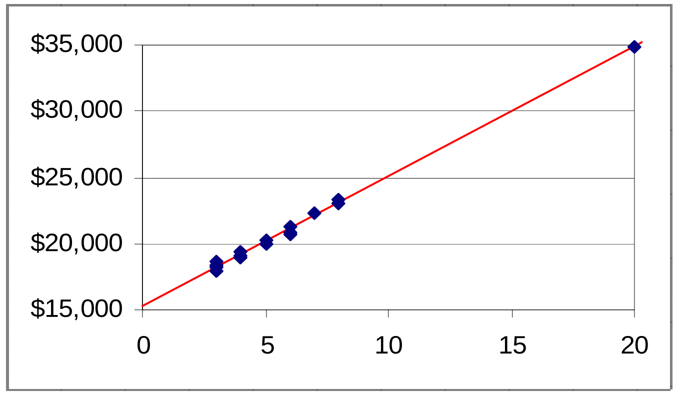


Now, consider three variations of the original sample data.

Case 1: An atypical (combination of) value(s) for the explanatory variable(s). Imagine, instead, that one more employee is included in the original sample: Tom, who has worked for the city for 20 years, and has an annual salary of $34,900.

Regression: Salary

	constant	Experience
coefficient	15228.34	983.65
std error of coef	108.85	15.21
significance	0.0000%	0.0000%
standard error of regression		242.93
adjusted coef of determination		99.64%

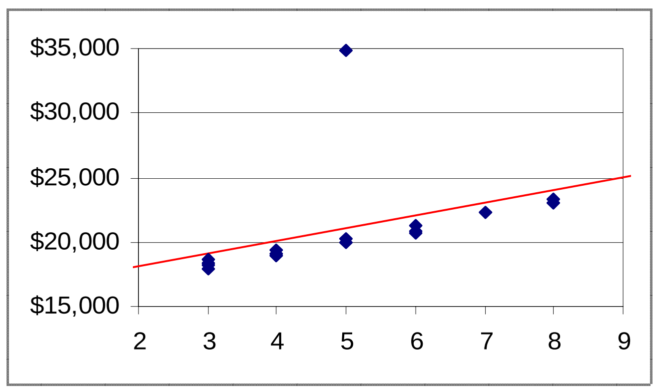


Tom's experience is quite atypical, yet his salary matches up with what the original prediction equation would have predicted. Consequently, the estimates of the model coefficients don't change by much when he's added to the sample. Indeed, since he fits the original model, his presence decreases the standard errors of the coefficients and raises the coefficient of determination.

Case 2: An atypical value for the dependent variable. Imagine that one more employee is included in the sample: Dick, who has worked for the city for 5 years, and has an annual salary of $34,900.

Regression: Salary

	constant	Experience
Coefficient	16148.30	984.09
std error of coef	3038.05	576.72
Significance	0.0109%	11.0014%
standard error of regression		3825.51
adjusted coef of determination		11.30%

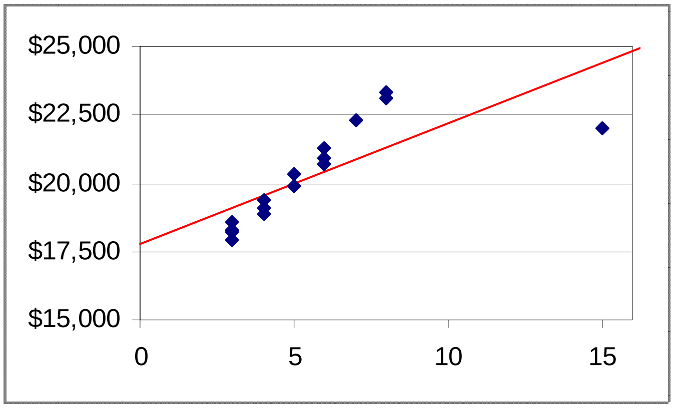


Dick's experience is as "typical" as it could possibly be – He has 5 years of experience, which happens to be the mean experience across all the individuals in the original sample. His atypical value for the dependent variable pulls the entire regression line upwards (raising our estimate of the constant term by nearly $1,000), without affecting its slope. However, his presence in the sample lowers the explanatory power of the model (the coefficient of determination drops from 97.95% to 11.30%), weakens the evidence that the "experience" effect is a real effect (the significance level of the t-ratio rises from nearly zero to 11.0014%), and increases the uncertainty in predictions made using the model (the standard error of the regression, which contributes to the standard error of the prediction, rises from roughly 252 to more than 3825).

Case 3: Atypical values for the explanatory variable(s) and the dependent variable. And finally, imagine that one more employee is included in the original sample: Harry, who has worked for the city for 15 years, and has an annual salary of $22,000.

Regression: Salary

	constant	Experience
coefficient	17784.85	440.47
std error of coef	644.98	101.66
significance	0.0000%	0.0689%
standard error of regression		1193.19
adjusted coef of determination		54.23%



Harry, atypical in his explanatory variable, and atypical, as well, in his dependent variable (relative to what we would have predicted from our original model), dramatically changes all of the regression statistics.

These three cases illustrate the three different ways that an individual can be an outlier, i.e., can differ substantially from the "typical" individuals in the sample.

Detecting Outliers

Clearly, it is important to be able to identify observations which are outliers in one of these three ways. However, if there are several independent variables in play (so a simple two-dimensional scatterplot can't capture the entire relationship), visual identification of outliers is very difficult. We need analytical tools to help us find the atypical individuals in our sample.

After performing a regression, KStat offers an option called "Model Analysis." This option expands and enhances the information displayed on the "Residuals" tab. The model analysis for the original sample data is displayed below.

Three new columns have been added to the "Residuals" output: the Studentized residual, leverage, and Cook's D for each observation.

Predicted values and residuals (original sample data)

			2.1604	0.1333	0.4814	
Salary	**predicted**	**residual**	**std'ized**	**leverage**	**Cook's D**	**Experience**
19100	19162.576	-62.57576	-0.2601	0.0894	0.0033	4
18600	18178.485	421.51515	1.8217	0.1576	0.3104	3
22300	22114.848	185.15152	0.8002	0.1576	0.0599	7
18300	18178.485	121.51515	0.5252	0.1576	0.0258	3
19900	20146.667	-246.6667	-1.0128	0.0667	0.0366	5
23100	23098.939	1.0606061	0.0049	0.2712	0.0000	8
20700	21130.758	-430.7576	-1.7906	0.0894	0.1574	6
18900	19162.576	-262.5758	-1.0915	0.0894	0.0585	4
18200	18178.485	21.515152	0.0930	0.1576	0.0008	3
20300	20146.667	153.33333	0.6296	0.0667	0.0142	5
20900	21130.758	-230.7576	-0.9592	0.0894	0.0452	6
23300	23098.939	201.06061	0.9342	0.2712	0.1624	8
21300	21130.758	169.24242	0.7035	0.0894	0.0243	6
19400	19162.576	237.42424	0.9869	0.0894	0.0478	4
17900	18178.485	-278.4848	-1.2036	0.1576	0.1355	3

Case 1: The **leverage** of an observation indicates the extent to which the observation is atypical due to the value(s) of the explanatory variable(s). In the original dataset, the two employees with 8 years of experience have somewhat higher leverage than the other employees, because their level of experience differs from the average experience (5 years) the most.

Predicted values and residuals (with Tom)

			2.1448	0.1250	0.4794	
Salary	**predicted**	**residual**	**std'ized**	**leverage**	**Cook's D**	**Experience**
19100	19162.932	-62.93209	-0.2697	0.0772	0.0030	4
18600	18179.284	420.71586	1.8218	0.0963	0.1769	3
22300	22113.876	186.12405	0.7932	0.0669	0.0226	7
18300	18179.284	120.71586	0.5227	0.0963	0.0146	3
19900	20146.58	-246.58	-1.0502	0.0659	0.0389	5
23100	23097.524	2.4760971	0.0106	0.0792	0.0000	8
20700	21130.228	-430.228	-1.8291	0.0625	0.1115	6
18900	19162.932	-262.9321	-1.1267	0.0772	0.0531	4
18200	18179.284	20.715862	0.0897	0.0963	0.0004	3
20300	20146.58	153.41996	0.6535	0.0659	0.0151	5
20900	21130.228	-230.228	-0.9788	0.0625	0.0319	6
23300	23097.524	202.4761	0.8686	0.0792	0.0324	8
21300	21130.228	169.772	0.7218	0.0625	0.0174	6
19400	19162.932	237.06791	1.0159	0.0772	0.0432	4
17900	18179.284	-279.2841	-1.2094	0.0963	0.0780	3
34900	34901.299	-1.299338	-0.0133	0.8382	0.0005	20

When Tom is added to the dataset, his substantial difference from the “typical” experience level is clearly signaled by his exceptionally-large leverage.

If a model involves two or more independent variables, it might be that each independent variable, on its own, is not atypical, but the combination of values is. For example, in a study of the determinants of the market value of homes, two explanatory variables might be the size of the lot, and the number of bedrooms. Most of the homes on smaller lots would typically have fewer bedrooms than the homes on larger lots. A house with many bedrooms, on a small lot, would be atypical and would have high leverage, even if neither the lot size nor the number of bedrooms was, by itself, atypical.

Case 2: The **Studentized residual** of an observation indicates the extent to which the dependent variable takes an atypical value, given the values of the explanatory variables.

Predicted values and residuals (with Dick)

			2.1448	0.1250	0.4794	
Salary	**predicted**	**residual**	**std'ized**	**leverage**	**Cook's D**	**Experience**
19100	20084.659	-984.6591	-0.2691	0.0852	0.0034	4
18600	19100.568	-500.5682	-0.1422	0.1534	0.0018	3
22300	23036.932	-736.9318	-0.2094	0.1534	0.0040	7
18300	19100.568	-800.5682	-0.2274	0.1534	0.0047	3
19900	21068.75	-1168.75	-0.3155	0.0625	0.0033	5
23100	24021.023	-921.0227	-0.2812	0.2670	0.0144	8
20700	22052.841	-1352.841	-0.3697	0.0852	0.0064	6
18900	20084.659	-1184.659	-0.3238	0.0852	0.0049	4
18200	19100.568	-900.5682	-0.2559	0.1534	0.0059	3
20300	21068.75	-768.75	-0.2075	0.0625	0.0014	5
20900	22052.841	-1152.841	-0.3151	0.0852	0.0046	6
23300	24021.023	-721.0227	-0.2202	0.2670	0.0088	8
21300	22052.841	-752.8409	-0.2058	0.0852	0.0020	6
19400	20084.659	-684.6591	-0.1871	0.0852	0.0016	4
17900	19100.568	-1200.568	-0.3411	0.1534	0.0105	3
34900	21068.75	13831.25	3.7341	0.0625	0.4648	5

When Dick is added to the dataset, his Studentized residual is quite large, signaling that his salary is very atypical for his level of experience.

The Studentized residual, in essence, tells us how many "standard deviations" away from the typical residual – 0 – the observed residual is. Technically, since the residual for each observation is only an estimate – based on our estimate of the true model – of the observation's true residual, some care must be taken in properly determining one standard-deviation's-worth of variability in the residuals. But intuitively, the Studentized residuals aren't much different from what we'd get if we simply divided the residuals by the standard error of the regression.

Case 3: In the first case, Tom had high leverage, but his dependent variable took a value close to what we would have predicted from the model based on the original dataset, so his leverage wasn't "exerted." Consequently, his inclusion didn't have much effect on our estimates of the coefficients of the model.

In the second case, Dick had a large Studentized residual, but little leverage. His inclusion somewhat changed our estimate of the constant coefficient, but had little effect on our estimate of the impact of the explanatory variable.

But when Harry was added to the dataset set, all of the coefficient estimates changed substantially: Harry had relatively high leverage, and that leverage was exerted as well.

Predicted values and residuals (with Harry)

			2.1448	0.1250	0.4794	
Salary	**predicted**	**residual**	**std'ized**	**leverage**	**Cook's D**	**Experience**
19100	19546.733	-446.7332	-0.3907	0.0817	0.0068	4
18600	19106.261	-506.2613	-0.4504	0.1125	0.0129	3
22300	20868.149	1431.8512	1.2485	0.0762	0.0643	7
18300	19106.261	-806.2613	-0.7173	0.1125	0.0326	3
19900	19987.205	-87.20508	-0.0756	0.0653	0.0002	5
23100	21308.621	1791.3793	1.5856	0.1034	0.1450	8
20700	20427.677	272.32305	0.2358	0.0635	0.0019	6
18900	19546.733	-646.7332	-0.5656	0.0817	0.0142	4
18200	19106.261	-906.2613	-0.8062	0.1125	0.0412	3
20300	19987.205	312.79492	0.2712	0.0653	0.0026	5
20900	20427.677	472.32305	0.4091	0.0635	0.0057	6
23300	21308.621	1991.3793	1.7626	0.1034	0.1792	8
21300	20427.677	872.32305	0.7555	0.0635	0.0194	6
19400	19546.733	-146.7332	-0.1283	0.0817	0.0007	4
17900	19106.261	-1206.261	-1.0731	0.1125	0.0730	3
22000	24391.924	-2391.924	-3.6633	0.7005	15.6970	15

Cook's D is a commonly-used statistic which combines the leverage and the Studentized residual of an observation, in order to measure the effect of that single observation on the estimates of the model coefficients. An observation with a high value of Cook's D is, all by itself, highly "influential" upon the results of the study.

Note that KStat flags (in red) exceptionally large values of the Studentized residuals (both positive and negative), and high values of leverage and Cook's D. The flagging is done by applying several "rules of thumb" that have been developed over the years – It is not definitive, but should serve as a guideline in focusing your attention on the most "extreme" observations.

Dealing with Outliers

So, you've performed a regression analysis, and detected some outliers in your dataset. What should you do?

"If it doesn't fit, throw it out." ***This is bad – evil – reprehensible!!!*** It might be okay when you're cleaning old clothes out of your closet, but this is NOT the way to deal with outliers!

The first course of action is to be sure that the data was typed correctly. Many outliers are simply the result of typographical errors. Can typos be costly? Read the article below.

Hint: For 10 Million, Write a 10, Then Six Zeros, Then a Little Dot

By ALBERT R. KARR
Staff Reporter of THE WALL STREET JOURNAL

(January 24, 1996) As small businesses competed in a recent government auction to provide new wireless personal communications services in the area of Norfolk, Va., one bidder seemed to get carried away. PCS 2000, a partnership based in Old San Juan, Puerto Rico, offered $180 million.

Granted, the new wireless phones promise to send low-cost messages, images and data to receivers the size of credit cards or even to Dick Tracy-style wrist-radios. But keeping in mind that a license for the Norfolk PCS market has an estimated value of $30 million – tops – PCS 2000 seemed more than generous when it eclipsed the previous high bid of $16,369,313 from DCR Communications, Inc. of Washington, D.C., by $163,630,687.

Even after deducting some small-business credits for which PCS 2000 qualifies, bringing the net bid down to $135 million, the offer ranked Norfolk near the top of 493 wireless-spectrum markets up for grabs in the Federal Communications Commission auction – behind New York, Los Angeles and Chicago but ahead of San Francisco and Philadelphia.

Something seemed wrong – and was it ever. PCS 2000 had misplaced a decimal point. It told the FCC it had really meant to bid $18 million; last week, it canceled the offer. But there is no certainty it can escape so easily. Under FCC rules, which FCC auction chief Kathleen Ham says were “pounded into the bidders” before the auction began,. PCS 2000 could face a penalty for backing out. A very big penalty, in fact: the difference between its $135 million net bid and whatever the winning offer turns out to be.

Faced with the prospect of forking over a sum that easily could top $100 million, PCS 2000 officials at first said the mistake was the FCC's, not theirs but now admit it probably was their goof. Ms. Ham says “we're very confident” the commission didn't err. The agency merely takes offers filed by computer, displays them on a screen and even lets bidders print copies; this gives all companies a chance to confirm whether they really want to. bid the amounts shown, she explains.

A simple default isn't a way out, either. The FCC then could suspend the company's licenses. So PCS 2000 is seeking a waiver of the penalty, something the FCC has never granted. But there is a hint of relief, Ms. Ham suggests: The agency may be sympathetic if PCS 2000 made a “clear error”- and admits it.

(Addendum: Eleven months later, the FCC reduced PCS 2000’s penalty to a “mere” $3.27 million.)

If the data was entered correctly, the next step is to examine the characteristics of the outlying individual. Consider Case 2, in which Dick is earning an extraordinarily large salary for someone with only five years of experience. Perhaps, upon further examination, you will learn that Dick is the mayor’s nephew, and that none of the other city employees in the sample were related to senior government officials.

How should this situation be handled? One approach is to redefine the population: Remove Dick from the dataset, and clearly indicate that the study covers only city employees who are not close relatives of senior city officials. Another approach is to modify the model to incorporate the newly-discovered explanatory variable: Include a “yes-no” variable which is 1 if an employee potentially benefits from nepotism, and 0 otherwise.

Regression: Salary

	constant	**Experience**	**Nepotism?**
coefficient	15226.21	984.09	14753.33
std error of coef	200.87	38.01	260.37
significance	0.0000%	0.0000%	0.0000%
beta-weight		0.4149	0.9080
adjusted coef of determination		99.61%	

In Case 3, Harry has an atypically-high level of experience. He could be excluded by redefining the coverage of the study to be only employees with less than ten years of experience. But, if he has nothing else “special” about him, you could also view him as a valuable observation which signals that, over the long run, the effect of experience on salary is nonlinear. In this particular case, if his salary were somewhat higher than that of the other employees, but below the line determined by the original dataset, I’d lean towards taking this latter approach, and would add $(\text{Experience})^2$ to my model. However, my judgment leads me to doubt that the “bend” in the relationship would be so extreme that the marginal effect of experience on salary eventually becomes negative, so I’d be more inclined – with the data as given – to either seek another explanation for Harry's low salary, or to take the explicit-exclusion approach.

Of course, when in doubt, a final, perfectly legitimate approach – if you're performing your study for the benefit of others – is to provide two alternative models, one including the outlying observation(s) and the other excluding them, and to indicate your uncertainty concerning which model is better. This is the approach I eventually took in the study attached below. Please note that a subsequent failure by those receiving the study to fully pursue an explanation for the outlying observations led to the demise of one company (Simon Marketing, with the associated loss of over 160 jobs), and over $25 million in eventual costs to another company (McDonald's).

KStat's "Model Analysis" tab contains a few additional statistical aids.

Predicted values and residuals (original sample data)

	0.1697	68.036%	**Breusch-Pagan heteroskedasticity test**			
	0.6039	73.937%	**Jarque-Bera non-normality test**			
	1.7064		**Durbin-Watson statistic**			**variance inflation**
			2.1604	0.1333	0.4814	1
Salary	**predicted**	**residual**	**std'ized**	**leverage**	**Cook's D**	**Experience**
19100	19162.576	-62.57576	-0.2601	0.0894	0.0033	4
18600	18178.485	421.51515	1.8217	0.1576	0.3104	3
…	…	…	…	…	…	…

The heteroskedasticity test will yield a significance level close to 0% if the residuals vary more widely for some values of the explanatory variables than for others. The non-normality test will yield a significance level close to 0% if the residuals fail to vary according to a normal distribution. The Durbin-Watson statistic is primarily of interest if the observations come from successive points in time, in which case a value near 0 signals that adjacent observations tend to have residuals of the same sign, and a value near 4 signals the opposite. Our original dataset appears to have none of these abnormal features. (If such features are present, they can suggest improvements to the model which lie beyond the scope of this course.)

The "variance inflation" factors (when there are several explanatory variables in the model) signal cases where colinearity (some of the explanatory variables seem to be telling the same "story" as some of the others) is present, and is reducing the precision of the coefficient estimates.

A Tale of 13 Outliers

In the fall of 1998, I was asked to predict the total retail value of prizes that would be claimed in a McDonald's promotional game to be run in Canada. (Backstory: For a number of years, I was the external "odds of winning"-calculator for Simon Marketing, the company that ran the McDonald's promotional games. Occasionally, they'd ask me to help with other issues. Canadian law required that an escrow account be set up to cover taxes due on prizes awarded in Canada. Since not all prizes are claimed, the goal of the study was to support the deposit of an escrow amount that would cover claimed prizes only, rather than all offered prizes.)

I was given three years of data, indicating the fraction of prizes of each value which were eventually claimed after similar games were run in the US between 1996 and 1998. I regressed redemption rate onto (the logarithm of) prize value in order to estimate the likelihood of prizes of different values being claimed in the Canadian promotion, and then used those estimates to predict the total retail value of the prizes that would be claimed.

Not surprisingly, larger prizes (such as automobiles) were more likely to be claimed than smaller prizes (such as free hamburgers). However, an examination of the data for outliers revealed a curious fact: Of the 13 largest prizes (worth $200,000 or more) offered in the previous three years, every one had been claimed. I found this a bit surprising, since many of the prizes required matching multiple tickets (representing, for example, properties in the board game "Monopoly") into a complete "set," and a consumer, not knowing that they'd acquired one of the "rare" tickets, might not match it with another ticket and realize they'd won a prize. I discussed this anomaly with the managers at Simon Marketing's Illinois office (located near McDonald's headquarters), who offered the explanation that people paid more attention to tickets that might win them large prizes. I was somewhat unsatisfied with this explanation, and pointed out that some prizes in the $50,000-$175,000 range had gone unclaimed – and that, personally, I'd pay as much attention to the chance of winning a $100,000 prize as I would to the chance of winning a $1,000,000 prize.

In my final report, I wrote:

Estimated Value of Claimed Prizes

Redemption-rate data from 1996-98 was used to estimate the likelihood of a prize of any given value being redeemed. A regression-based model yielded the following relationship:

Redemption rate = 0.282947 + 0.096909 * log(prize ARV)

The (B) columns on the next page display the expected value of prizes claimed in a game offering the 1998 prize categories using this model.

However, all 13 of the largest prizes offered in the three games were redeemed. Using the model above, the chance of this happening would be about 7.5% - not impossible, but somewhat improbable.

The data was reanalyzed, using the assumption that all prizes in the $200,000+ categories would be redeemed, and estimating redemption rates for smaller prize categories after excluding the large-prize data from the sample. The resulting relationship for predicting redemption rates in the smaller prize categories is:

Redemption rate = 0.367524 + 0.062011 * log(prize ARV)

The (A) columns on the next page display the expected value of prizes claimed using this more-conservative model.

Note that the (A) model predicts that 61.77% of all offered prize value would be claimed (see bottom of next page), while the (B) model predicts 56.90%. In actuality, 60.66% of the 1996-1998 total prize value was claimed.

Bob Weber

November 4, 1998

As you can see, I was sufficiently uncomfortable that I offered two different regression models, one of which ASSUMED that, for whatever reason, all of the $200,000+ prizes would be claimed.

Nearly three years later, the real explanation came to light.

WASHINGTON, Aug. 21, 2001

AP

(CBS) The FBI has arrested eight people in connection with what federal investigators call a scheme that fraudulently netted more than $13 million worth of McDonald's game prizes.

Authorities said Tuesday a criminal ring involved Simon Marketing, Inc., a company responsible for McDonald's game security. Among those arrested was an employee of the company's security department, based in Lawrenceville, Ga. The employee embezzled winning game pieces, officials charged.

The head of security at Simon's Georgia headquarters – the person who manually inserted the rare, large-prize-winning tickets into randomly-selected stacks of "common" tickets – had been systematically stealing the highest-value rare tickets and conspiring with confederates to claim all of the large prizes. The subsequent scandal drove Simon Marketing out of business (it exists today as a shell company – SWWI – still dealing with legal liability issues), and cost McDonald's well over $25 million in customer restitution efforts.

If only they had taken the outliers more seriously!

Here's an interesting column from the Chicago Tribune. You might find it worthy of discussion during your "Crisis Management" course.

`Monopoly' turns into game of risk for McDonald's

by David Greising

Published August 22, 2001

It was late spring. FBI agents were swarming all over McDonald's promotional giveaway games. The evidence indicated an official from Simon Marketing Inc., a promotional-game firm, was fixing McDonald's "Who Wants to Be a Millionaire?" game.

"Millionaire" had run its course. McDonald's marketers wanted to relaunch the successful "Monopoly" game. More important, the Feds wanted McDonald's to run "Monopoly," too.

McDonald's Chief Executive Jack Greenberg faced a difficult decision.

Should he run the game, despite strong evidence that Simon security employee Jerome Jacobson for as many as six years had fixed McDonald's promotions by arranging to send the big-money game pieces to phony winners? The FBI agent running the investigation insisted he needed one more game to gather wiretap evidence.

Or should Greenberg kill "Monopoly"? To run the game in the face of alleged fraud would invite lawsuits and might damage the McDonald's name.

Greenberg consulted with McDonald's lawyers. He did not consult with his board of directors. Then he did what he considered the only right thing: He green-lighted the game.

"I had to do what was right. If you're sitting in my chair, I think you'd do the same thing," Greenberg told me.

Of course, I'm not sitting in Greenberg's chair. As I quizzed Greenberg on the subject, I was sitting in the far more comfortable seat of the Tuesday second-guesser.

I thought Greenberg made the wrong move. Consider the circumstances. McDonald's has had a rough year. The mad cow scare has hammered European sales. The strong dollar has hurt McDonald's outside the United States. Domestic business is sagging.

And now, the FBI wants Greenberg to launch a corrupted game likely to damage McDonald's image and expose the company to lawsuits?

I thought Greenberg should stick to running McDonald's and leave the law enforcement to the feds. Gumshoes are notoriously risk-averse. They want reams of evidence when a few solid scraps can convict.

Even before the "Millionaire" launched July 11, the FBI already had phone and surveillance evidence against Jacobson and at least three key accomplices. The evidence covered 13 suspects and winnings of $7.7 million going as far back as 1995.

If I were Greenberg, I would have congratulated the feds for cracking a big case. I would have refused to put my customers or my company at risk. I would have pulled the plug on "Monopoly."

Here's what happened because Greenberg made the opposite decision: Just as the "Monopoly" game launched, the FBI tripped across an entirely new branch of its investigation. It found Andrew Glomb, allegedly one of the people who recruited fake "winners" for Jacobson.

The FBI allegedly taped Glomb and Jacobson discussing the transfer of a winning game piece. They allegedly discussed whether they could trust their crooked "winner" to keep their dirty secret.

Because Greenberg gave the FBI an extra go-ahead, the feds nailed Glomb, and have evidence against seven other winners of at least $6 million in prizes, dating to 1999.

"I would do it again," Greenberg says. "What we found out allowed the FBI to complete its investigation."

To make amends to customers, Greenberg plans to count up the ripped-off prize money and launch another game. "Every dollar that has been stolen, our customers will have a chance to win again," Greenberg says.

Makes you wonder what they'll call the new game. "Crooked Cash"? Or maybe "Jailbirds' Jackpot"?

Greenberg knows his decision wasn't simple. He understands customers, shareholders and directors may question the relaunch of a "Monopoly" game he knew was crooked. He won't be surprised if lawsuits result.

I wouldn't have made the call Greenberg did. I wouldn't have put my company's name at risk. And, it turns out, I apparently wouldn't have helped catch some extra crooks and recover some extra cash.

I didn't see the issue as black and white. But Greenberg did. And thanks to him, the feds hope to sew some black and white outfits for Jacobson and his gang.