Machine Learning in Finance

Overview

Overview

FORDHAM

THE JESUIT UNIVERSITY OF NEW YORK

Gabelli School of Business

Neural Net Specificaton for Quantile Regression, Part I

Neural Net Specificaton for Quantile Regression, Part I

- Quantile regression is a regression based on minimizing the sum of the absolute deviations of the target observations from quantile τ. For median regression, of course, $\tau = .05$.
- We took the negative of the weighted returns (by market capitalization for 19 of the 20 GSIB's banks and regressed them on on the returns of the other banks plus nine controls.
- We took the negative so that the quantile at $\tau = .95$ represents losses at or below the 5% range. We also did a quantile regression for the media, with $\tau = .05$.
- We define $\triangle Covar$ for bank i: $\triangle Covar_{t,i} = \hat{y}_{95,t} - \hat{y}_{50,t}$
- We did quantile regressions for both linear specification as well as with polynomial expansions and time-varying estimation

Neural Net Specificaton for Quantile Regression, Part I

- With neural nets we normalize all of the variables, both targets and covariates, with the same squashier function:

$$x^* = \frac{2x}{\max(x) - \min(x)} - \frac{\min(x) + \max(x)}{\max(x) - \min(x)}$$

- Each neuron is then transformed by either a logsigmoid or a hyberbolic tangent (aka tansig function):

$$n_{j,t}^1 = \omega_{j,0}^1 + \sum_{i=1}^{k} \omega_{j,i}^1 x_{i,t}^*, j = 1, ...j^*$$

$$Logsigmoid : N_{i,t}^1 = \frac{1}{1 + e^{-n_{i,t}^1}}$$

$$Tansig : N_{i,t}^1 = \frac{2}{1 + e^{-2n_{i,t}^1}} - 1$$

Neural Net Specificaton for Quantile Regression, Part I

- In a simple or shallow neural net with one hidden layer, and m=1,...m* target variables, each target in the output layer, is a linear function of the j^* neurons in the first hidden layer:

$$y_i^m = \omega_0^{o,m} + \sum_{j=1}^{j^*} \omega_j^{o,m} n_{j,t}^{m-1}$$

- In a deep neural network, each layer in layers 2,3,...n hidden layers is a linear function of the previous hidden layer neurons
- These linear combinations of neurons in turn can be transformed by logsigmoid or tansig functions or RELU (Relative Linear United)unctions, in which the value has a lower limit of zero and an upper limited of one
- Each hidden layer also as a bias (constant term).
- You can see that there are lots of options for specifying the complexity of a neural network.

FORDHAM

THE JESUIT UNIVERSITY OF NEW YORK

Gabelli School of Business

Neural Net Specificaton for Quantile Regression, Part II

Neural Net Specificaton for Quantile Regression, Part II

- One strategy is to do a full sample estimation and select the best neural net specification on the basis of the traditional regularization criteria, the Akaike Information Criteria (AIF), Schwartz Information Criterion (BIF) and the Hanan-Quinn Information Criterion
- Basic idea is to select the network structure on the basis of the Sum of Squired Errors (SSE), the number of observations (n) and the number of parameters. The three criteria are ranking criteria, choose the model with the lowest value.

$$AIF = log(SSE)/n + 2k/n$$

$$BIF = log(SSE)/n + k[log(n)]/n$$

$$HQIF = log(SSE)/n + k[log(log(n))]/n$$

Neural Net Specificaton for Quantile Regression, Part II

- Note that the values of the three criteria have no particular meaning in themselves. So having these values for one regression model is not particularly helpful. They are used for comparing outcomes of different regressions.
- They are useful for **ordering** the outcomes of alternative regressions models based on reducing the squared errors with the smallest number of parameters.
- The AIC criterion penalizes the number of parameters by only a factor of $2/n$, so is judged too weak, while the BIF penalizes by a factor of $\log(n)/n$, perhaps too strong a penalty.
- The HQIF stands in the middle with a penalty factor for too many parameters of $\log(\log(n))/n$.

Neural Net Specificaton for Quantile Regression, Part II

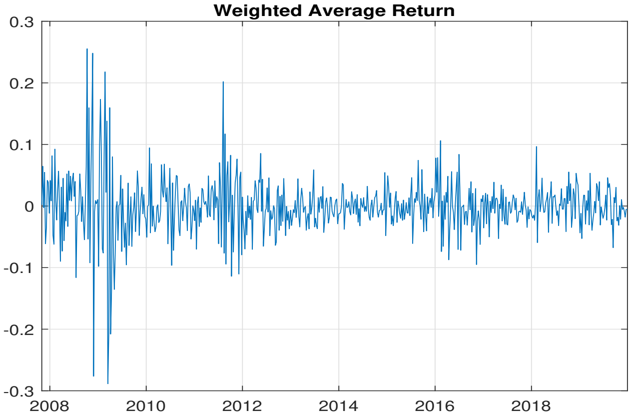
- Note that the values of the three criteria have no particular meaning in themselves. So having these values for one regression model is not particularly helpful. They are used for comparing outcomes of different regressions.
- They are useful for **ordering** the outcomes of alternative regressions models based on reducing the squared errors with the smallest number of parameters.
- The AIC criterion penalizes the number of parameters by only a factor of $2/n$, so is judged too weak, while the BIF penalizes by a factor of $\log(n)/n$, perhaps too strong a penalty.
- The HQIF stands in the middle with a penalty factor for too many parameters of $\log(\log(n))/n$.

FORDHAM

THE JESUIT UNIVERSITY OF NEW YORK

Gabelli School of Business

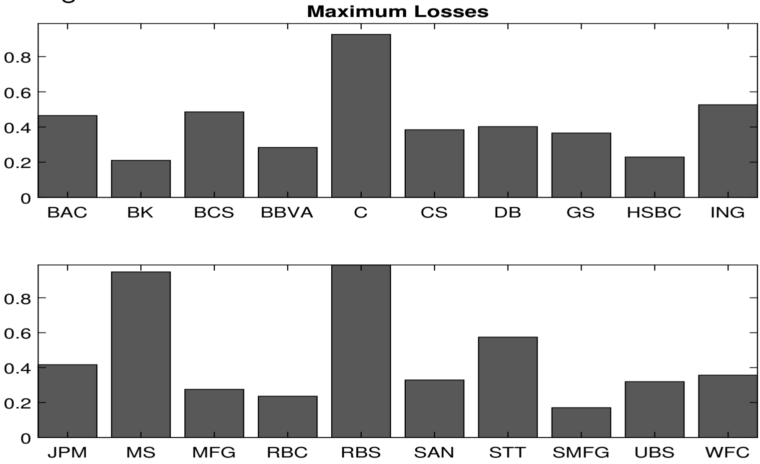
Quantile Regression with Neural Networks, Part I



- We see that most of the large values take place at the beginning of the sample.

Quantile Regression with Neural Networks, Part I

- Maximum losses in market returns over the sample period. We see the C, MS, and RBS lead the pack in taking the largest losses.



FORDHAM

THE JESUIT UNIVERSITY OF NEW YORK

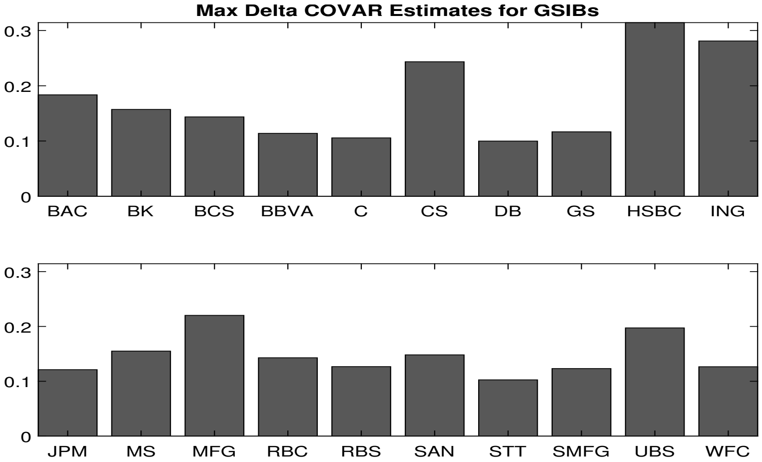
Gabelli School of Business

Quantile Regression with Neural Networks, Part II

:Criterion:	Neural Net Structure: 3	20	[20-10-5]	[20-15-10-5]
AIC	**0.231**	1.516	3.091	4.760
BIC	**0.750**	4.897	9.978	15.365
HQIF	**0.215**	1.413	2.881	4.438

- We see that the simplest single-hidden layer network with 3 tansig neurons outperforms the more complex nets based on all three information criteria.
- Logsigmoid transformations yield practically identical results.

Quantile Regression with Neural Networks, Part II



- We see that HSBC leads the pack, following by ING, CS, and MFG, for maximal negative impact values on the financial system.
- We note that the banks with the maximum impact on systemic risk are not identical to the banks which experienced

FORDHAM
THE JESUIT UNIVERSITY OF NEW YORK

Gabelli School of Business

Quantile Regression with Neural Networks, Part III

Quantile Rregression with Neural Networks, Part III

- Dates of maximum contagion effects by GSIB bank. HSBC and ING, the two largest transmitters, generate their effects later, several months later, than the others.

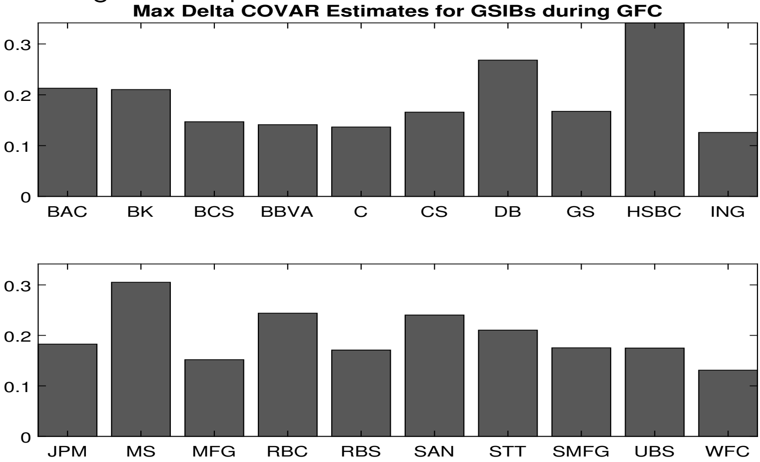
Date	GSIB Bank	Date	GSIB Bank
10/24/2008	BAC	10/17/2008	JPM
10/31/2008	BK	1/26/2009	MS
9/26/2008	BCS	11/28/2008	MFG
9/26/2008	BBVA	10/17/2008	RBC
9/26/2008	C	1/26/2009	RBS
11/28/2008	CS	11/21/2008	SAN
5/25/2009	DB	12/12/2008	STT
2/23/2009	GS	12/19/2008	SMFG
4/6/2009	HSBC	10/10/2008	UBS
2/23/2009	**ING**	9/26/2008	WFC

FORDHAM | Gabelli School of Business

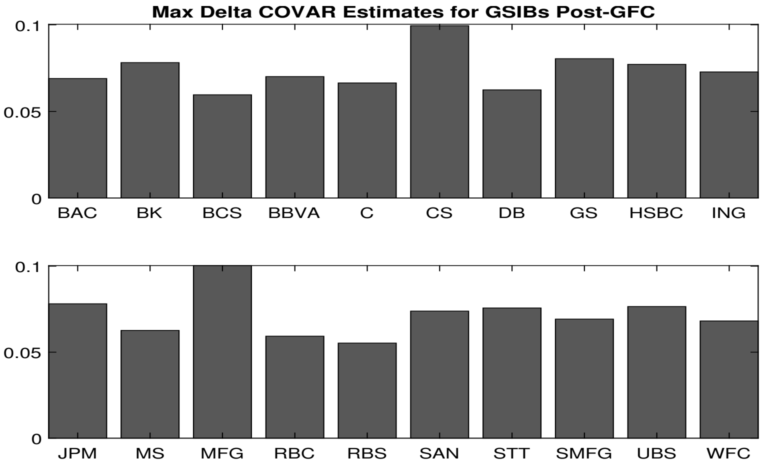
THE JESUIT UNIVERSITY OF NEW YORK

Splitting the Sample: Before and After the GFC

- We see that the HSBC while DB and MS are not far behind during the GFC period.



Splitting the Sample: Before and After the GFC



- In the post-GFC period, the $\triangle$Covar effects are smaller that the max values of the GFC period.
- The leaders are CS and MFG but HSBC is not far behind

FORDHAM | Gabelli School of Business

THE JESUIT UNIVERSITY OF NEW YORK

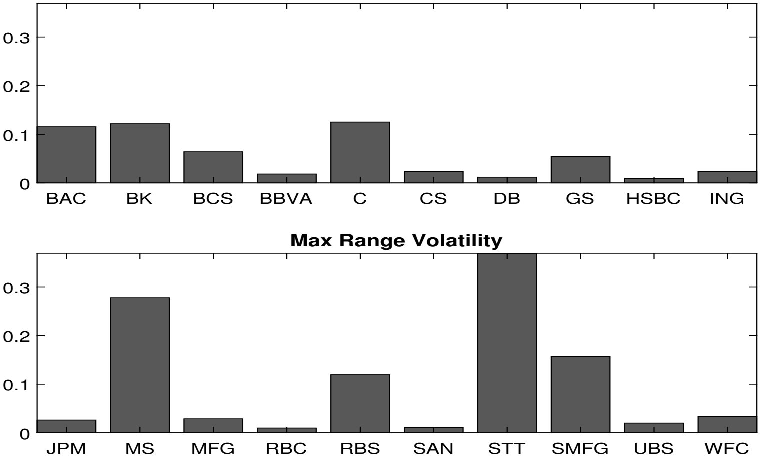
Comparison: Neural Net VAR-X Specification, Part I

Comparison: Neural Net VAR-X Specification, Part I

- We model daily range volatility of each bank as a function of five lags of the full set of 20 banks as well as a set of nine control variables for various measures of aggregate risk.
- The main difference is that the 109 input variables (20 banks with 5 lags, plus 9 controls) are mapped into neurons in a second hidden layer.
- For deep learning networks, the first hidden layer is mapped into a second hidden layer.
- We will assess shallow networks as well as deep-learning networks.

Comparison: Neural Net VAR-X Specification

- We see the max range volatility values of the 20 banks over the sample period. STT, MS amd SMFG led the pack.



Comparison: Neural Net VAR-X Specification

- Most of the banks experienced their maximum range volatility values near the time of GFC,with the exception of JPM, which experienced its maximum range volatility in 2015.

Date	GSIB Bank	Date	GSIB Bank
2/20/2009	BAC	**8/24/2015**	**JPM**
9/18/2008	BK	9/18/2008	MS
1/16/2009	BCS	10/29/2008	MFG
5/10/2010	BBVA	10/9/2008	RBC
11/21/2008	C	10/13/2008	RBS
10/8/2008	CS	10/28/2008	SAN
9/18/2008	DB	9/18/2008	STT
9/18/2008	GS	10/7/2008	SMFG
3/3/2009	HSBC	9/17/2008	UBS
10/10/2008	ING	2/20/2009	WFC

FORDHAM
THE JESUIT UNIVERSITY OF NEW YORK
Gabelli School of Business

Comparison: Neural Net VAR-X Specification, Part II

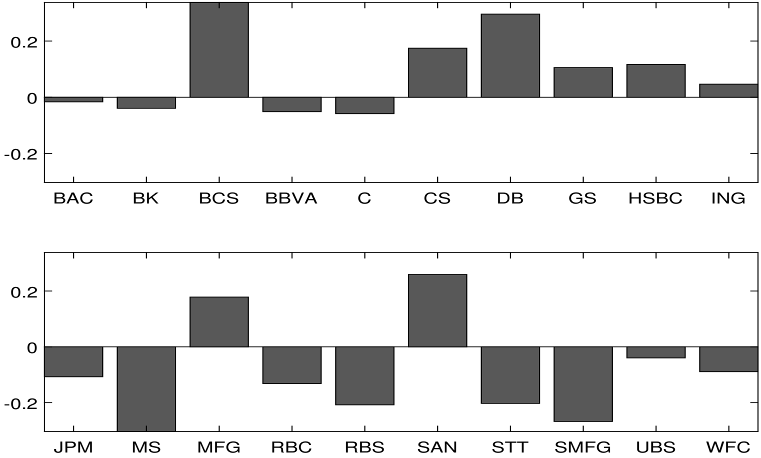
Comparison: Neural Net VAR-X Specification, Part II

- We see by the Hannan-Quinn Information Criterion that the relatively shallow neural net does slightly better than the more extensive networks.
- Put another way, the more elaborate networks do not add significantly more explanatory power than the shallow single hidden-layer network.

	Network Structure:			
Criterion:	20	[20-10]	[20-15-10]	[20-15-10-5]
HQIF	2.269	2.2768	2.4567	2.4262

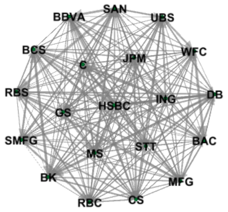
Comparison: Neural Net VAR-X Specification, Part II

- We see that BCS, DB and SAN are the largest net transmitters. MS and SMFG are the largest net receptors.



Comparison: Neural Net VAR-X Specification, Part II

- Network connectedness. Note that HSBC is in the dead center of the network.



FORDHAM

THE JESUIT UNIVERSITY OF NEW YORK

Gabelli School of Business