RhoAI

Sports Media, Motorsports, etc.

Andrew Maness, Rho AI

January 30, 2020

Outline for Today's Class

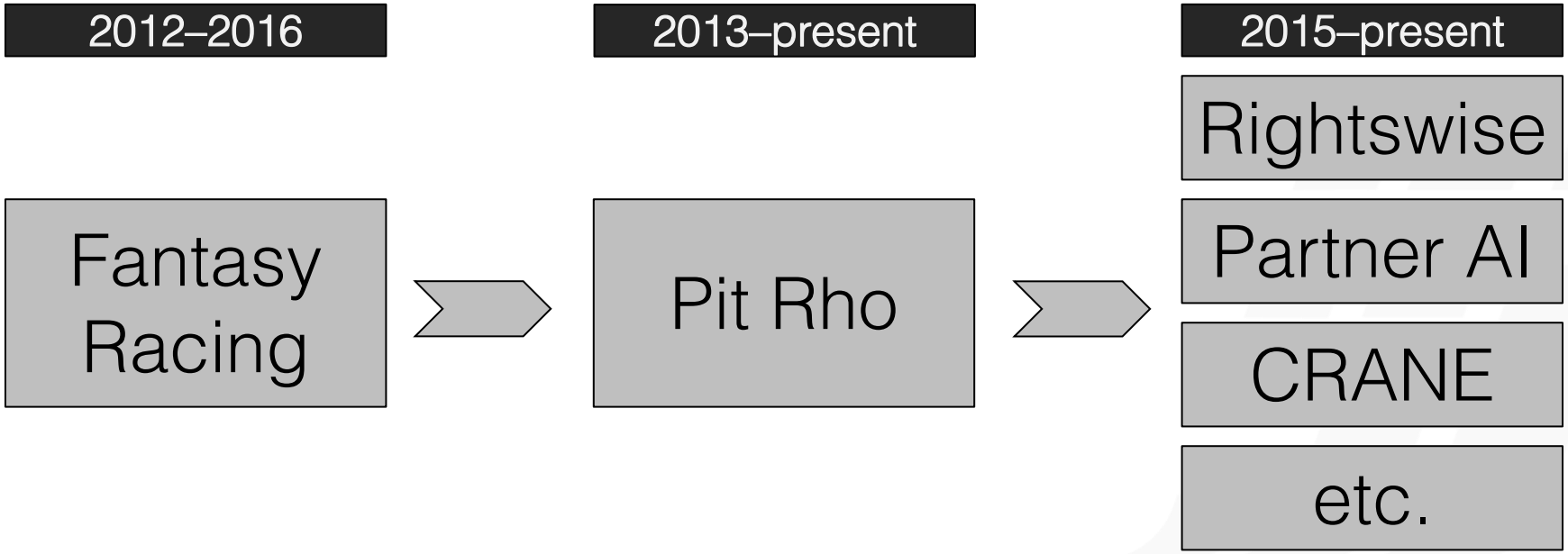
- My background / Rho AI (~3 min.)
- Rightswise (~10 min.)
- Motorsports issues (~2 min.)
- Pit Rho (~10 min.)
- Questions, comments, etc.

About Andrew Maness (pre-Rho)

- Kansas State University, 2006–2010
  - Mathematics
  - Statistics
- Wichita State University, 2010–2012
  - Economics
  - Finance
- Federal Reserve Bank, 2012–2014
  - Housing markets
  - Financial stress–testing
- RACINGnomics (formerly NASCARnomics)
  - Quant.–driven research of auto racing industry
  - Web-based essays, lite consulting

About Rho AI

- Rho AI (formerly Pit Rho), 2014–present
  - Founded in 2012 (4 people – now 39; no outside capital)
  - Next-level insight through accurate, unique, fast data
  - *e.g.*, next-level clean water distribution, smart investing
- My role: Tech. director, previously senior / analyst



Data Science and Sports

Rho AI draws from the roots and leadership of the MIT Baseball Hedge Fund, Cargill, MC10, the Houston Rockets, Red Bull Racing, the MIT Sloan Sports Analytics Conference, and leading venture capital firms and startups.

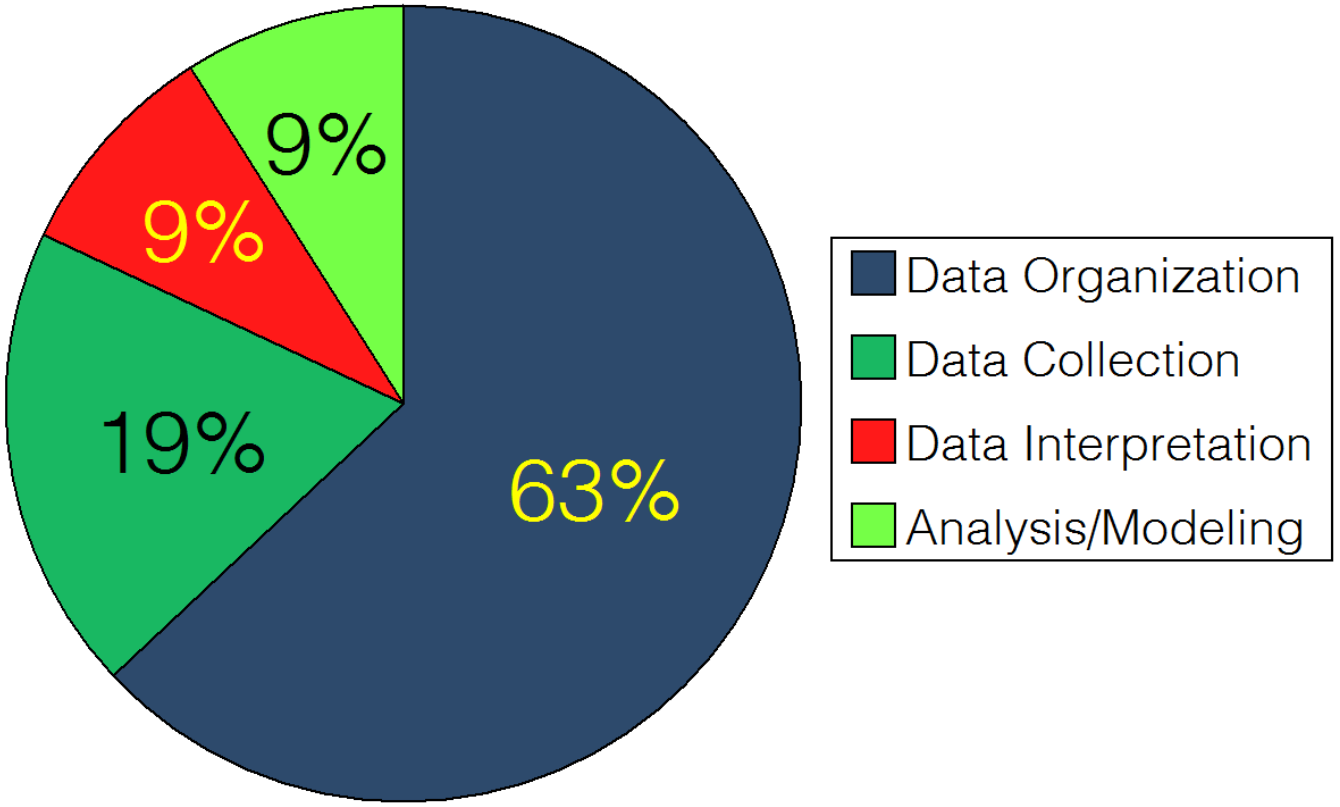
Rho AI was built to develop custom, predictive analytics solutions that provide actionable insights in real-time over distributed hardware.

This is underpinned with tools and capabilities hardened by the demanding environment of real-time motorsports strategy calculations.

Current Partnerships

How a Data Scientist Spends Time

- ~82% devoted to data collection/organization
- Source: [2016 Data Science Report]

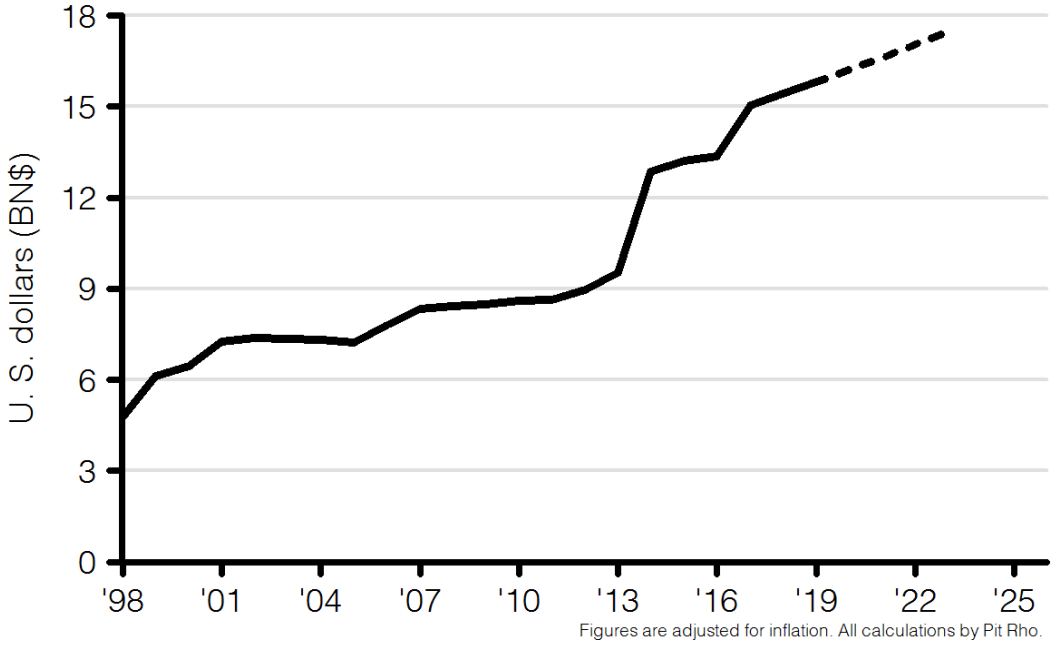


Rightswise Terminal

- Live valuations of sports media rights contracts
  - Ex.: ESPN pays ~$2.1BN/year for NFL's "MNF"
  - Analogous to stock–market services (*e. g.*, Google Finance)
  - Updates theoretical market price of contracts every minute
- Terminal is proprietary; we'll discuss framework
- Why is the Rightswise valuable / important?
  - Sports media rights market is based mostly on intuition
  - Complements industry experts with rigorous models based off historical events to determine factors of pricing
  - There's a lot of cash involved in this industry currently
  - Lots of changes – trying to objectively make sense of it

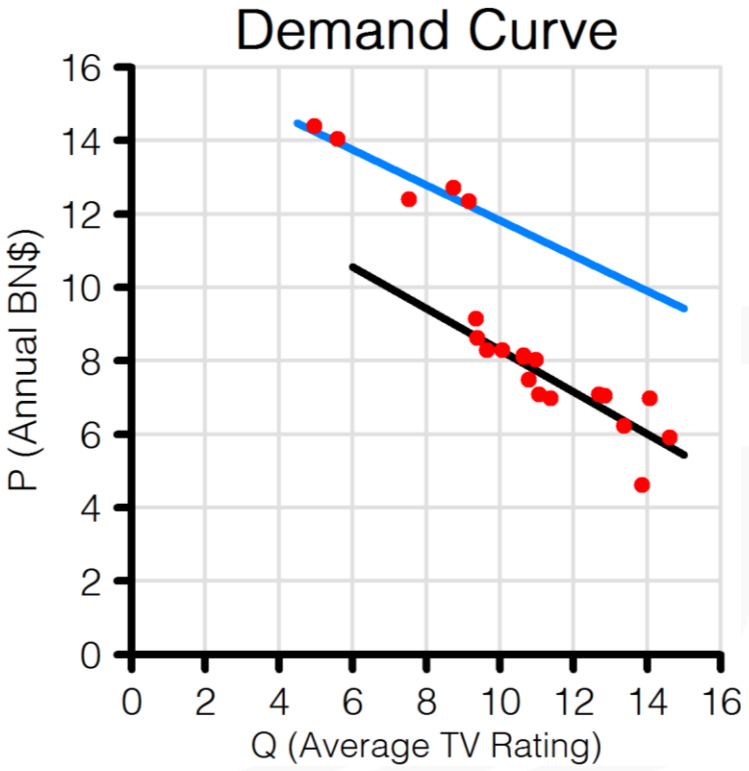
Size of Sports Media Rights Market

- Total money (BN$) paid by networks in exchange for broadcasting rights
- Across 21 major sports series; national contracts only, i.e. not international or local



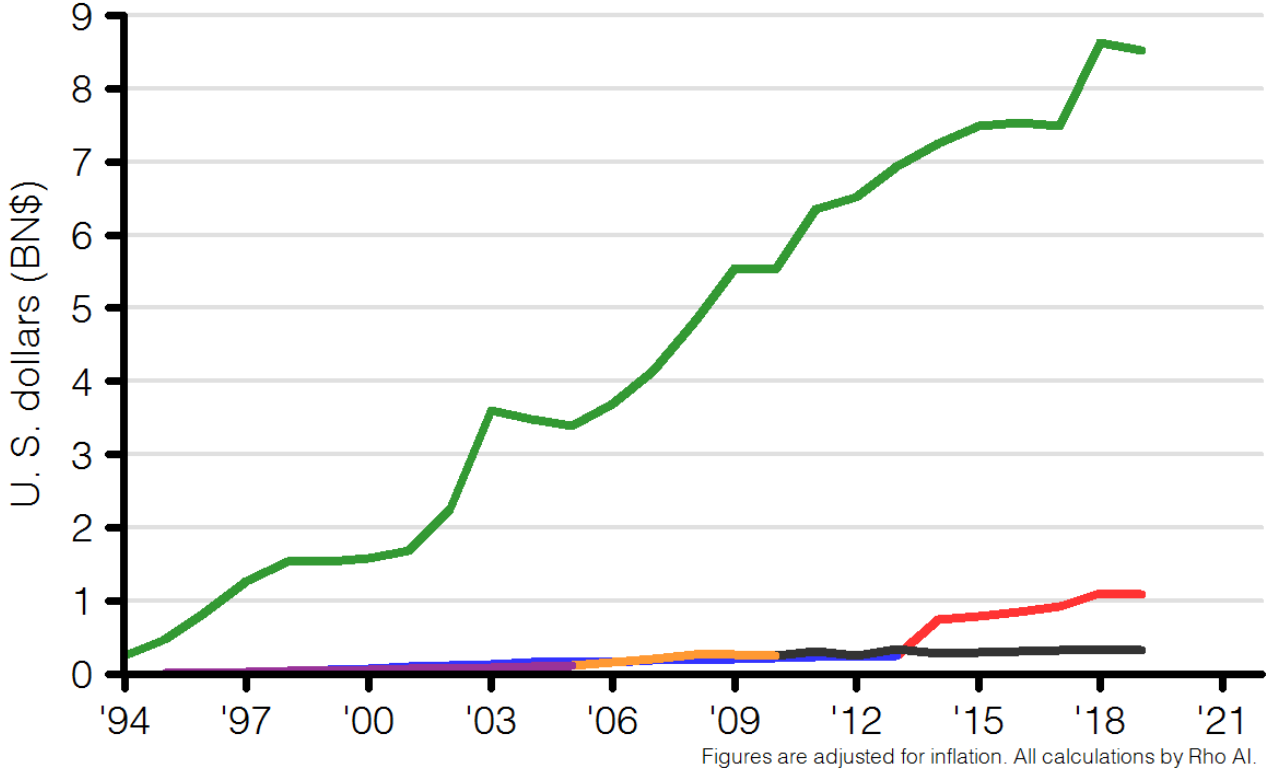
What's Driving This Increase? (I)

- Increase in payments for broadcasting rights
- Sharp increase in size of market, c. 2013
  - Ex.: Big 12 received ~$48MM from Disney/FOX in 1996; ~$217MM in 2013
- "Law of demand"; demand curve shifts right
  - Income ↑: Buyers' (networks') income increases steadily
  - Change in taste and preferences: "More relevant to people"
  - # buyers ↑: FOX Sports 1 and NBC Sports Network



What's Driving This Increase? (II)

- Sub. revenue (BN$) by network and year
- ESPN (green), Speed (blue) / FOX Sports 1 (red), OLN (purple) / VS (orange) / NBCSN (blk)



Modeling Contract Prices

- Not all contracts are created equally
  - Sports media rights market is sum of all deals
  - Individual contracts are heterogeneous; not “one size fits all”
  - Difficult to estimate demand of each contract
- Hedonic regression modeling
  - HRM approaches demand from the margins
  - Treats each feature of contract independent of other elements
  - Additive: sum of marginal prices equals contract price
- Annual price of contract = f(local, env.)
  - Annual price of an individual contract is influenced by local and environmental factors
  - Need data on networks, sports series, and contracts to model

Gathering Data: Networks

- Buyers of traditional rights contracts are nets.
  - Ex.: ESPN purchases right to air NFL, NBA games
  - Recently, streaming has become new element in contracts
- Examine reach, fees of 41 nets.; 1994-2019
  - Reach: # households with subscription to network
  - Fee: Per-month-per-household fee paid for network
  - Infer monthly subscription revenue: reach × fee
- 4,793 data points describing buyers / networks
  - Gathered through 544 sources; available upon request
  - Data collected manually (nontrivial)

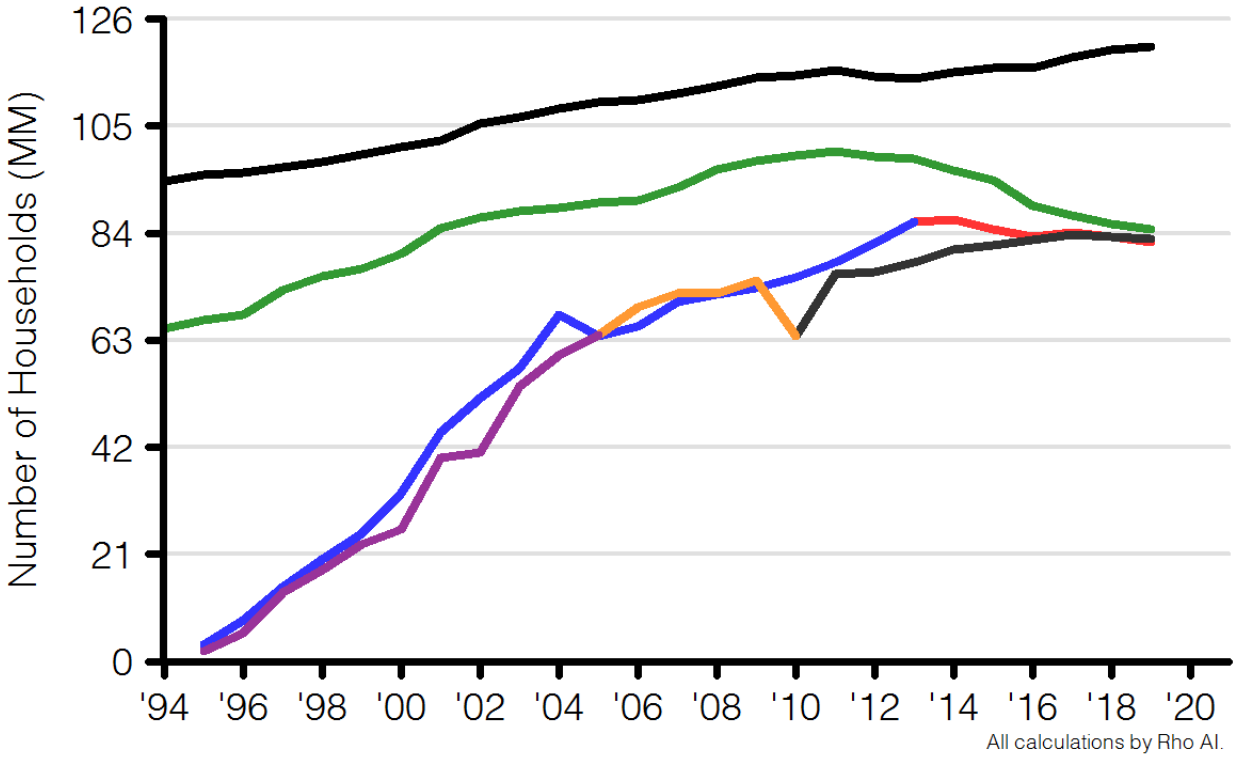
List of Networks Examined

Network	Years of Data	Availability
ABC	1994–2018	over-the-air
AXS TV	2012–2018	cable
Big 12 Network	1994–2013	syndication
Big Ten Network	2006–2018	cable
CBS	1994–2018	over-the-air
CBS Sports Network	2008–2018	cable
CNBC	1994–2018	cable
CNN Sports Illustrated	1996–2002	cable
College Sports TV	2002–2007	cable
ESPN	1994–2018	cable
ESPN+	1994–2005	syndication
espn2	1994–2018	cable
ESPNU	2005–2018	cable
FOX	1994–2018	over-the-air
FOX Soccer Channel	1996–2012	cable
FOX Sports 1	2012–2018	cable
FOX Sports 2	2013–2018	cable
FOX Sports Net	1996–2018	cable
Fuel TV	2003–2012	cable
FX	1994–2018	cable
FXX	2013–2018	cable

Network	Years of Data	Availability
Golf Channel	1995–2018	cable
HDNet	2001–2011	cable
MLB Network	2008–2018	cable
NBA TV	1999–2018	cable
NBC	1994–2018	over-the-air
NBC Sports Network	2011–2018	cable
NFL Network	2003–2018	cable
NHL Network	2007–2018	cable
Outdoor Life Network	1995–2005	cable
Pac-12 Network	2011–2018	cable
Raycom	1994–2018	syndication
SEC Network	2013–2018	cable
Speed Channel	1995–2011	cable
Spike TV	2003–2018	cable
TBS	1994–2018	cable
The Tennis Channel	2003–2018	cable
TNN	1994–2002	cable
TNT	1994–2018	cable
USA Network	1994–2018	cable
Versus	2006–2010	cable
ACC Network	2019-2019	cable

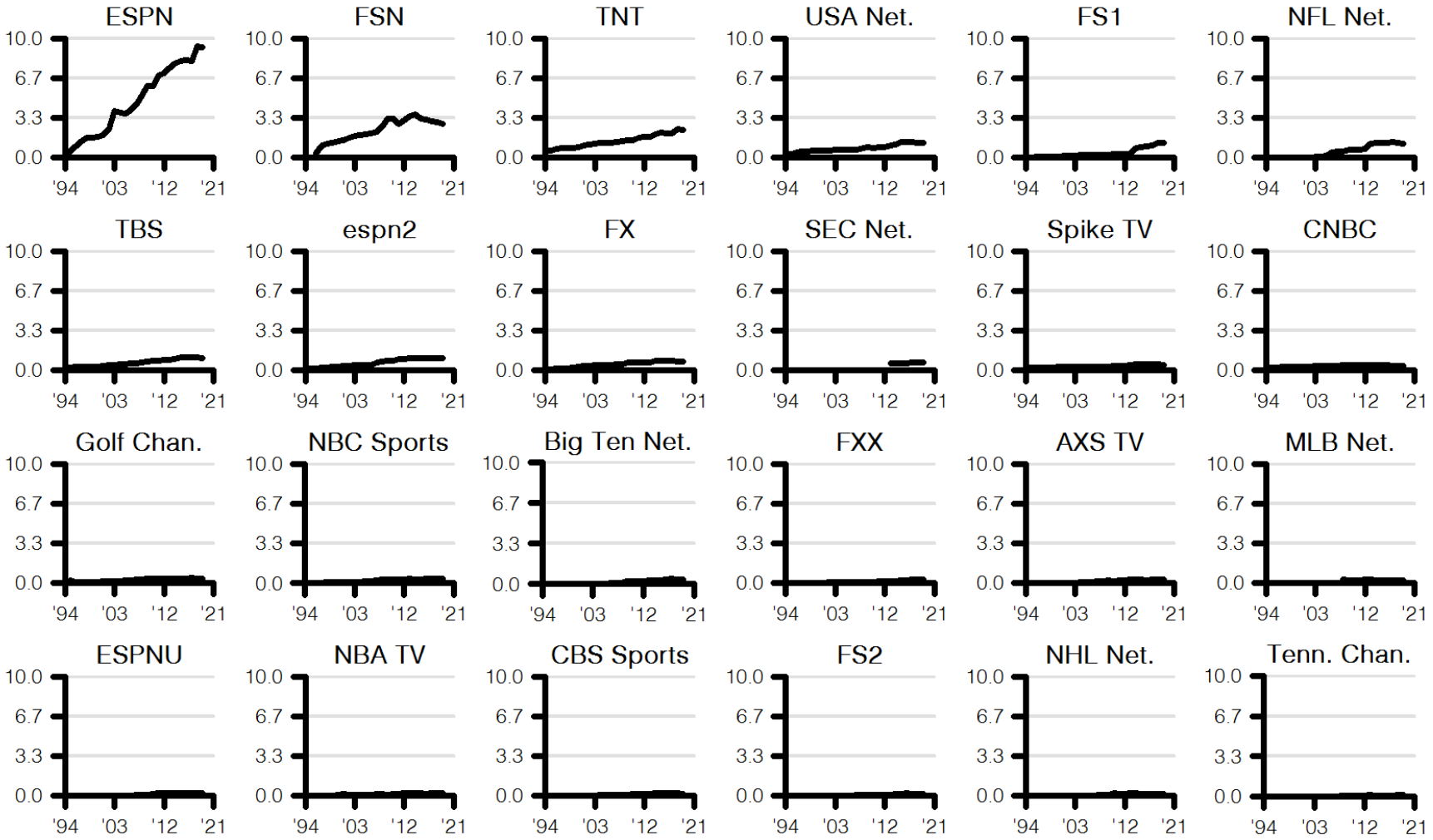
Examining Reach of Networks

- Reach by network and year; total number of T. V. households in black
- ESPN (green), Speed (blue) / FOX Sports 1 (red), OLN (purple) / VS (orange) / NBCSN (blk)



Subscription Revenue of Networks

- Annual rev. ($BN) from sub. fees by cable net.

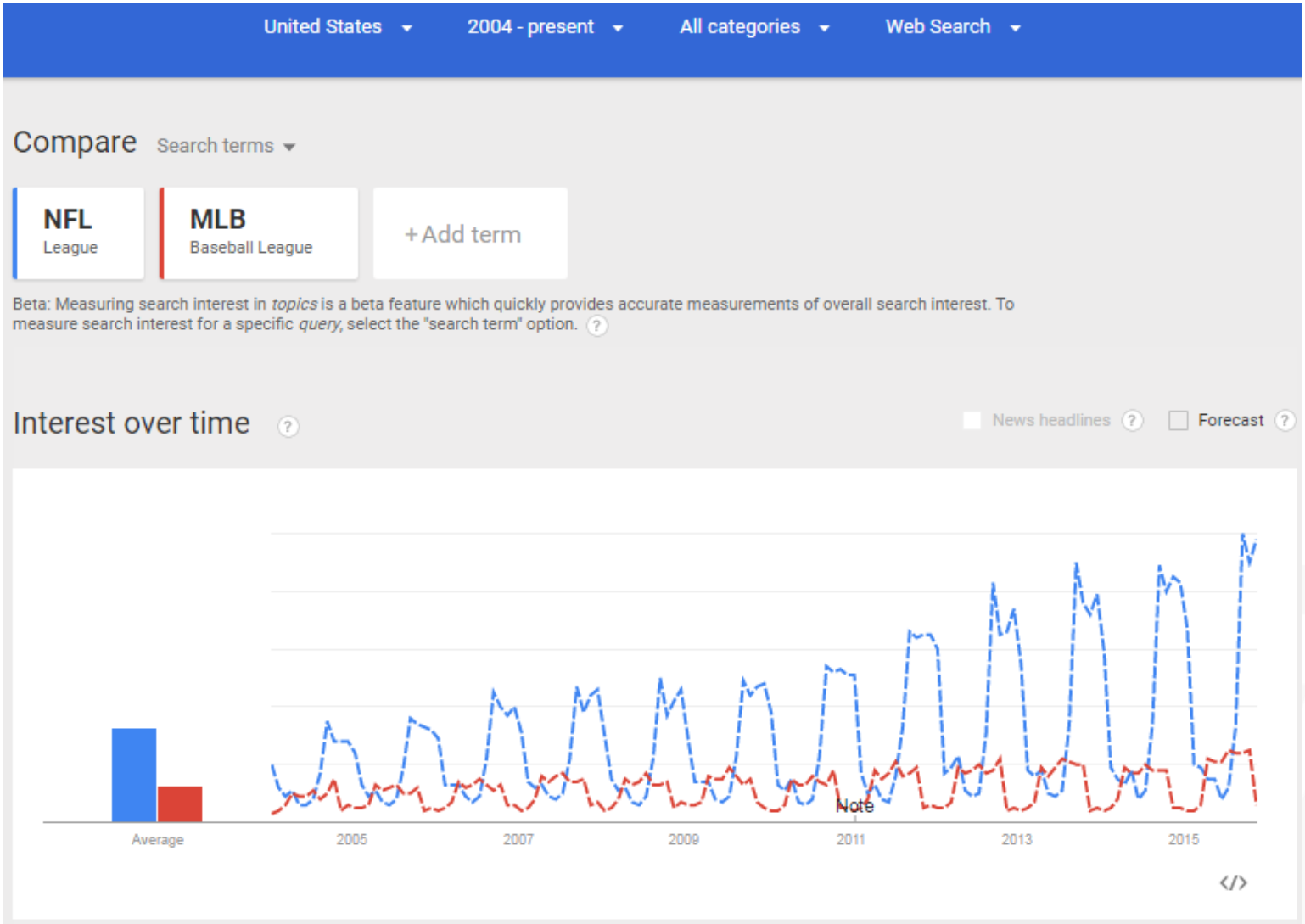


Figures are adjusted for inflation. All calculations by Rho AI.

Gathering Data: Sports Series

- Sellers of traditional rights contracts are sports series
  - Ex.: NBA holds agreements with Turner, Disney to air events
  - Some series produce coverage for own events, e. g. MLB Network
  - Occasionally, series pay networks to show their own programming
- Examine broad interest of 54 sports series
  - Scrape info. from Google Trends (nontrivial)
  - Create “Interest Index” – web interest in sports series / topics
  - “Series *Y* has *X* times greater / lesser interest than series *Z*”
- Example of NFL and MLB

Google Trends Data (I)



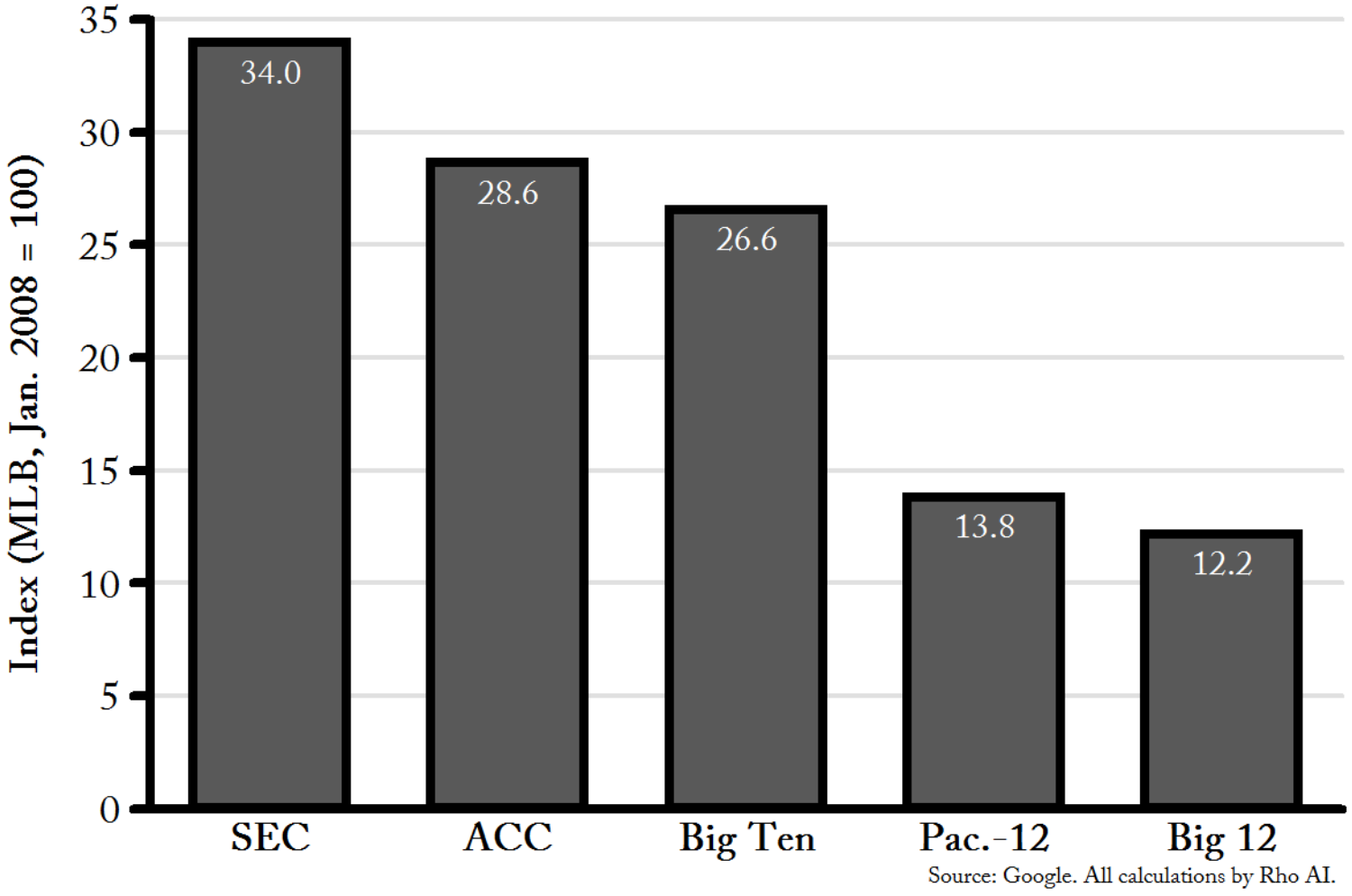
Google Trends Data (II)

- Scrape info. from Google Trends
  - Automated weekly pull for each topic + baseline topic
  - # queries expressed as share of wk. queries by topic in sample
  - Greatest share normalized to 100, others scaled from 0 to 100
- Adjustments made to Trends figures
  - Each topic compared to MLB
  - Data are shifted s. t. MLB's first figure in 2008 is 100
  - Account for seasonality with one-year moving average of adj. index
  - Apply other smoothing mechanisms

Week of:	Raw Index		Adjusted Index	
	MLB	NFL	MLB	NFL
1/6/2008	4	29	100	725
1/13/2008	4	31	100	775
1/20/2008	3	23	75	575
1/27/2008	4	12	100	300
2/3/2008	4	20	100	500
2/10/2008	4	9	100	225
2/17/2008	5	8	125	200
2/24/2008	5	12	125	300
3/2/2008	6	12	150	300
3/9/2008	6	9	150	225

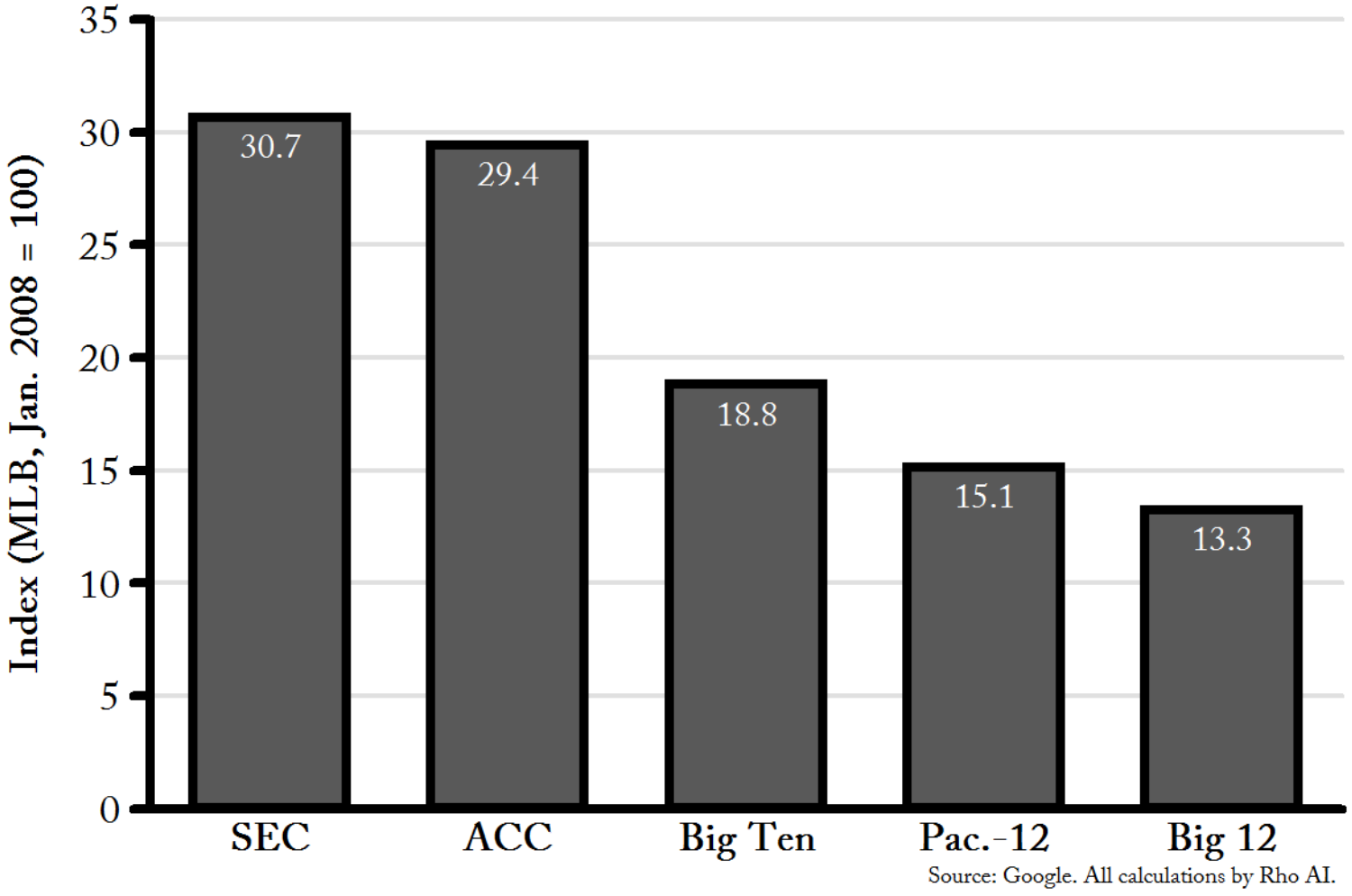
Interest in Coll. Conferences (Old!)

- Chart of interest in five major college conferences as of Oct. 2018
- Measured as one-year moving average of Google Trends' interest index



Interest in Coll. Conferences (New!)

- Chart of interest in five major college conferences as of Jan. 2020
- Measured as one-year moving average of Google Trends' interest index



Interest in “Top” Sports Series

- Chart of interest in major sports series, Jan. ‘20
- 1-yr m.a. of Google Trends’ interest index

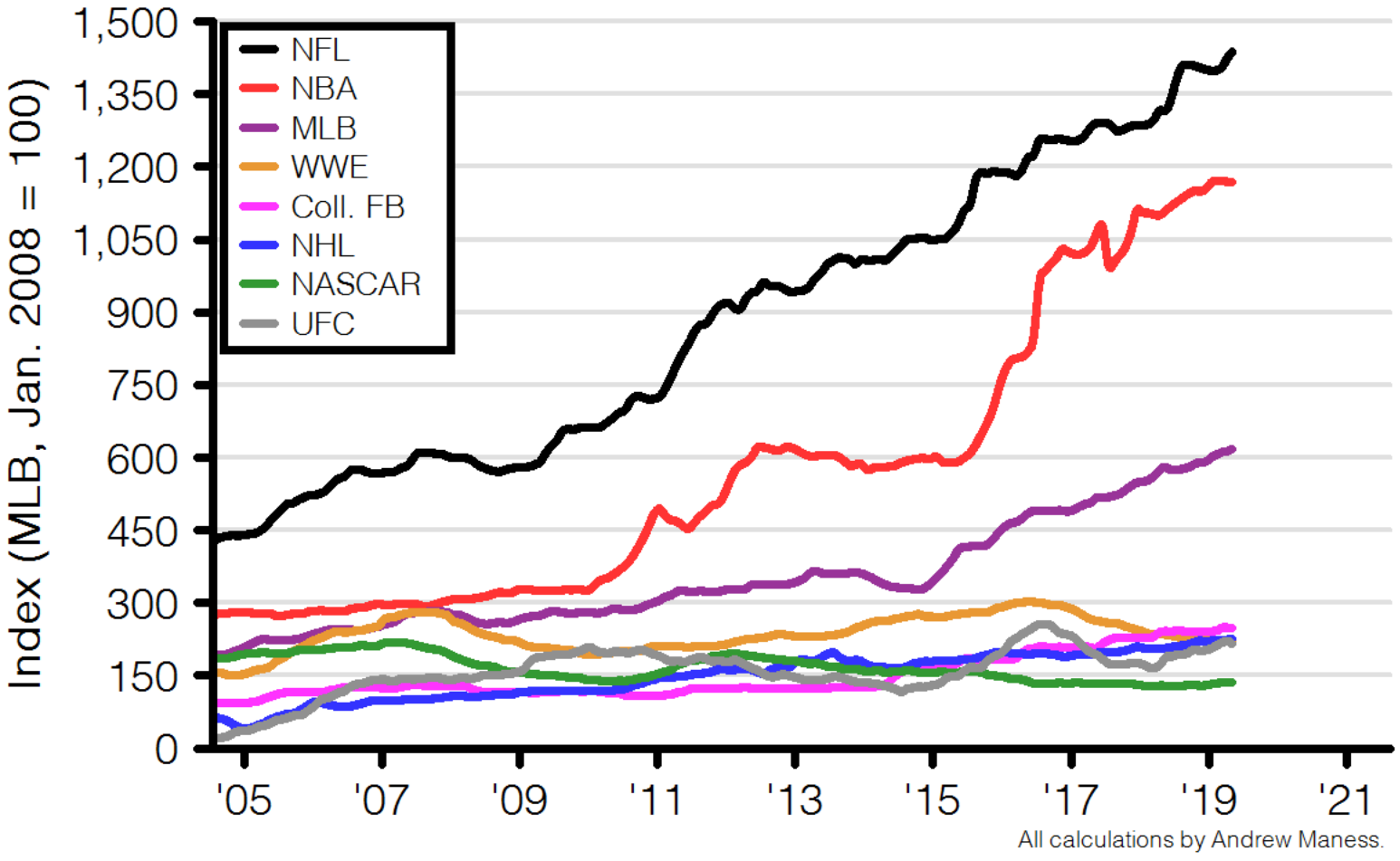


Table: Interest in 54 Sports Series

- Sorted by most to least interesting, per Google

Sports Series	Interest Index	CAGR since '04
NFL	1,437.0	+8.5%
NBA	1,168.8	+10.8%
MLB	618.5	+7.5%
College football	248.4	+6.5%
National Hockey League	225.9	+10.9%
WWE	219.9	+2.1%
Ultimate Fighting Championship	215.4	+11.4%
NASCAR	137.2	–2.6%
Premier League	135.7	+15.7%
College basketball	124.3	+5.3%
UEFA Champions League	103.5	+18.3%
PGA TOUR	99.9	+1.5%
World Cup	81.2	+3.8%
US Open	70.2	+1.8%
Formula 1	68.1	–1.6%
Wimbledon	57.8	+5.4%
Olympic Games	56.0	–3.6%
MLS	36.8	+0.8%
U.S. Open	34.5	+3.5%
Australian Open	31.4	+4.0%
Open Championship	31.0	+1.9%
Big Ten Conference	30.7	+3.7%
Southeastern Conference	29.4	+5.2%
Monster Jam	21.9	+11.7%
Drag racing	21.4	–5.8%
College World Series	20.1	+1.6%
WNBA	19.0	+2.1%

Sports Series	Interest Index	CAGR since '04
Atlantic Coast Conference	18.8	–0.6%
Tour de France	18.5	–11.5%
Minor League Baseball	17.7	–1.2%
Big 12 Conference	15.1	–0.3%
Monster Energy AMA Supercross An FIM Wc	14.2	+0.8%
Pacific-12 Conference	13.3	+2.8%
ATP Tour	12.2	–1.5%
FIBA	11.9	+11.3%
Women's Tennis Association	11.9	–0.9%
Professional Bull Riders	11.5	+0.6%
NBA G League	9.1	+7.5%
IndyCar	8.5	–1.8%
Augusta National Golf Club	8.3	+0.1%
Canadian Football League	8.0	–0.2%
American Hockey League	8.0	–3.0%
International Motor Sports Association	7.4	–0.7%
Australian Football League	6.3	–4.1%
Triple Crown of Thoroughbred Racing	5.4	+0.6%
College ice hockey	5.0	–6.4%
USA Track & Field	4.4	–3.6%
National Rugby League	4.4	–4.8%
Major League Lacrosse	4.1	–4.1%
U.S. Figure Skating	3.6	–6.9%
Arena Football League	3.1	–10.3%
America's Cup	2.9	–4.5%
World Baseball Classic	1.7	–0.1%
North American Soccer League	1.5	–6.8%

SOURCE: Google. Online queries in U. S. only. Individual daily indices are scaled to a common baseline: 100 = MLB interest on January 1st, 2008. "Interest Index" is a one–year moving average of the daily index for each sports series. Events which are held every four years are presented as four–year moving averages. "CAGR" is the compounded annual growth rate for the interest index from fifteen years ago. Table is current through December 31st, 2019. Calculations by Rho AI.

Gathering Data: Contracts

- Collect contracts between networks, sports series
  - Sample includes 224 contracts across 21 sport series, 1994–2018
  - Tabulate 16 quantifiable characteristics for each contract
  - Over 5,000 data points aggregated
- Factors tallied include:
  - Contract date
  - Sports series
  - Network(s)
  - # events to air
  - Contract start date
  - Contract end date
  - Contract cancelation date
  - Total price of contract
  - Contract renewal?
  - # champions from prev. 10 yrs.
  - Peak T. V. rating from prev. 3 yrs.
  - Avg. T. V. rating from prev. 3 yrs.
  - Streaming rights included?
  - # international events
  - # overnight hours
  - # overall programming hours

Ex. Contract (NBA/TNT, 2014)

- NBA signed contract with Turner on 10/6/14
- Turner will pay NBA ~$12.6BN, 2017–2025
- Renewal from parties' previous contract
- # unique champions, prev. 10 yrs.: 5
- Peak T. V. rating, prev. 3 yrs.: 10.3
- Avg. T. V. rating, prev. 3 yrs.: 1.3
- Turner pays for rights to stream games
- ~318 hours of events per yr.; 28 overnight, 0 int'l
- ~64 games aired on TNT per year

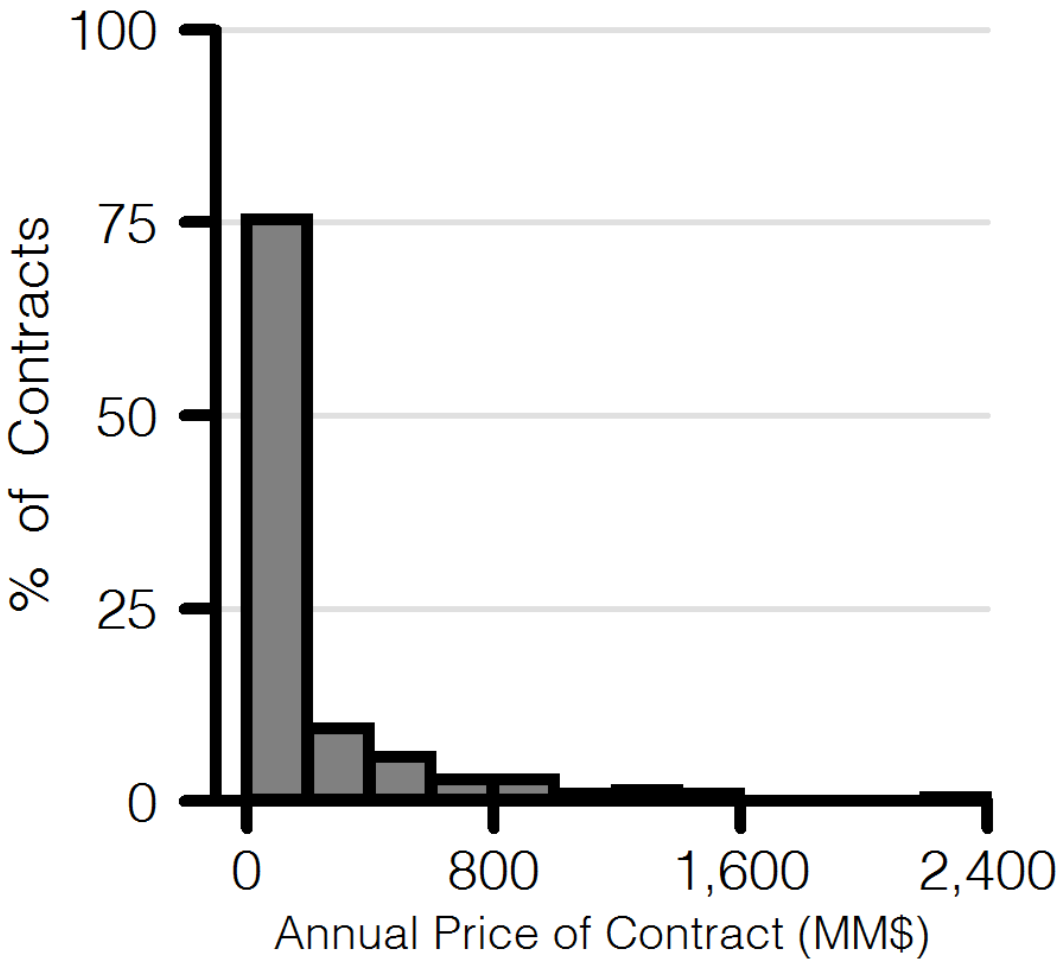
Descriptive Statistics for Contracts

Variable	Minimum	Mean	Median	Maximum	St. Dev
Total value of contract (MM$)	$0.0	$1,329.2	$214.1	$16,582.5	$2,563.5
Length of contract (years)	1.0	6.2	6.0	25.0	3.6
Annual value of contract (MM$)	$0.0	$170.1	$36.0	$2,072.8	$308.0
Number of hours per annum in contract	3.5	102.2	63.0	1,000.0	119.5
Value of each hour (MM$)	$0.00	$2.80	$0.71	$40.14	$5.60
Anticipated average television rating	0.0	3.3	1.6	16.2	3.6
Anticipated top television rating for a given event	0.0	6.9	5.0	35.1	6.9
Annual number of households-by-hours (1,000s)	100	187,053	103,474	1,281,448	228,796
Value of each household-hour	$0.00	$0.94	$0.40	$16.68	$1.72
Annual growth rate of U. S. GDP	-2.78%	2.73%	2.69%	4.69%	1.51%
Nielsen television ratings decay value	8.53	11.12	10.34	15.42	2.20
International programming scheduled	0.0%	7.8%	0.0%	91.7%	21.3%
Programming scheduled to air on over-the-air television	0.0%	54.3%	51.0%	100.0%	44.6%
Programming scheduled to air overnight	0.0%	3.4%	0.0%	38.4%	7.6%
Number of champions from previous 10 season	3.0	6.2	6.0	10.0	1.5
Value of S&P 500 Index (stock market) at time of contract	823.1	1,409.5	1,391.7	2,030.4	283.4
1-year MA of sports series' interest index at time of contract	7.8	123.0	46.5	882.9	176.3
Does the contract include a "big four" sport? No=0, Yes=1	0.00	0.22	0.00	1.00	0.41
Does the contract include a college conference? No=0, Yes=1	0.00	0.29	0.00	1.00	0.46
Were live Internet streaming rights included? No=0, Yes=1	0.00	0.28	0.00	1.00	0.45
Was network an incumbent rights-holder? No=0, Yes=1	0.00	0.73	1.00	1.00	0.45

N=443

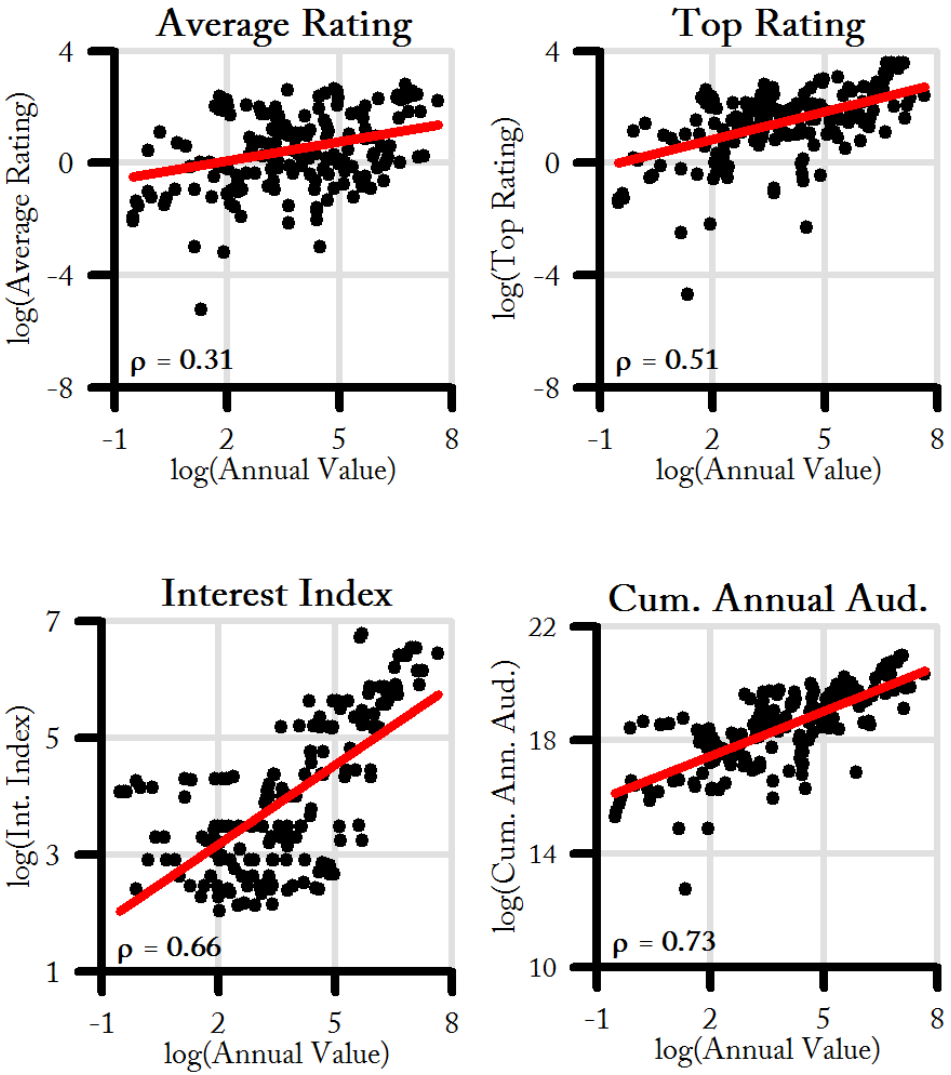
Histogram of Contract Prices

- Over 75% of contracts in sample exchange less than $200MM annually
- Annual prices follow a negative binomial distribution



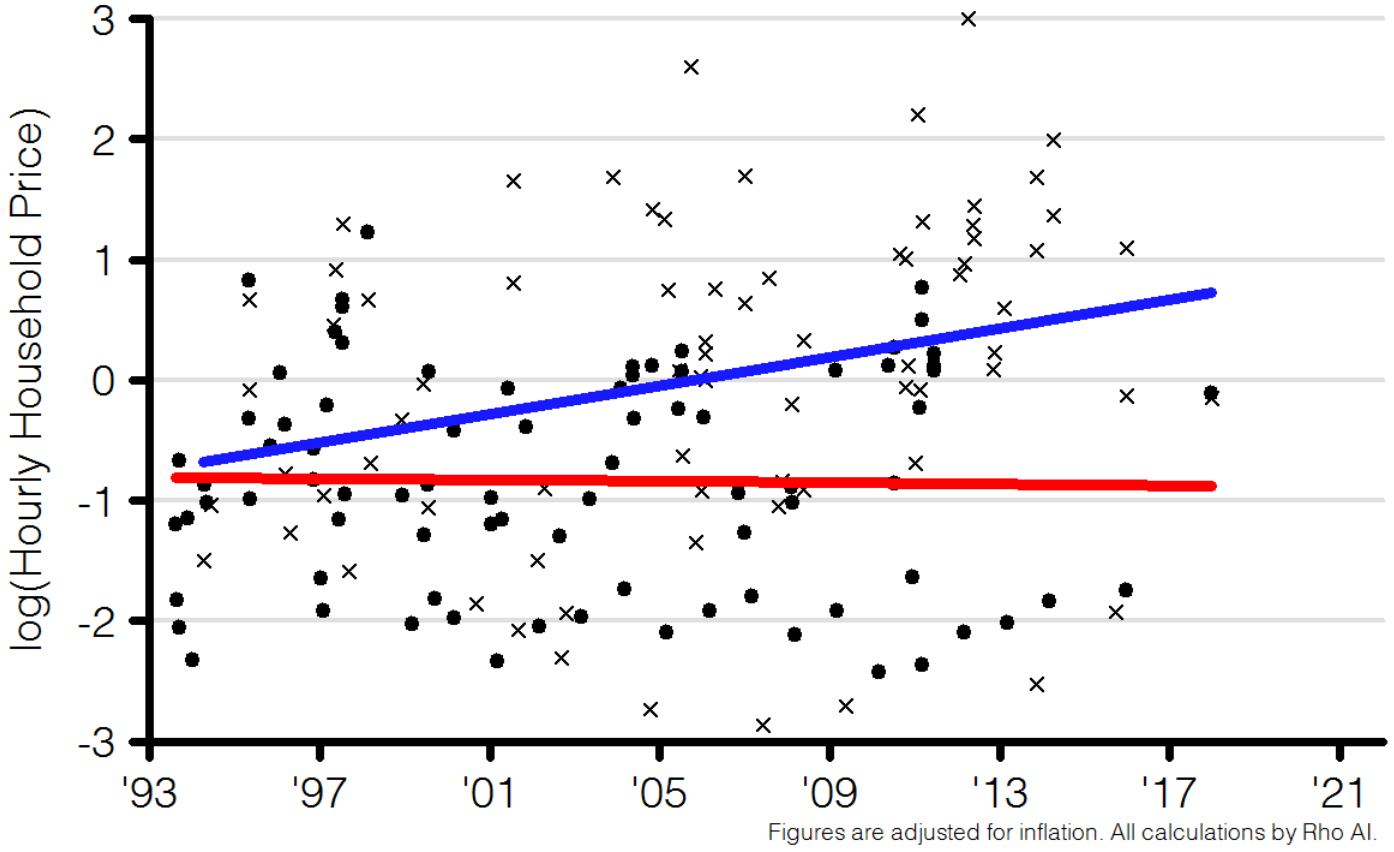
Analytics: Interest Metrics, Prices

- All "interest metrics" exhibit positive correlation with annual prices
- Cumulative annual audience strongest
  - Events' annual rating multiplied by number of event hours
  - Suggests that price is influenced by quality *and* quantity of events



Analytics: Premium on Cable?

- Hourly household price: price paid for each anticipated household-hour
- Cable (x / blue) ↑; broadcast (o / red) ↔



GLM / Negative Binomial Model

- Dependent variable follows neg. bin. distrib.
  - Annual price of contract = f(local, environment)
  - Generalized linear model (GLM) with HRM regression
- Theorizations on factors of annual prices
  - Consulted industry experts; must be intuitive
  - Potential factors:

Characteristic	Theory
Average primetime T. V. rating among top 30 programs	−
Length of Contract	+
Live and Online Streaming	+
Cumulative Annual Audience	+
Most Events on Over-the-air Networks	−

Characteristic	Theory
Economic Growth	+
International Events	−
Google Trends Interest Index	+
College Conference	−
Big Four Sports Series	+

GLM / Negative Binomial Results

- Generalized linear model, neg. bin. family
- Most intuitive, meaningful model tested

Dependent Variable: Log Annual Value	Coef.	Std. Err.	t-score	p-value	95% confidence interval	
log(Ratings Decay)	-1.854	(0.351)	-5.28	0.00	[-2.542 ,-1.166]	***
log(Length of Contract)	0.403	(0.124)	3.26	0.00	[0.160 ,0.645]	**
Live streaming rights	0.269	(0.134)	2.01	0.04	[0.007 ,0.531]	*
log(Cumulative Annual Audience)	0.597	(0.050)	11.88	0.00	[0.499 ,0.695]	***
% events that are over-the-air	-0.224	(0.108)	-2.08	0.04	[-0.435 ,-0.013]	*
U. S. growth domestic product	0.081	(0.035)	2.33	0.02	[0.013 ,0.149]	*
% events that are international	-2.098	(0.278)	-7.54	0.00	[-2.643 ,-1.553]	***
log(Interest Index)	0.259	(0.045)	5.71	0.00	[0.170 ,0.348]	***
College conference	-0.414	(0.126)	-3.28	0.00	[-0.662 ,-0.167]	**
Big Four Sports Series	0.810	(0.127)	6.41	0.00	[0.563 ,1.058]	***
Constant	-4.558	(1.221)	-3.73	0.00	[-6.952 ,-2.164]	***
Pseudo R-squared	0.1971					
Log-Likelihood	-836.22					
Observations	224					

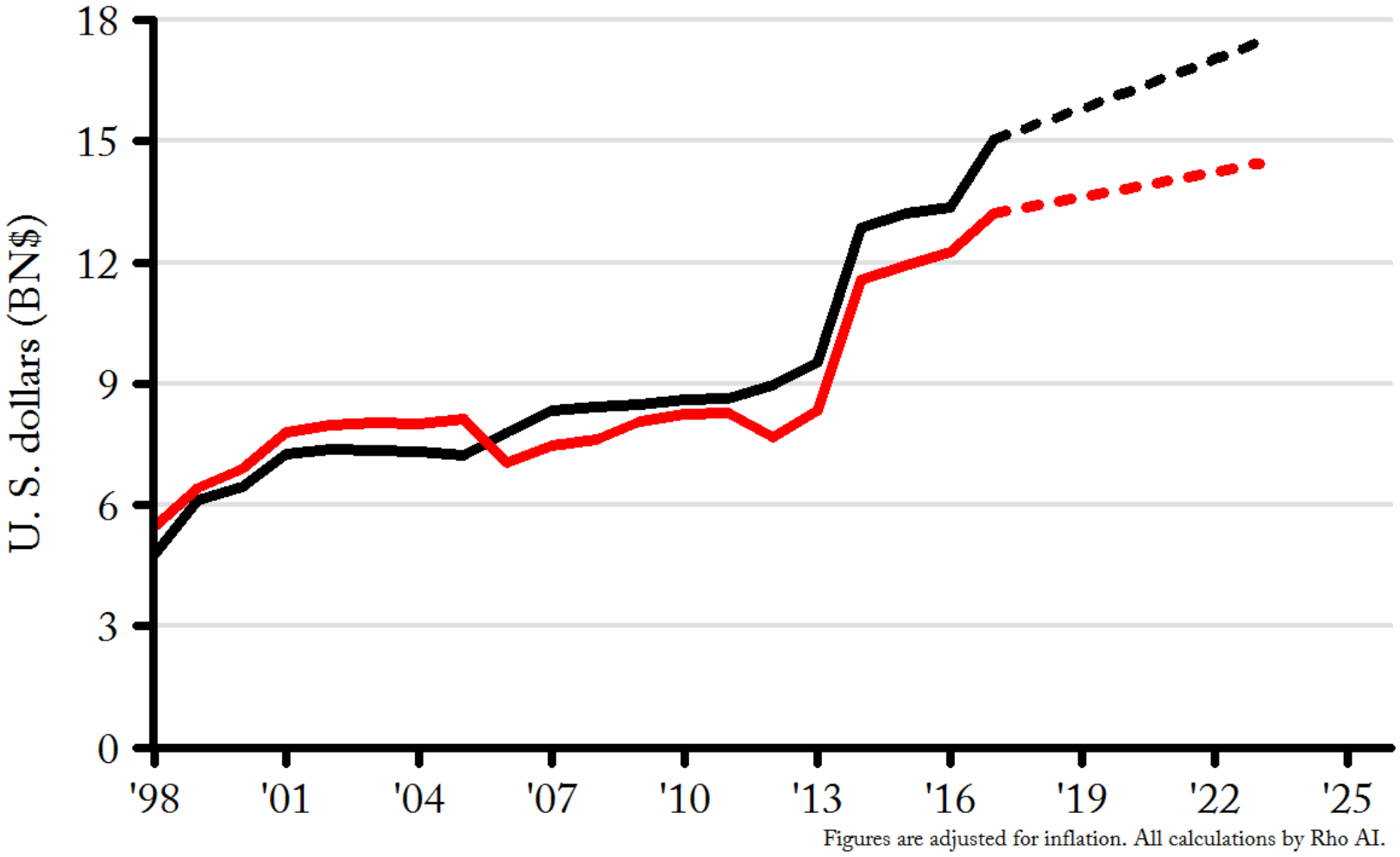
* Significant at 5%, ** Significant at 1%, *** Significant at 0.1%.

Analysis of Results

- Model holds well to industry experts' beliefs
- Notable effects, *ceteris paribus*:
  - Rate of annual price is about 30% greater with streaming
  - Contracts which emphasize free T. V. are priced ~20% less
  - Annual price runs ~35% less when sellers are college conferences
  - "Big 4" series command a rate >120% than others
  - As anticipated cumulative T. V. audience increased by 10%, expect an annual price about 5% greater
  - A good economy correlates with strong annual prices
  - Longer-term deals associate with an annual price premium
- Forecast / fit each contract based on tested factors

Fitting Contracts

- Comparing size of market (black) to forecasted size of market (red)
- Steady increase apparent, but model predicts a little flatter

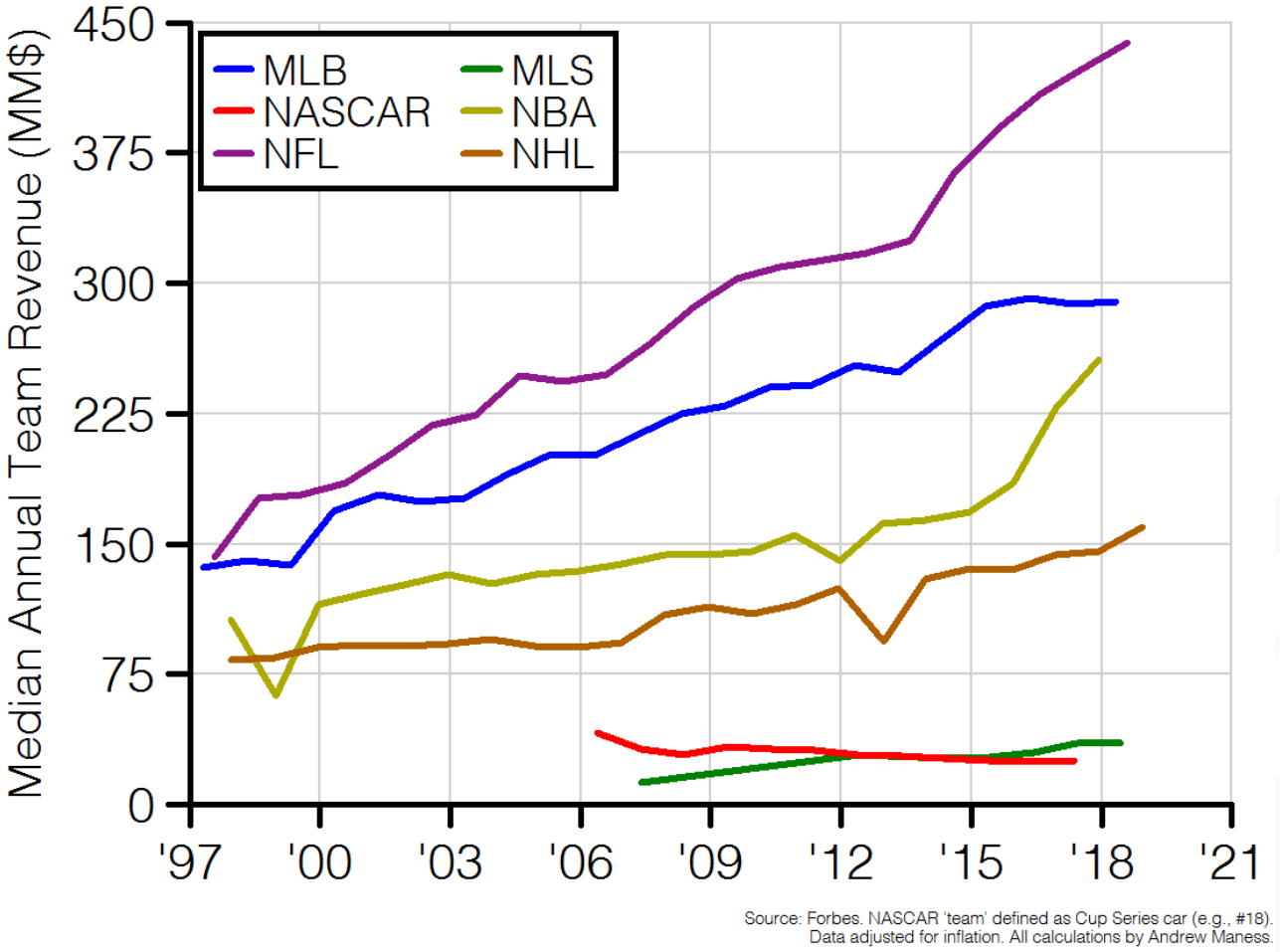


Conclusions

- Shortcomings of model
  - Does not capture all nuances of contracts
  - Small sample size
  - Assumes television-only universe within one section of time
- Markets might be overheated; why?
  - Buyers' misjudging magnitude of demand shift
  - Sellers' giving away rights not accounted for in model
  - Adaptive, not rational, expectations
- Is there a bubble?
  - Most networks locked-in for near- and medium-terms
  - Series, networks trying to increase revenue in other areas
  - Next inflection point: college conference renewals in ~2 yrs.

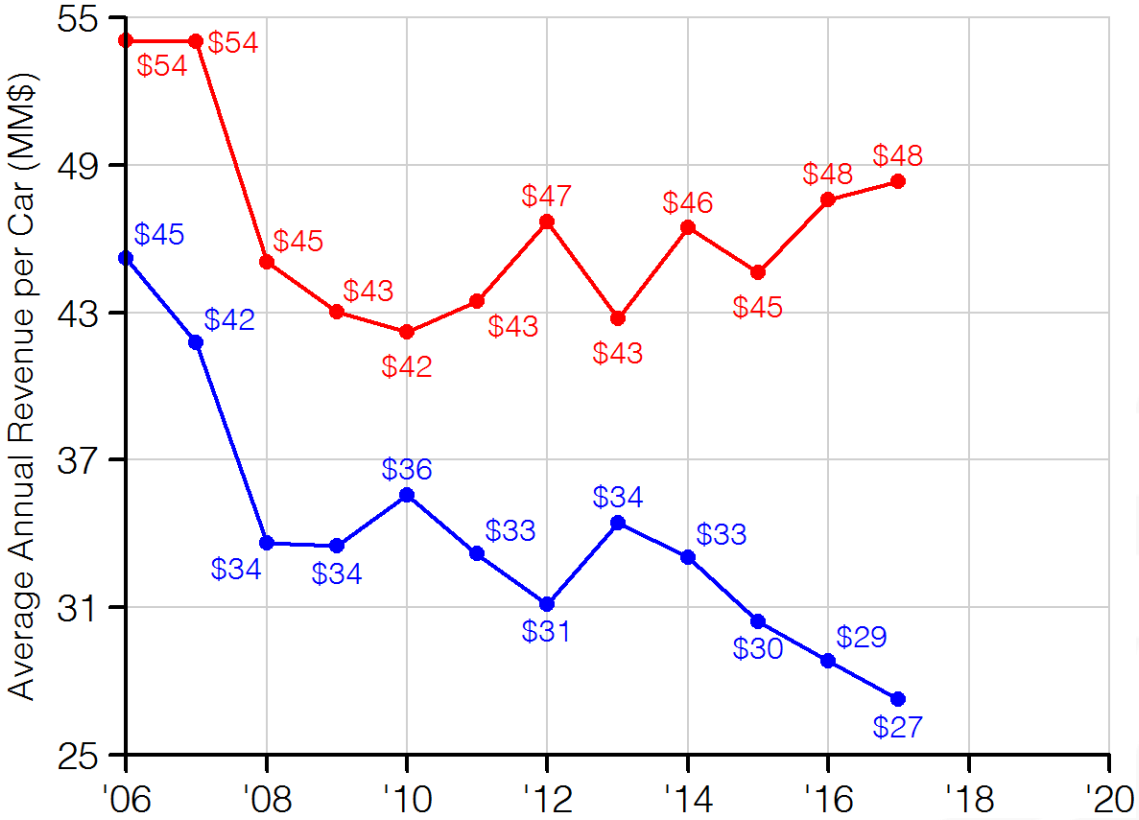
Revenue Over Time

- Median Annual Team Revenue by League



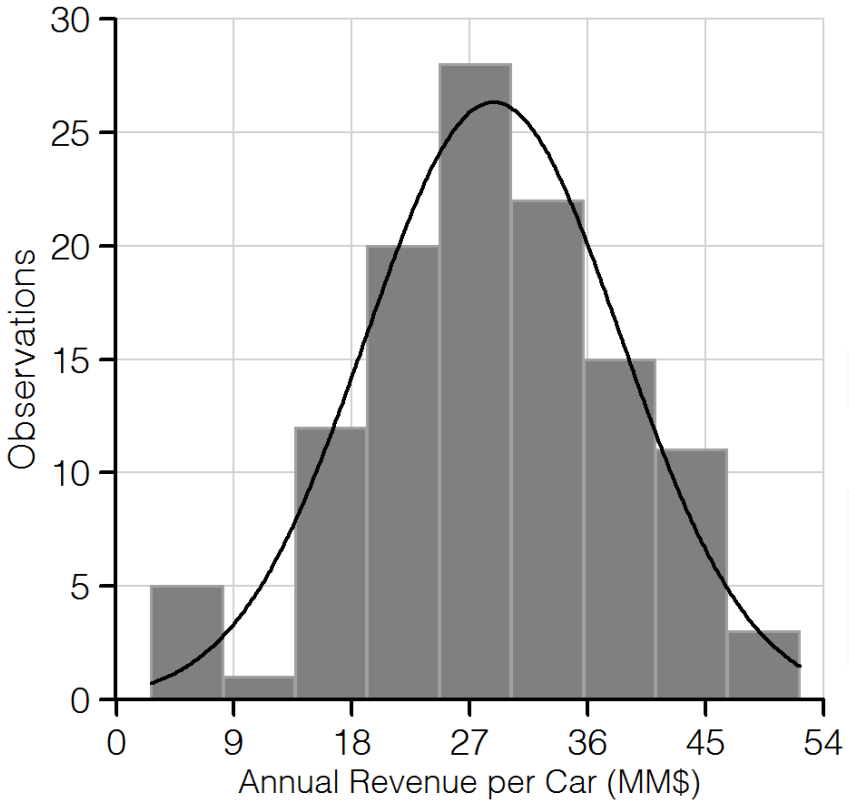
Revenue Over Time (cont'd)

- Median Annual Team Revenue, 2006–2017
  - All forms of revenue (i.e., the Forbes values)
  - 10pctile in red, 50pctile (median) in blue



Data Modeling / Analysis

- Annual non-prize revenue per car is normal-ish
  - Data are rarely perfect
  - Be sure to note how far from perfect it is



- Kurtosis: 3.06
- Skewness: -0.30

Data Modeling / Analysis

- Based on knowledge, anecdotes, analysis – what are the expected results of this model?
- What influences non-prize revenue?

Characteristic	Theory
Engine Leasing	+
Technical Alliance	+
Average Viewership from Previous Year	+
Average Finish (Less Wins) from Previous Year	–

Characteristic	Theory
Number of Wins from Previous Year	+
Championship from the Previous Year?	+
Any Popular Drivers?	+
Other ideas?	?

Data Modeling / Analysis

- The actual model (but we have to explain it!)

Dependent Variable: Annual Non-Prize Revenue	Coef.	Std. Err.	t-score	p-value	95% confidence interval
Engines Leased per Race	3.342	(0.881)	3.80	0.00	[1.597, 5.088] ***
Technical Alliances per Race	2.391	(2.388)	1.00	0.32	[-2.343, 7.125]
Average Viewership per Race ^	2.281	(0.420)	5.43	0.00	[1.448, 3.114] ***
Average Finish (less wins) ^	-0.724	(0.162)	-4.46	0.00	[-1.046, -0.402] ***
Number of Wins ^	0.271	(0.162)	1.67	0.10	[-0.051, 0.593]
Championship ^	3.495	(1.673)	2.09	0.04	[0.178, 6.812] *
Dale, Jr., on the team?	24.847	(6.182)	4.02	0.00	[12.594, 37.101] ***
Constant	2.488	(5.754)	0.43	0.67	[-8.919, 13.894]
Adjusted R-squared	0.9714				
Log-Likelihood	-487				
Observations	117				

* Significant at 5%, ** Significant at 1%, *** Significant at 0.1%.

^ Lagged by one year

Includes fixed effects for number of entries

Data Modeling / Analysis

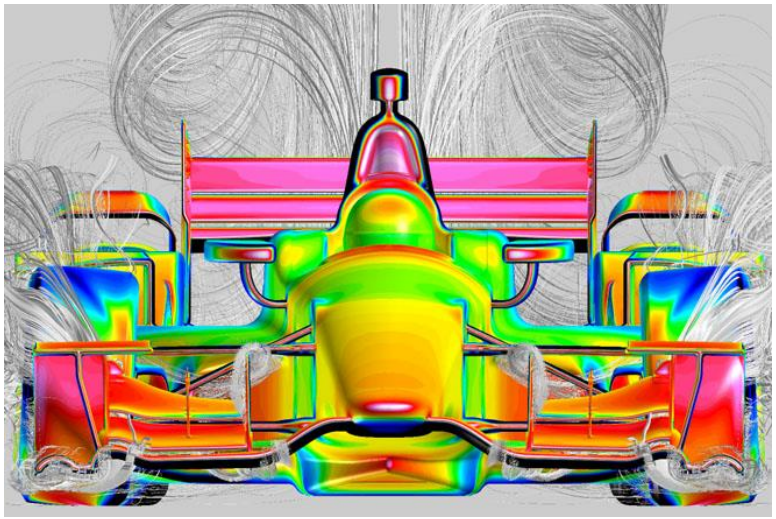
- Each engine lease is worth ~$3.3MM / year
  - This is consistent with expert information
- Technical alliances not statistically significant
  - Should be included; key component to revenue in reality
  - Still positive directionally – imperfect, but passes sniff test
- For every additional 1mil vwrs / year, expect $2.3MM additional revenue in following year
- Team performance is important!
  - Average finish position, wins, and titles matter
- Having Dale, Jr., gets you a full season of sponsorship (no other driver impacts like this)

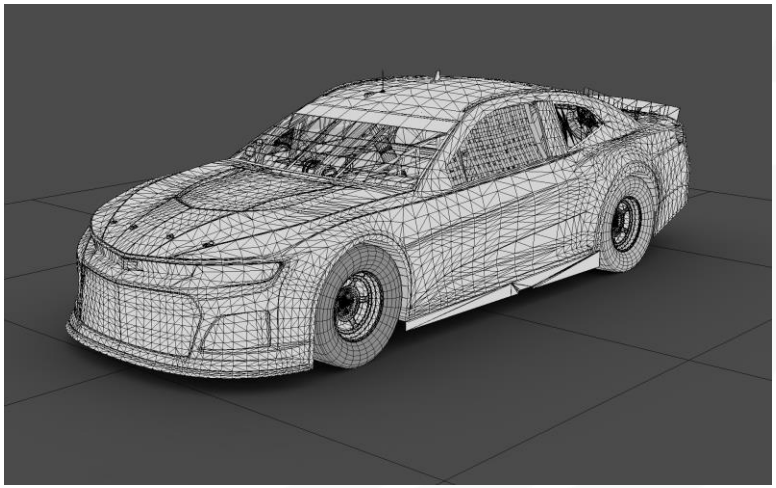
Actionable Insight

- Median annu rev per car ↓1.21 MM$/yr since ‘06
- Rev. gap btwn. 10pctle and 50pctle growing
- Team performance still matters, though
  - Improve your average finish position
  - Win some races and maybe a championship
- This is where Pit Rho can help teams
  - Improves your average finish by ~1.7 positions
  - More affordable than other potential alternatives
  - ∴ Relatively great return on investment
- Bottom line: This can still work financially despite broader struggles in the industry

Motorsport: A High-Tech Endeavor

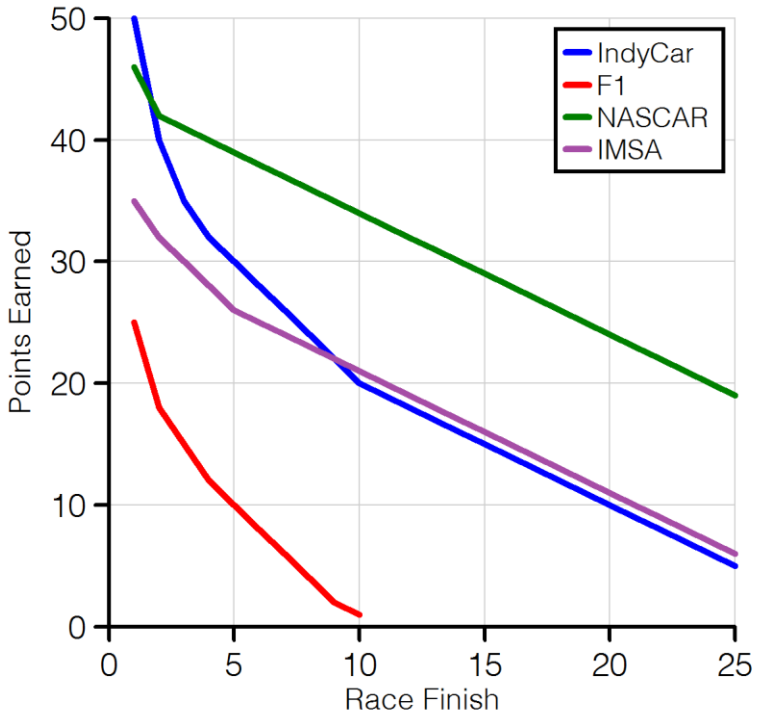
- Computational fluid dynamics
- Wind tunnel technology
- Full vehicle simulators
- Custom materials development
- Precision measurement, monitoring
- Strategy should be high-tech, too!





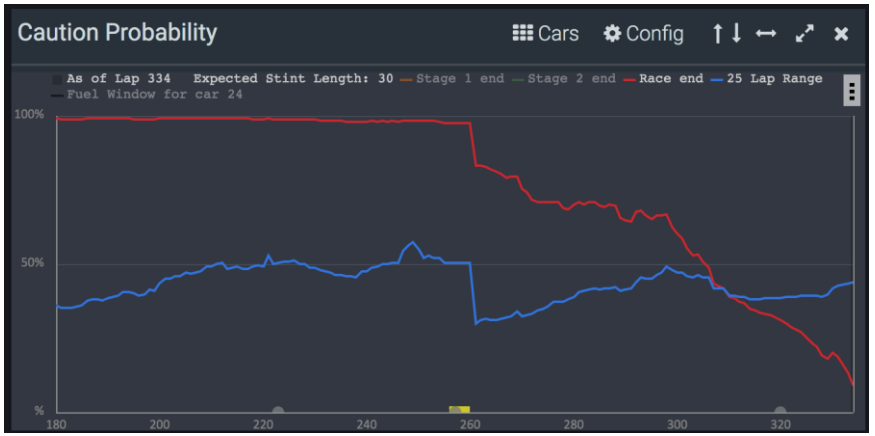
Motorsport: Team Objectives

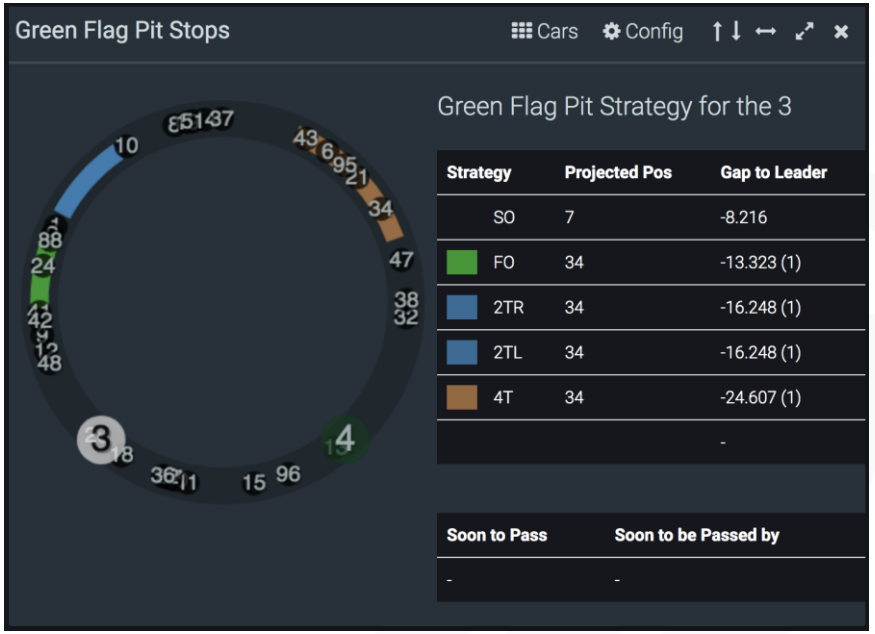
- Win races (proper mindset – find a way)
- More specifically, score best finish possible
- Better finishes → more championship points
- Team with most points at end of season earns title
  - Major incentive to battle for each position
  - Like a marathon; fast, but finish—ultimate engineering exercise

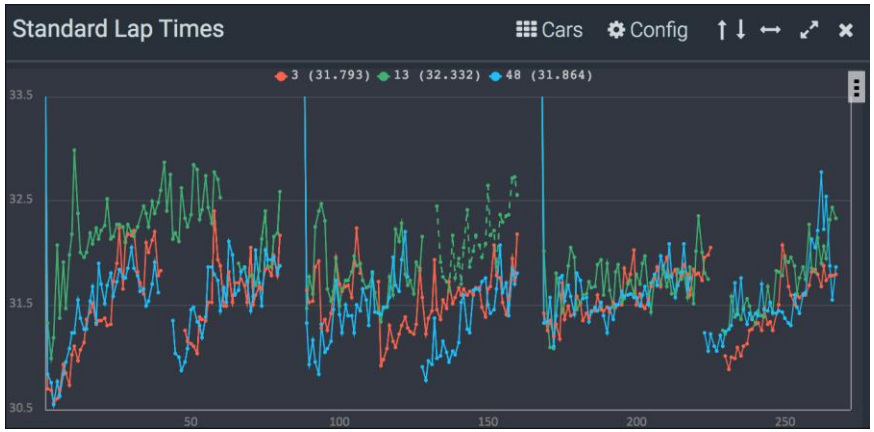


NASCAR: Strategy Basics

- Likelihood that a caution comes out
- Trying to exit pitstop with “clean” track
- Minimizing lap times, slow laps, etc.

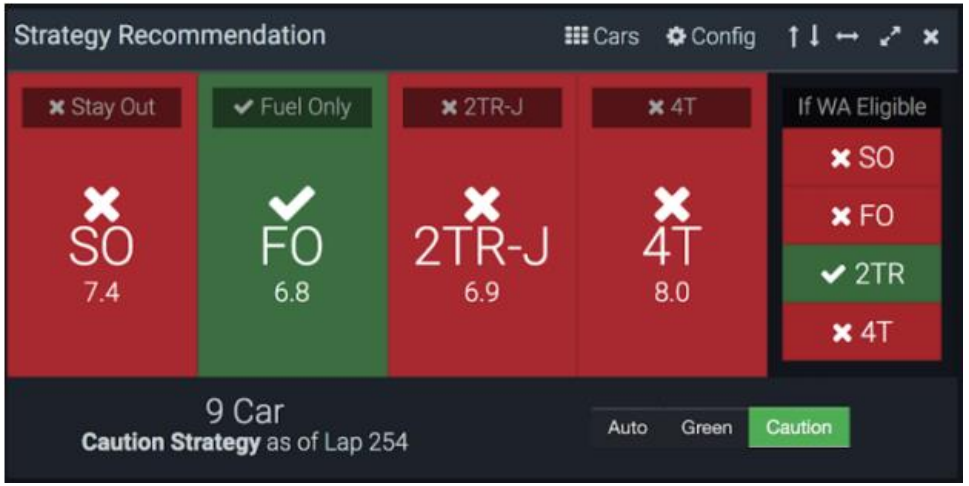


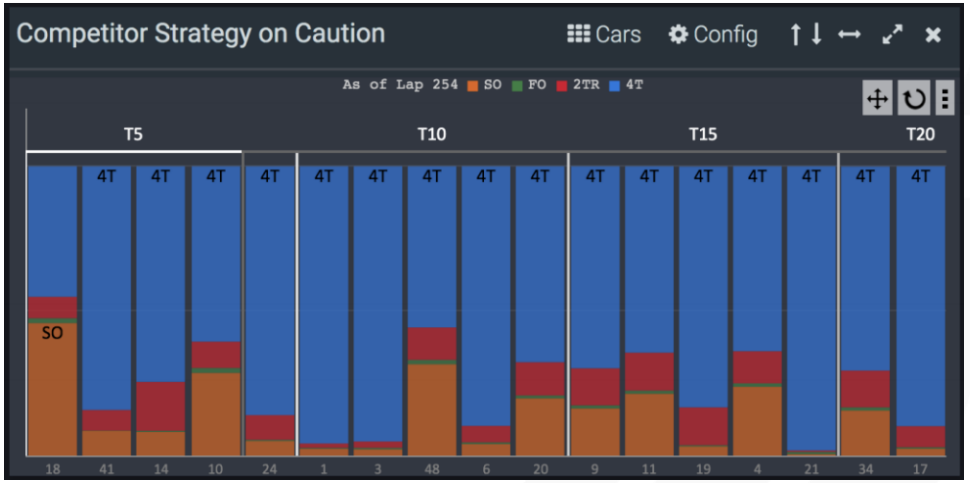




NASCAR: Strategy Decisions

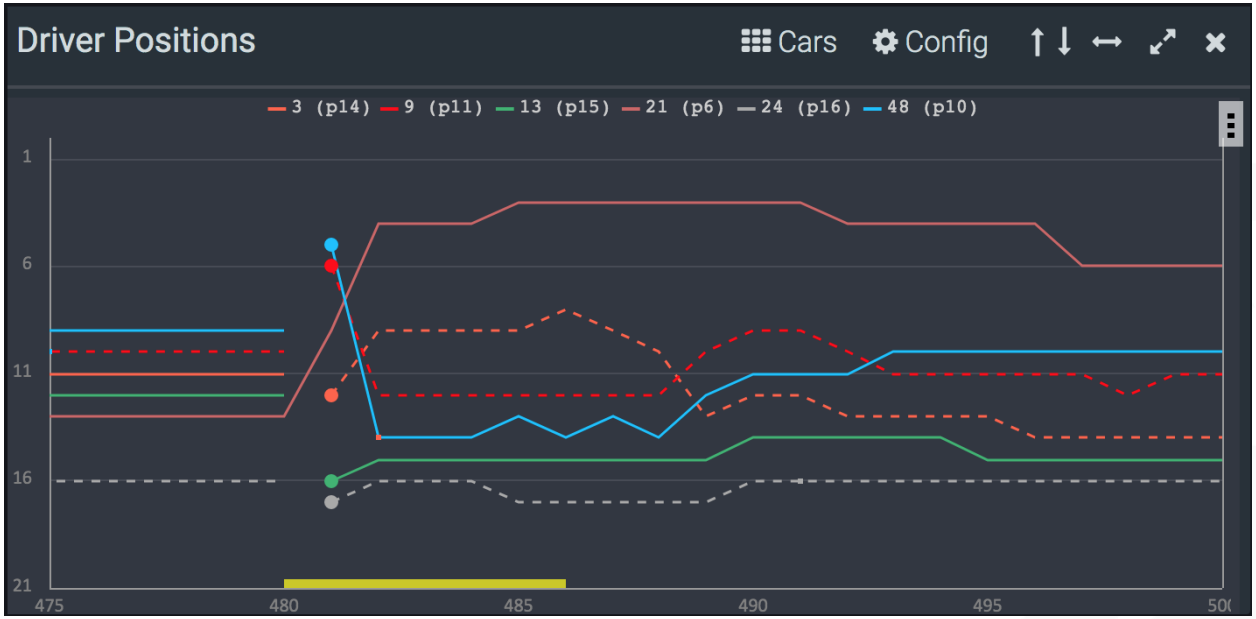
- Strategy Recommendation: There's one right answer
- Largely depends on what you think your competition will do
- Risk versus reward!





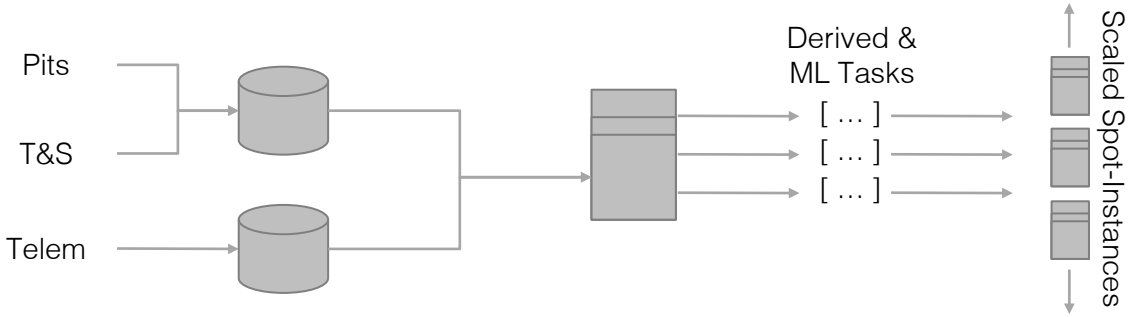
19BRI1 video example

- Race strategy is critical - changes outcome of race
- Quick example: 19BRI1
  - Caution comes out; decisions are made: 3:03:30-3:06:39
  - The final sprint to the finish: 3:10:16-3:13:57



Application Flow Overview

Data Input	Data Mgmt.	Data Processing	Data Analysis
Scraped Pit Data	Pit Data Correction & Augmentation	Low Level Timing & Scoring	Data Analysis Modes: Live - Replay - Offline
At-Track Pit Data	Optional Auto-Ingestion of At-Track Feeds	Pit Data Processing & Historic Lap Data Revision	Advanced Caching for Full Race Replay for All Races
Timing & Scoring	Timing & Scoring Cleanup and Error Correction	Derived Statistics & Predictive Strategy	>20 Unique Data Views & Custom Configuration
Telemetry	Telemetry Cleanup and Error Correction	Raw & Derived Telemetry	Cross Device Support: PC - Mac - Tablet



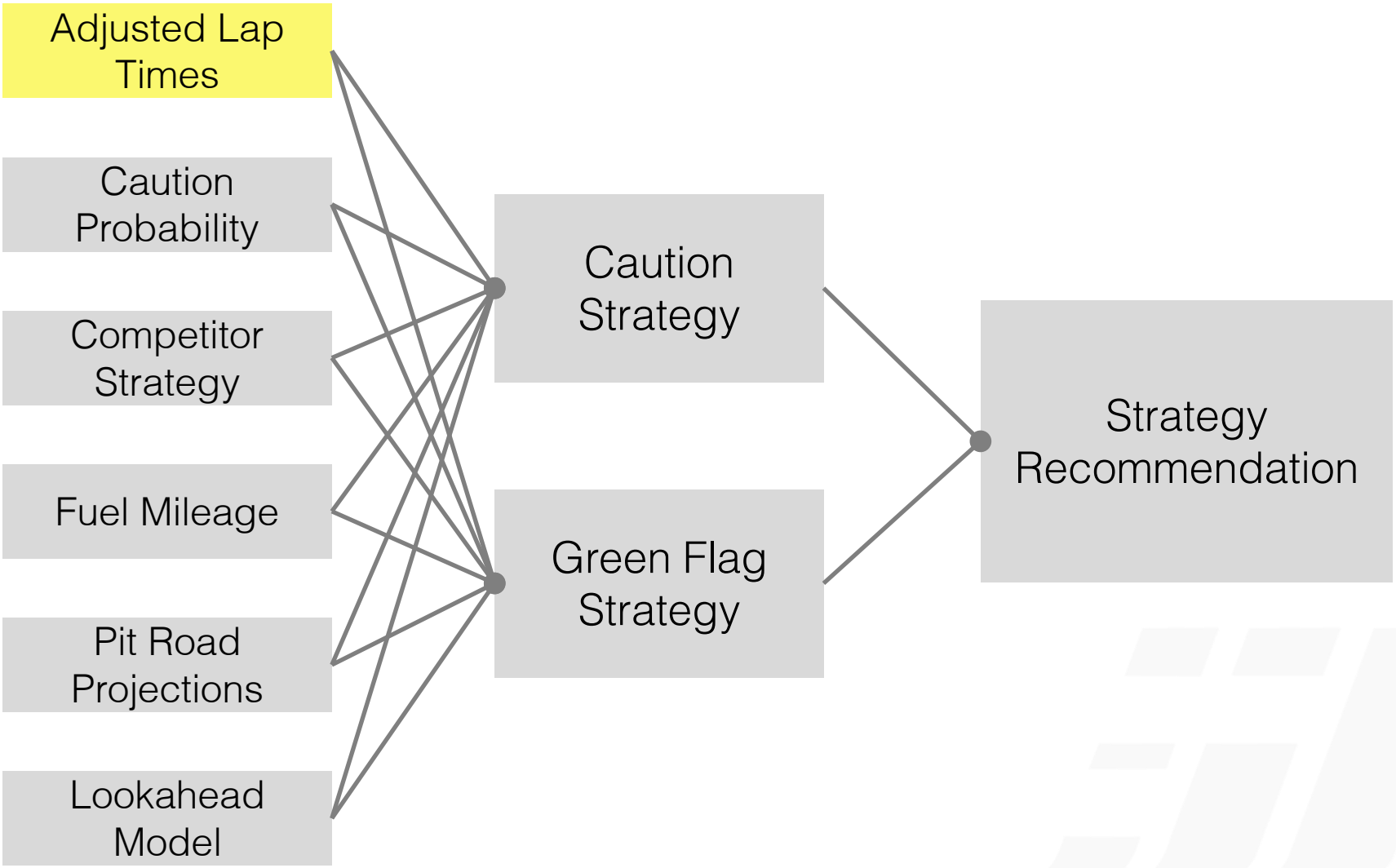
In-Race Compute Power

21 Cores; 38 GB RAM

During Week Compute Power

4 Cores; 7.5 GB RAM

Calculation Schematic



Brief Overview of Components

- Adjusted Laptimes (bulk of discussion)
- Caution Probability
- Competitor Strategy
- Fuel Mileage (major 2019 focus; why?)
- Pit Road Projections
- Lookahead Model
- Caution Strategy
- Green Flag Strategy
- Strategy Recommendation (actionable insight)

Adjusted Lap Times (Overview)

Inputs	Key Model Attributes	Outputs
Actual lap time	• Bayesian regression framework that automatically learns throughout race • Customized model for each race based on the specific characteristics of that track • Model is updated throughout race, automatically incorporating information about new tires, etc. • Performance of every car on every lap is continually updated and validated	Adjusted lap time
Position		
Laps since tires		Relative speed comparison
Tires last pit		
Traffic density		Predicted future lap times
Laps since restart		Tires vs. position trade-off

Walter, Gero; Augustin, Thomas (2009). "Bayesian Linear Regression—Different Conjugate Models and Their (In)Sensitivity to Prior-Data Conflict" (PDF). Technical Report Number 069, Department of Statistics, University of Munich.
Gelman, Andrew, Carlin, John B., Stern, Hal S. and Rubin, Donald B. (2003). Bayesian Data Analysis, Second Edition. Boca Raton, FL: Chapman and Hall/CRC. ISBN 1-58488-388-X.

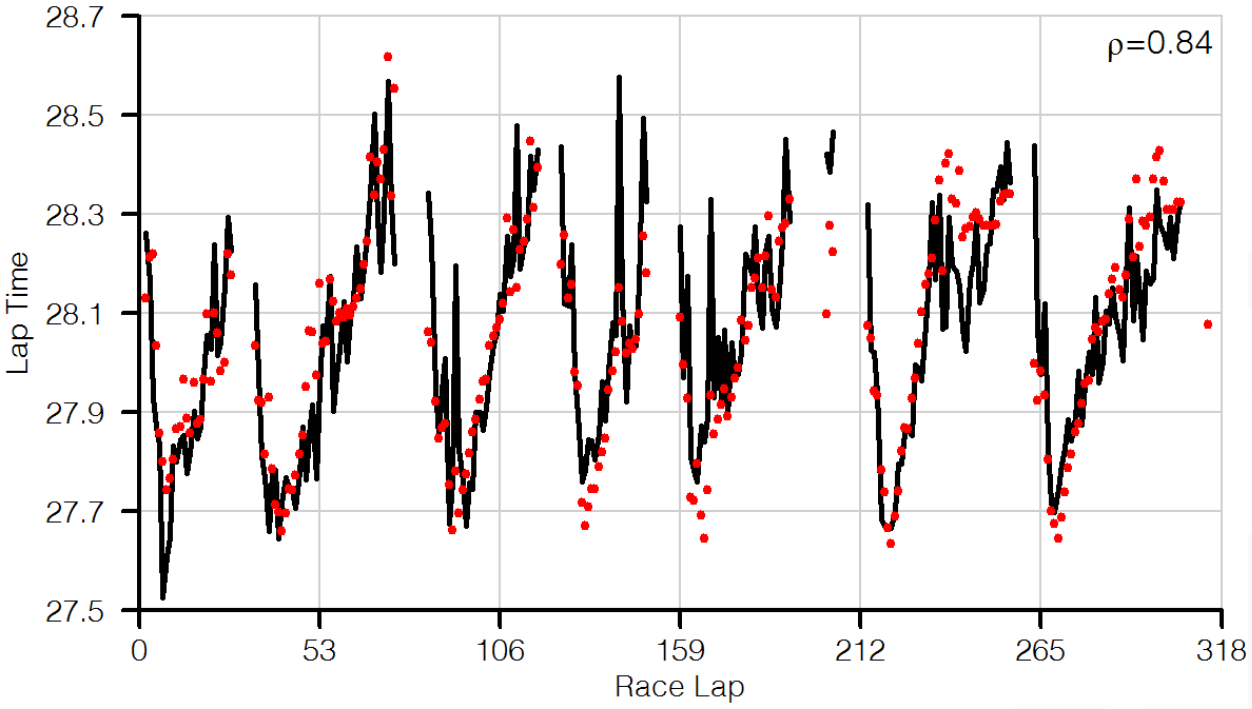
Adjusted Lap Times (Details)

$$
\begin{aligned}
\mu_n &= (X^TX + \Lambda_0)^{-1}(X^Ty + \Lambda_0\mu_0), \text{ where} \\
\mu_n &= \text{current parameter estimate} \\
\mu_0 &= \text{pre-race parameter estimate} \\
\Lambda_0 &= \text{pre-race weighting} \\
X^TX,\ X^Ty &= \text{in-race data (data} \uparrow \Rightarrow \text{weight} \uparrow)
\end{aligned}
$$

- Bayesian framework enables race-specific model to update throughout event
  - Naturally learns as the race progresses
  - Tuned to maximize predictive ability from past event data
  - Learning rate is fastest early in the race, continues throughout race
- Traffic density, “passing friction,” time to pass
  - Reviewing to include possible driver-specific passing skills
  - Methodical; ensure that it improves already-robust model
- Restart dynamics new as of 2019
- Advanced statistical techniques to separate effect of car and position (multicollinearity; not solvable with OLS)

Adjusted Lap Times (Accuracy)

- Predicted (lines) vs. actual lap times (dots)
- Positive correl. btwn. predicted, actual laptimes
- 17PHX1, #31



17PHX1, #31 (vs. #18 and #42)

- Entire race (in case you are bored)
- Caution comes out — 3:11:00-3:12:07
- Final strategy decisions — 3:12:23-3:13:31
- Restart and final two laps — 3:16:49-3:19:32
- Pit Rho analysis in real-time workspace
  - Race Order
  - Competitor Strategy on Caution
  - Primary and Contingency Strategy Recs.
  - CSS Recommendation (what is this?)
  - Lane Position Gain
  - Driver Positions, Speed Comparison – “mic drop” moment

2019: New Rules Package

- Full announcement here

2019: New Rules Package

- What changed engineering-wise? (Ahem, JB.)
- Using statistics / linear regression from VEG test to predict major changes
  - Raw data (from NASCAR T&S)
  - Video clips to understand qualitatively
- Using test (and/or practice) data to “prepare” the advanced in-race models
  - Prepares teams for potential fundamental changes in race strategy (see: 19TEX1 example)
  - Even though we have that safety net of Pit Rho keeping up with track conditions, we have greater peace of mind

2019: New Rules Package (Plan)

- Scrape NASCAR T&S data in real-time
  - Great practice for intake when it matters in-season
  - Parse by practice/test session
- Remove cars that are not competitive
- Analyze only relevant test session(s)
- Review data to determine “position”
  - Adjust position based on # of cars in group
- Do the data/model match expectations?

2019: New Rules Package (Model)

- OLS regression model on laptimes
  - Again, imperfect model – but still OK; why?
  - Rough adj. model produces sensible adjusted laptimes
  - No significant tire falloff! Jives with physics of new package

Dependent Variable: Lap Time	Coef.	Std. Err.	t-score	p-value	95% confidence interval
Position	0.051	(0.019)	2.64	0.01	[0.013 ,0.090] **
# Cars Competing	-0.042	(0.026)	-1.62	0.11	[-0.092 ,0.009]
Laps on Tires	-0.002	(0.007)	-0.33	0.74	[-0.016 ,0.011]
Constant	31.093	(0.375)	82.88	0.00	[30.354 ,31.832] ***
Adjusted R-squared	0.2932				
Log-Likelihood	-11				
Observations	236				

* Significant at 5%, ** Significant at 1%, *** Significant at 0.1%.

Includes fixed effects for teams

2019: New Rules Package (Pace)

- Fixed effects in model for each car's pace
  - Rough, imperfect – but it passes the sanity check
  - Well-funded, traditionally fast cars rank toward the front
  - Not much difference in adjusted pace across field (as expected)

Car	Laps	Avg. Pos.	Avg. Raw Laptime	Avg. Adj. Laptime
2	24	7.625	30.778	30.833
6	24	8.042	30.918	30.953
18	24	3.333	30.686	30.962
14	23	3.696	30.708	30.970
48	24	3.208	30.704	30.986
21	23	6.174	30.857	30.991
95	24	8.458	30.998	31.011
47	16	6.250	30.885	31.045
3	24	1.167	30.681	31.068
1	15	9.467	31.084	31.084
43	24	5.875	31.078	31.286
13	21	3.143	31.015	31.368

19TEX1, #9 (vs. #11, #24, #48)

- Entire race
- Final caution decisions — 2:52:10-2:52:41
- #9's final pitstop — 3:13:50-3:14:37
- Pit Rho analysis in real-time workspace
  - Race Order
  - Competitor Strategy on Caution
  - Primary Recommendation
  - Caution Strategy
  - CSS Recommendation
  - Green Flag 2TR vs. 4T
  - Speed Comparison (#9 vs. #11 – wow)