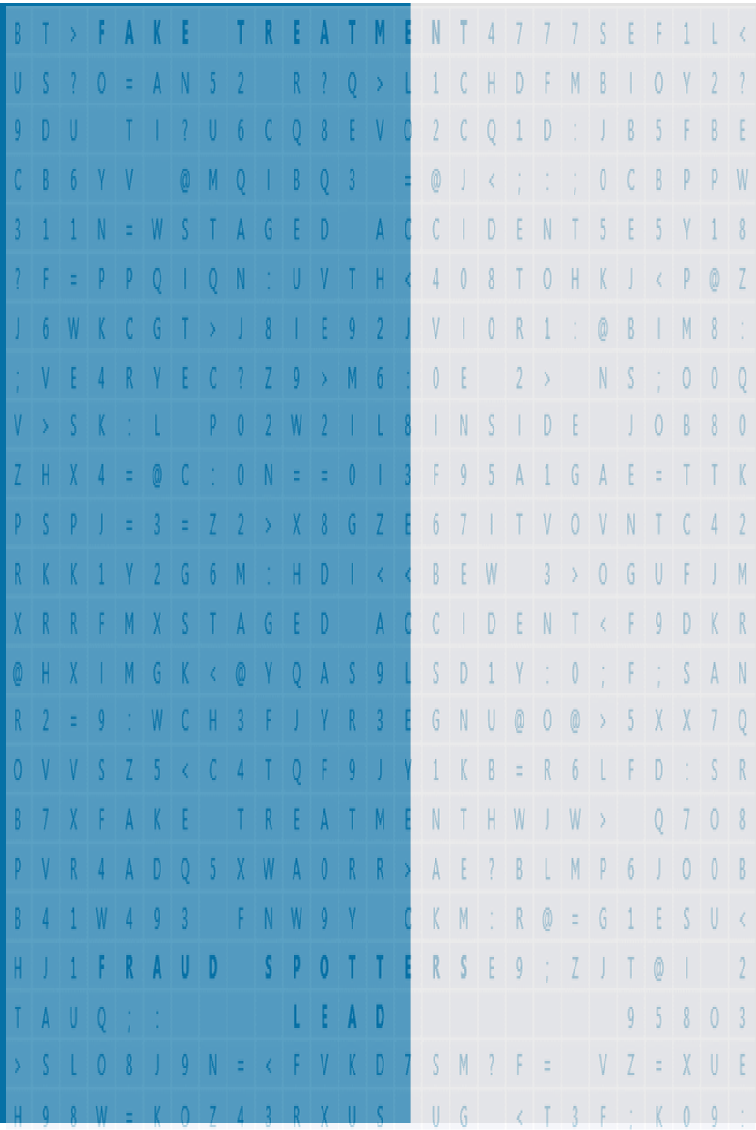
Session 5: Fraud Analysis

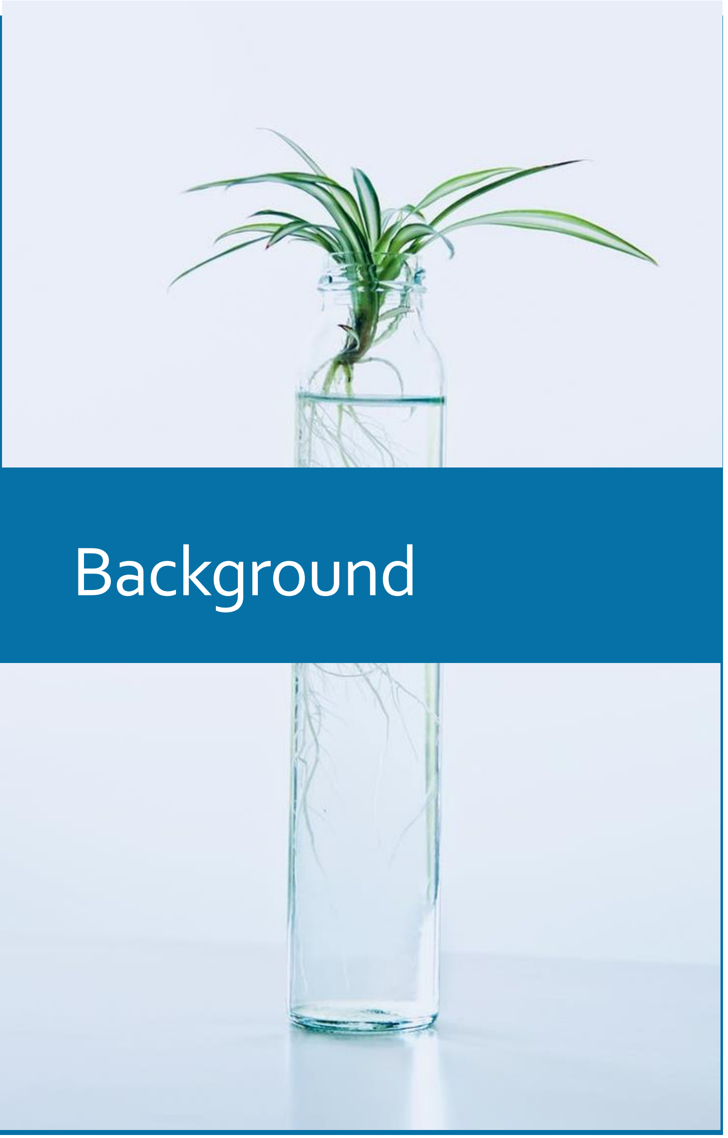
Some applications of math related to insurance fraud; how I got into this; and similar applications

By Ben Turner



Agenda

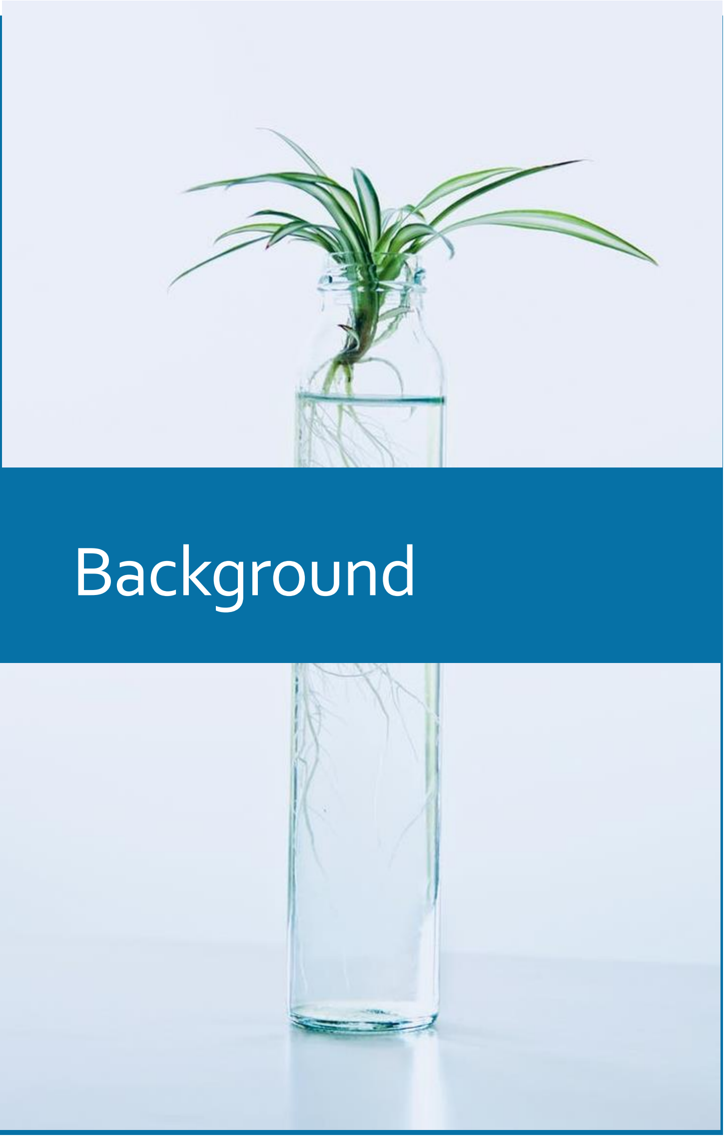
- Background
- Texas Mutual
  - Provider scorecard tool
- President of Windhaven
  - System conversion
  - Lawsuits
  - Massive fraud
  - Fraud fighting tool
- Election integrity analysis
  - Sandwich equation
  - Results (see www.fraudspotters.com/evaluation-of-red-flags-in-fraud/)
- Relevance of External Data
  - Provider tool (see www.fraudspotters.com/cagny user: cagny, password: 2021)
- Questions at the end
  - Please no political commentary or political questions
- We are on slide 2 or 53



Background

- Undergraduate in economics
- Worked at Mercer as pension actuary
- Returned to graduate school and earned a law degree, MBA degree, and passed the California bar.
- At approximately the same time published an article in the 2004 CAS Ratemaking Journal article on credibility weighted segmentation for ratemaking and became an ACAS.

- Personal life: my wife became severely ill, and I stopped talking exams.



Background

- Performed ratemaking generalized linear modeling for the following companies:
  - Farmers Insurance
  - Bristol West Insurance
  - Windhaven Insurance*
  - Texas Mutual Insurance*
- *For the last two roles I was the chief actuary

Texas Mutual

Texas Mutual:

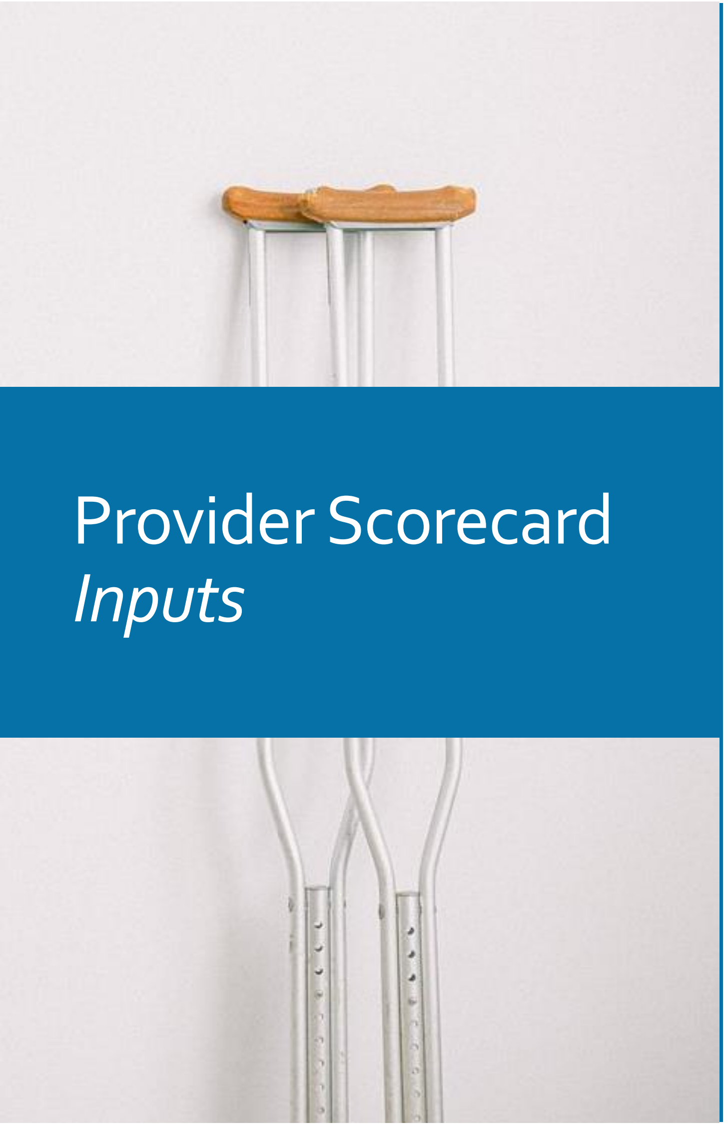
- Largest workers compensation carrier in Texas
- In existence since 1994
- Services 20% or more of the workers compensation claims of Texas
- Texas does NOT allow claim settlements
- Every medical bill on a covered injury is covered for life

Texas Mutual
Actuary department

- Reserves, pricing, reinsurance, ERM, and all other traditional actuarial work
- We also did a number of non-traditional things:
  - Premium fraud
  - Medical fraud
  - Internal fraud analysis
  - Customer analysis
  - Small business underwriting/pricing
  - Analysis of workers compensation legislation
  - Premium projections
  - Etc.
- Most importantly, I was involved in a provider scorecard tool

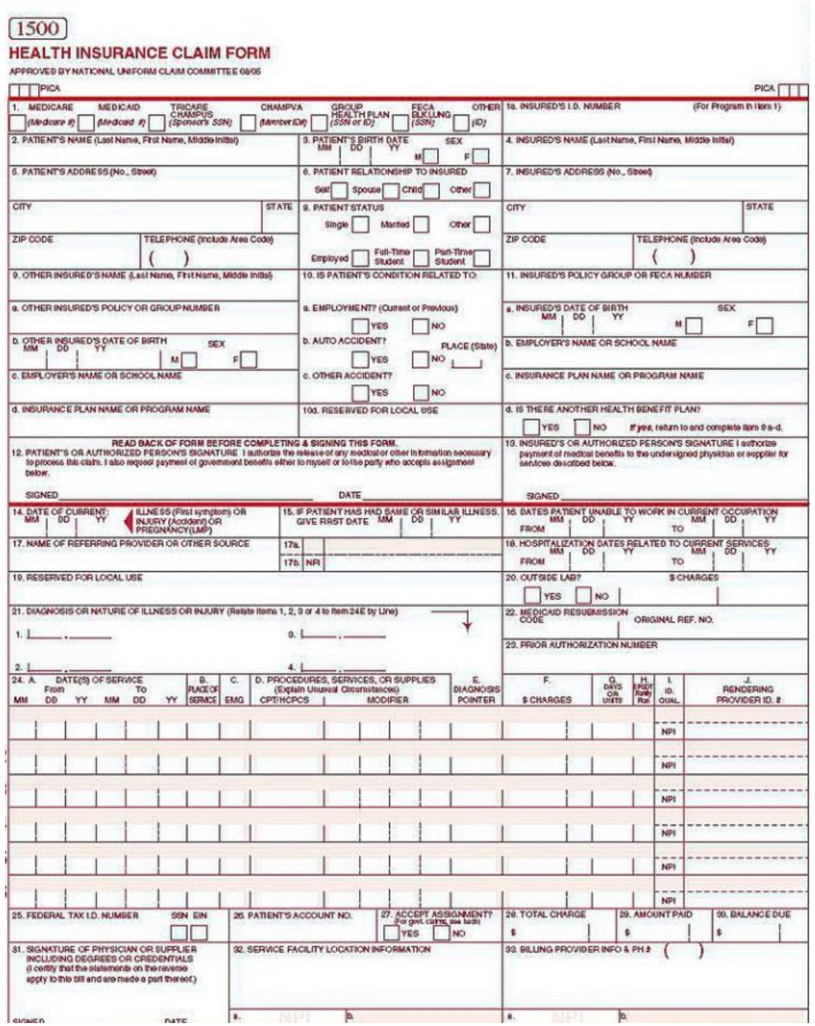
Provider Scorecard
Physician Network

- Texas allows workers compensation carriers to set up a network of doctors
- Give a discount to employers who participate in the network
- Our network:
  - Preferred provider of about 50 providers
  - We routed claimants to the best doctors
  - Best doctors based on statistics related to getting back to work
  - Best doctors paid more, not less
- We had 15 years of data *prior* to the formation of the network



Provider Scorecard
Inputs

HCFA 1500



1500

HEALTH INSURANCE CLAIM FORM

APPROVED BY NATIONAL UNIFORM CLAIM COMMITTEE 08/05

PICA PICA

1. MEDICARE MEDICAID TRICARE CHAMPUS CHAMPVA GROUP HEALTH PLAN FECA BLK LUNG OTHER
(Medicare #) (Medicaid #) (Sponsor's SSN) (Member ID#) (SSN or ID) (SSN) (ID)

1a. INSURED'S I.D. NUMBER (For Program in Item 1)

2. PATIENT'S NAME (Last Name, First Name, Middle Initial)

3. PATIENT'S BIRTH DATE MM DD YY SEX M F

4. INSURED'S NAME (Last Name, First Name, Middle Initial)

5. PATIENT'S ADDRESS (No., Street)

6. PATIENT RELATIONSHIP TO INSURED Self Spouse Child Other

7. INSURED'S ADDRESS (No., Street)

CITY STATE

8. PATIENT STATUS Single Married Other

CITY STATE

ZIP CODE TELEPHONE (Include Area Code) ( )

Employed Full-Time Student Part-Time Student

ZIP CODE TELEPHONE (Include Area Code) ( )

9. OTHER INSURED'S NAME (Last Name, First Name, Middle Initial)

10. IS PATIENT'S CONDITION RELATED TO:

11. INSURED'S POLICY GROUP OR FECA NUMBER

a. OTHER INSURED'S POLICY OR GROUP NUMBER

a. EMPLOYMENT? (Current or Previous) YES NO

a. INSURED'S DATE OF BIRTH MM DD YY SEX M F

b. OTHER INSURED'S DATE OF BIRTH MM DD YY SEX M F

b. AUTO ACCIDENT? PLACE (State) YES NO

b. EMPLOYER'S NAME OR SCHOOL NAME

c. EMPLOYER'S NAME OR SCHOOL NAME

c. OTHER ACCIDENT? YES NO

c. INSURANCE PLAN NAME OR PROGRAM NAME

d. INSURANCE PLAN NAME OR PROGRAM NAME

10d. RESERVED FOR LOCAL USE

d. IS THERE ANOTHER HEALTH BENEFIT PLAN? YES NO If yes, return to and complete item 9 a-d.

READ BACK OF FORM BEFORE COMPLETING & SIGNING THIS FORM.

12. PATIENT'S OR AUTHORIZED PERSON'S SIGNATURE I authorize the release of any medical or other information necessary to process this claim. I also request payment of government benefits either to myself or to the party who accepts assignment below.

SIGNED DATE

13. INSURED'S OR AUTHORIZED PERSON'S SIGNATURE I authorize payment of medical benefits to the undersigned physician or supplier for services described below.

SIGNED

14. DATE OF CURRENT: MM DD YY ILLNESS (First symptom) OR INJURY (Accident) OR PREGNANCY(LMP)

15. IF PATIENT HAS HAD SAME OR SIMILAR ILLNESS. GIVE FIRST DATE MM DD YY

16. DATES PATIENT UNABLE TO WORK IN CURRENT OCCUPATION MM DD YY FROM TO

17. NAME OF REFERRING PROVIDER OR OTHER SOURCE

17a.

17b. NPI

18. HOSPITALIZATION DATES RELATED TO CURRENT SERVICES MM DD YY FROM TO

19. RESERVED FOR LOCAL USE

20. OUTSIDE LAB? YES NO $ CHARGES

21. DIAGNOSIS OR NATURE OF ILLNESS OR INJURY (Relate Items 1, 2, 3 or 4 to Item 24E by Line)

1. 3.

2. 4.

22. MEDICAID RESUBMISSION CODE ORIGINAL REF. NO.

23. PRIOR AUTHORIZATION NUMBER

24. A. DATE(S) OF SERVICE From To MM DD YY MM DD YY | B. PLACE OF SERVICE | C. EMG | D. PROCEDURES, SERVICES, OR SUPPLIES (Explain Unusual Circumstances) CPT/HCPCS MODIFIER | E. DIAGNOSIS POINTER | F. $ CHARGES | G. DAYS OR UNITS | H. EPSDT Family Plan | I. ID. QUAL. | J. RENDERING PROVIDER ID. #

NPI
NPI
NPI
NPI
NPI
NPI

25. FEDERAL TAX I.D. NUMBER SSN EIN

26. PATIENT'S ACCOUNT NO.

27. ACCEPT ASSIGNMENT? (For govt. claims, see back) YES NO

28. TOTAL CHARGE $

29. AMOUNT PAID $

30. BALANCE DUE $

31. SIGNATURE OF PHYSICIAN OR SUPPLIER INCLUDING DEGREES OR CREDENTIALS (I certify that the statements on the reverse apply to this bill and are made a part thereof.)

SIGNED DATE

32. SERVICE FACILITY LOCATION INFORMATION

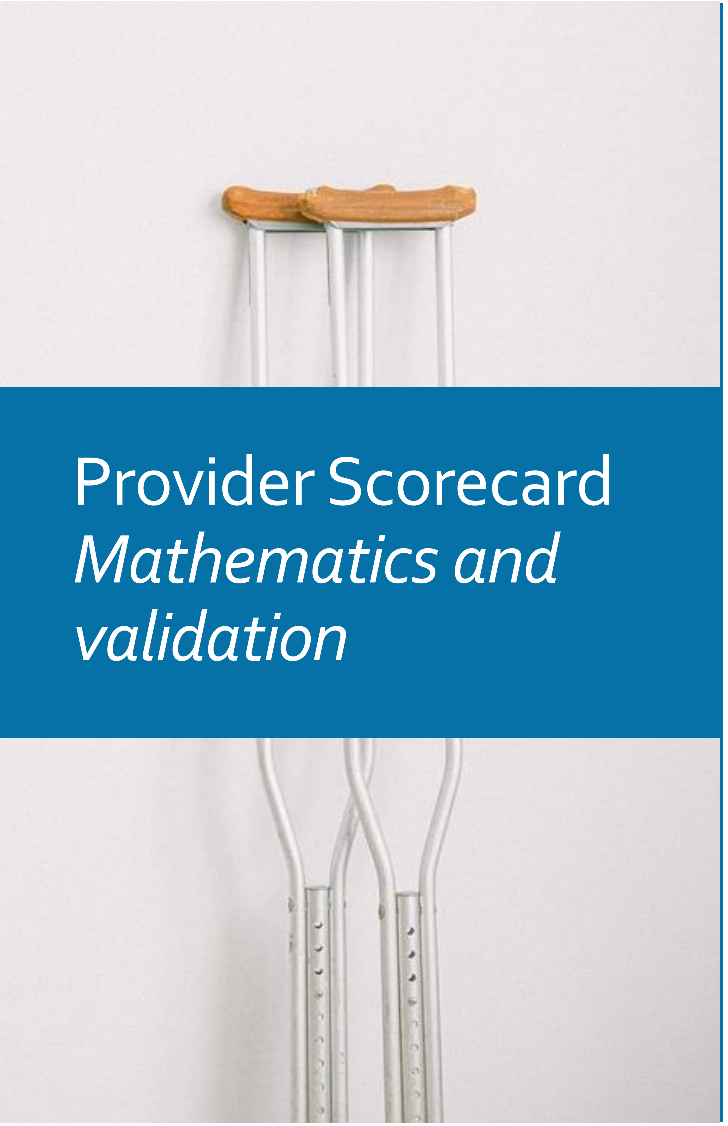
a. NPI b.

33. BILLING PROVIDER INFO & PH # ( )

a. NPI b.

All of this data plus:

- Hospital bills
- Pharmacy
- Medical Examinations
- Claims adjuster notes



Provider Scorecard *Mathematics and validation*

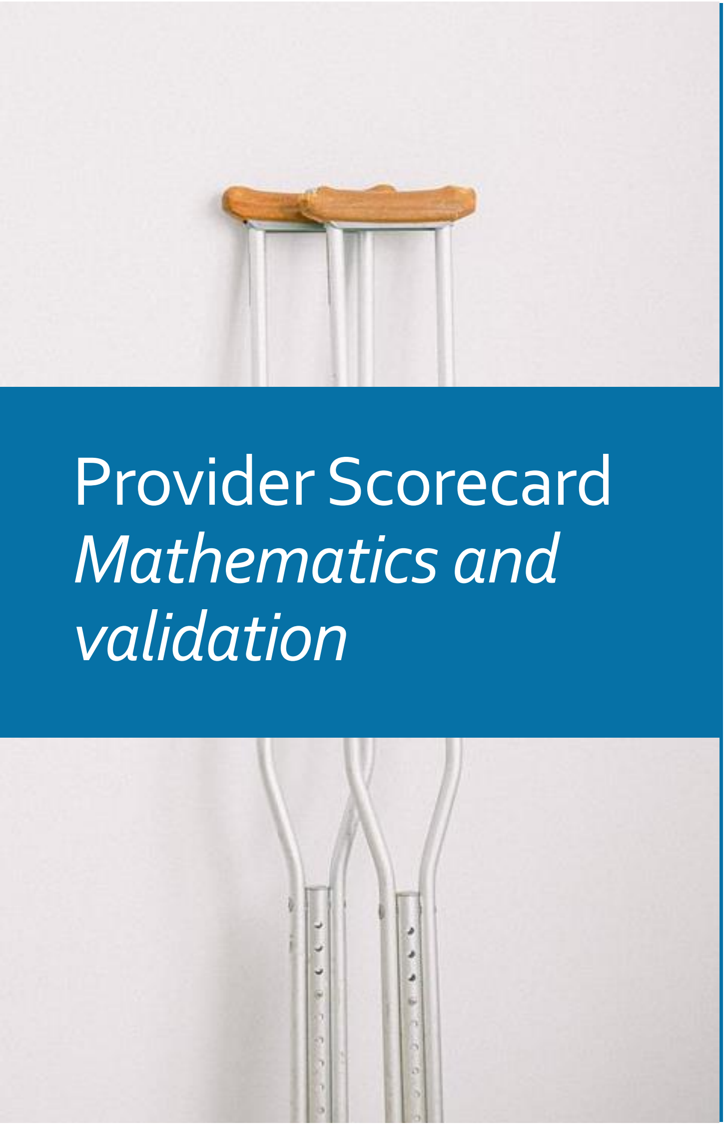
- Clustering techniques to group small data groups into larger
- General linear modeling, using the Gamma distribution, to make predictions about various types of outcomes
- All analysis was done on training and then compared to withheld datasets

Provider Scorecard *Outputs*

For every claimant

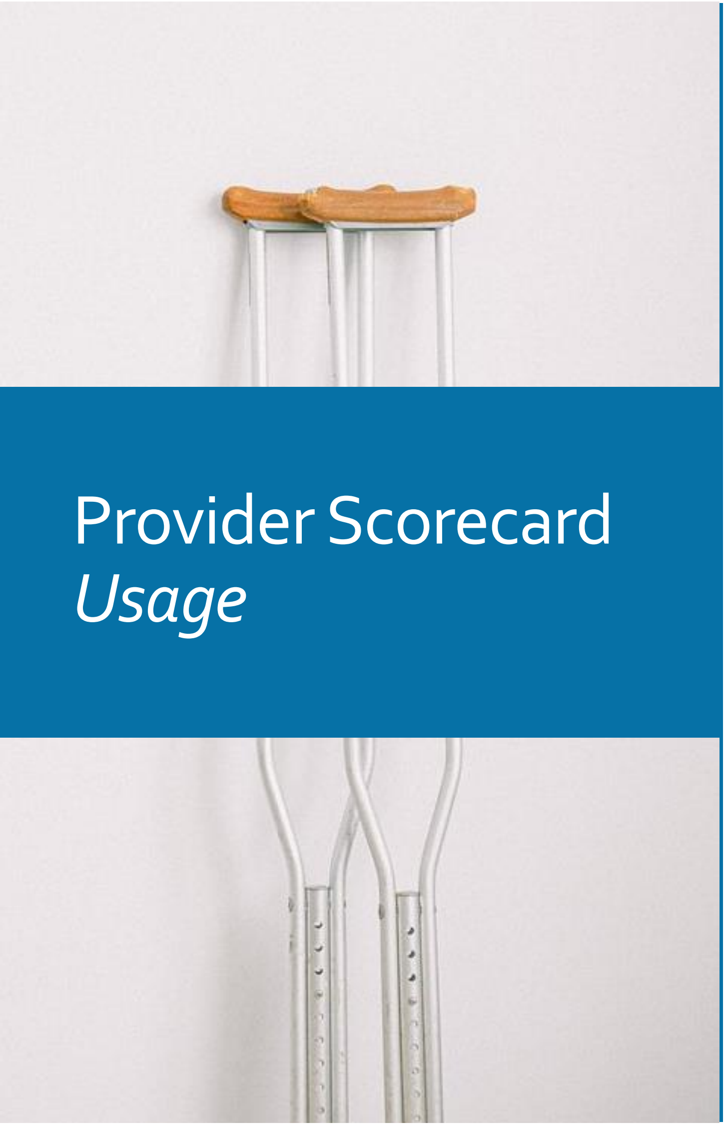
- Based on first 30 days of treatment we predicted the total:
  - Number of visits to doctor
  - Days off of work
  - Types of treatments
  - Medications used
  - % final disability score

This information was then aggregated by each physician in Texas, creating a scorecard for every physician (600+ physicians with credible data)



Provider Scorecard *Mathematics and validation*

- Results were sent to medical experts.
  - A small subset sent to validate
  - Multiple doctors / bill reviewers
  - Ranked the doctors
  - Our rank matched



Provider Scorecard
Usage

- 95% of our customers had chose the “network” option
- We developed a methodology to direct our claimants to the appropriate doctor given the circumstances
- Some doctors were removed from our network
- We believe we saved $35 million dollars a year!!!
- Helped a lot of people get the best treatment
- If I were injured I would want access to this tool

Provider Scorecard *Observations*

- Pre-existing commercial medical bill review / fraud software and processes were inadequate
- In house, insurance company, medical doctors are good for large claims, but do not have the data skills to police small to medium size claims.

Provider Scorecard *Observations*

- There are some very terrible doctors who somehow manage to:
  - See their patients wayyyyy more often then expected
  - Have patients that do not get better
- There are some remarkable doctors who:
  - Somehow help their patients get back to work quickly

Provider Scorecard
More Observations

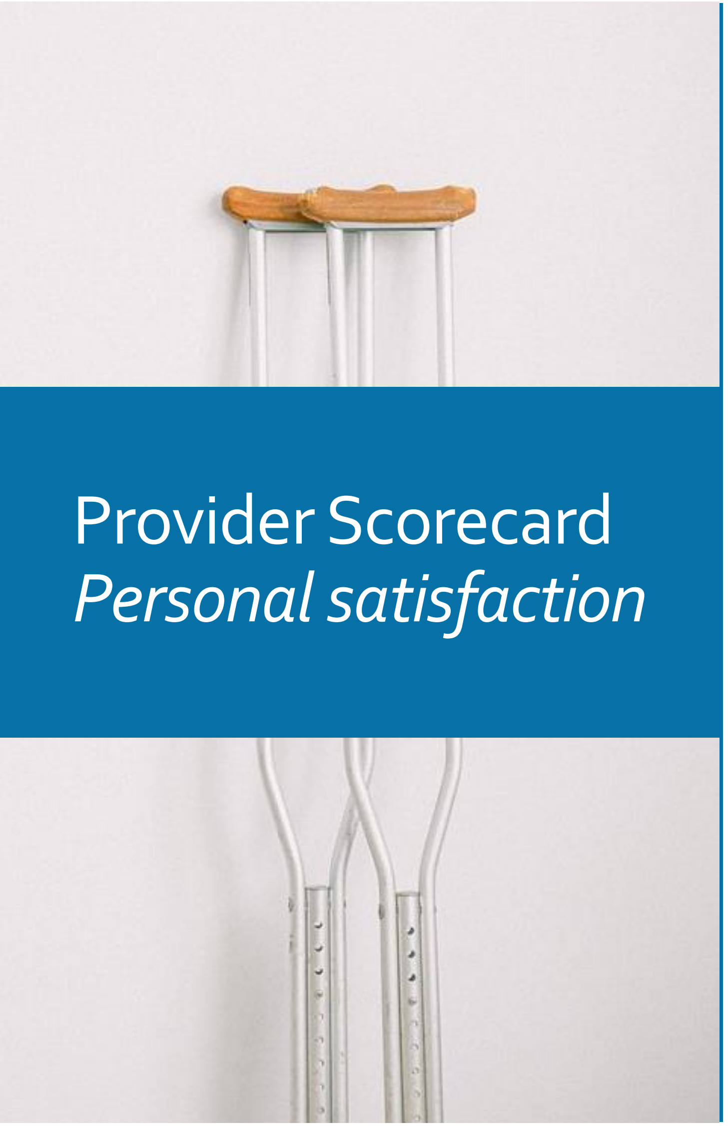
- Specialists:
  - Back surgeons/neurosurgeons—there are some very good surgeons (and bad ones)
  - Knee and ankle surgeons—be skeptical of "the state of the art".
- Just because a doctor has been a successful marketer, it does NOT mean the doctor is actually good at getting people back to work

Provider Scorecard
More Observations

- Some doctors upcode.
- Beware implant companies with ties to doctors
- Oxycontin is not good (everyone knows this now)
- Beware compound pharmacies

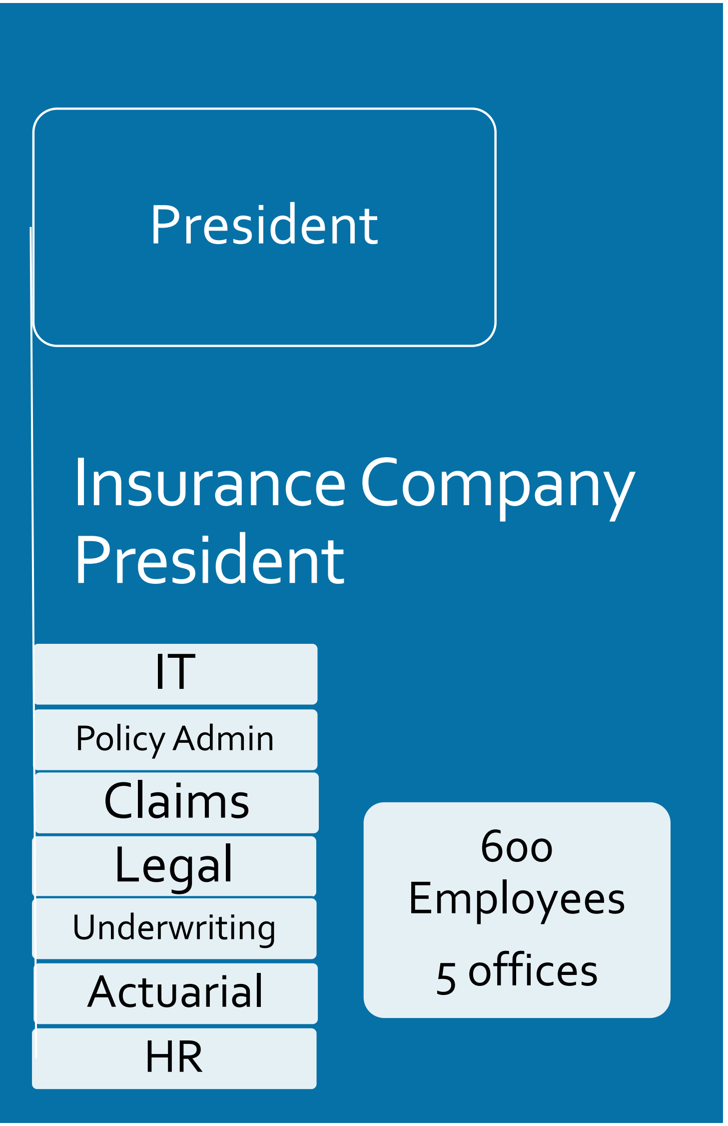
Provider Scorecard
More Observations

- Some doctors are fraudsters



Provider Scorecard
Personal satisfaction

- It was truly amazing to see that routing an injured person from doctor A to doctor B could provide noticeably better outcomes.
- I felt I had done a universal good.



President of Windhaven Insurance Company

- I was asked to be president, primarily because of my technical / IT competence
- I personally ran the IT department because the CIO departed, and we did not have time to get a replacement
- We performed a full system conversion in Texas and then Florida
- Eventually I assumed all roles seen to the right

Insurance Company President
Legal issues

Towards the end of our conversion, I became aware of some serious problems

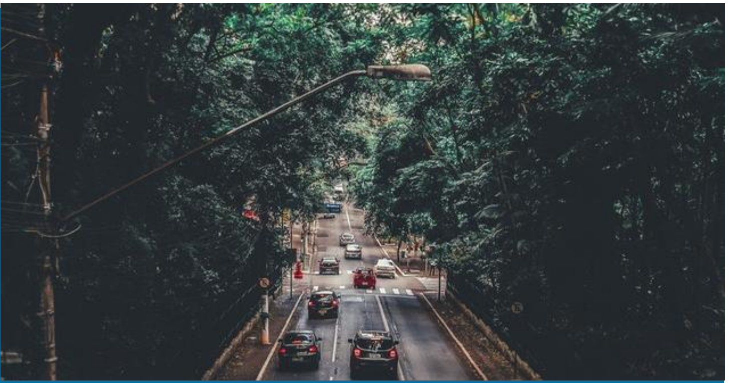
- Our chief legal officer quit
- I ran the legal department until we found a replacement
- There were many lawsuits for which our defense was challenged

Lawsuits in Florida *by year against insurance carriers*

Served Year	# Lawsuits
2016	192,598
2017	229,188
2018	278,739
2019	322,171
2020	400,943
2021	177,554
Total	1,601,193
2021 Projected	532,662

Lawsuits in Florida
Top carriers being sued

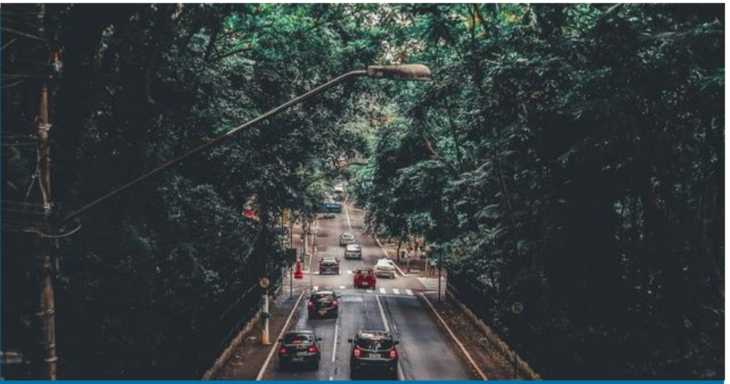
Defendant	**# Lawsuits**
GEICO	266,899
Progressive	215,983
State Farm	163,164
All State	71,348
Citizens Property Insurance Corporation	52,372
USAA	51,252
Universal Property & Casualty Insurance Company	38,018
Infinity Auto Insurance Company	34,022
Windhaven Insurance Company*	**33,568**
Blue Cross And Blue Shield Of Florida, Inc	30,794
Direct General Insurance Company	28,290
Ocean Harbor Casualty Insurance Company	25,024
Liberty Mutual Insurance Company	20,263
Infinity	18,261
United Services Automobile Association	16,378
United Automobile Insurance Company	16,161
Heritage Property & Casualty Insurance Company	15,679
Security National Insurance Company	13,891
United Property & Casualty Insurance Company	13,353
Tower Hill	12,467



PIP Fraud

PIP = auto insurance protection, aka, "no fault insurance"

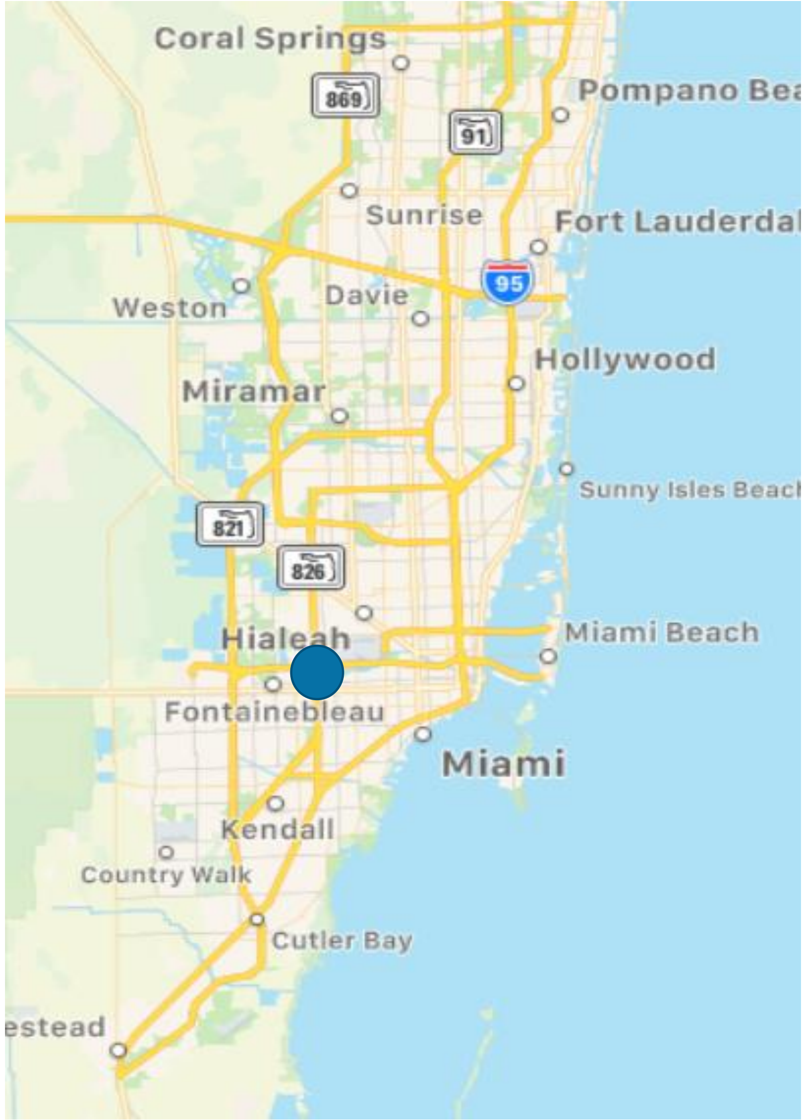
- After addressing lawsuits, it seemed like all was well
- However, our loss ratio was getting worse.
- An exhaustive effort was made to find what was driving the loss ratio
- GLMs were performed on every aspect of pricing, underwriting, territory, credit, etc.
- I determined that whatever was driving our loss ratio issue was independent of our various underwriting and pricing efforts
- The only thing that I could determine was that the location of the accidents and location of the medical clinics was unusual
- I literally created a geographic map which pinpointed where I believed fraud was occurring

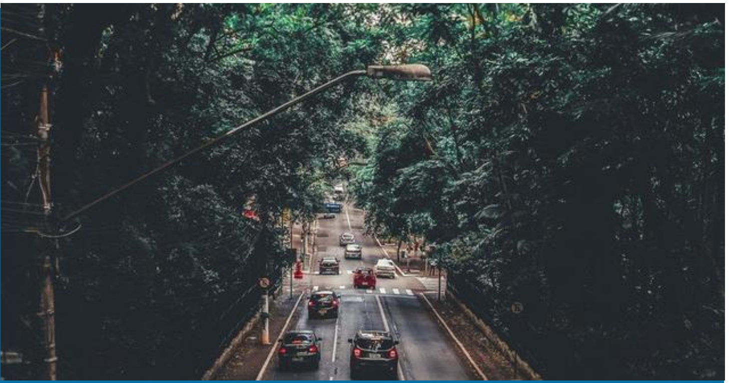


PIP Fraud

PIP = auto insurance protection, aka, “no fault insurance”

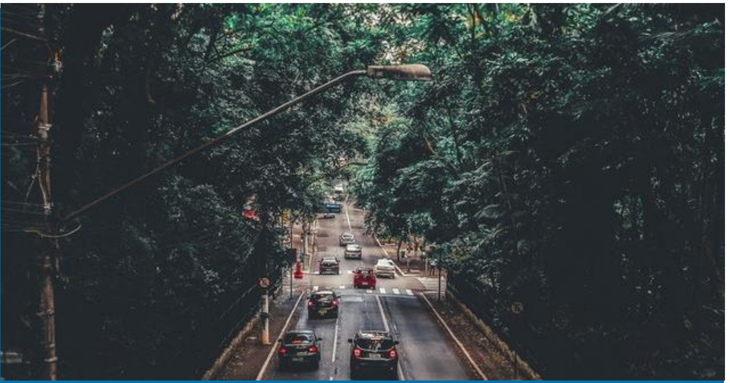
- I literally created a geographic map which pinpointed where I believed fraud was occurring





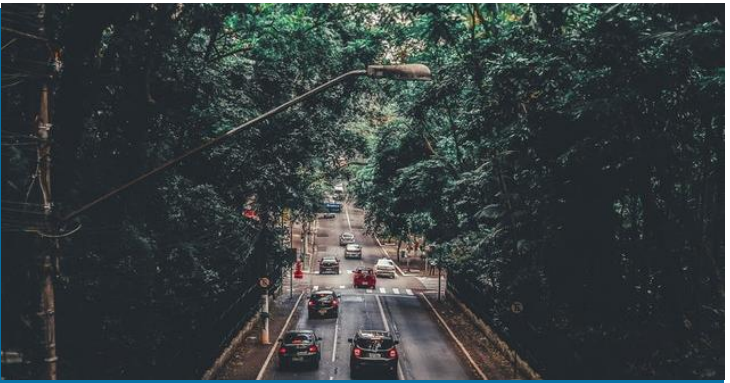
PIP Fraud
Scheme revealed

- We got a break in a case in North Florida (not Miami)
- We got a full-on confession from someone involved in a fraud scheme
- He explained how it works
- I will pause here and explain, but these are bullet points:
  - Individuals were recruited and
    - Lightly crashed their cars
    - Called the police to make a police report
    - Visited a clinic controlled by the crime syndicate
    - Signed forms indicated multiple days of visits
    - Received $5k per person who was in the car and participated
    - Recruited a friend to be the next person in a fake accident
  - Clinic then
    - Billed insurance carrier for many visits and medical activity that never happened



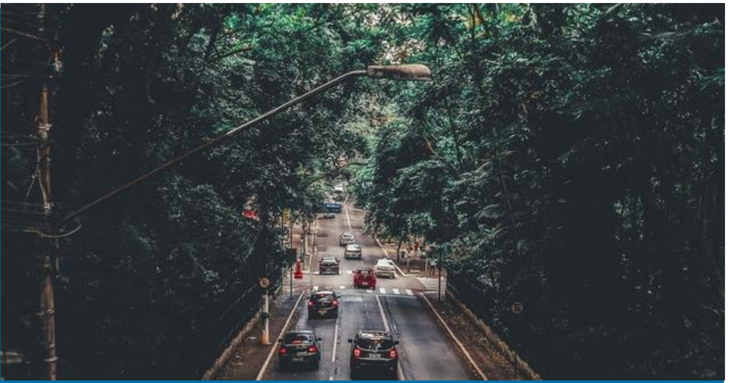
PIP Fraud
Carrier data

- I obtained all information from claims coming from the clinic with a confession
- The most interesting aspect of the clinic was its medical bills:
  - No other clinics involved in treatment of individual
  - Improbably similar treatment profile of each individual
  - Individuals who were in the other car were universally treated at a different clinic in the same strip mall as the clinic in question



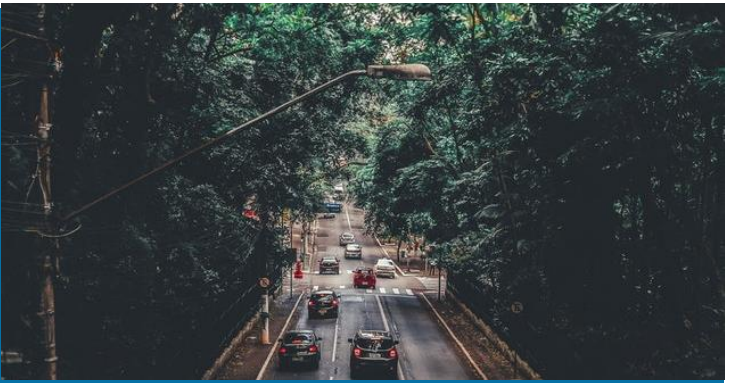
PIP Fraud
Medical Bills

- We obtained all electronic medical bills associated with all claims submitted to this carrier
- I built statistics similar to what I did at a workers compensation carrier, but focused on finding behavior similar to the clinic with the signed confession
- We found:
  - 15 clinics with statistical profiles WORSE than the clinic with signed confession
  - We found clinics where the CPT codes indicated that individual doctors were working more than 24 hours a day, just on our claimants!
  - We found that these clinics were associated with people who had been sued by other insurance carriers for years. They simply closed down old clinics and made new clinics.
  - A significant portion of our claims were coming from these 15 clinics



PIP Fraud
Dealing with known fraud

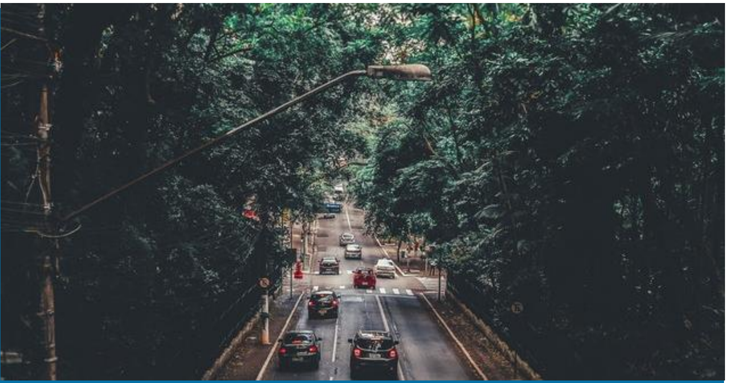
- While the police were heavily involved in our situation with a confession; THEY DECLINED TO DO ANY PROSECUTIONS
- Our company was being financially drained and traditional SIU (strategic investigations units) were of limited value
- Traditional SIU focuses on:
  - Convictions of criminals
  - After the fact lawsuits against clinics
  - This was of limited value in our circumstances
- Instead of traditional SIU, we did the following
  - Built statistical metrics on all clinics in Florida
  - Built metrics associated with individual claims that were suspect
  - Ran a nightly job of all incoming first notice of loss (FNOL) and all incoming bills
  - Flagged claims that met specific metrics
  - Withheld payments on claims (there was/is a legal provision for doing this in Florida) until the injured claimant submitt4ed to examinations under oath



PIP Fraud
Dealing with known fraud

- Instead of traditional SIU, we did the following
  - Withheld payments on claims (there was/is a legal provision for doing this in Florida), where we had very strong evidence of fraud, until the injured claimant submitted to examinations under oath

FL 627.736-6(g) An insured seeking benefits under ss. 627.730–627.7405, including an omnibus insured, must comply with the terms of the policy, which include, but are not limited to, submitting to an examination under oath. The scope of questioning during the examination under oath is limited to relevant information or information that could reasonably be expected to lead to relevant information. **Compliance with this paragraph is a condition precedent to receiving benefits.** An insurer that, as a general business practice as determined by the office, requests an examination under oath of an insured or an omnibus insured without a reasonable basis is subject to s. 626.9541.

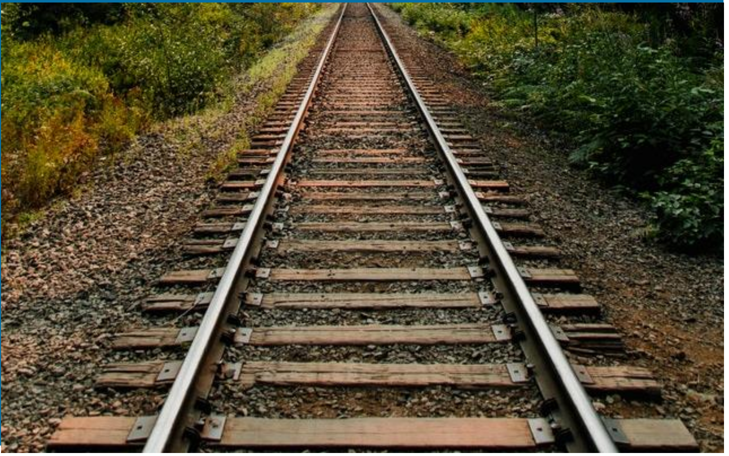


PIP Fraud
Examinations under oath

- We prepared a tool that would provide for SIU investigators to understand exactly what was alleged to have happened in the medical clinics
- When claimants showed up for an examination under oath, our investigators knew more details about the "accident" and its "treatment" then did the claimant
- Armed with more information than the claimant, the claimant would realize he was being caught under oath, and many actually confessed

- **As the confessions mounted, <u>total</u> incoming PIP claims dropped precipitously, 25 to 50%**
- These clinics and many others suddenly no longer had claimants
- You cannot sue a carrier if you have not filed a claim!

Transition to New Life



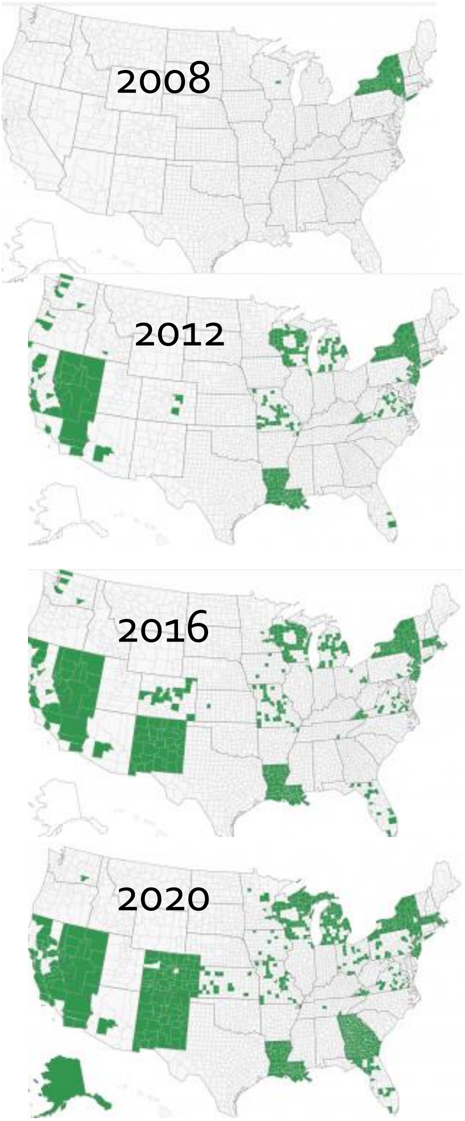
- Our company was damaged financially too much to continue
- In bankruptcy proceedings, pieces of the company were bought by investment capital
- I was asked to be the president of the new company
- I had the opportunity of running all aspects, including marketing and finance
- I worked for this new company for five months, but decided the difference in priorities and styles was too different and I was wasting my time working with them
- I resigned

- While doing freelance work, multiple people approached me and asked me about our process of preventing fraud
- Right as this was happening, the US presidential election happened

U.S. Election

- Everyone here knows the U.S. election was very contentious
- I assume everyone here knows that some people alleged fraud occurred during the election
- I heard a fraud theory related to election brand machines that seemed very unlikely, and I decided to investigate it
- To do this I did my own work, but ended up submitting my work to multiple professionals (actuaries and PhDs) to critique my work
- I used many of the skills I had used in past fraud investigations, but I also learned some techniques that I find interesting and useful for the actuarial community

U.S. Election



Did the introduction of this machine affect election outcomes?

- Cleanest way to test is to compare 2008 vs 2020
- Change in 2008 to 2020 was

Dominion Counties vs. Non-dominion Counties

Dominion	Number of Counties	Average of Difference
No	2,393	-9.0%
Yes	657	-6.4%
Total	3,050	-8.4%

Simple Linear Regression: Ordinary Least Squares

Variable	Coefficient	P-value
Intercept	-8.98%	0.000%
Dominion	2.55%	0.000%

Traditional Regression

- A traditional regression is defined as follows:
- $Yi = B_{1i}*X_{1i} + B_{2i}*X_{2i} + \ldots + Bni*X_{ni} + Error_i$
- If we define U as a column of errors, then in matrix form:

$$Y = X\beta + U.$$

- Beta is estimated in this manner

$$\widehat{\beta}_{\mathrm{OLS}} = (\mathbb{X}'\mathbb{X})^{-1}\mathbb{X}'\mathbb{Y}.$$

- The variance and standard deviation of Beta are used for hypothesis testing
- Overall the variance of beta can be found as

$$V\left[\widehat{\beta}_{\mathrm{OLS}}\right] = V[(\mathbb{X}'\mathbb{X})^{-1}\mathbb{X}'\mathbb{Y}] = (\mathbb{X}'\mathbb{X})^{-1}\mathbb{X}'\Sigma\mathbb{X}(\mathbb{X}'\mathbb{X})^{-1}$$

where $\Sigma = V[u]$.

- However, nearly all people are taught to assume that the sample errors have equal variance and σ2 and are uncorrelated, which reduces the formula to be:

$$v_{\mathrm{OLS}}\left[\widehat{\beta}_{\mathrm{OLS}}\right] = s^2(\mathbb{X}'\mathbb{X})^{-1}, \quad s^2 = \frac{\sum_i \widehat{u}_i^2}{n-k}$$

- Which, in practical terms means, for each beta, you take the diagonal $(\mathbb{X}'\mathbb{X})^{-1}$ of and multiply it by estimated variance of the error term.

Traditional Regression

A different way of looking at this

A traditional regression has these assumptions

- The regression model is linear in the coefficients and the error term
- The error term has a population mean of zero
- All independent variables are uncorrelated with the error term
- Observations of the error term are uncorrelated with each other
- The error term has a constant variance (no heteroscedasticity)
- No independent variable is a perfect linear function of other explanatory variables
- The error term is normally distributed (optional)

If this is all met, then (normally distributed option):

$$v_{\mathrm{OLS}}\left[\widehat{\beta}_{\mathrm{OLS}}\right] = s^2(\mathbb{X}'\mathbb{X})^{-1}, \quad s^2 = \frac{\sum_i \widehat{u}_i^2}{n-k}$$

Heteroscedasticity-consistent standard errors using Sandwich Equation

However, we can cross out one assumption, like this:

- The regression model is linear in the coefficients and the error term
- The error term has a population mean of zero
- All independent variables are uncorrelated with the error term
- Observations of the error term are uncorrelated with each other
- ~~**The error term has a constant variance (no heteroscedasticity)**~~
- No independent variable is a perfect linear function of other explanatory variables
- The error term is normally distributed (optional, only needed if for confidence testing)

And then, we are back to the equation I showed two slides ago:

$$V\left[\widehat{\beta}_{\text{OLS}}\right] = V[(\mathbb{X}'\mathbb{X})^{-1}\mathbb{X}'\mathbb{Y}] = (\mathbb{X}'\mathbb{X})^{-1}\mathbb{X}'\Sigma\mathbb{X}(\mathbb{X}'\mathbb{X})^{-1} \quad \text{where } \Sigma = V[u].$$

and where V[u] = a diagonal matrix where the diagonal is made up of the square of each observation

Then we take the diagonal of the big equation above, also known as “Sandwich Equation”, to obtain the standard error of each beta.

The actual estimated betas are exactly the same using either technique, only the confidence interval changes

Why wouldn’t you <u>always</u> drop the assumption of no heteroscedasticity?

Sandwich Equation

- Can be computed in Excel with this formula, and the diagonal of the matrix produced is the standard error of each beta:

$$(X^{t}X)^{-1}X^{t}\mathrm{diag}(ee^{t})X(X^{t}X)^{-1}$$

```
= MMULT(MMULT(MMULT((XtX)-1,TRANSPOSE(ColumnOf1s))
*TRANSPOSE(ColumnOfErrorTerms*ColumnOfErrorTerms),(ColumnOf1s),(XtX)-1)
```

- You can also compute this type of regression in R using this package:
  - Library("sandwich")
- And running this code after performing a regression:
  - coeftest(regression, vcov = sandwich)
- The code can be run using weighted least squares, and Excel can be modified for weighted least squares as well.
- I have sample code on my website, and alternatively, I would be happy to email you sample code.

Machine Fraud *Theory*

Multiple Linear Regression: Weighted Least Squares, Two types of P-values

Variable	**Coefficient**	**P-value**	**P-value Consistent**
Intercept	-5.36%	0.00%	0.00%
RuralUrbanContinuumCode2013	-0.83%	0.00%	0.00%
ManufacturingDependent2000	-2.69%	0.00%	0.00%
HiAmenity	-0.51%	0.00%	27.41%
HiCreativeClass2000	5.74%	0.00%	0.00%
Low_Education_2015_update	3.00%	0.00%	0.00%
PopChangeRate1019	0.19%	0.23%	0.00%
Net_International_Migration_Rate _2010_2019	0.18%	0.34%	29.70%
Dominion	1.55%	0.00%	0.11%

Machine Fraud
What are we testing?

- We are testing the theory that a county using a particular brand machine will have results that are systematically different than what would otherwise be expected
- P-values are helpful, and they were very important to some of my reviewers, but the real questions are:
  - Do the counties of a particular machine perform different than randomly assigned counties?
  - If so, does this occur even when adjusting for demographic variables?
- Note that although demographic variables are included in this analysis:
  - They are there to control for demographic effects.
  - They are not in competition with the machine variables.
  - We are not looking for the “best fit” for predicting counties.
- We are looking to understand if a given theory about fraud appears to hold true after adjusting for known factors.
- This is similar to doing a fraud study where you use known, understood, insurance variables such as credit score, territory, age, etc., to adjust, but the real question is whether the fraud theory performs better than randomness.
- We can compare the fraud theory to randomness using
  - p-values or
  - inserting random variables as a comparison
- We will do a little of both.

Machine Fraud *Theory*

- The machine fraud theory is that a particular machine is somehow compromised (either via a malicious hack or intention of developers) and altered the election outcome
- The machine that was singled out, was not prevalent in 2008, so as a control, we use 2008 as a starting point
- The formula being tested is:

$$\frac{Dem.Pres.Vote.2020}{Dem.Pres.Vote.2020+Rep.Pres.Vote.2020} - \frac{Dem.Pres.Vote.2008}{Dem.Pres.Vote.2020+Rep.Pres.Vote.2008} = a + b1\times X1 + b2\times X2 + b3\times X3 + e$$

- Each observation is a U.S. County

Machine Fraud Theory
Demographic and random variables

- I included 155 demographic variables, all obtained from a government website at: https://www.ers.usda.gov/data/ruralATlas/
- These demographic variables contain things like you would expect such as:
  - Race
  - Education
  - Population Density
- Where possible I created variables to imitate groupings similar to the size of the block of counties that a machine brand has
- Where possible I created variables to test changes in demographics over time
- I added 24 random variables as follows:
  - B1, B2, B3, B4, B5, B6: Binary fields of 0 or 1 according to a Bernoulli trial with probability of 0.25 (B1, B2), 0.50 (B3, B4), 0.75 (B5, B6).
  - C1, C2, C3, C4: The C1 flag indicates that the first letter of the county is among A through L. C2, indicates the same thing for the 2nd letter of the county. C3 and C4 follow this pattern but for the 3rd and 4th letter.
  - N1, N2, N3, N4, N5: These are all unit normal random variables.
  - S1, S2, S3, S4: These are the same as C1, C2, C3, C4, except the operation is done on state names instead of county names.
  - U1, U2, U3, U4, U5: These are all uniform distributed random variables from 0 to 1.

Machine Fraud Theory
Machine Variables

Machine Flag	Counties Flagged	% of Total 2008 + 2020 Vote
ESS	**1456**	**44.50%**
Dominion	**629**	**28.50%**
Democracy Live	**436**	**24.80%**
VotingWorks	273	11.90%
Unisyn Voting Solutions	248	4.00%
Total (counties can have more than one brand)		113.70%

Machine Fraud Theory *Testing*

- I tested many ways, but the method I think is most interesting is when I iterated over every possible 3 variable combination.
- For example, one permutation could be:
  - Change in Vote Share = a*intercept + b1*Age65AndOlderPct2010 + b2*PopDensity2010 + b3***WhiteNonHispanicPct2010** + error term.
- Then the next permutation could be the same as the prior with one change:
  - Change in Vote Share = a*intercept + b1*Age65AndOlderPct2010 + b2*PopDensity2010 + b3***BlackNonHispanicPct2010** + error term.
- This can be tested 185*184*183 = 6,093,160 times, or since this should not be order dependent, it can be tested 185*184*183/3! = 1,038,220 times, for same answer.

Machine Fraud Theory
Random variable results

Coef. Name	# Times Regressed	Sign	% Inconsistent Sign	Percent Consistent and Significant at 5%	Sandwich: Percent Consistent and Significant at 5%
U4	16,836	1	0%	99%	78%
B1	16,836	1	0%	95%	40%
B4	16,836	1	0%	92%	18%
N2	16,836	1	1%	88%	2%
B6	16,836	1	1%	87%	2%
N1	16,836	1	1%	80%	1%
N5	16,836	1	0%	80%	17%
N4	16,836	-1	2%	66%	0%
S3	16,836	-1	21%	54%	25%
B2	16,836	1	3%	53%	2%
S4	16,836	1	19%	48%	12%
U5	16,836	-1	6%	46%	1%
S1	16,836	1	25%	45%	12%
N3	16,836	1	5%	45%	0%
C2	16,836	1	12%	43%	1%
B3	16,836	1	3%	39%	0%
U3	16,836	-1	9%	37%	0%
C1	16,836	1	9%	32%	0%
U1	16,836	-1	42%	31%	0%
U2	16,836	1	17%	23%	2%
S2	16,836	-1	35%	21%	6%
B5	16,836	-1	49%	19%	0%
C4	16,836	1	34%	15%	1%
C3	16,836	-1	44%	13%	0%

Machine Fraud Theory
Results for random variables, resorted

Coef. Name	# Times Regressed	Sign	% Inconsistent Sign	Percent Consistent and Significant at 5%	Sandwich: Percent Consistent and Significant at 5%
U4	16,836	1	0%	99%	78%
B1	16,836	1	0%	95%	40%
S3	16,836	-1	21%	54%	25%
B4	16,836	1	0%	92%	18%
N5	16,836	1	0%	80%	17%
S4	16,836	1	19%	48%	12%
S1	16,836	1	25%	45%	12%
S2	16,836	-1	35%	21%	6%
N2	16,836	1	1%	88%	2%
B6	16,836	1	1%	87%	2%
B2	16,836	1	3%	53%	2%
U2	16,836	1	17%	23%	2%
N1	16,836	1	1%	80%	1%
U5	16,836	-1	6%	46%	1%
C2	16,836	1	12%	43%	1%
C4	16,836	1	34%	15%	1%
N4	16,836	-1	2%	66%	0%
N3	16,836	1	5%	45%	0%
B3	16,836	1	3%	39%	0%
U3	16,836	-1	9%	37%	0%
C1	16,836	1	9%	32%	0%
U1	16,836	-1	42%	31%	0%
B5	16,836	-1	49%	19%	0%
C3	16,836	-1	44%	13%	0%

Machine Fraud Theory

Results for machines and random variables

Coef. Name	Sign	# Times Regressed	% Inconsistent Sign	Percent Consistent and Significant at 5%	Sandwich: Percent Consistent and Significant at 5%	Sandwich: Median P-value
Dominion	**1**	**16,836**	**0.00%**	**99.90%**	**86.90%**	**0.40%**
U4	1	16,836	0.10%	99.10%	77.70%	2.60%
Unisyn Voting Solution	-1	16,836	0.00%	100.00%	62.10%	3.60%
B1	1	16,836	0.00%	95.50%	39.60%	7.90%
S3	-1	16,836	21.20%	54.00%	24.90%	24.30%
Hart InterCivic	1	16,836	3.00%	88.70%	22.40%	10.90%
B4	1	16,836	0.10%	91.70%	18.10%	12.80%
N5	1	16,836	0.20%	80.00%	16.70%	12.00%
S4	1	16,836	19.40%	47.90%	12.40%	27.70%
Democracy Live Flag	**1**	**16,836**	**4.60%**	**69.60%**	**11.70%**	**22.60%**
S1	1	16,836	24.90%	45.30%	11.60%	38.90%
ESS	**-1**	**16,836**	**36.40%**	**37.10%**	**6.00%**	**43.80%**
S2	-1	16,836	34.60%	21.20%	5.60%	58.00%
VotingWorks	-1	16,836	37.90%	19.10%	5.30%	64.40%
B2	1	16,836	3.20%	53.10%	2.10%	35.00%
***						

Public Data is Great!

- Bringing this back to insurance, in some cases fraud analysis can be done by publicly available data
- The following screenshots are from a tool I created based on publicly available data to help small Florida PIP (personal injury protection, aka no-fault) carriers ascertain the threat level when they receive a new bill.
- For every incoming bill, even for very small companies with no credible data
  - The company can enter in an NPI
  - The company immediately learns how litigious the doctor is.
- Here is a link provided for participants in this meeting, if you would like to test it for yourself:
- www.fraudspotters.com/CAGNY/
- user: cagny, password: 2021

Provider Radar Example 1 (only to discuss if we can't show the tool live)

Based on the foregoing, on a scale of 1 to 10, this is how we rate this NPI:	
Method	**Score**
NPI Lawsuits	1
NPI Taxonomy	1.00
NPI Attorneys	1
Overall	1

Provider Radar Example 2 (only to discuss if we can't show the tool live)

Based on the foregoing, on a scale of 1 to 10, this is how we rate this NPI:	
Method	**Score**
NPI Lawsuits	1
NPI Taxonomy	5.56
NPI Attorneys	1
Overall	5.56

Provider Radar Example 3 (only to discuss if we can't show the tool live)

Based on the foregoing, on a scale of 1 to 10, this is how we rate this NPI:	
Method	**Score**
NPI Lawsuits	1.02
NPI Taxonomy	1.06
NPI Attorneys	2.17
Overall	2.17

Provider Radar Example 4 (only to discuss if we can't show the tool live)

Based on the foregoing, on a scale of 1 to 10, this is how we rate this NPI:	
Method	**Score**
NPI Lawsuits	9.33
NPI Taxonomy	1.15
NPI Attorneys	5.30
Overall	9.33

Public Data

- Public data, cleverly joined and thoughtfully analyzed, can provide:
  - Instant analysis of potential issues with claims
  - Independent of the size of the book of the carrier

Recap / Last Slide

We talked about the following

- Provider scorecard tool
- Lawsuits in Florida
- Using medical bills to identify PIP fraud
- Sandwich equation
- Using public data and random variables to evaluate issues such as the election
- Using public data to triage incoming claims

Questions?

- Please no political questions or comments

Contact

- https://www.linkedin.com/in/benjamin-j-turner/
- https://www.fraudspotters.com/contact/