CIC Seminar: Econometrics and Computation

Felix Chan [1]

[1]School of Economics and Finance, Curtin University

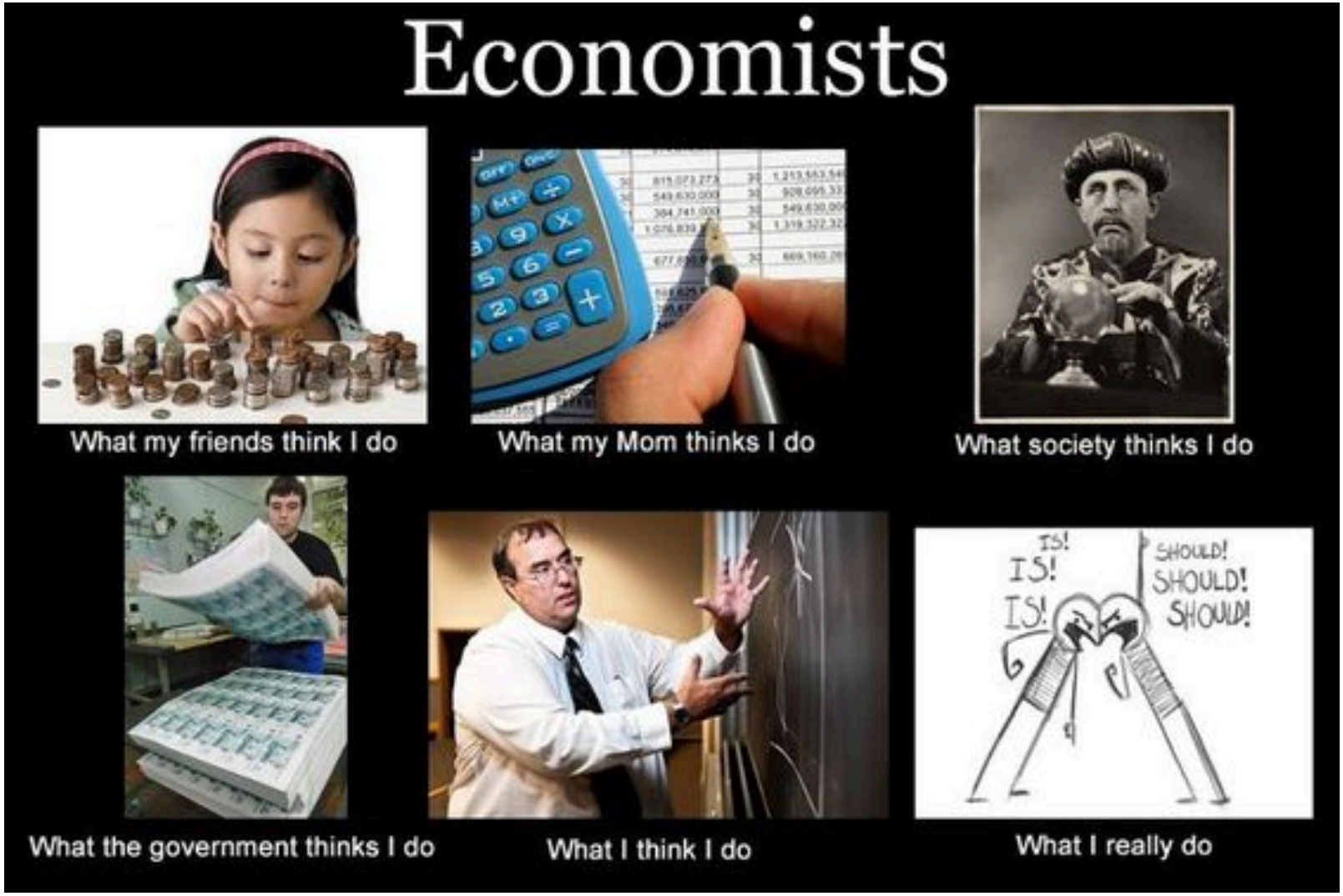
October 2016

Outline

Who are we and what do we do?

- What is economics?
- What do economists do?

What do we do?


Economists
What my friends think I do
What my Mom thinks I do
What society thinks I do
What the government thinks I do
What I think I do
IS!
IS!
IS!
SHOULD!
SHOULD!
SHOULD!
What I really do

Who are we and what do we do?

- "Economics is the only field in which two people can share a Nobel Prize for saying opposing things."

Who are we and what do we do?

- "Economics is the only field in which two people can share a Nobel Prize for saying opposing things."

Who are we and what do we do?

- "Economics is the only field in which two people can share a Nobel Prize for saying opposing things."

- That is probably why they also gave the prize to Lars Peter Hansen, an econometricians, in 2013.

Econometrics

- Economic theories are mostly creation of individuals, exist only in their minds.
- Our job is to examine these theories with data.
- Data sources are limited. Why?

Econometrics

- Economic theories are mostly creation of individuals, exist only in their minds.
- Our job is to examine these theories with data.
- Data sources are limited. Why?
- We generally cannot do experiment.
  - Not a good idea to increase interest rate by 10% and see how the economy reacts.
  - Ethic committee would not allow us to "remove" 20% of the population and see how may that affect unemployment.

This Talk is about...

- A quick (and incomplete) survey on different types of econometric problems.
- Solved, partially solved and unsolved.
- The link to computation should be obvious given the audience.
- It attempts to follow the two golden rules

This Talk is about...

- A quick (and incomplete) survey on different types of econometric problems.
- Solved, partially solved and unsolved.
- The link to computation should be obvious given the audience.
- It attempts to follow the two golden rules
  - Popularity of a talk is inversely proportional to the number of equations.
  - But directly proportional to the number of pictures and animations.

Disclaimer

My knowledge is limited!

Problem 1: A Variable with 'Super Power'

- Monetary economists working for central banks often use money supply to target inflation.
- So modelling money demand, specifically M1 and its relation to other variables, is an important component to understand macro-economics and monetary policy.
- Armed with regression analysis with OLS, t-statistics (p-value) and R^2, macro-economists discovered a variable with super-(explanatory) power some 40 years ago.

$$M1_t = \alpha + \beta \text{Dark\_Knight}_t + \text{other stuffs} + u_t$$

- Just like an undergraduate student will tell you - p-value should be less than 0.05, R^2 is preferred to be close to 1.

Super-variable

So we regress monthly U.K. Money Demand against the 'dark knight'.

	coef	**std err**	**t**	**P>|t|**	**[95.0% Conf. Int.]**
Intercept	12.2239	0.009	1348.474	0.000	12.206 12.242
dark_knight	0.0001	8.97e-07	139.637	0.000	0.000 0.000

Super-variable

So we regress monthly U.K. Money Demand against the 'dark knight'.

	coef	**std err**	**t**	**P>|t|**	**[95.0% Conf. Int.]**
Intercept	12.2239	0.009	1348.474	0.000	12.206 12.242
dark_knight	0.0001	8.97e-07	139.637	0.000	0.000 0.000

$R^2 = 0.986$.

But wait... the power continues...

It turns out that the "dark knight" can also explain US GDP at the quarterly frequency.

	coef	std err	t	P>|t|	[95.0% Conf. Int.]
Intercept	8.8718	0.007	1221.024	0.000	8.857 8.886
dark_knight	5.909e-05	7.2e-07	82.087	0.000	5.77e-05 6.05e-05

But wait... the power continues...

It turns out that the "dark knight" can also explain US GDP at the quarterly frequency.

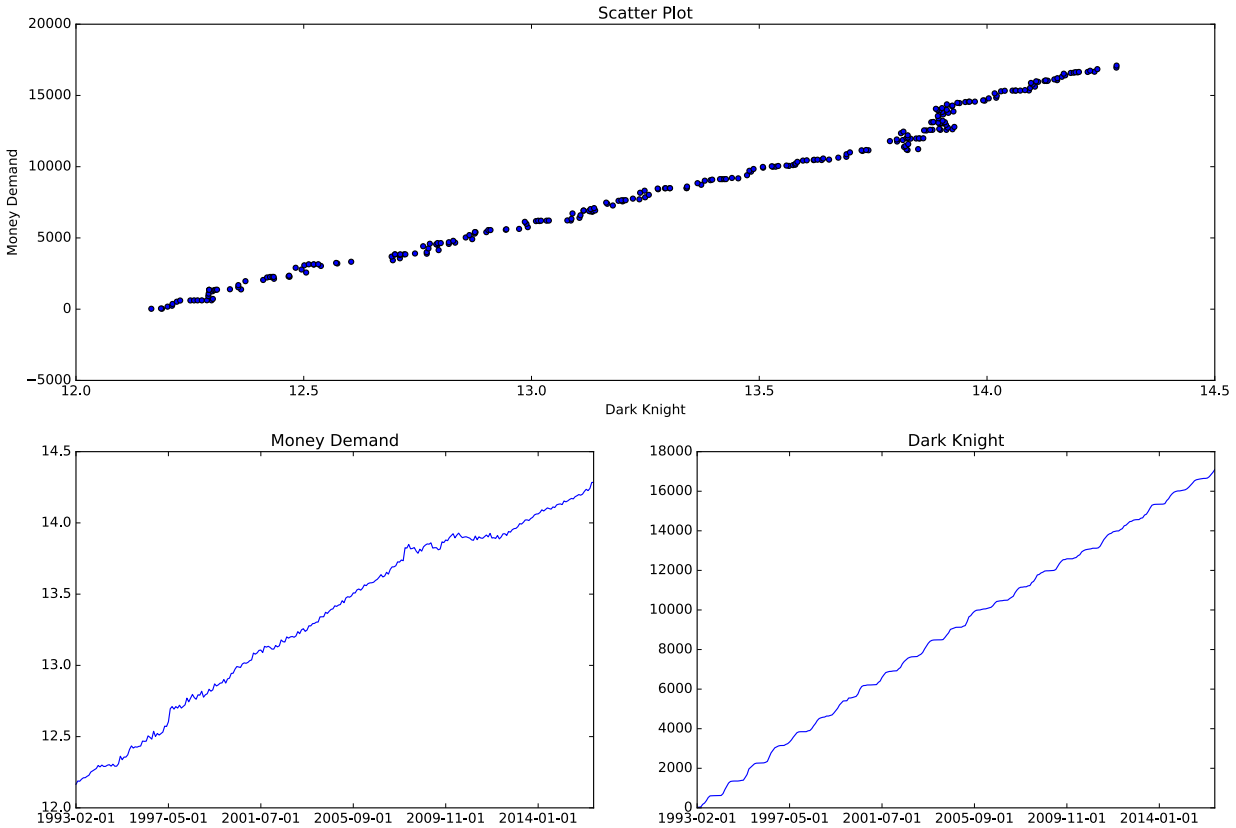
	coef	**std err**	**t**	**P>|t|**	**[95.0% Conf. Int.]**
Intercept	8.8718	0.007	1221.024	0.000	8.857 8.886
dark_knight	5.909e-05	7.2e-07	82.087	0.000	5.77e-05 6.05e-05

$R^2 = 0.987$ - It explains US GDP even better!!!

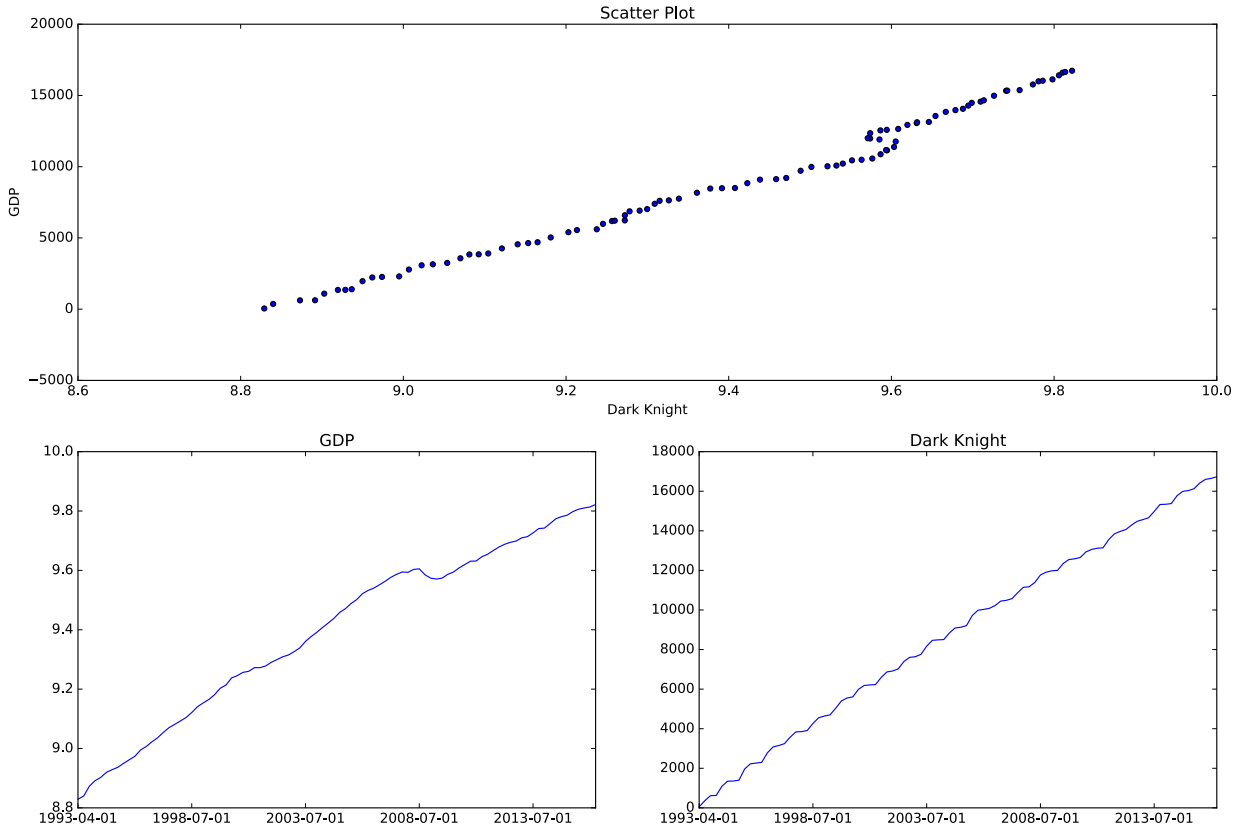
How can one variable explain so much?

- First rule about econometrics: if it is too good to be true, it is usually false.
- Probably a good idea to plot some graphs

Trend explains a lot



Trend explains a lot



Trend does not explain everything...

Let us de-trend all variables and run the regression again. For Money Demand we get this

	coef	**std err**	**t**	**P>|t|**	**[95.0% Conf. Int.]**
Intercept	-1.04e-14	0.004	-2.75e-12	1.000	-0.007 0.007
dt_dk	0.0003	1.76e-05	16.958	0.000	0.000 0.000

$R^2 = 0.508$. 'Explanatory power' clearly drops but the coefficient is still 'significant'!

Trend does not explain everything...

Same with de-trended GDP.

	coef	**std err**	**t**	**P>|t|**	**[95.0% Conf. Int.]**
Intercept	3.97e-15	0.003	1.44e-12	1.000	-0.005 0.005
dt_DK	0.0001	1.24e-05	12.021	0.000	0.000 0.000

$R^2 = 0.616$. Again, 'explanatory power' drops but the coefficient is still significant.

Should we be worried?

- Should we be worried? After all, we might have actually found a useful variable to explain both UK Money Demand and US GDP....
- Let's review the identity of this super variable.

Should we be worried?

- Should we be worried? After all, we might have actually found a useful variable to explain both UK Money Demand and US GDP....
- Let's review the identity of this super variable.

Steady rain boosts Perth's monthly average

Updated 23 May 2014, 7:45am

Steady rain from a cold front moving over the southern half of the state overnight has seen Perth pass its monthly average rainfall for May.

Perth City recorded 38 millimetres of rain and Jandakot almost 50 millimetres.

This month, Perth recorded more than 123 millimetres of rainfall. The average for May is 117.

Since November, the city has been experiencing an extended dry period.

The State Emergency Service (SES) received more than 20 calls for help overnight, mostly for minor leaks and water damage to property.

PHOTO: Rain has been falling steadily over Perth. (Andrew O'Connor)

MAP: Perth 6000

Topics: weather, perth-6000

First posted 23 May 2014, 7:21am

Print | Email | Facebook | Twitter | More

More stories from Western Australia

What's going on...?

- It turns out that if the variables are non-stationary, then regressing one against the other often leads to *spurious regression* as noted in Granger and Newbold (1974).
- The standard regression theory doesn't really apply in this case.
- See the demo.

The Solution...

- As it turns out, most macro-economics variables are non-stationary, most likely $I(1)$. So inference based on standard theory with OLS isn't quite right.
- Two things need to be done: (i) Identify if variables are non-stationary and (ii) if so, need to apply a different sets of estimation techniques, for example, co-integration, Engle and Granger (1986) and Johansen (1988)

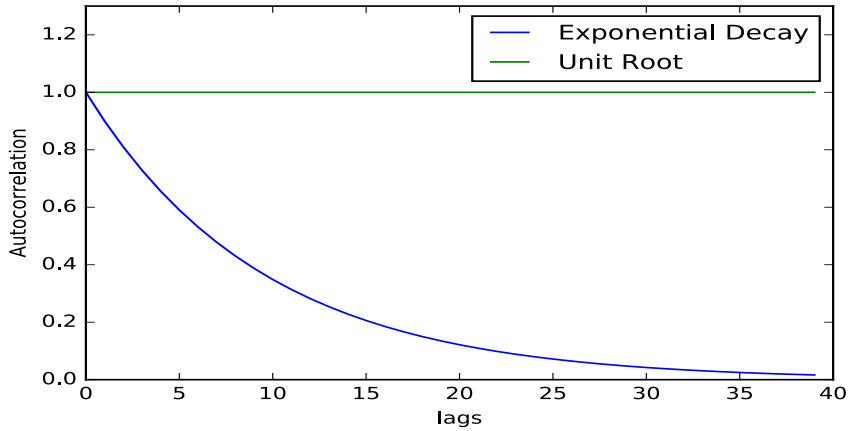
Simulation is inevitable

- Theory in this literature involves abstract mathematical objects such as Wiener process (Brownian Motion). It can be shown that both (i) and (ii) require simulation procedures in the case of $I(1)$ for purposes of valid inference.

$$T^{-1/2}{}_{t(\hat{\beta})} \sim \frac{\int_0^1 V(t)W(t)dt - \int_0^1 V(t)dt \int_0^1 W(t)dt}{\left\{\left[\int_0^1 V^2(t)dt - \left(\int_0^1 V(t)dt\right)^2\right]\left[\int_0^1 W^2(t)dt - \left(\int_0^1 W(t)dt\right)^2\right] - \left[\int_0^1 W(t)V(t)dt - \int_0^1 W(t)dt \int_0^1 V(t)dt\right]^2\right\}^{1/2}}$$

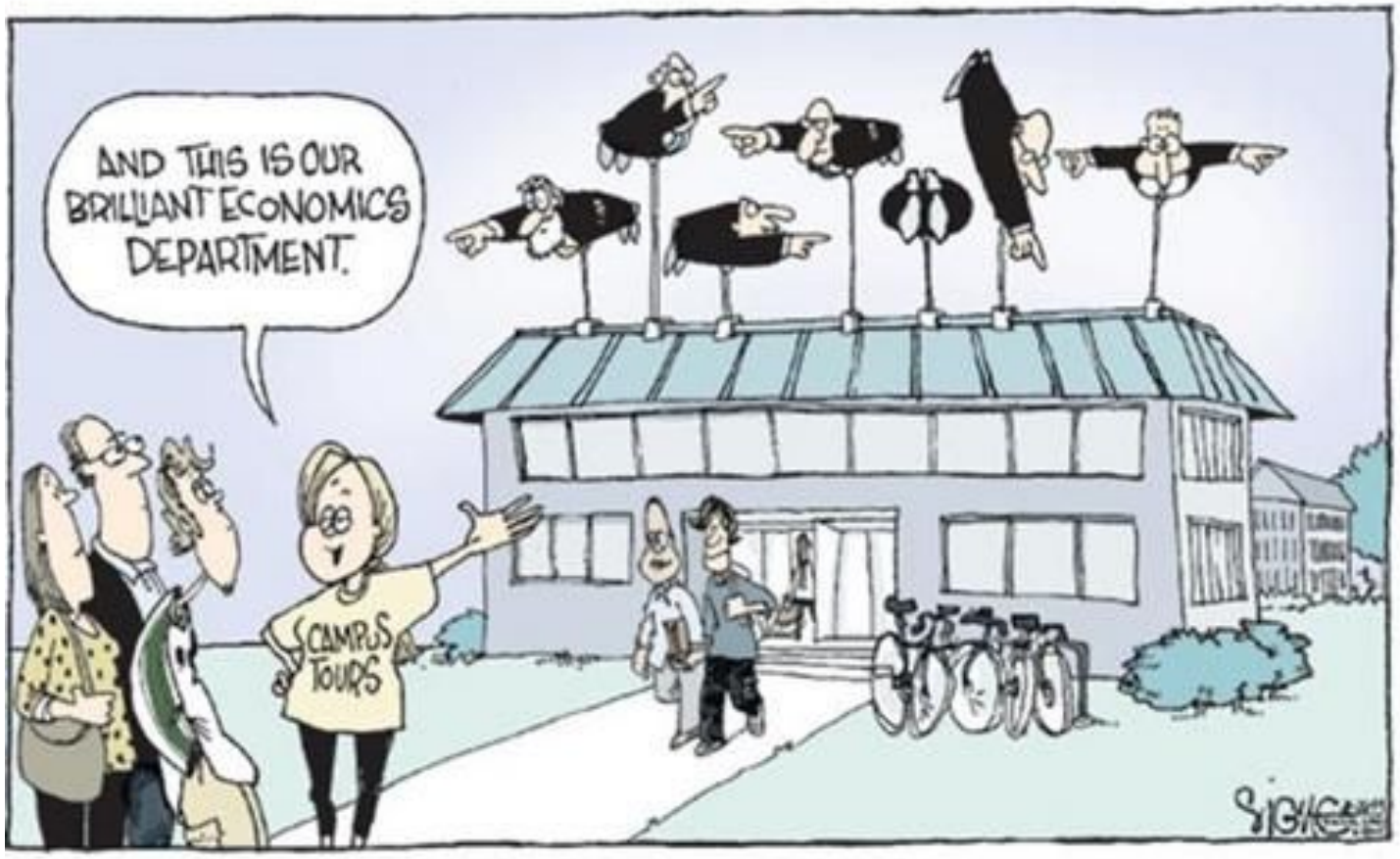
Results are limited

- Theory so far also focus on $I(1)$ process, we don't know as much about fractionally integrated process or any potential properties for co-fractional-integration.



- Some evidence to suggest that prices may be $I(2)$. Liquidation discount (fire sale) of stocks exhibited fractionally integrated behaviour (Singh et. al 2016).
- Most of these are likely to lead to non-standard theories. Computation and simulation are indispensable in these investigations.
- Techniques require stationarity should be applied with cautions.

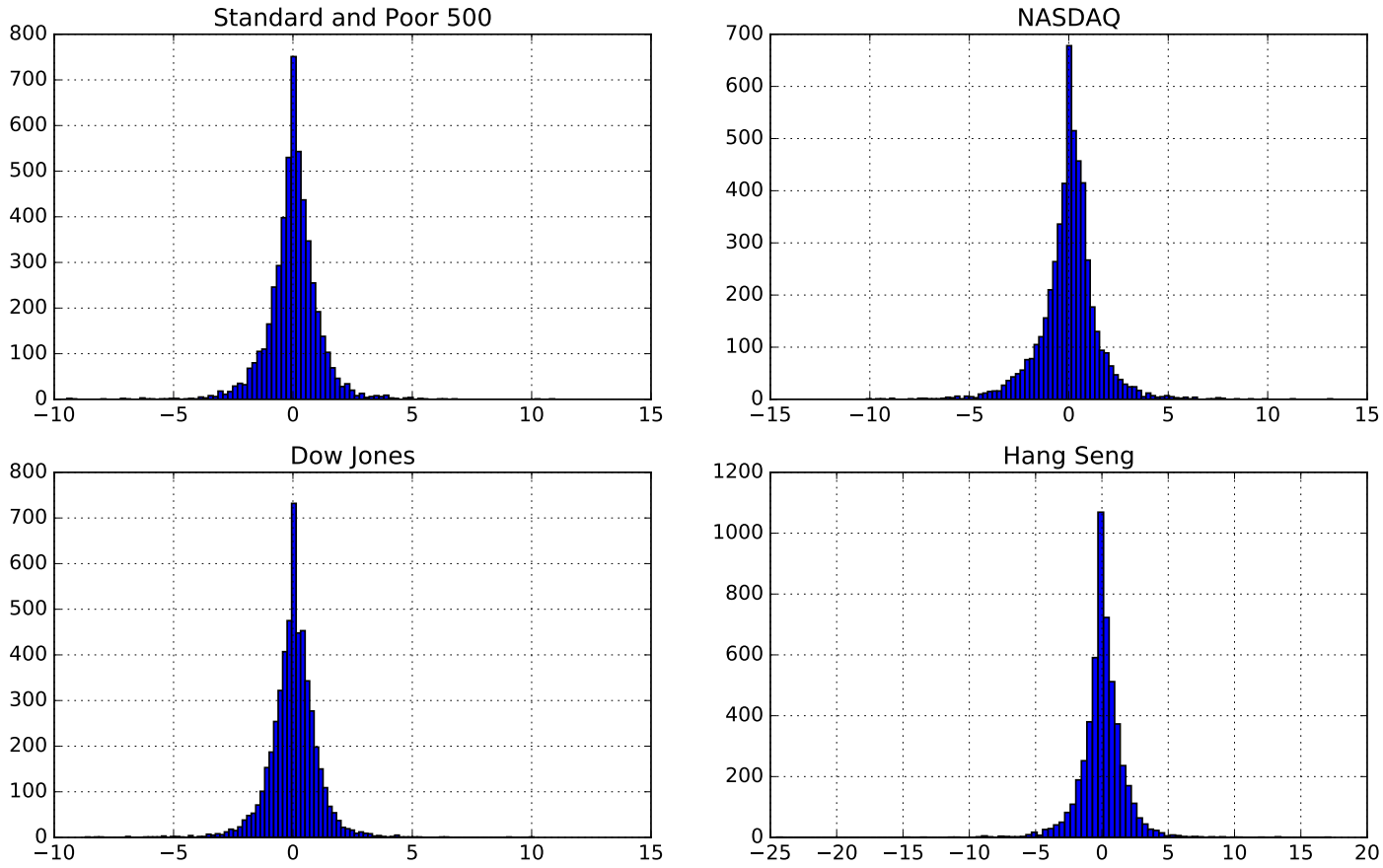
Problem 2: Tales of two tails



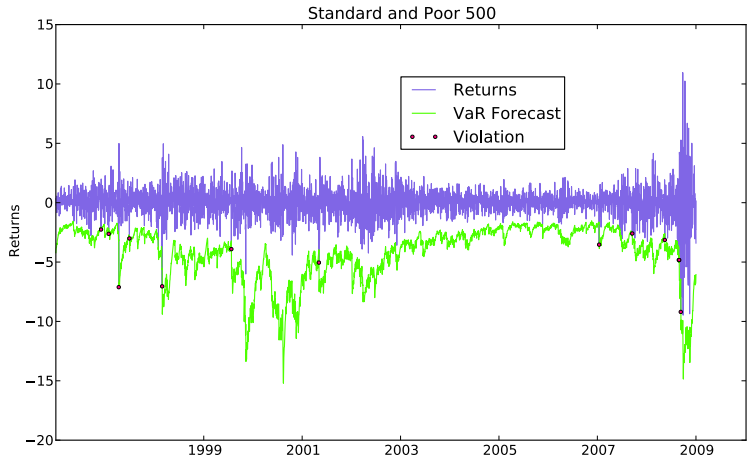
Forecast

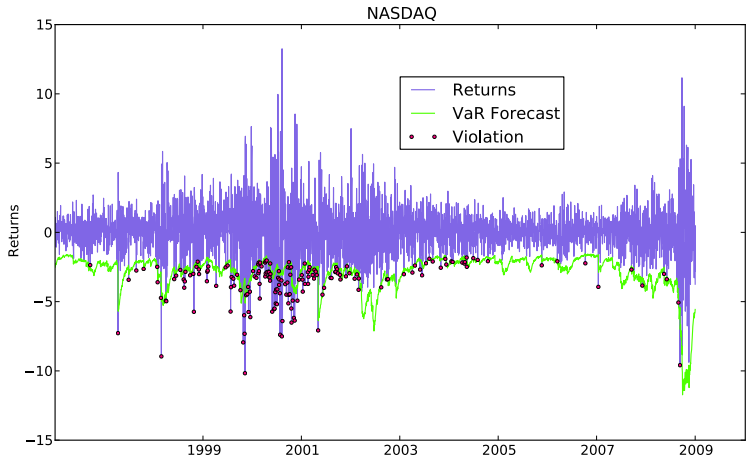
- Economic forecast is hard not only because some variables may be non-stationary with long-memory.
- Individuals use economic forecasts as information to make decisions. So their actions can actually change the outcomes based on the forecast.
- This generally doesn't happen to weather forecast in the short run.
- Standard forecast criteria, such as MSE, MAD, MAPE may not be all that appropriate.
- Historically, we also focus on forecasting the mean, but often we need to forecast the tail.... for example, we need to know how bad can something get....

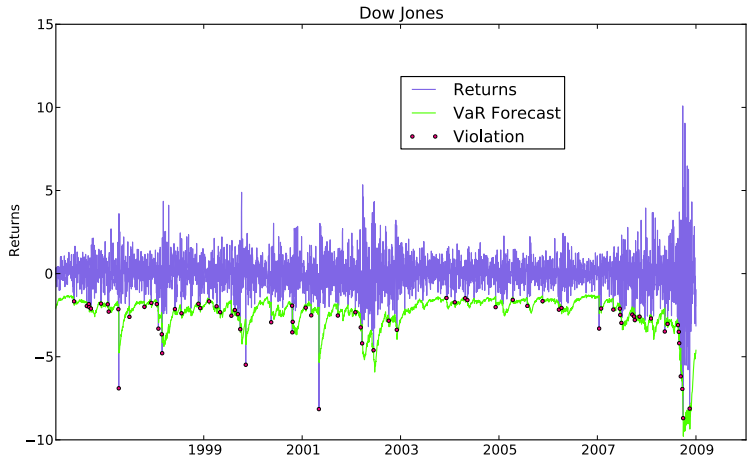
Risk Forecasting

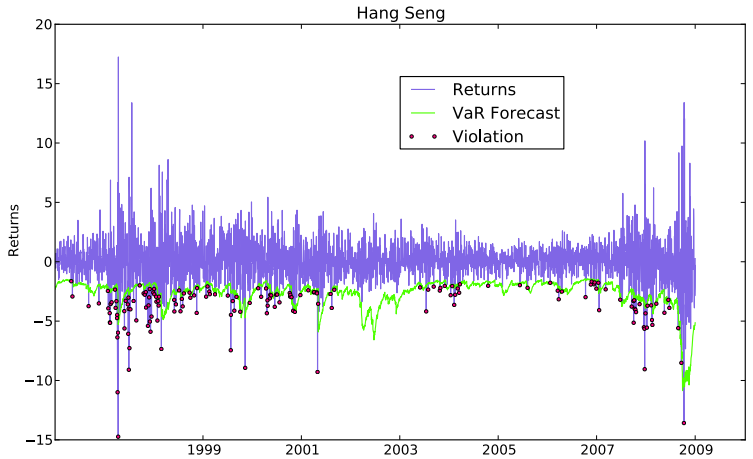


Risk Forecasting









A few fun facts

- Fat tails and clustering
- and most importantly...

A few fun facts

- Fat tails and clustering
- and most importantly...
- some evidence to suggest that the conditional variance may be changing over time.
- This means past information can help us to model variance more accurately, which leads to better forecast of the tail (VaR).
- The idea of conditional variance led to GARCH type models, initiated by Engle (1982). He won the Nobel Prize for this in 2003 (shared with Sir Clive Granger).

$$
\begin{aligned}
r_t &= \mathbb{E}\left(r_t | \Im_{t-1}; \boldsymbol{\phi}\right) + \varepsilon_t \\
\varepsilon_t &= \eta_t \sqrt{h_t} \qquad \eta_t \sim (0,1) \\
h_t &= \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}
\end{aligned}
$$

Loading Clip

How to use GARCH?

- Computationally simple.
- Given the data of returns, which we observed.
- Estimate the parameter vectors $\boldsymbol{\theta} = (\boldsymbol{\alpha}', \boldsymbol{\beta}'\omega)'$ and $\boldsymbol{\phi}$ using (Q)MLE, ie maximising the log-likelihood function

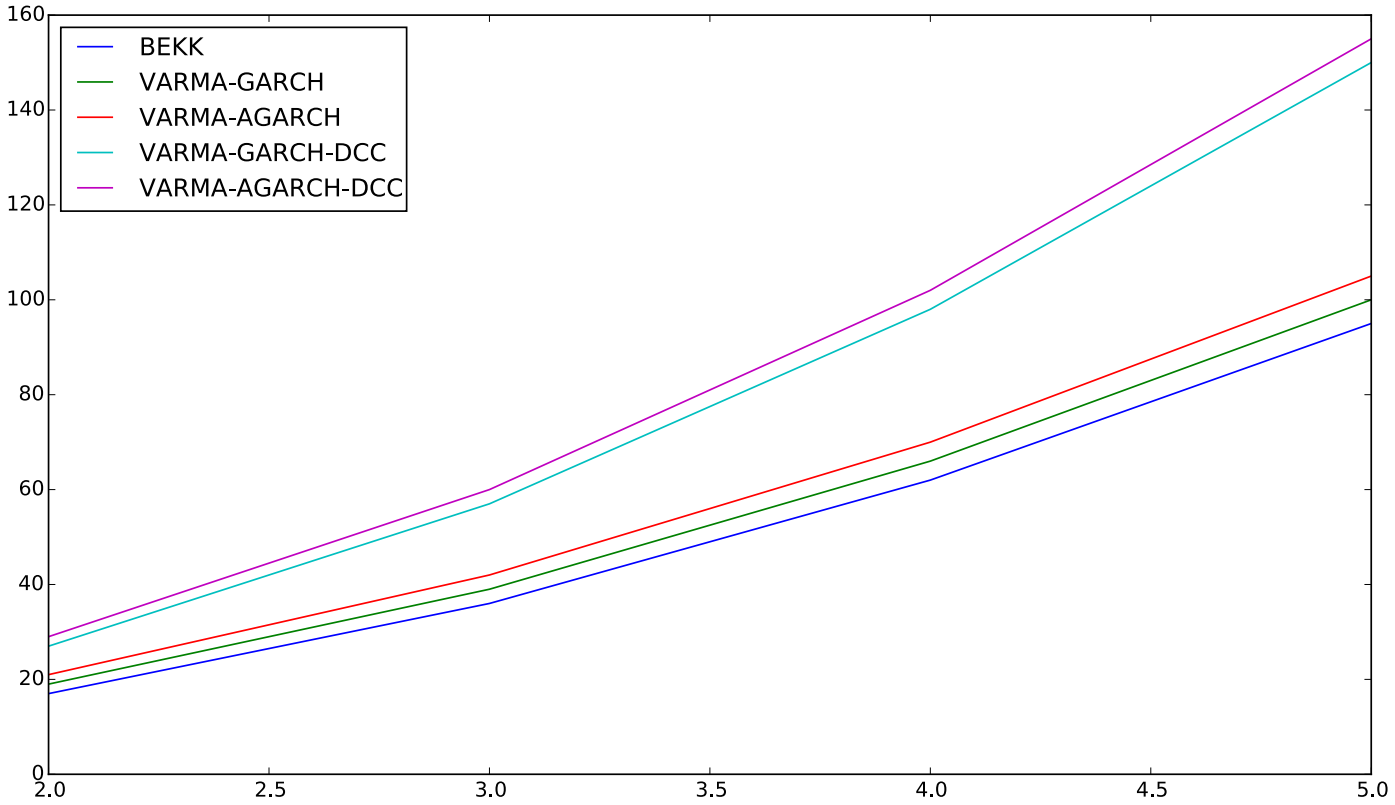
$$l(\boldsymbol{\theta}, \boldsymbol{\phi}) = -\frac{1}{2} \sum_{t=1}^{T} \left(\log h_t(\boldsymbol{\theta}) + \frac{\varepsilon_t^2(\boldsymbol{\phi})}{h_t(\boldsymbol{\theta})} \right)$$

Portfolio tends to have more than one asset...

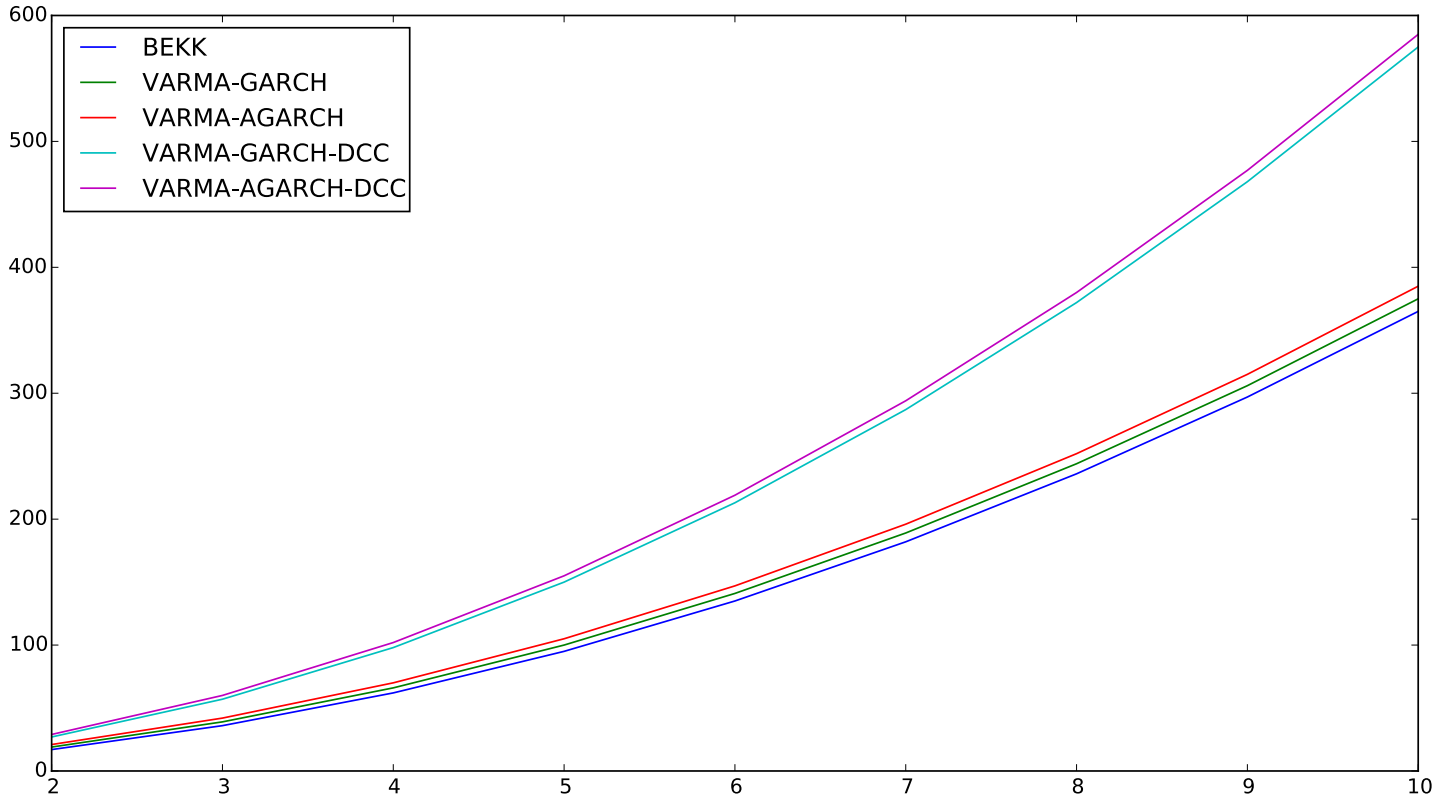
- One can extend the GARCH model to handle multivariate case but the number of parameters grow... very quickly.
- The log-likelihood are also getting 'bigger' and its maximisation becomes more and more unstable.

$$l(\boldsymbol{\theta}, \boldsymbol{\phi}) = -\frac{1}{2} \sum_{t=1}^{T} \left(\log \mathbf{H}_t(\boldsymbol{\theta}) + \boldsymbol{\varepsilon}_t(\boldsymbol{\phi})' \mathbf{H}_t^{-1}(\boldsymbol{\theta}) \boldsymbol{\varepsilon}_t(\boldsymbol{\phi}) \right)$$

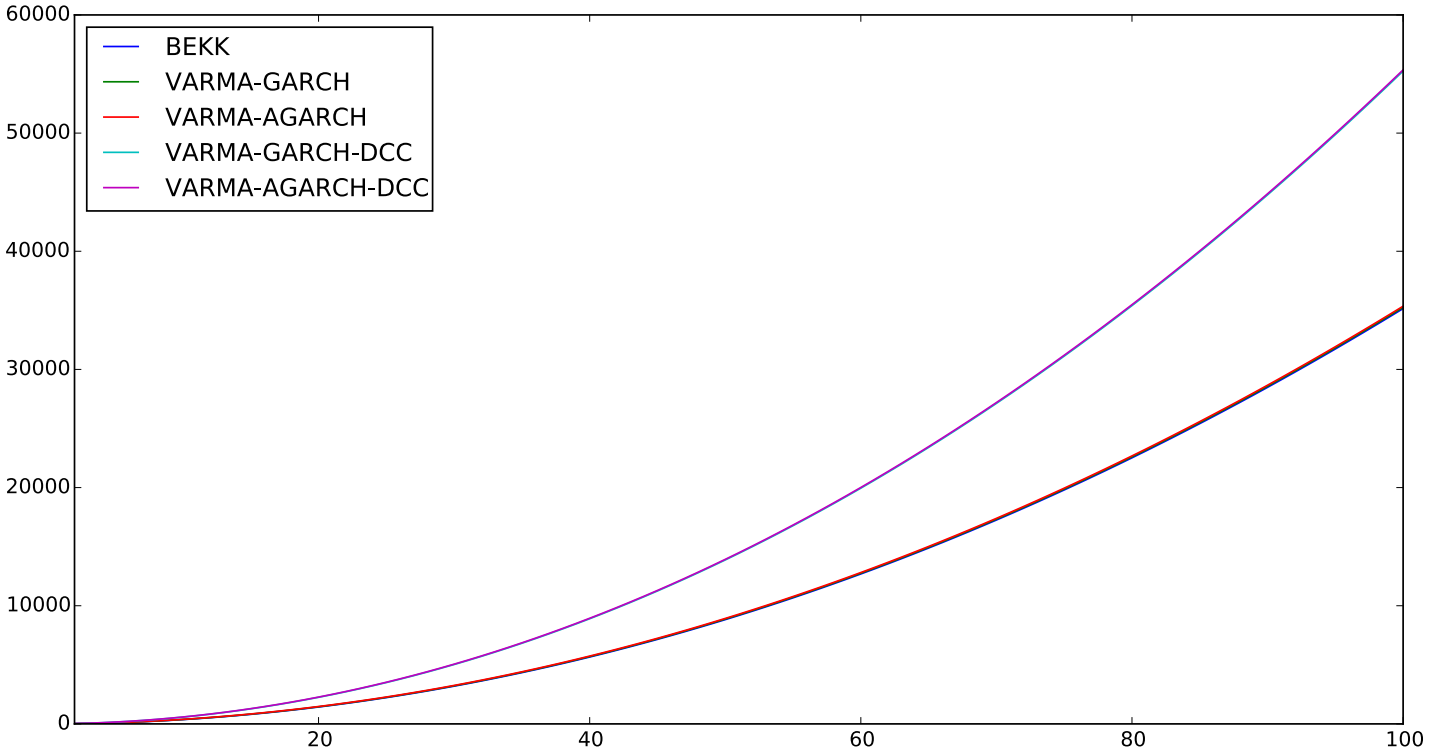
Curse of Dimensionality



Curse of Dimensionality



Curse of Dimensionality



More and more...

- As far as I know, the curse of dimensionality is an unsolved problem in the literature.
- Most applied researchers attempted to get around it rather than solving it.
- But this will only get worse.... why?

More and more...

- As far as I know, the curse of dimensionality is an unsolved problem in the literature.
- Most applied researchers attempted to get around it rather than solving it.
- But this will only get worse.... why?
- GARCH type models do pretty well with data at daily frequency but now we have intra-daily data.

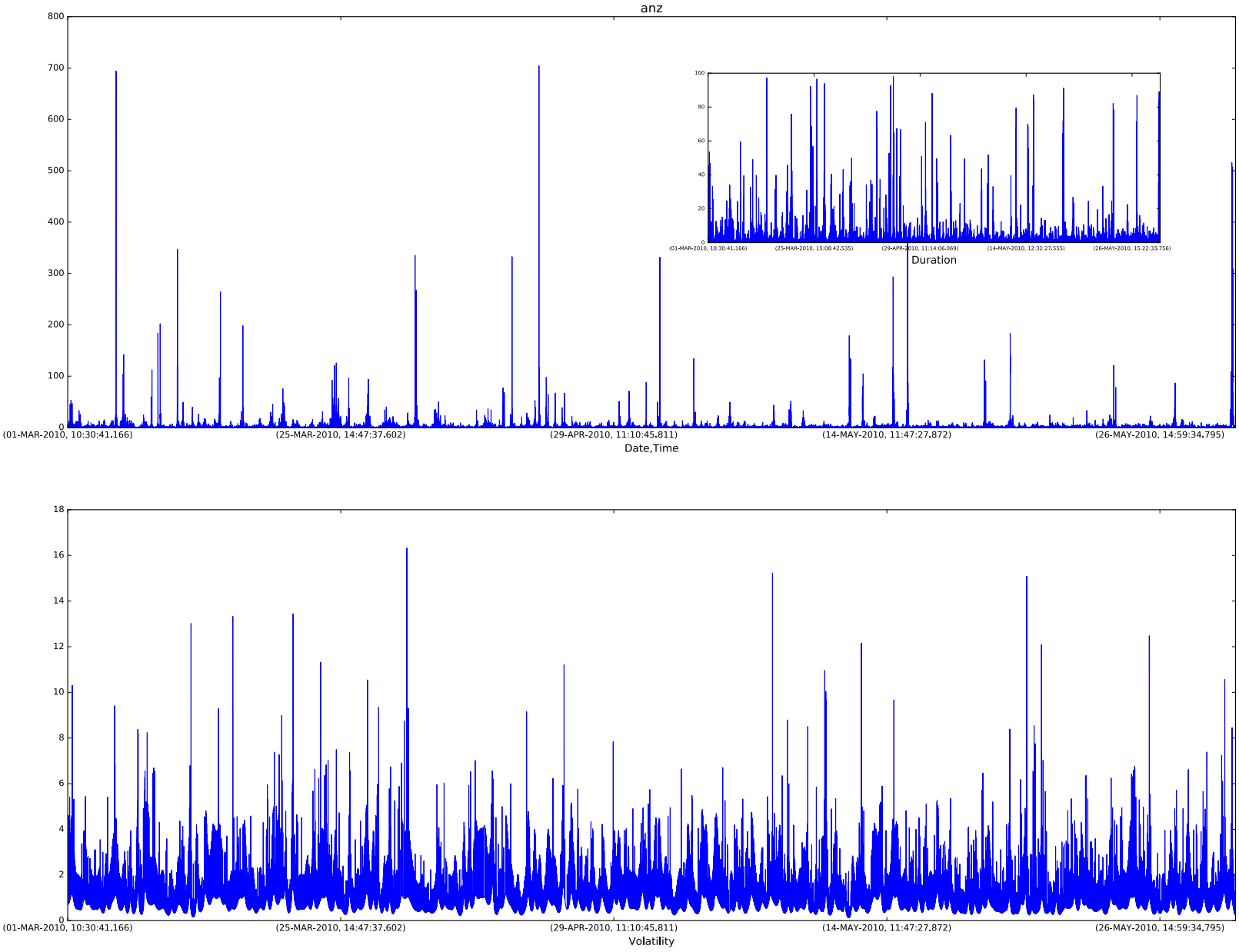
More and more...

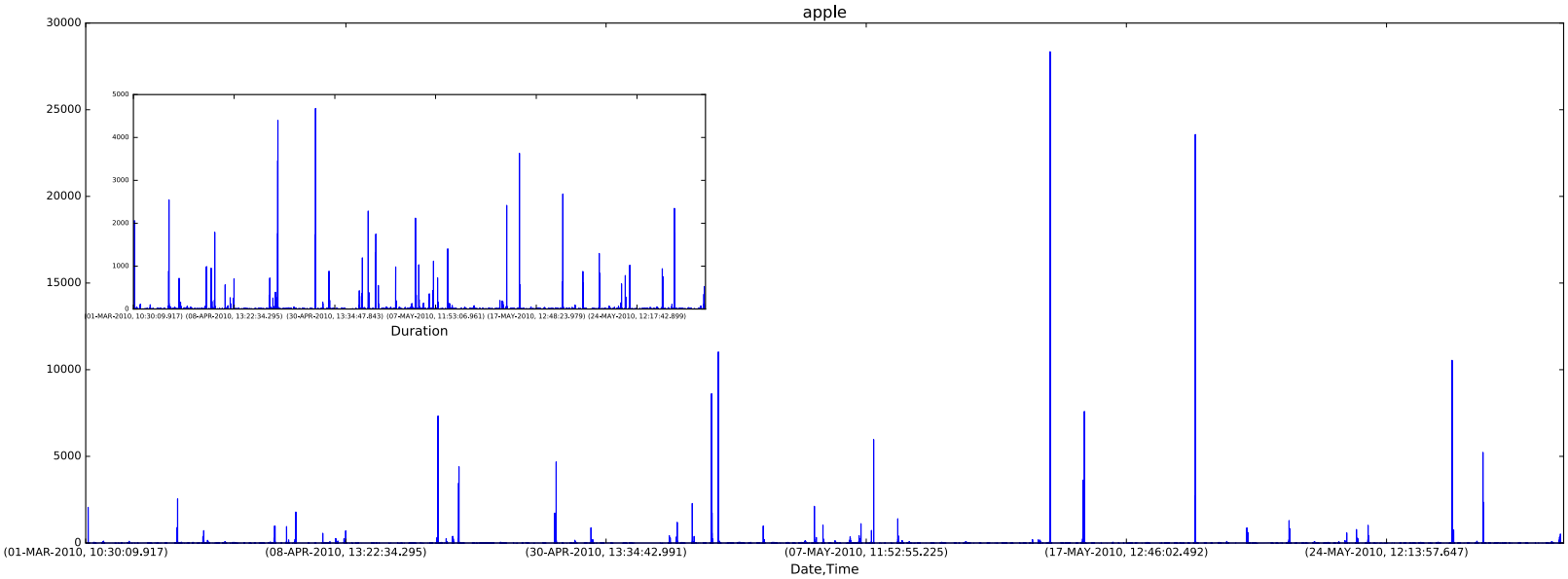
- As far as I know, the curse of dimensionality is an unsolved problem in the literature.
- Most applied researchers attempted to get around it rather than solving it.
- But this will only get worse.... why?
- GARCH type models do pretty well with data at daily frequency but now we have intra-daily data.
- So we want more.

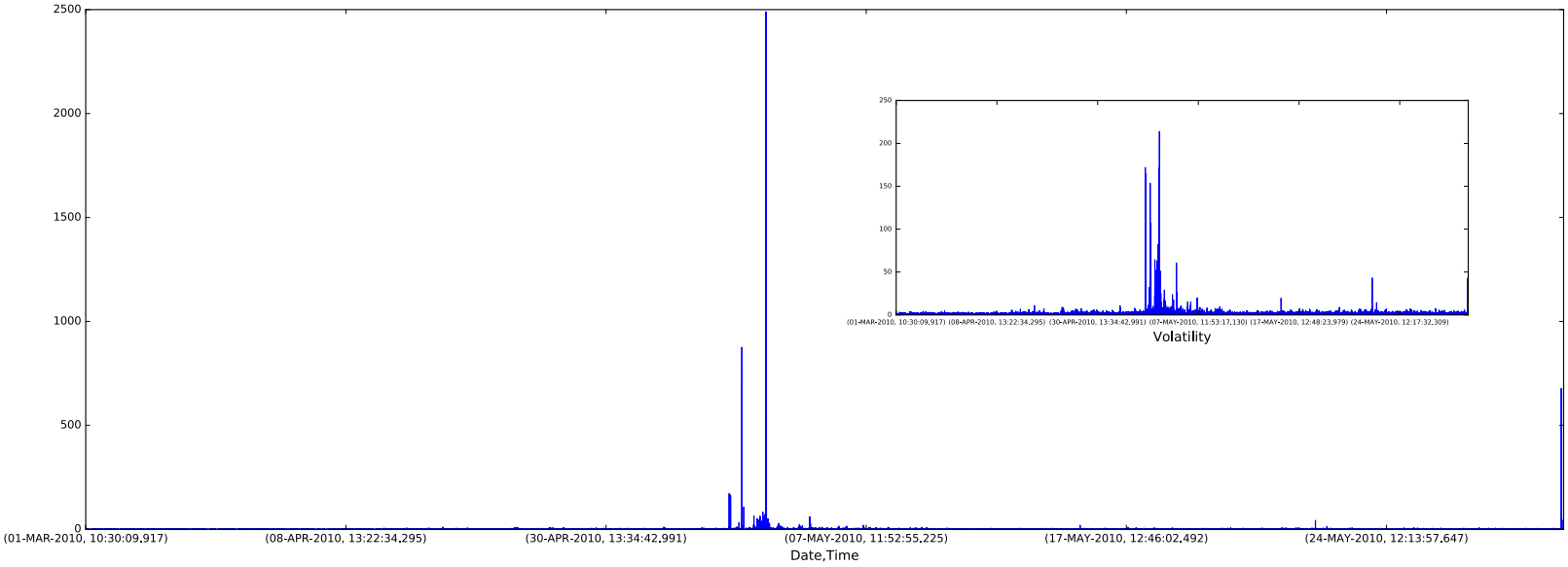
Being Greedy

- We have tick-by-tick data.
- 8GB data per stock with around 27 variables such as bid, ask, volume... etc (around 100 million observations).
- This means we can perhaps (i) predict when the next price change will happen and (ii) how much it will change, based on past history.

Loading Clip


anz
800
700
600
500
400
300
200
100
0
(01-MAR-2010, 10:30:41.166)
(25-MAR-2010, 14:47:37.602)
(29-APR-2010, 11:10:45.811)
(14-MAY-2010, 11:47:27.872)
(26-MAY-2010, 14:59:34.795)
Date,Time
100
80
60
40
20
0
(01-MAR-2010, 10:30:41.166)
(25-MAR-2010, 15:08:42.535)
(29-APR-2010, 11:14:06.069)
(14-MAY-2010, 12:32:27.555)
(26-MAY-2010, 15:22:33.756)
Duration
18
16
14
12
10
8
6
4
2
0
(01-MAR-2010, 10:30:41.166)
(25-MAR-2010, 14:47:37.602)
(29-APR-2010, 11:10:45.811)
(14-MAY-2010, 11:47:27.872)
(26-MAY-2010, 14:59:34.795)
Volatility


apple
30000
25000
20000
15000
10000
5000
0
5000
4000
3000
2000
1000
0
(01-MAR-2010, 10:30:09.917) (08-APR-2010, 13:22:34.295) (30-APR-2010, 13:34:47.843) (07-MAY-2010, 11:53:06.961) (17-MAY-2010, 12:48:23.979) (24-MAY-2010, 12:17:42.899)
Duration
(01-MAR-2010, 10:30:09.917)
(08-APR-2010, 13:22:34.295)
(30-APR-2010, 13:34:42.991)
(07-MAY-2010, 11:52:55.225)
(17-MAY-2010, 12:46:02.492)
(24-MAY-2010, 12:13:57.647)
Date,Time


2500
2000
1500
1000
500
0
250
200
150
100
50
0
(01-MAR-2010, 10:30:09,917) (08-APR-2010, 13:22:34,295) (30-APR-2010, 13:34:42,991) (07-MAY-2010, 11:53:17,130) (17-MAY-2010, 12:48:23,979) (24-MAY-2010, 12:17:32,309)
Volatility
(01-MAR-2010, 10:30:09.917)
(08-APR-2010, 13:22:34.295)
(30-APR-2010, 13:34:42.991)
(07-MAY-2010, 11:52:55.225)
(17-MAY-2010, 12:46:02.492)
(24-MAY-2010, 12:13:57.647)
Date,Time

Joint Distribution between Duration and Return

Loading Clip

We are too relax...

Loading Clip

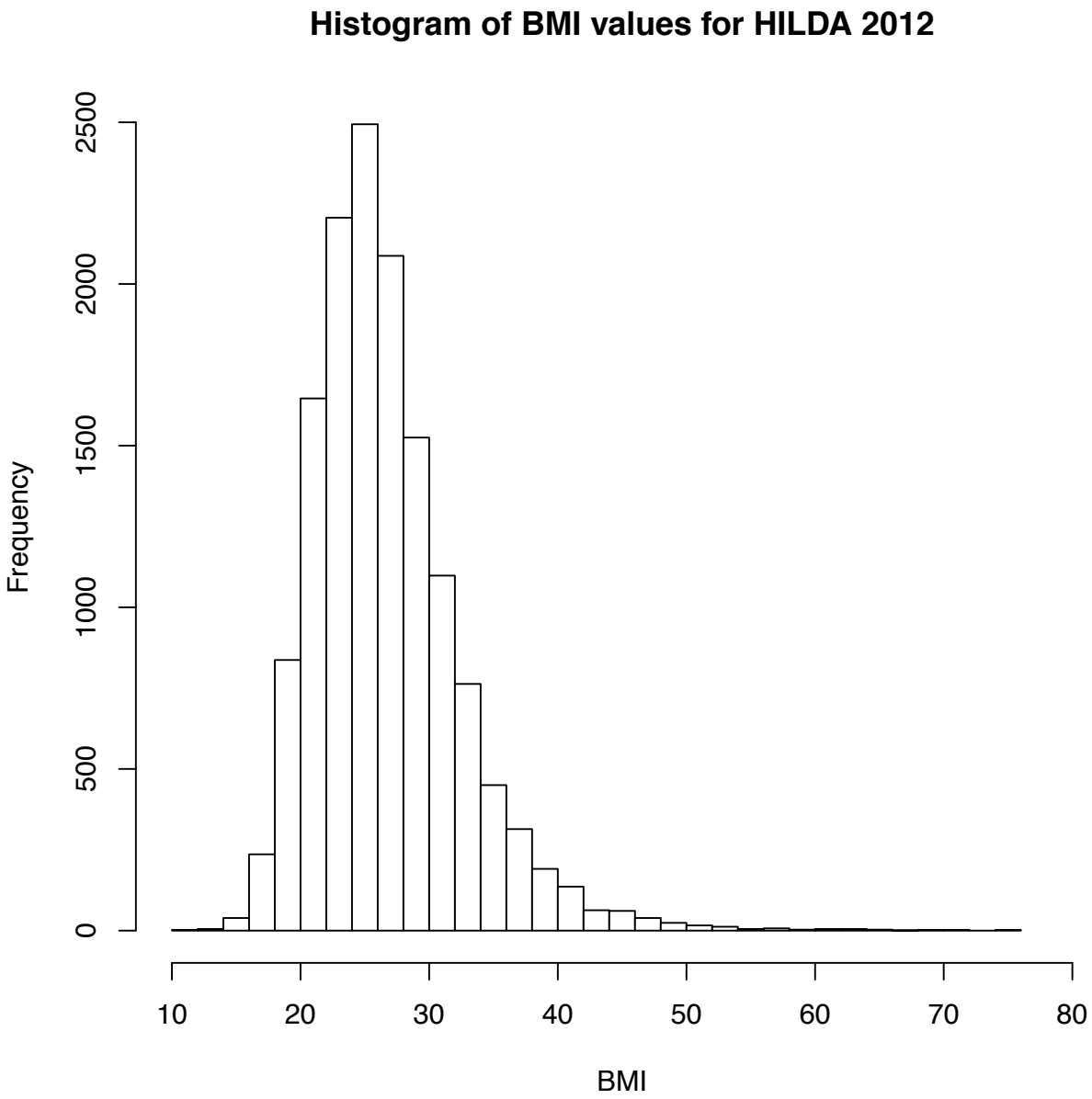
It is not just about money...

- Household, Income and Labour Dynamics in Australia (HILDA) Survey
  - National level panel data set 2001-2012
  - Approximately 10,000 households containing 26,000 individuals
  - Contains information on health and subjective well-being, economic well-being, employment status, family attributes, level of education etc.

It is not just about money...

- Household, Income and Labour Dynamics in Australia (HILDA) Survey
  - National level panel data set 2001-2012
  - Approximately 10,000 households containing 26,000 individuals
  - Contains information on health and subjective well-being, economic well-being, employment status, family attributes, level of education etc.
- 2006 - Individual height and weight.
- BMI and Obesity.

BMI Distribution



Results

$$f(y|\mathbf{X_i}) = Q_i^{-1} \exp\left[(\beta \mathbf{X}_i)^T \ \mathbf{y}\right]$$

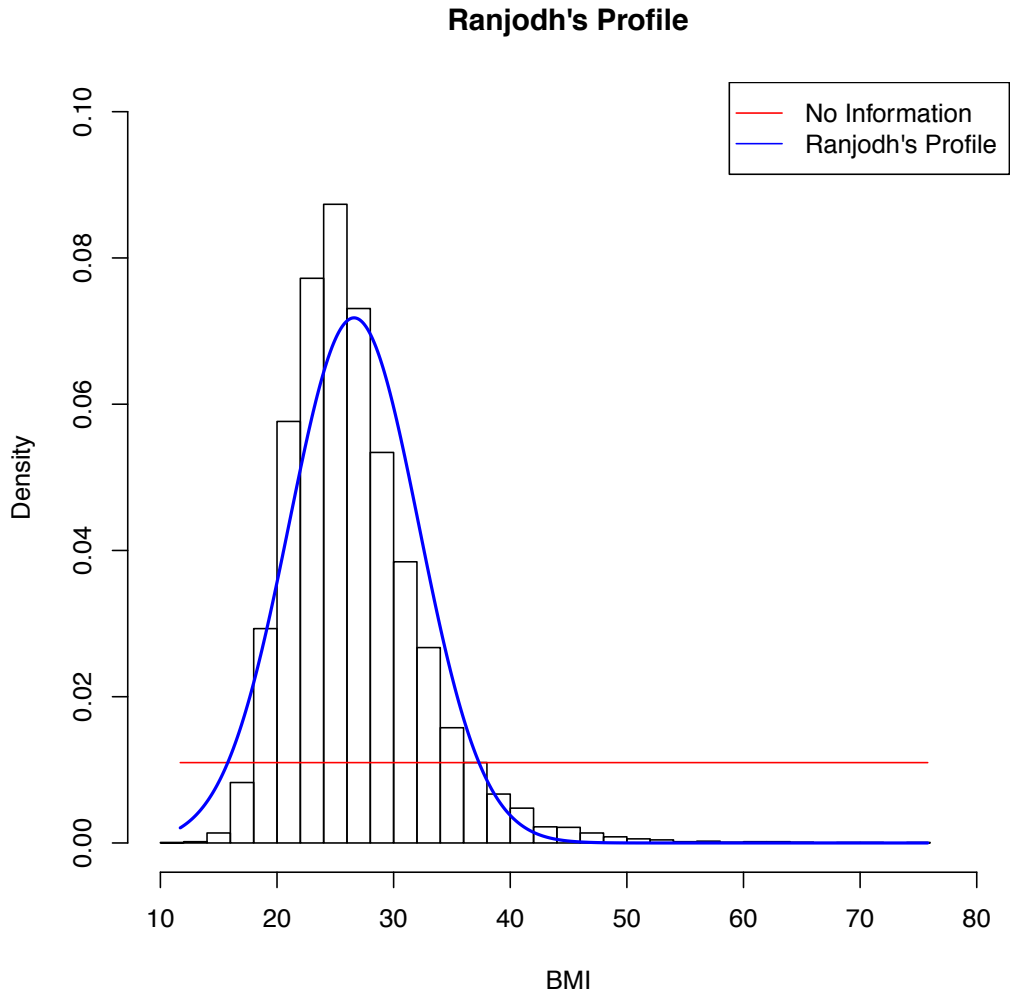
- Variable Selection
  - *Zhang and Wang (2004)* BMI - Gender, Socio-economic inequality, Age and Ethnicity
  - *Houle 2014* BMI - Gender, Ethnicity and Education
  - *Bottai et al. (2014)* BMI - Physical activity

Results - Model

$$
\begin{aligned}
\lambda_1 &= \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \\
\lambda_2 &= \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 \\
\lambda_3 &= \beta_7 X_7 \\
\lambda_4 &= \beta_8 X_8 .
\end{aligned}
$$

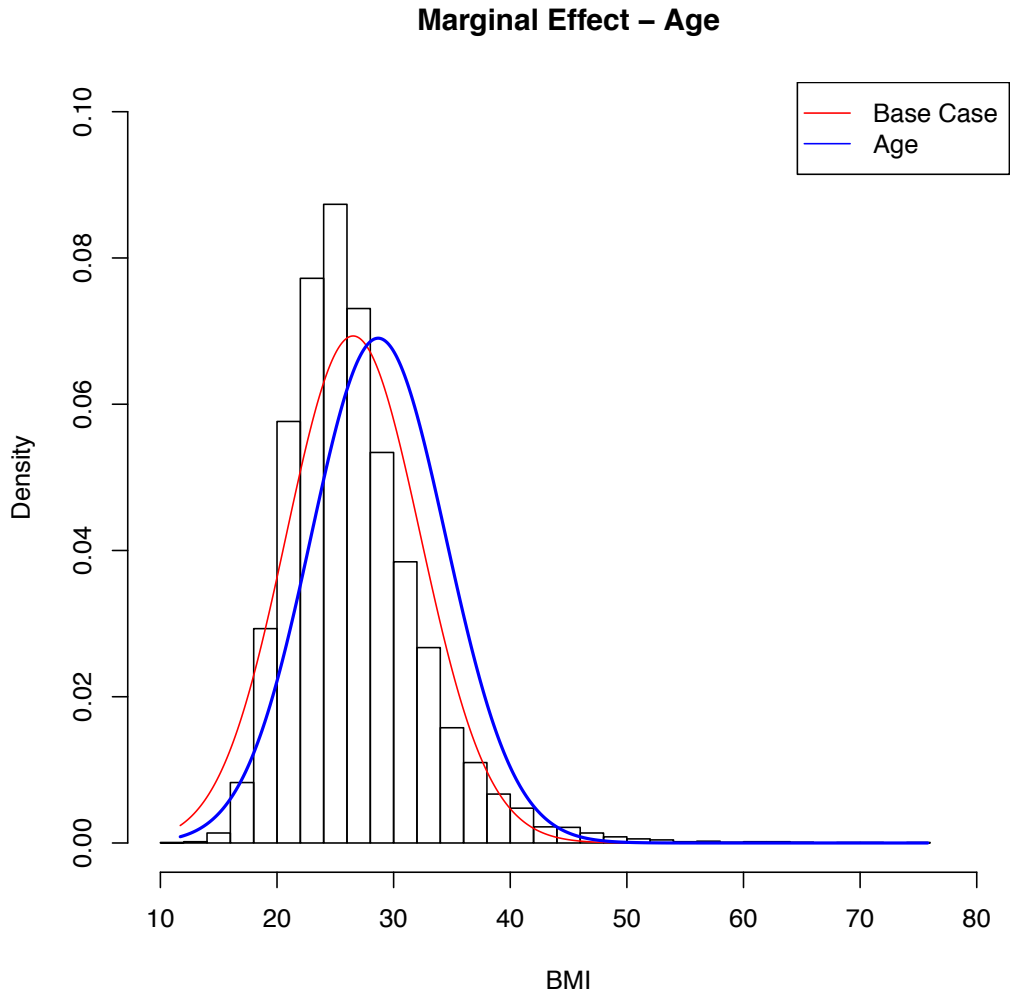
Covariate	Estimate
(Log of) Age (X_1)	0.238735
Active (X_2)	0.100000
Married (X_3)	0.000012
Male (X_4)	-0.000036
(Log of) Household Income (X_5)	-0.001385
Number of Children (X_6)	-0.000557
Employment (X_7)	0.000017
University Education (X_8)	-0.000003

Model - Profiling



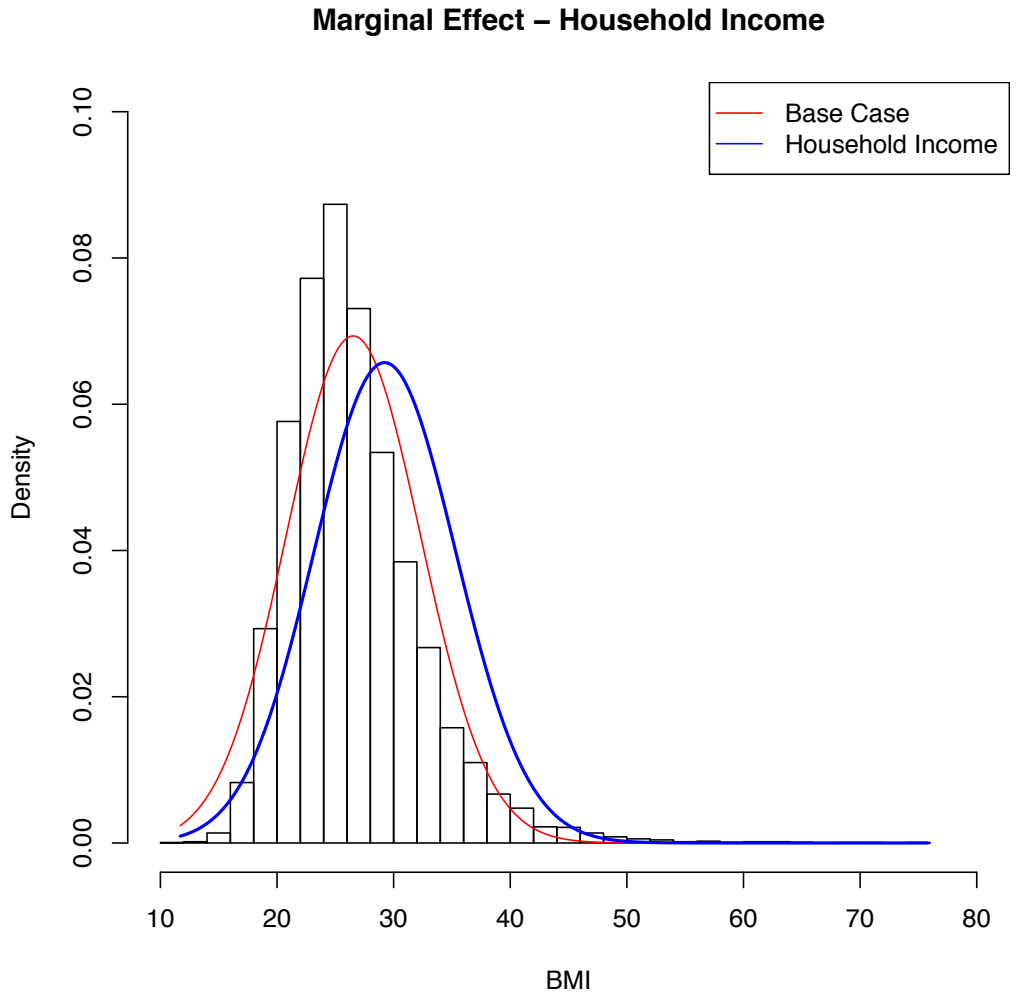
Underweight	Normal	Overweight	Obese
6.14%	32.67%	34.31%	26.89%

Marginal Effect - Age



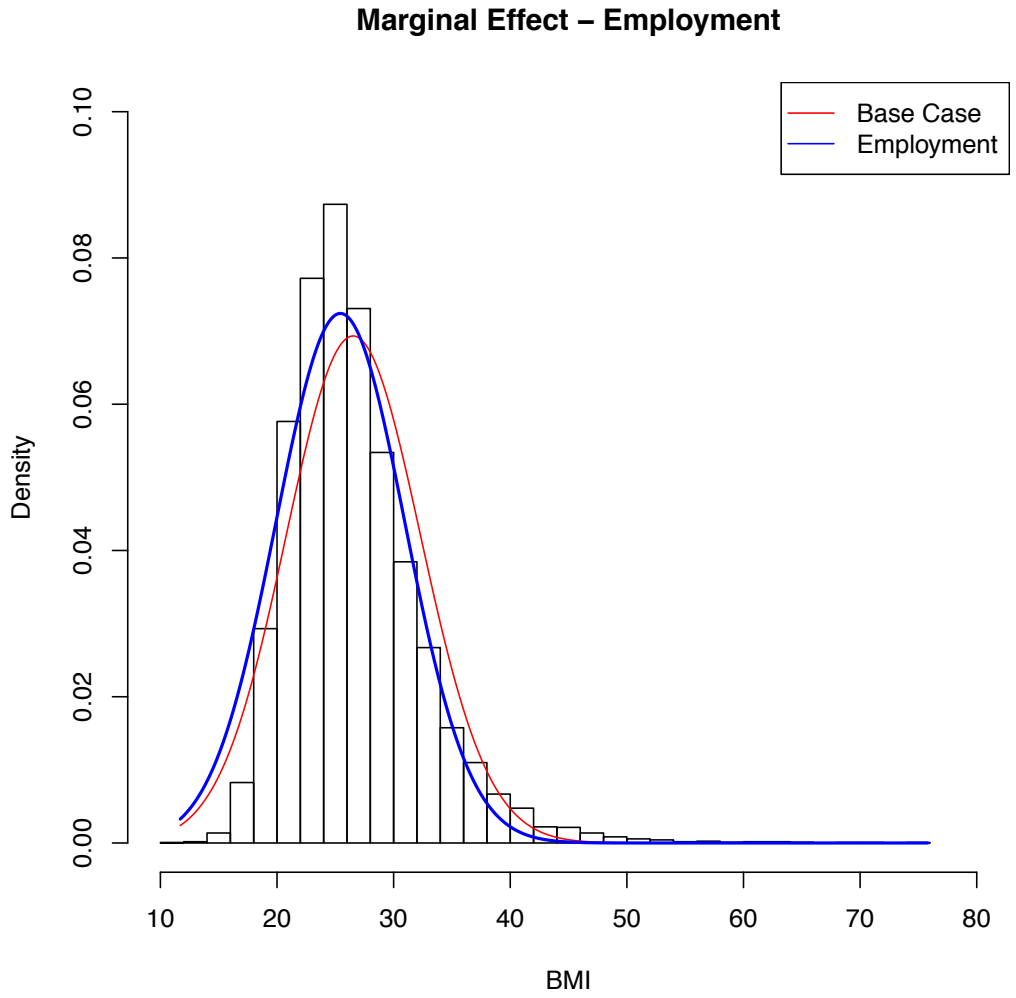
	Underweight	Normal	Overweight	Obese
Base Case	6.66%	32.56%	33.17%	27.62%
Age	3.09%	22.89%	32.82%	41.20%

Marginal Effect - Household Income



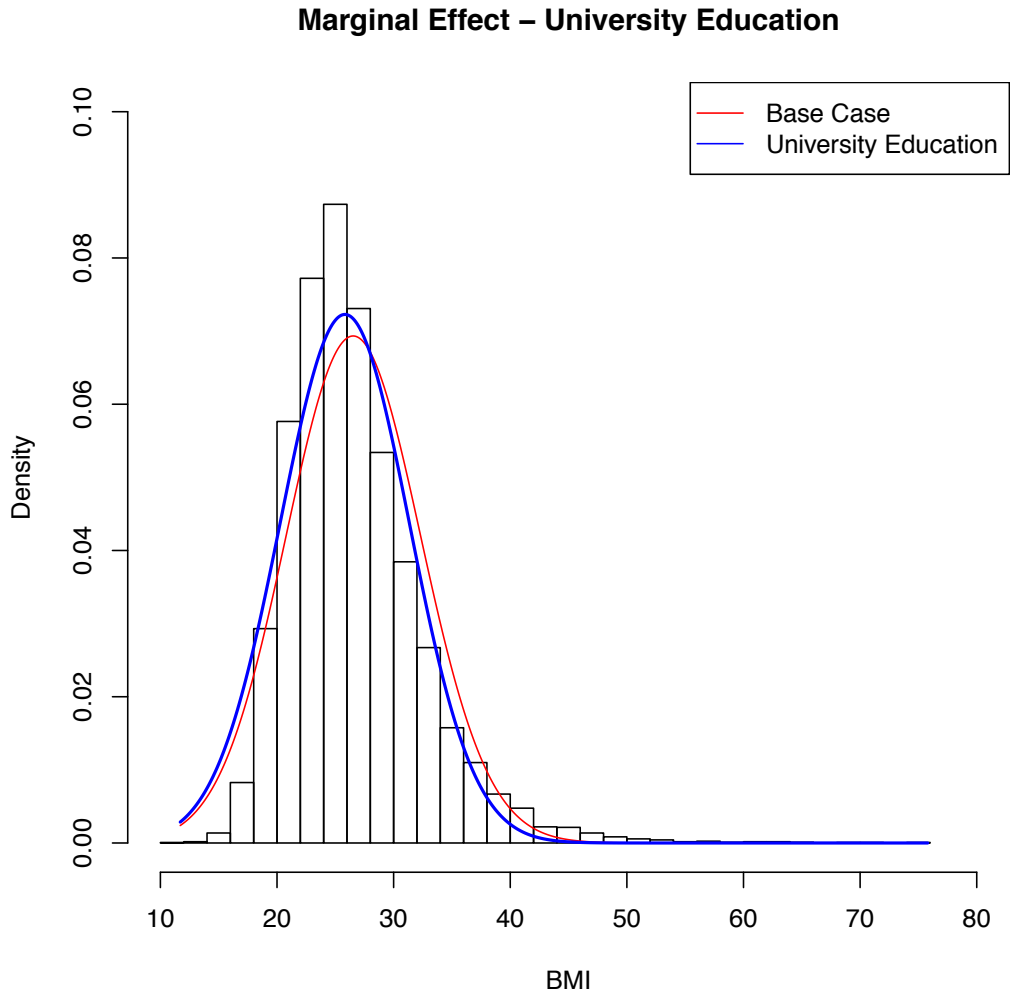
	Underweight	Normal	Overweight	Obese
Base Case	6.66%	32.56%	33.17%	27.62%
Household Income	3.05%	20.96%	30.74%	45.25%

Marginal Effect - Employment



	Underweight	Normal	Overweight	Obese
Base Case	6.66%	32.56%	33.17%	27.62%
Employment	8.76%	38.02%	32.79%	20.43%

BMI Distribution



	Underweight	Normal	Overweight	Obese
Base Case	6.66%	32.56%	33.17%	27.62%
University Education	7.87%	36.34%	33.47%	22.32%

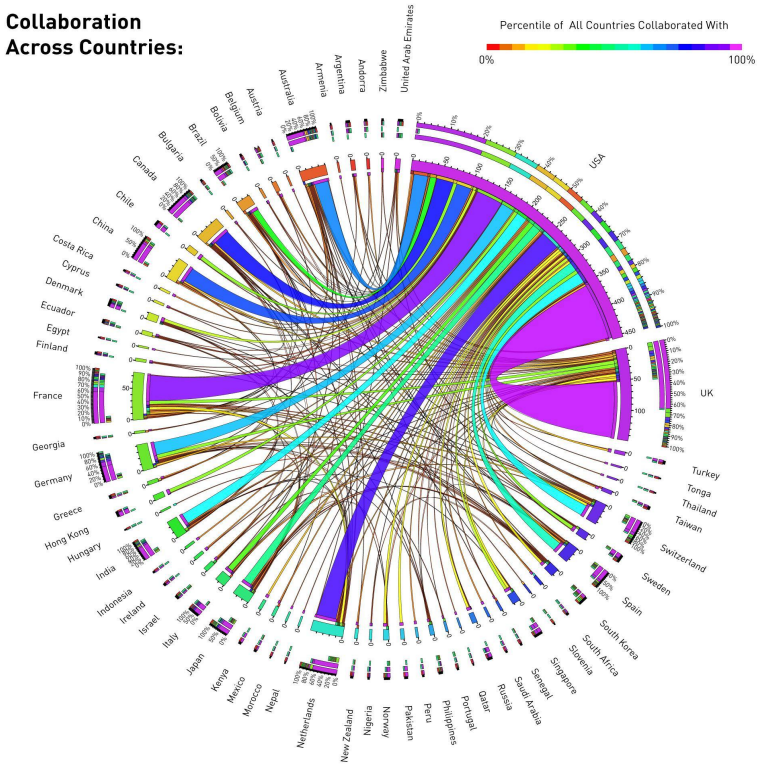
Conditional Distribution

- Recent developments lead to conditional distribution with higher moments.
- Various theoretical and computation issues arise.
- Existence of moments and maximisation with highly nonlinear function.

Problem 3: Black Cats in a Dark Room

- Searching for structural break(s).
- Evaluating policy impacts can be formulated as a structural break problem with (unknown) breakpoint. Why unknown?
- Relationships between variables may change in anticipation to the policy change.
- An example is the gravity model of trade and the introduction of Euro on 1 January 1999 (and of course the Brexit may renew this line of research).

Gravity Model of Trade



$$e_{ijt} = \beta_1 GDP_{it} + \beta_2 GDP_{jt} + \beta_3 dist_{ijt} + \underbrace{\mathbf{x}_{ijt}'\boldsymbol{\beta}}_{\text{other observable stuff}} + u_{ijt}$$

$$u_{ijt} = \alpha_{ij} + \gamma_{it} + \lambda_{jt} + \varepsilon_{ijt}$$

Search through them all....

- Is $(\boldsymbol{\beta}', \beta_1, \beta_2, \beta_3)'$ the same before and after the introduction of Euro? If not, when did the change happen?
- Some empirical evidence suggested that economies made adjustment before the official introduction. This means the break point may be earlier than 1999.
- How to look for these break(s)?

Search through them all....

- Is $(\boldsymbol{\beta}', \beta_1, \beta_2, \beta_3)'$ the same before and after the introduction of Euro? If not, when did the change happen?
- Some empirical evidence suggested that economies made adjustment before the official introduction. This means the break point may be earlier than 1999.
- How to look for these break(s)?

- Search through them....Andrews (1993)

Some Empirical Results

More shameless self promotion:

Table 9: euro's trade effect

Estimator(s) in S_i^v	Presumed/Estimated break date	Value of Z-statistic	p-value
	Known common break		
$S_i^v(\widehat{\Theta}_i)$	1998 Q1	3.45	<.01***
$S_i^v(\widehat{\Theta}_i, \widehat{\Sigma}_i)$	1998 Q1	3.56	<.01***
$S_i^v(\widehat{\Theta}_i, \bar{\mathbf{M}}_{\mathbf{W}})$	1998 Q1	1.23	.219
$S_i^v(\widehat{\Theta}_i, \widehat{\Sigma}_i, \bar{\mathbf{M}}_{\mathbf{W}})$	1998 Q1	2.12	.034**
	Unknown individual break		
$S_i^v(\widehat{\Theta}_i)$	1997 Q3	3.08	.002***
$S_i^v(\widehat{\Theta}_i, \widehat{\Sigma}_i)$	1997 Q2	17.33	<.01***
$S_i^v(\widehat{\Theta}_i, \bar{\mathbf{M}}_{\mathbf{W}})$	1998 Q1	2.69	.007***
$S_i^v(\widehat{\Theta}_i, \widehat{\Sigma}_i, \bar{\mathbf{M}}_{\mathbf{W}})$	1997 Q3	13.40	<.01***

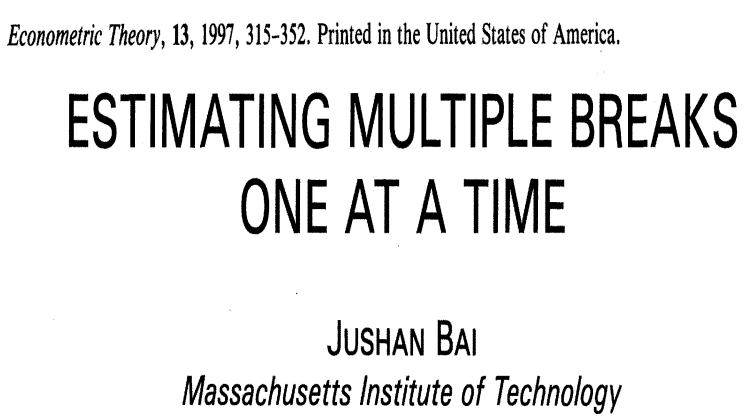
Note: $S_i^v(\widehat{\Theta}_i, \widehat{\Sigma}_i)$ uses the estimated variance-covariance matrix as in equation (8) and $S_i^v(\widehat{\Theta}_i, \widehat{\Sigma}_i, \bar{\mathbf{M}}_{\mathbf{W}})$ uses the CCE estimator constructed as in section 4 with the aforementioned variance-covariance matrix. The average of the estimated break dates is calculated as $\widehat{\bar{T}} = \frac{1}{N}\sum_{i=1}^{N} \widehat{T}_{0i}$. The break search interval starts in 1992 Q1 until 2004 Q4. */**/*** indicate 10%/5%/1% level of significance.

But there may be more than one break...

- Andrews' results work well when there is only one break.
- What happened with multiple breaks? Searching through all possible combinations is not practically feasible. Not to mention we don't know how many break points in general.

But there may be more than one break...

- Andrews' results work well when there is only one break.
- What happened with multiple breaks? Searching through all possible combinations is not practically feasible. Not to mention we don't know how many break points in general.


Econometric Theory, 13, 1997, 315-352. Printed in the United States of America.

ESTIMATING MULTIPLE BREAKS ONE AT A TIME

JUSHAN BAI
Massachusetts Institute of Technology


Journal of Statistical Planning and Inference 141 (2011) 3367–3381

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Review

Detection of structural breaks in a time-varying heteroskedastic regression model

Cathy W.S. Chen [a,*], Richard Gerlach [b], Feng-Chi Liu [a]

[a] *Department of Statistics, Feng Chia University, Taiwan*
[b] *Discipline of Operations Management and Econometrics, University of Sydney, Australia*

Slow progress... an open area with many more questions.

The Unobserved

$$e_{ijt} = \beta_1 GDP_{it} + \beta_2 GDP_{jt} + \beta_3 dist_{ijt} + \underbrace{\mathbf{x}_{ijt}'\boldsymbol{\beta}}_{\text{other observable stuff}} + u_{ijt}$$

$$u_{ijt} = \alpha_{ij} + \gamma_{it} + \lambda_{jt} + \varepsilon_{ijt}$$

- No reason to believe that the unobserved is uncorrelated with the observable regressors. In fact, quite the opposite.
- If all we care about are the $\boldsymbol{\beta}$'s then we can use the *dummy variables* approach. Works well when it is only (i, t) on the index space.

The Unobserved

$$e_{ijt} = \beta_1 GDP_{it} + \beta_2 GDP_{jt} + \beta_3 dist_{ijt} + \underbrace{\mathbf{x}_{ijt}'\beta}_{\text{other observable stuff}} + u_{ijt}$$

$$u_{ijt} = \alpha_{ij} + \gamma_{it} + \lambda_{jt} + \varepsilon_{ijt}$$

- No reason to believe that the unobserved is uncorrelated with the observable regressors. In fact, quite the opposite.
- If all we care about are the β's then we can use the *dummy variables* approach. Works well when it is only (i, t) on the index space.
- The number of dummy variables (the number parameters) increase as the number of countries increases in the (i, j, t)!
- $N(N-1)$ trade pairs. $N = 33$ with $T = 40$ we have $33 \times 32 = 1056$ dummy variables in the time invariant case. This makes the calculation of the OLS estimator problematic $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$.
- Need something else.

Go beyond three dimensions

- More generally, econometricians are moving towards multidimensional data.