A Tour of Financial Modeling

Dale W.R. Rosenthal[1]

1 June 2018
R/Finance 2018

[1] info@q36llc.com

Q36

Introduction

Who am I?

- Head partner, Q_{36}: quant finance firm.
- PhD, Statistics, U. Chicago; BS, Elec. Eng., Cornell.
- Was: finance professor at UIC; now teach MSF class at Notre Dame.
- Presented @ ECB, FDIC, CFTC, BdF, BoFinland, NBSlovakia, RBNZ.
- 3 years: Proprietary algo trader, Equity Trading Lab, Morgan Stanley
- 5 years: Equity derivatives strategist, Long-Term Capital Management
- Interned at Goldman Sachs
- Co-organizer of R/Finance conference since start in 2009.

Q_{36}

Roadmap

- We often talk about fundamentals and modeling.
- Must unite these to make business/trading decisions.
- Over the next hour we are going to explore a few markets.
- The goal: see how to build a model and infer business knowledge.
- A beautiful aspect of financial theory is its wide reach.
- We are going to take an idea and see how far we can reach.
- We will try modeling different aspects of a spread:
  - A "brew" spread. Thoroughly.
- Will try to tease out spread, how it has changed.
- Maybe see implications for trading, IO, valuation?

Q_{36}

The Brew Spread

- A brew spread (or "corn crush") is ethanol less inputs.
- Typically, we look at the two biggest inputs:
  - Corn and natgas (to distill corn mash into ethanol).
- Typical inputs: 1 bushel of corn and 72.8k BTUs of natgas
- Typical outputs: 2.8 gal ethanol, 17 lbs of DDGs, 0.7 lbs corn oil.
- Researchers at UIUC (Farm Doc!) suggests 42 cents/gal of costs[2]
- Ethanol, corn, natgas: all traded futures; DDGs no longer.
- Does ratio change over time or with corn price levels?

[2] `farmdocdaily.illinois.edu/2017/02/the-profitability-of-ethanol-production-in-2016.html`

Q_{36}

Data

- To answer these questions, we will look at US daily data.
- Specifically: Futures end-of-day settlement prices.
- Obviously, daily data ignores intraday spread dynamics.
- Data limitations and contract changes:
  - Ethanol contract: Slow switch, Railcar to Platt's Jun–Aug 2017.
  - Natgas e-mini contract introduced in 2014; still low-volume

Q_{36}

Contract Conversions?

- Must know units for our prices vs. units for contracts.

Contract	Sym	Quoted	Contract
Corn	C	¢/bu	5000 bu
Natgas	NG	\$/mmBTU	10000 mmBTU
Mini Natgas	QG	\$/mmBTU	2500 mmBTU
Railcar Ethanol	EH	\$/gal	29000 gal
Platt's Ethanol	CU	\$/gal	42000 gal

- Why 42000 gallons? Equal to 1000 barrels (bbl).
- So maybe we should trade 15–5–2 corn–ethanol–mini natgas?

Q_{36}

Loading Data

We load the data using *R* and Quandl. Blend ethanol prices:

```
library(xts)
library(Quandl)
futures.tickers <- paste("CHRIS/CME_", c("NG1", "C1", "CU1", "EH1"), sep="")
settle.column <- "Settle"
futs.raw <- Quandl(futures.tickers[1], type="xts")[,settle.column]
for (i in 2:length(futures.tickers)) {
    tmp <- Quandl(futures.tickers[i], type="xts")[,settle.column]
    futs.raw <- cbind(futs.raw, tmp)
}
colnames(futs.raw) <- c("natgas", "corn", "platts.ethanol", "railcar.ethanol")

## Make smooth transition from railcar ethanol to Platt's
transition.length <- length(index(futs.raw["20170601/20170831"]))
transition.wts <- (1:transition.length)/transition.length
futs.raw$ethanol <-
    rbind(futs.raw$railcar.ethanol["/20170531"],
          futs.raw$railcar.ethanol["20170601/20170831"]*(1-transition.wts)+
          futs.raw$platts.ethanol["20170601/20170831"]*transition.wts,
          futs.raw$platts.ethanol["20170901/"])
futs <- futs.raw["20111001/"]
```

Q_{36}

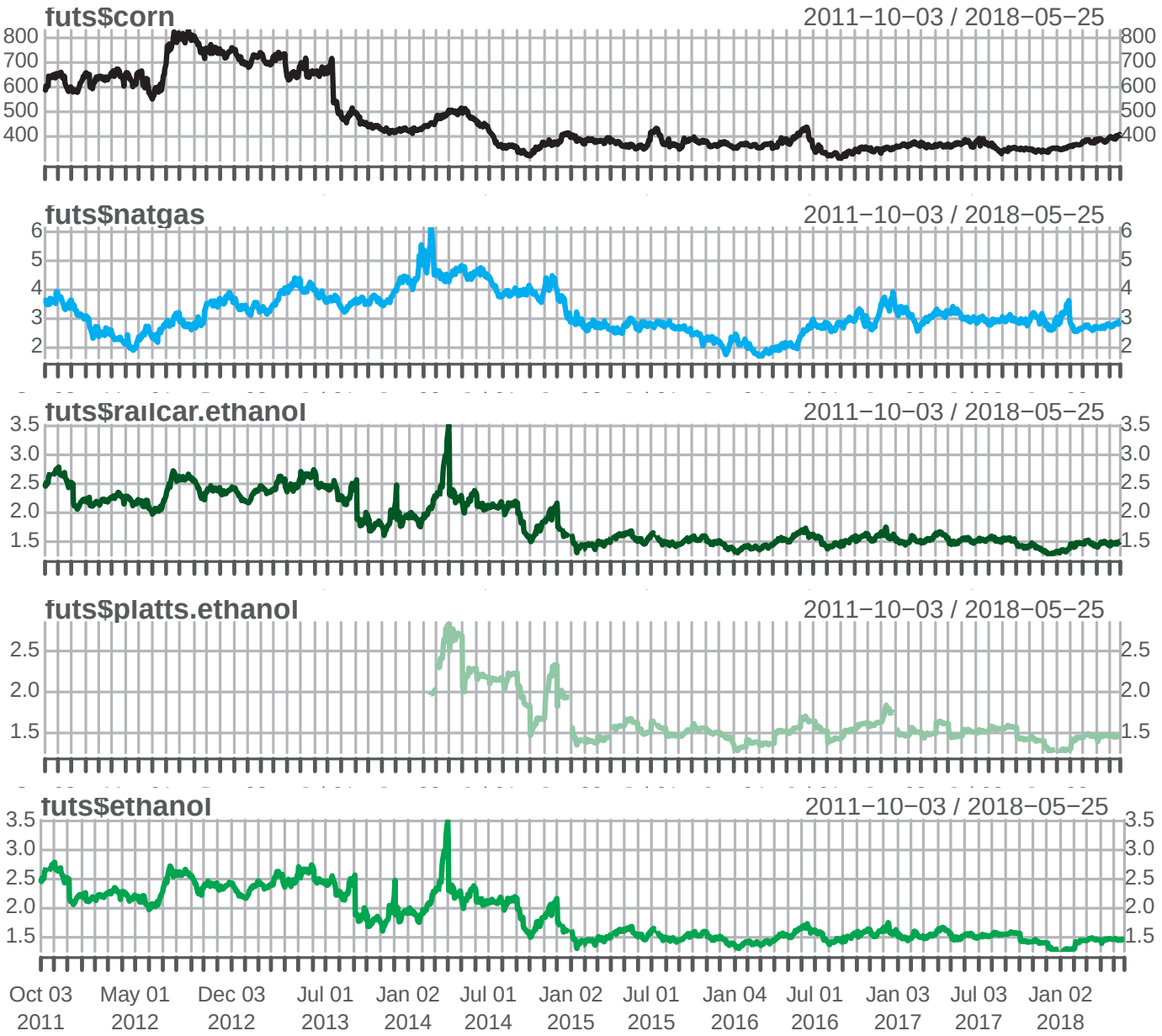
Observation

- The first step is always to look at the data.
- Are there spikes? data errors? comovement?
- For spreads, it helps to plot them on similar scales:

```
plot(futs$corn, col="black")
plot(futs$natgas, col="blue")
plot(futs$railcar.ethanol, col="darkgreen")
plot(futs$platts.ethanol, col="lightgreen")
plot(futs$ethanol, col="green")
```

Q_{36}

Observation: Plot



Q_{36}

Observation: Commentary

- We can notice a few things from these plots.
- Corn prices were higher up to mid-2013, now lower. (Tech shock?)
- Natgas has weird spikes in Jan+Feb (Polar Vortex) and late-2014.
- Spikes in ethanol about 1.5 months later; inventory implications.
- Spike transmission also suggests many ethanol producers unhedged.
- Natgas spike in late 2017 does not affect ethanol prices. Hedged?
- Blending of ethanol prices does not seem to have caused any issues.

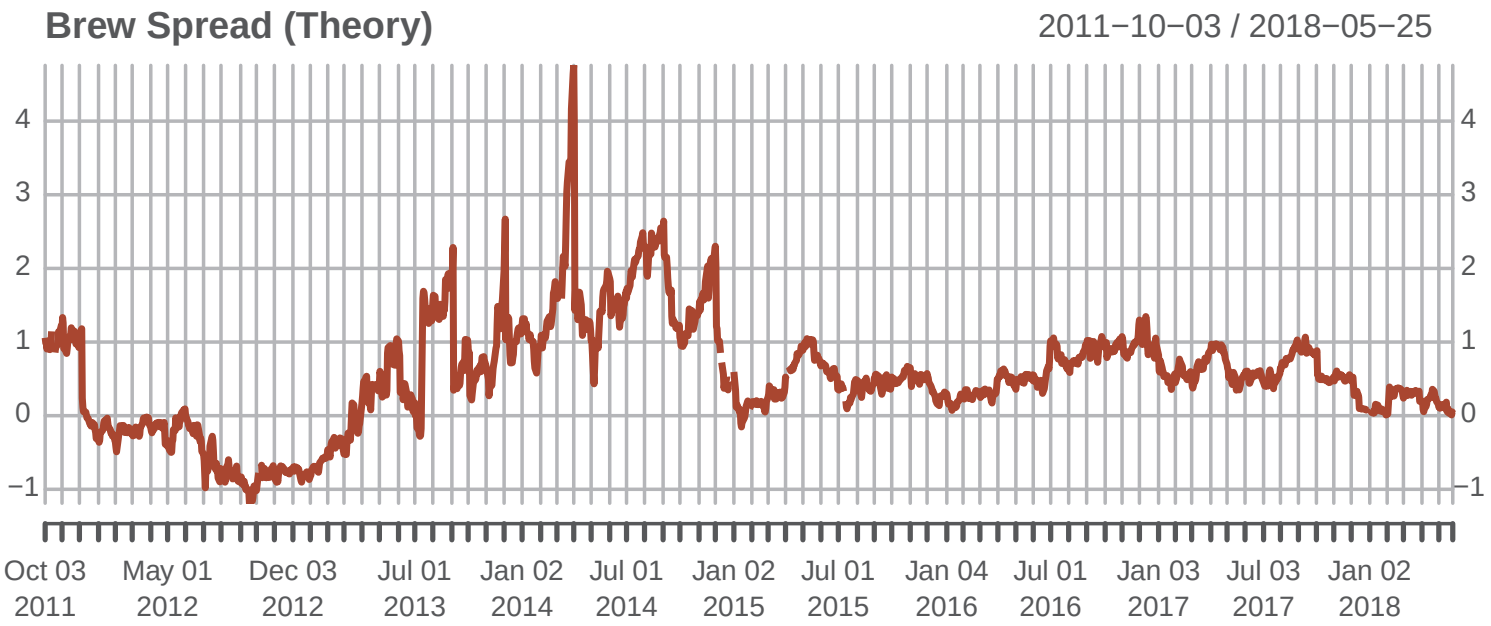
Q_{36}

Traded Brew Spreads

- We noted earlier that we could trade a brew spread: 15-5-2.
- Theoretical spread (IA State) would be 15-5-2.18 (almost identical).
- However, prices have different ratios. Plot brew spreads:[3]

```
brew.spread <- 2.8*futs$ethanol-futs$corn/100-72.8/1000*futs$natgas/15
plot(brew.spread, col="brown", main="Brew Spread (Theory)")
```

- Oscillated mostly b/w -$1/bu and $2/bu except in 2014.
- More recently has largely stayed b/w 0 and $1/bu.;



[3] In brown, since maybe it is a nice brown ale.

Estimating an O-U Model

- Recall the Ornstein-Uhlenbeck SDE:

$$dS_t = \gamma(\mu - S_t)dt + \sigma dW_t. \tag{1}$$

- Solving this and discretizing time, we get:

$$S_{i+1} = \mu(1 - e^{-\gamma\Delta t}) + S_i e^{-\gamma\Delta t} + \sigma\sqrt{\frac{1 - e^{-2\gamma\Delta t}}{2\gamma}}\epsilon_i \tag{2}$$

  where $\epsilon_i \sim N(0, 1)$.
- So we can fit a linear model: $S_{i+1} = a + bS_i + \eta_i$.
- Then we convert to O-U parameters:

$$\gamma = \frac{-\log(b)}{\Delta t} \qquad \mu = \frac{a}{1-b} \qquad \sigma = \hat{\sigma}_\eta\sqrt{\frac{-2\log(b)}{\Delta t(1-b^2)}} \tag{3}$$

- Implied half-life of convergence: $\log(2)/\gamma$.

Q36

Estimating an O-U Model: Code

- Fitting this linear model is easy with *R* and xts:

```
ou.brew <- lm(brew.spread~lag(brew.spread))
summary(ou.brew)
> ...
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)       0.008847   0.004190   2.111   0.0349 *
lag(brew.spread) 0.982637   0.004600 213.636   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1379 on 1650 degrees of freedom
  (24 observations deleted due to missingness)
Multiple R-squared:  0.9651, Adjusted R-squared:  0.9651
F-statistic: 4.564e+04 on 1 and 1650 DF,  p-value: < 2.2e-16
```

- Good to see both intercept and slope are significant.
- R^2 = *terrible* metric, but good we explain 96.5% of variation.

Q_{36}

Estimating an O-U Model: Parameters

- We then use these to solve for brew spread O-U parameters:[4]

Spread	$\hat{a}$	$\hat{b}$	$\hat{\sigma}_\eta$	$\hat{\mu}$	$\hat{\gamma}$	$\hat{\sigma}$	HL
Theory	0.0088	0.9826	0.138	0.510	0.0175	0.139	39.6

- *Half-life*: mean time to revert halfway to μ.
- Half-life of about 40 *business days* = 2 calendar months.
- If spread not converging by half-life: structural change?
- Is the theoretical spread best? Tough to say. Let's look deeper.

Q_{36}

[4] For $\Delta t = 1$ — which means our units will be in days.

Finding Data Implied Spreads: Cointegration

- We start by asking if the prices are cointegrated.
- Use the `urca` (Unit-Root Cointegration Analysis) package.
- Must tell the procedures which model form we care about.
- Usually: want "transitory" form, "constant" spread.

```
library(urca)
vars2study <- c("ethanol", "corn", "natgas")
ci.model <- ca.jo(futs[,vars2study], type="trace",
                  ecdet="const", spec="transitory")
summary(ci.model)
```

Q_{36}

Finding Data Implied Spreads: Cointegration?

```
Test type: trace statistic , without linear trend and constant in cointegration
Eigenvalues (lambda):
[1] 2.870931e-02 4.026871e-03 1.398593e-03 1.718671e-18
Values of teststatistic and critical values of test:
          test 10pct  5pct  1pct
r <= 2 |  2.32  7.52  9.24 12.97
r <= 1 |  9.03 17.85 19.96 24.60
r = 0  | 57.41 32.00 34.91 41.07
```

- Is there a cointegrating relationship? Yes, just 1.
- Reject test for 0 relationships: $57.41 > 34.91, 41.07$.
- Can't reject ≤ 1 relationships. ($9.03 < 19.96$)
- So there does seem to be a cointegrating relationship.

Q_{36}

Finding Data Implied Spreads: Cointegration...

```
cajorls(ci.model)
> $beta
>                    ect1
> ethanol.l1  1.000000000
> corn.l1    -0.002377774
> natgas.l1  -0.236428486
> constant    0.002311395
```

- What does the data suggest? Strangeness.
- Long 1 gal ethanol, short 0.24 bu corn, short 236kBTU natgas.
- So brewers used 2/3 of typical corn and 3.25 as much heat?
- Should we trade this? Probably not. You have a brain; use it.

Q_{36}

Finding Data Implied Spreads: Cointegration Co-problems

- Cointegration doesn't know that corn + natgas $\implies$ ethanol.
- When a model fails, your responsibility is to use your head.
- Economics drives the industry. Maybe the data is wrong.
- Data are settlement prices; set at end-of-day.
- People might try to manipulate settle prices.
- Is this likely? If settle prices are off, how to arb them?
- More likely: maybe the industry has changed over time.

Q_{36}

Time Varying Spreads? Code

- Perhaps the crack spread is varying with time.
- Break things up by year to investigate.

```
years <- paste(2011:2018)
for (yr in years) {
  tmp.brew <- lm(brew.spread[yr]~lag(brew.spread[yr]))
  gamma.brew <- -log(coef(tmp.brew)[2])  # -log of the slope
  mu.brew <- coef(tmp.brew)[1]/(1-coef(tmp.brew)[2])
  print(sprintf("%s Brew: sd=%0.3f mu=%0.3f gamma=%0.3f hl=%0.2f",
                yr, sd(brew.spread[yr], na.rm=TRUE),
                mu.brew, gamma.brew, log(2)/gamma.brew))
}
```

Q_{36}

Time Varying Spreads? Results

Year	sd	$\hat{a}$	$\hat{b}$	$\hat{\sigma}_\eta$	$\hat{\mu}$	$\hat{\gamma}$	$\hat{hl}$
2011	0.519	-0.015	0.955	0.149	-2.753	0.005	128.89
2012	0.332	-0.012	0.980	0.070	-0.614	0.020	34.69
2013	0.726	0.037	0.943	0.230	0.646	0.059	11.81
2014	0.688	0.080	0.949	0.226	1.556	0.053	13.18
2015	0.242	0.013	0.969	0.056	0.414	0.032	21.96
2016	0.297	0.017	0.978	0.062	0.763	0.023	30.47
2017	0.208	0.014	0.972	0.056	0.502	0.029	24.19
2018	0.118	0.019	0.903	0.052	0.194	0.102	6.78

- Seems brew spread has decreased in variance, mean; faster reverting.
- Structural break after 2012? Check the news!
- Sure enough: Heavy losses; consolidation and plant closures.
- Reversion half-lives of 7–34 biz days $<$ 39 biz days.
- Why larger full-sample half-life estimate? Structural breaks!

Q_{36}

Estimating Industry Profitability: Data

- Now let's switch to looking at industry profitability.
- Tough to get industry-wide refining profits; we will proxy.
- Use profits (calendar year, in millions) of ethanol refining firms:

Year	2011	2012	2013	2014	2015	2016	2017
ADM	3997	3623	3889	4768	3966	3619	3506
ANDE	353	358	365	397	376	346	319
BG	2627	2573	2760	2621	2693	2410	1764
CHSCP	1734	1917	1603	1842	1311	900	910
GPRE	172	97	173	426	207	281	261
INGR	1126	1238	1131	1115	1242	1402	1473
MGPI	7	25	21	28	59	65	76
PEIX	19	-20	36	112	10	54	6
REGI	127	58	239	161	111	172	84
REX	35	14	64	142	51	71	44
Total	10197	9883	10309	11612	10024	9319	8442

- Sum these by year, see what predictive power spread has.
- Also need to compute average spread

Q_{36}

Mean Spreads — and Other Stats?

- Now calculate yearly mean and sd of brew spreads.

```
## clean up period.apply tagging w/EOY date
brew.mean <- period.apply(brew.spread, endpoints(brew.spread,"years",1),
                          mean, na.rm=TRUE)["/2017"]
brew.sd <- period.apply(brew.spread, endpoints(brew.spread,"years",1),
                        sd, na.rm=TRUE)["/2017"]
pl$mean.spread <- xts(coredata(brew.mean), order.by=pl.years)
pl$sd.spread <- xts(coredata(brew.sd), order.by=pl.years)
```

- This yields the following brew spread stats:

Year	2011	2012	2013	2014	2015	2016	2017
Mean	0.72	-0.49	0.51	1.60	0.45	0.65	0.60
Std Dev	0.52	0.33	0.73	0.69	0.24	0.30	0.21

- Add post-2014 dummy; spread dynamics seem to change then.

Q_{36}

Estimating Industry Profitability: Models?

- Profits often quadratic in mispricing; size ↑ as profits ↑.
- We can fit a linear model with post-2014 dummy, mean-squared:

```
pl$post2014 <- index(pl) >= as.POSIXct("20150101", format="%Y%m%d")
lm(total ~ brew.mean + brew.sd, data=pl)
lm(total ~ brew.mean + brew.sd + post2014, data=pl)
lm(total ~ brew.mean + I(brew.mean^2) + brew.sd, data=pl)
```

- Unfortunately, results are not good:

	Model 1		Model 2		Model 3	
Variable	Est.	t-stat	Est.	St-stat	Est.	t-stat
Intercept	8.45 bn	12.8	9.76 bn	5.4	8.65 bn	14.4
Spread Mean	0.22 bn	0.4	0.63 bn	0.8	-0.50 bn	-0.7
Spread Std Dev	3.24 bn	2.0	0.65 bn	0.2	2.59 bn	1.8
Post-2014			-1.02 bn	-0.8		
Spread Mean^2					0.76 bn	1.4

- What went wrong?? Look at the companies we chose.
- ADM and Bunge are large multi-commodity traders.
- Ingredion, however, was once known as Corn Products.

Q_{36}

Estimating Industry Profitability: Models w/Better Data?

- So create a total without ADM and BG, refit.

```
pl$revtotal <- pl$total - pl$ADM - pl$BG
```

- These results are better:

	Model 1		Model 2		Model 3	
Variable	Est.	t-stat	Est.	St-stat	Est.	t-stat
Intercept	3.02 bn	12.6	3.79 bn	7.0	3.13 bn	42.0
Spread Mean	0.04 bn	0.2	0.28 bn	1.2	-0.36 bn	-4.1
Spread Std Dev	1.22 bn	2.1	-0.30 bn	-0.3	0.87 bn	4.7
Post-2014			-0.60 bn	-1.5		
Spread Mean^2					0.42 bn	6.4

- So maybe we have found an OK model: Model 3.
- Post-2014 dummy does not seem to be helpful.
- If we had profits/gallon, results would be much stronger.
- 2018? Brew spread so far: mean=0.60, sd=0.21
- $\Longrightarrow$ \$3.2 bn profit for firms\{ADM,BG}

Q36

Valuing a Simple Refinery

- Consider building a simple ethanol refinery (100 mn gal).
- How would we estimate the value of a refinery?
- If we had a model of profits/gal, we could use that.
- We sort of do: FarmDoc (UIUC) suggests costs around \$0.42/gal.

$$\text{profit/gal} = (\bar{S}_{yr,brew} - 0.42)^{+}. \tag{4}$$

- But this is inherently an Asian option-like payoff.
- We can simulate the mean-reverting spread as:

$$S_T = \hat{\mu} + e^{-\gamma T}(S_0 - \hat{\mu}) + \sigma\sqrt{\frac{1 - e^{-2\gamma T}}{2\gamma}} Z_T \tag{5}$$

where $Z_T \overset{iid}{\sim} N(0,1)$

Q36

Valuing a Simple Refinery: Code

- Use overall O-U numbers (since post-2014 dummy not significant).
- So: $\hat{\mu} = \$0.51$, $\hat{\gamma} = 0.0175$, $\hat{\sigma} = 0.139$.
- Price as of 1 Jan 2017; use $S_0 = \$0.822$. (Or 2018: $S_0 = 0.07\ldots$)
- But this is an Asian option; can approximate:
  - $\bar{S}_{0\to T} \approx S_T - \frac{\sigma^2}{2}$; $\sigma_{\bar{S}} \approx \sigma\sqrt{\frac{1}{3}}$.

```
T <- 1; r <- 0.02  # 1 year option, 2% int. rates
s0 <- 0.822-0.139^2/2; mu <- 0.51; gamma <- 0.0175
sigma.brew <- 0.139/sqrt(3); K <- 0.42
num.sims <- 100000  # 100,000 simulations
z <- rnorm(num.sims)  # simulated Z_T values
s.sim <- mu + exp(-gamma*T)*(s0-mu) +
    sigma.brew*sqrt((1-exp(-2*gamma*T))/(2*gamma))*z
profit.sim <- pmax(s.sim - K, 0)
opt.value <- mean(profit.sim)*exp(-r*T)
opt.sd <- sd(profit.sim*exp(-r*T))
```

Q_{36}

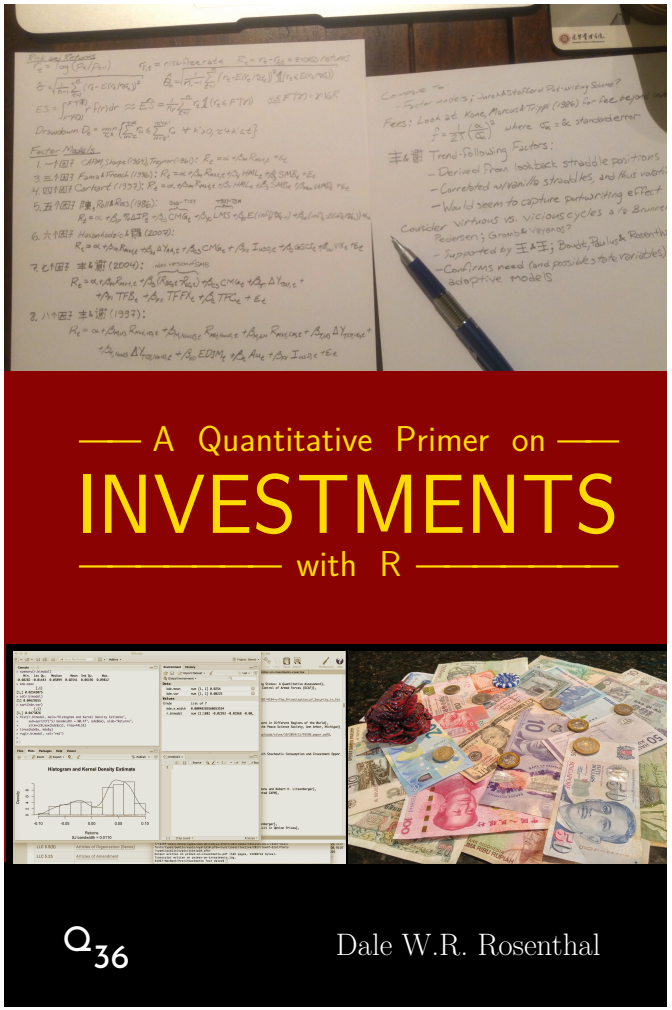
Valuing a Simple Refinery: Real Option Results

- For Ornstein-Uhlenbeck with Asian option approximation:
- Refinery value: \$0.379/gal; sd: \$0.078/gal = \$37.89 mn/year.
- For 100 mn gal/year refinery: profit=\$37.9 mn/year; sd=\$7.8 mn.
- \$37.9 mn perpetuity value: $\frac{\$37.9mn}{r_f=0.02}$ = \$1.9 bn, sd \$391 mn
- Compare to DCF: $\frac{0.822-0.42}{WACC=0.08?} \times 100$ mn = \$503 mn
- If we price as of 1 Jan 2018: $S_0 = 0.07\ldots$ value is 0.
- Why? Hard to model politicians breaking long-established policy.

No time to look at CAPM, factor models; Jensen's Inequality correction!

Q36

Want More?



- Buy the book! Amazon, B&N, etc.
- Come to the book launch! Email firm (info@q36llc.com)
- Check out the slides: `www.q36llc.com/think/investments`

Q36