FUNDAÇÃO GETULIO VARGASESCOLA BRASILEIRA DE ECONOMIA E FINANÇASCURSO DE GRADUAÇÃO EM CIÊNCIAS ECONÔMICAS

ECONOMETRIA II

Ánalise no preço de CDS dos principais Bancos de Investimentos.

Diogo David Sánchez Lima

Rio de Janeiro

2022

1. Introdução

O mercado de capitais costuma a criar mecanismos financeiros tanto para aumentar a exposição dos seus agentes a determinados investimentos, quanto para "hedgear" certos investimentos.

Nesse trabalho vamos investigar a relação entre séries temporais e o preço de um hedge muito conhecido, o *Credit Default Swap* (CDS), que ficou muito conhecido fora do mercado financeiro graças ao filme *The Big Short* (2015) que retrata a Crise do Subprime em 2008, e como o *Michael Burry* gestor de um fundos de investimento, consegue aumentar sua rentabilidade comprando CDS para assegurar default no mercado imbiliário. Quando a probabilidade de não-pagamento da parte devedora do titulo imobiliário aumenta, ele começa a rentabilizar e chega a +400%.

Um CDS é um contrato bilateral, em que os dois lados possuem responsabilidade e direitos, a ponta compradora do CDS, possui o dever de pagar cupons periodicamente a ponta vendedora, já a ponta vendedora tem o dever de assegurar a ponta compradora a eventuais riscos de default, desde que, esses riscos de default sejam nos ativos que estão no escopo do contrato de CDS.

Ao longo do trabalho vamos começar estudando a possibilidade de um modelo de precificação simples AR 2 para CDS Bancário, depois vamos evoluir para intervalos de confiança “naturais” para o preço de de CDS, depois passamos a ver um modelo AR 2 para calcular o descolamento de um CDS contra o seu setor.

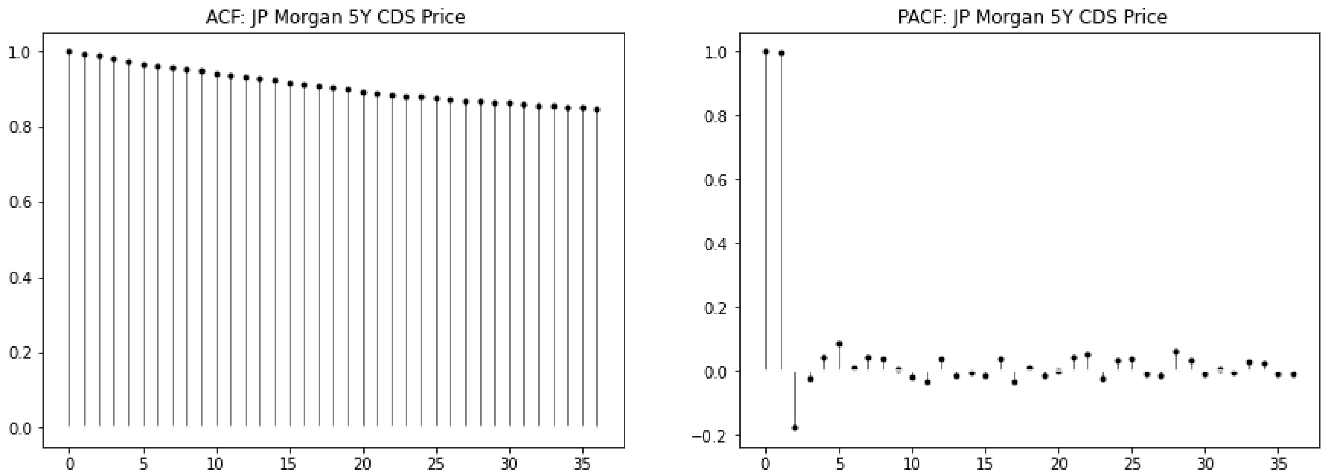
Na segunda parte vamos colocar dados fundamentalistas do balanço dos bancos em um modelo VAR e a distância entre spreads em contratos de CDS de periodos longos, e vamos utilizar o modelo para entender a primeira derivada do preço.

2. Modelo autoregressivo para preço de CDS.

Nesse trabalho vamos investigar uma categoria especifica de CDS, o CDS que assegura o risco de default de bancos de investimentos. Na nossa base de dados, temos, os CDS para JPMorgan, SocGen, MorganStanley, BankofAmerica, Citi, GoldmanSachs e DeustcheBank; dentre o periodo de Janeiro 2005 até Setembro de 2022 e em todas as durações de contrato (6 meses até 10 anos).

A seguir o gráfico com preço de um CDS de JP Morgan com duração de 5 anos, outros contratos de outros bancos ou com durações diferentes teriam formatos parecidos, com aumento de preço em periodos como 2008, 2020 e 2012.

A seguir o gráfico com ACF e PACF do ativo escolhido com exemplo:



Embora pareça que é um AR 2, vamos testar qual AR tem melhor desempenho via *Baysian Information Criterion* em um modelo.

BIC TABLE

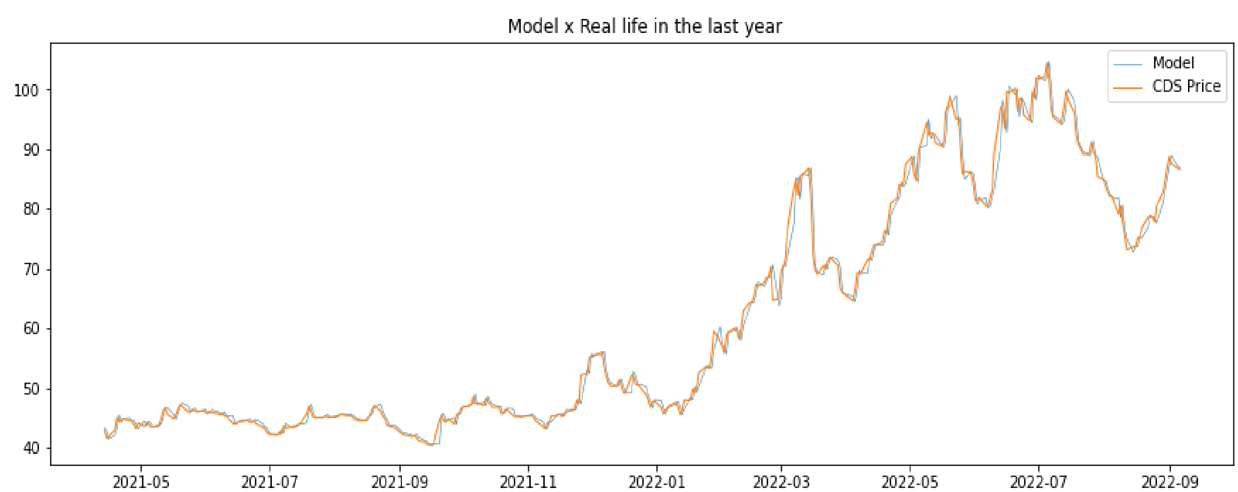
	Q = 0	**Q = 1**	**Q = 2**	**Q = 3**	**Q = 4**
P = 0	45659.0	39875.0	35647.0	32672.0	30707.0
P = 1	24577.0	24453.0	24443.0	24451.0	24426.0
P = 2	24441.0	24448.0	24451.0	24432.0	24424.0
P = 3	24447.0	24455.0	24462.0	24427.0	24430.0
P = 4	24447.0	24424.0	24460.0	24433.0	24439.0

Como podemos perceber pela tabela de BIC por número de variaveis, um AR(2) e um ARMA(4,1) parecem boas opções, vou optar por um AR(2).

Utilizando esse número de parametros para o nosso modelo autoregressivos, teremos o seguintes output para o summary.

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                      y   No. Observations:                 4599
Model:                 ARIMA(2, 0, 0)   Log Likelihood              -12203.822
Date:                Wed, 30 Nov 2022   AIC                          24415.645
Time:                        21:52:28   BIC                          24441.379
Sample:                             0   HQIC                         24424.702
                               - 4599
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const         66.3214     12.482      5.313      0.000      41.857      90.785
ar.L1          1.1696      0.006    209.768      0.000       1.159       1.181
ar.L2         -0.1757      0.005    -32.160      0.000      -0.186      -0.165
sigma2        11.8033      0.050    236.649      0.000      11.706      11.901
==============================================================================
Ljung-Box (L1) (Q):                   0.08   Jarque-Bera (JB):         494296.14
Prob(Q):                              0.78   Prob(JB):                      0.00
Heteroskedasticity (H):               0.27   Skew:                         -0.12
Prob(H) (two-sided):                  0.00   Kurtosis:                     53.79
==============================================================================
```

A seguir a diferença entre o predict do modelo versus o que foi realmente observado.



Olhando os parametros phi_1 e phi_2 do modelo pode surgir o questionamento de raiz unitária sobre esse modelo, entretando o p-valor de um teste de Dickey-Fuller é de 6.46% onde rejeitariamos a hipótese nula de raiz unitárias para niveis de significancia maiores do que 10%.

3. Modelo Autoregressivo para estimar intervalo de confiança de preços

Uma coisa que pode ser muito interessante no monitoramento de preços de CDS - tanto para entender o quanto o mercado ta vendo de risco no banco quanto o provável risco que você esta se sujeitando em ser contraparte desse banco - é conseguir traçar um intervalo de confiança para discriminar “desvios brancos” de desvios causados por fatores externos.

Dito isso, para traçar o intervalo de confiança de preços futuros com base nos preços passados iremos assumir o modelo autoregressivo utilizado anteriormente e iremos transforma-lo em um modelo VAR (1) e abrilo como somatório, que fica escrito da seguinte forma:

$$Y_t = \sum (\begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix}^i \times E_{t-i})$$

A seguir podemos escrever o valor esperado das observações futuras como:

$$Y_{t+s} = \sum (\begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix}^i (_{1,1}) \times E_{t+s-i})$$

$$Y_{t+s} = \sum_{i=0}^{i=s-1} (\begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix}^i (_{1,1}) \times E_{t+s-i}) + \sum_{i=s} (\begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix}^{i+s} (_{1,1}) \times E_{t-i})$$

$$\begin{bmatrix} Y_{t+s} \\ Y_{t+s-1} \end{bmatrix} = (\sum_{i=0}^{i=s-1} (\begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix}^i \times \begin{bmatrix} E_{t+s-i} \\ 0 \end{bmatrix})) + (\begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix}^s \times \begin{bmatrix} Y_t \\ Y_{t-1} \end{bmatrix})$$

Para facilitar as contas unimos a contante μ com o termo de erro $\mathcal{E}_t$ de forma que o termo utilizado nas equações é $E_t \sim N(\mu, \sigma^2)$, ou seja, não é um ruído branco.

Agora que dividimos essa equação no somatório de duas parcelas, vamos chamar a primeira parcela dessa equação de Parte A e estimar qual a contribuição dessa parte para o valor esperado e o seu desvio padrão.

$$\sum_{i=0}^{i=t+s-1} k * E_{t+s-i}$$

$$\sigma^2_{Y_{t+s}} = \sum_{i=0}^{i=t+s-1} k^2 * \sigma^2_E$$

$$\mathbb{E}[Y_{t+s}] = \sum_{i=0}^{i=t+s-1} k * \mu$$

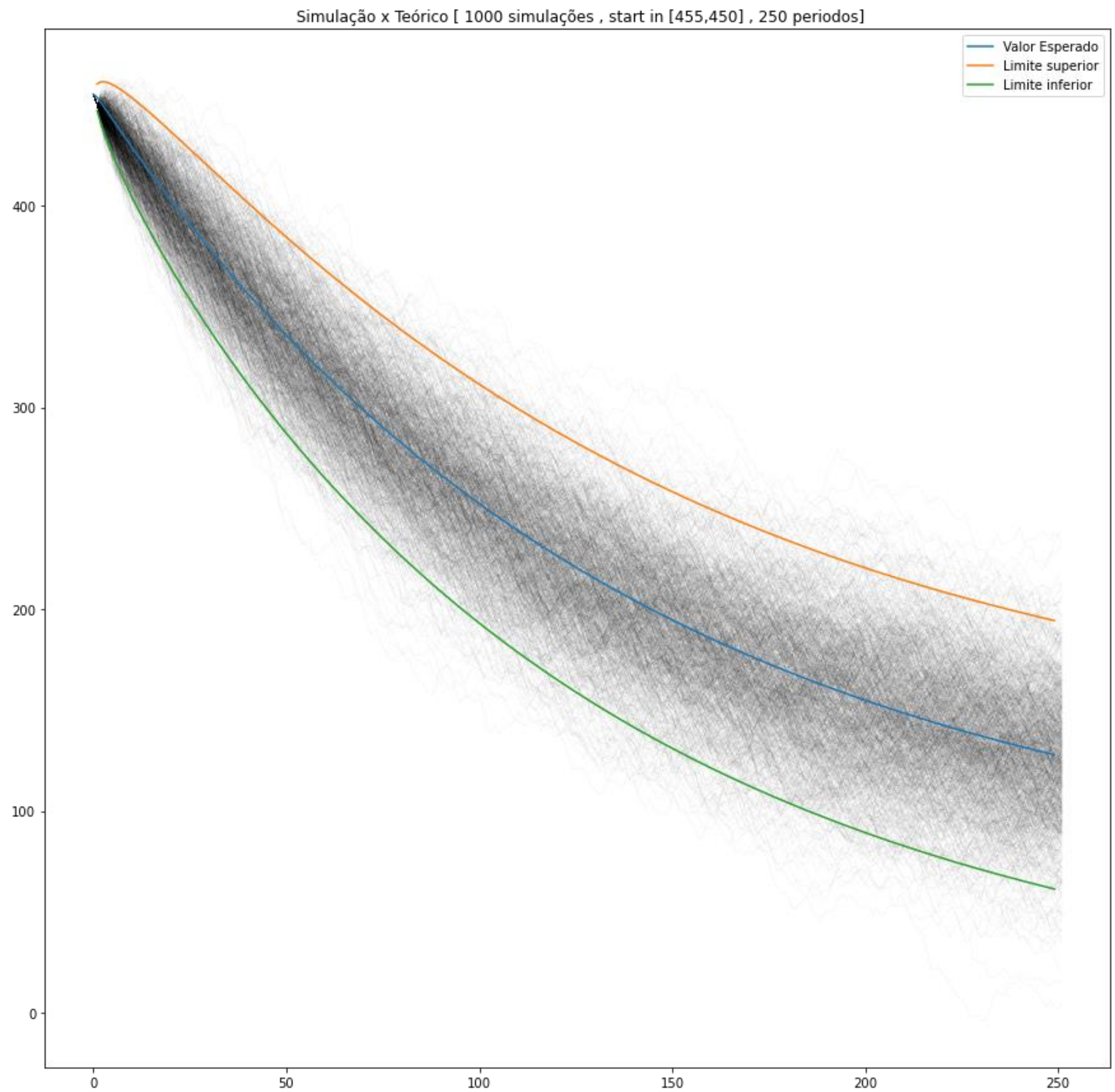
Em relação a Parte B da equação temos:

$$[A^s * Y_t](1,1) \text{ (Primeiro termo desse vetor)}$$

$$\sigma^2_{Y_{t+s}} = 0$$

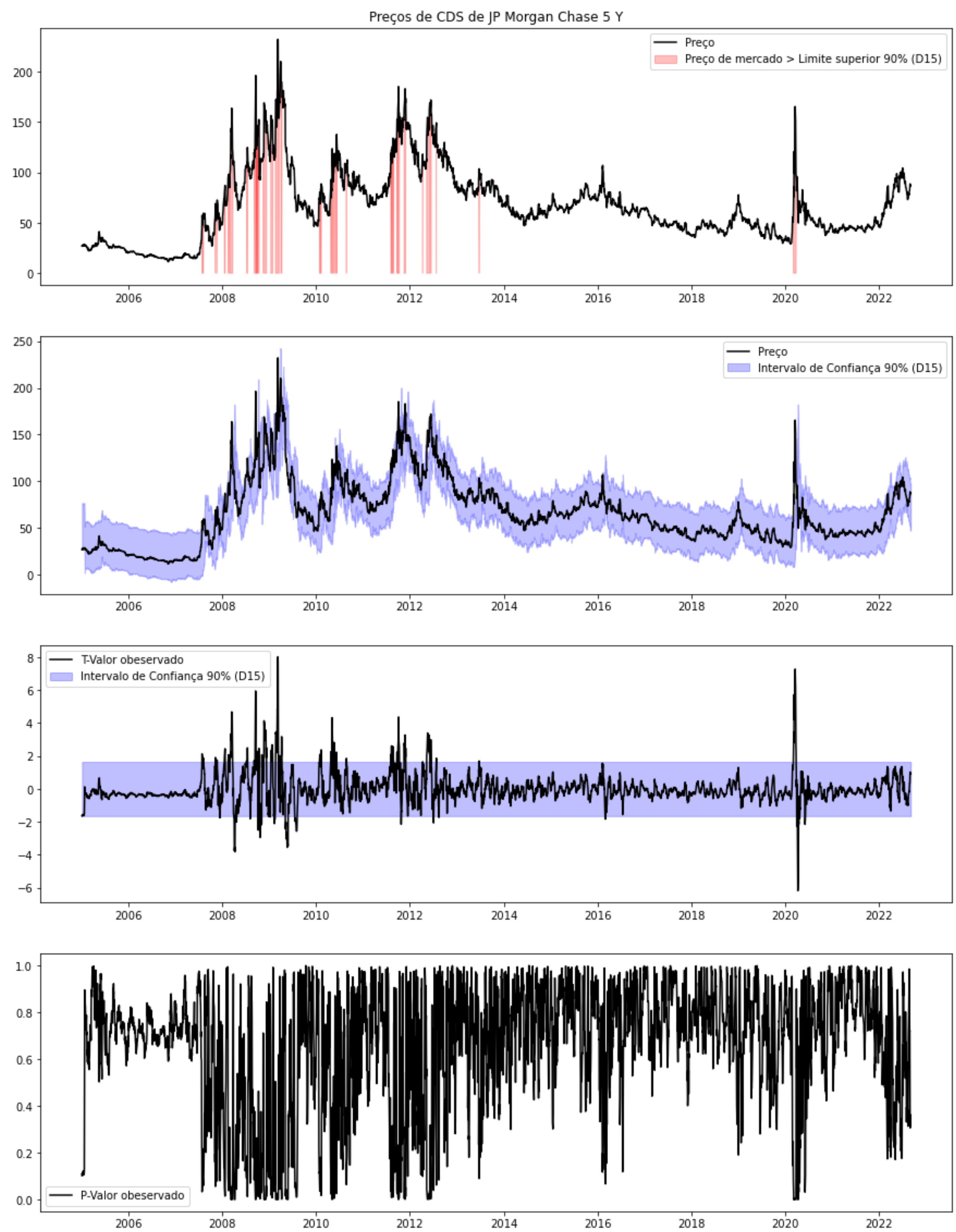
$$\mathbb{E}[Y_{t+s}] = [A^s * Y_t](1,1)$$

Unindo essas duas equações conseguimos estimar qual o valor esperado de Y_{t+s} com base nos parametros Φ_1 e Φ_2. A seguir o resultado de 1000 simulações via *Python* de todos os possiveis resultados de um AR 2 com os parametros Φ_1 e Φ_2 encontrados no último modelo para estimar o *CDS 5Y* de *JP Morgan*.



Cada linha preta representa uma simulação e o intervalo utilizado é o intervalo teórico utilizando as equações que estimamos anteriormente para o nivel de confiança de 95% com s=1 até s=250.

A seguir 4 gráficos comparando o que foi observado no preço do CDS que estamos estimando com um intervalor de confiança estimado com 15 dias de defasagem (s=15):

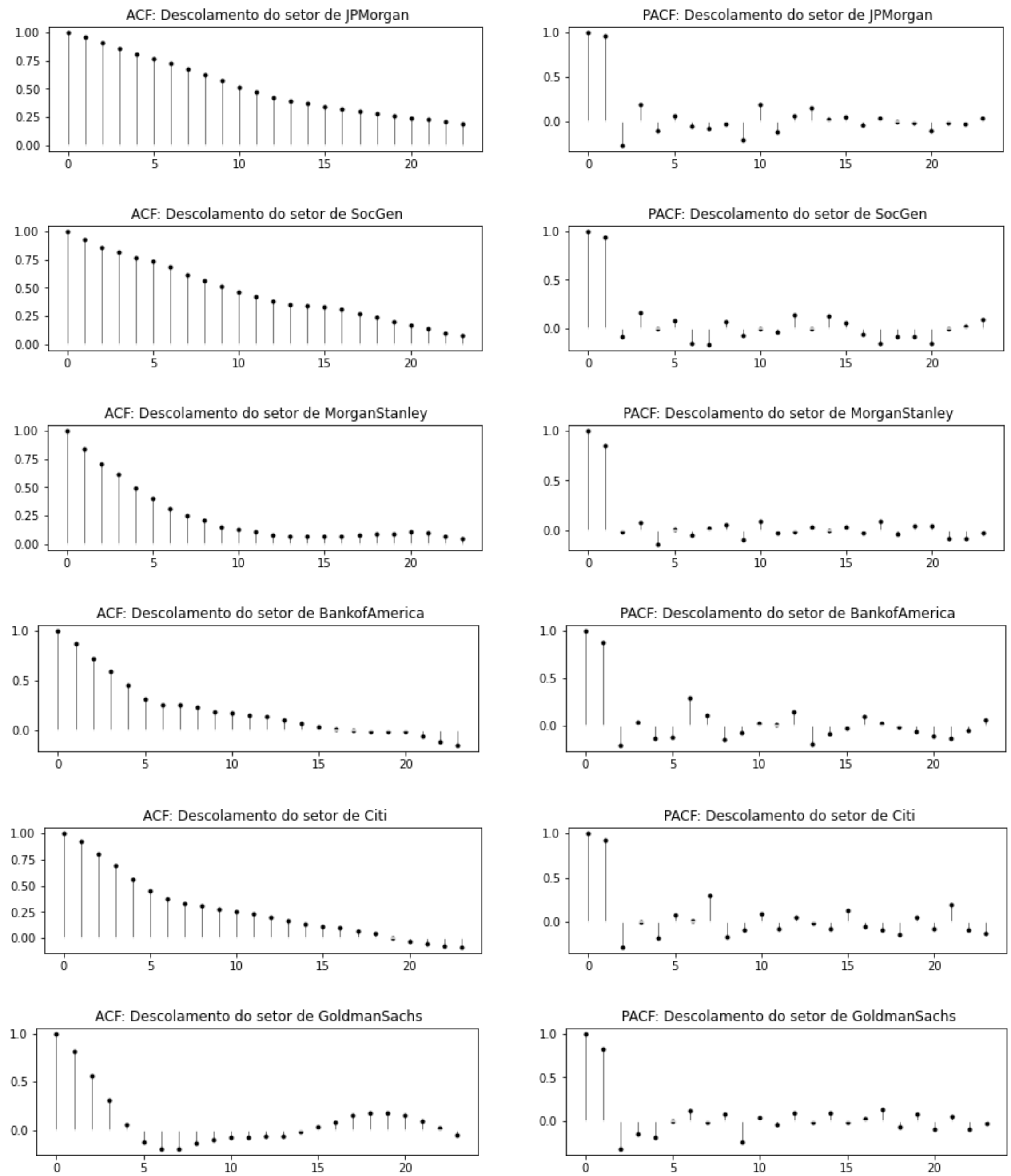


Como podemos observar os principais desvios do preço ocorreram em periodos de crise, como por exemplo em 2008 na crise do subprime, além disso conseguimos perceber que a subida inesperada dos preços tendem a não ser aleatória, na maioria das observações realmente houveram eventos exógenos.

4. Modelo Autoregressivo para estimar desvio indiocincráticos de preços

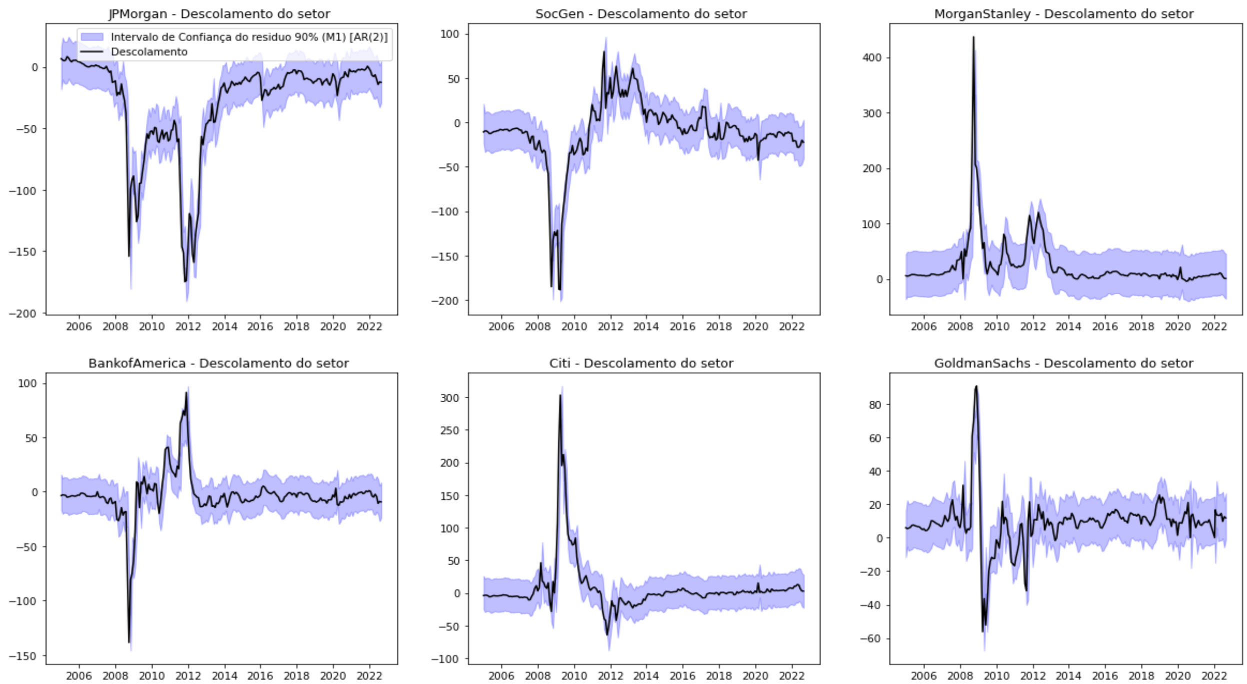
É claro que dentro dessas flutuações existem difentes fatores, os riscos do setor e os riscos indicincráticos do Banco escolhido, buscando tirar as flutuações do setor e olhar apenas para os riscos indiocincráticos, vamos criar um índice que será uma média aritmética dos preços de CDS de 5 anos dos nossos dados, como o intuito é observar a diferença de um banco contra o seu setor não faz sentido fazer uma média ponderada pelo tamanho do banco, visto que, isso pode acabar superestimando alguns grandes bancos como o próprio JP Morgan e excluindo os bancos um pouco menores.

Nesse tópicos vamos observar o comportamento do descolamento do preço do CDS de um banco com a média do mercado, o calculo é bastante simples, basta pegar o 'índice' que utilizamos e calcular a diferença dele com cada um dos bancos da base, assim teremos o 'descolamento' de cada banco com o 'índice', abaixo segue a ACF e PACF para cada banco dessa variavel de estudo.

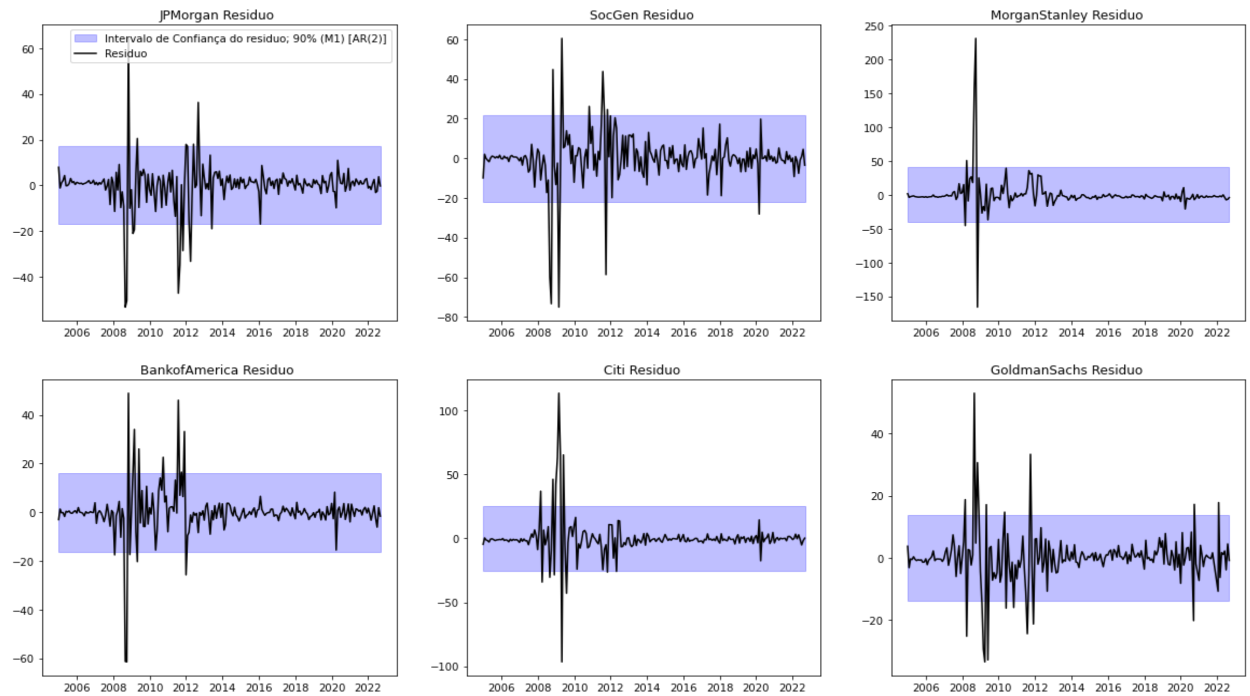


Para generalizar os modelos e olhando o formato das PACF irei optar por um modelo AR (2) novamente e utilizaremos dados mensais de preço.

A seguir o desvio médio de preço mensal do CDS de cada Banco contra o preço médio do setor e um intervalo de confiança a partir de uma defasagem de um mês.



A seguir o residuo do modelo contra o seu intervalo de confiança:

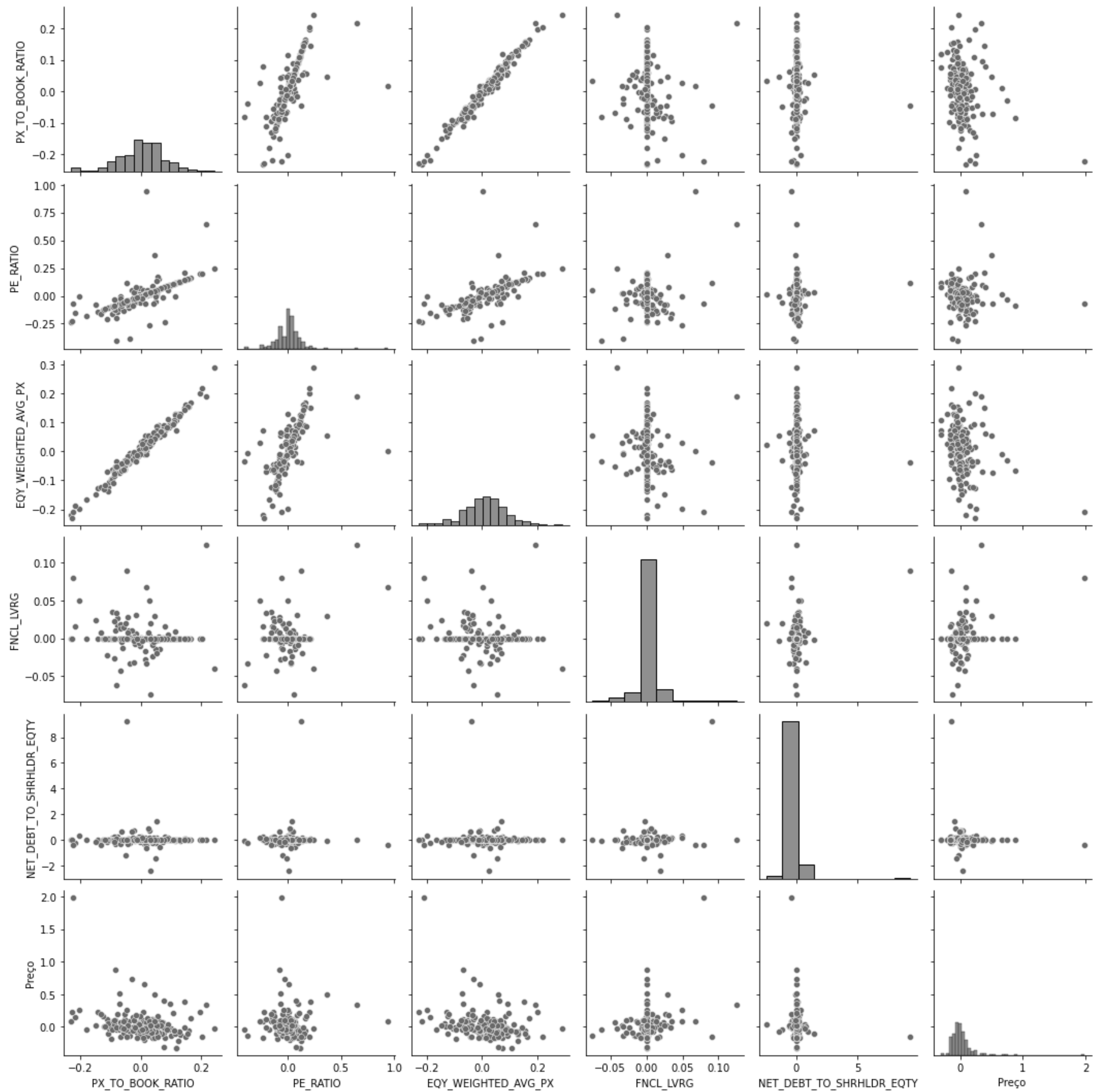


Como podemos perceber, exceto em momentos conturbados, o residuo do modelo segue bastante comportado estimando na maioria dos meses com boa margem um intervalo de confiança adequado para para os desvios nos preços do CDS.

5. Utilizando modelo VAR com variáveis fundamentalistas das empresas para explicar as flutuações no preço de CDS

Vamos avançar um pouco e partir para um modelo mais elaborado, vamos utilizar um modelo VAR com as seguintes variaveis: Price to Book, P/E ratio, Volatility, Financial Lavarage, Net Debt to Sharholder Equity para explicar o preço do CDS 5Y de cada banco.

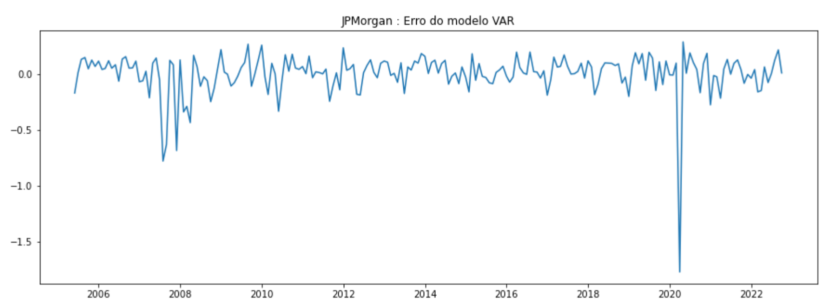
A seguir um Pair Plot com os dados referentes ao JP Morgan:



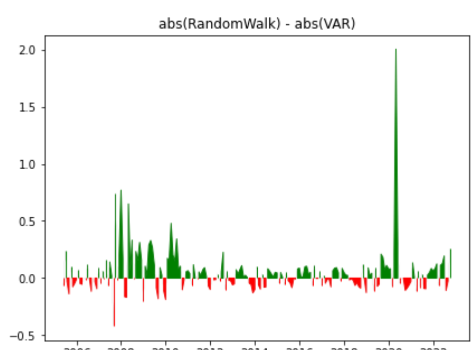
Como podemos perceber algumas variaveis podem ter um certo grau de relação umas com as outras, a seguir temos os parametros estimados para o preço do CDS 5y de JP Morgan via modelo VAR com 1 grau de diferenciação e com número de parâmetros escolhidos via AIC.

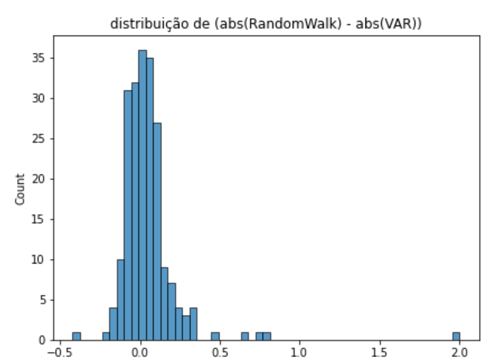
Params	Valor	P- Valor
const	0.052964	0.023
L1.PX_TO_BOOK_RATIO	1.403502	0.313
L1.PE_RATIO	-0.211340	0.241
L1.EQY_WEIGHTED_AVG_PX	-2.097563	0.120
L1.FNCL_LVRG	0.172722	0.865
L1.NET_DEBT_TO_SHRHLDR_EQTY	-0.004602	0.831
L1.Preço	-0.137356	0.092
L2.PX_TO_BOOK_RATIO	0.413982	0.779
L2.PE_RATIO	-0.340024	0.057
L2.EQY_WEIGHTED_AVG_PX	-0.357589	0.806
L2.FNCL_LVRG	1.136048	0.260
L2.NET_DEBT_TO_SHRHLDR_EQTY	-0.015513	0.468
L2.Preço	-0.101856	0.202
L3.PX_TO_BOOK_RATIO	1.167271	0.394
L3.PE_RATIO	0.019353	0.912
L3.EQY_WEIGHTED_AVG_PX	-1.392409	0.309
L3.FNCL_LVRG	0.326914	0.746
L3.NET_DEBT_TO_SHRHLDR_EQTY	-0.025118	0.245

A seguir o erro do modelo VAR com um grau de diferenciação.

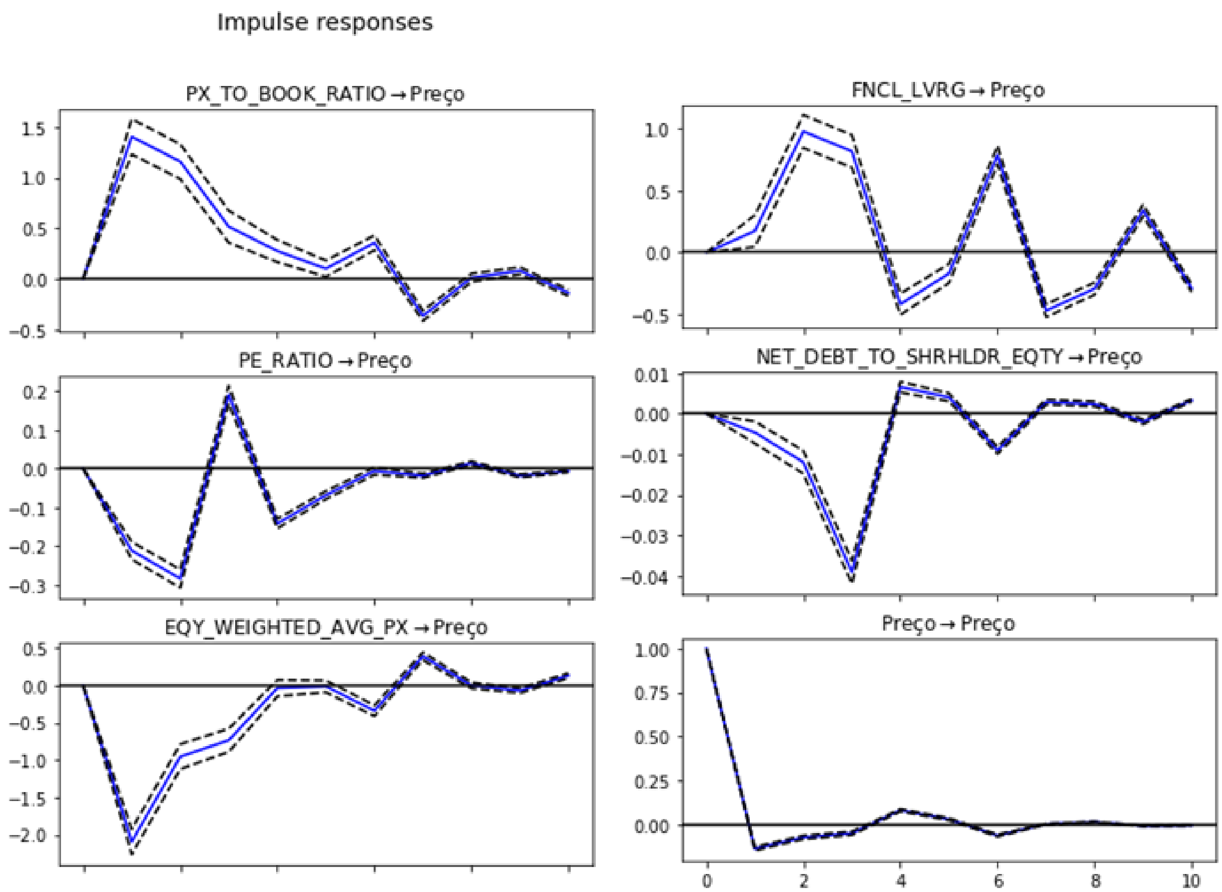


A seguir uma comparação do modelo VAR contra simplemente repetir a última variação observada. Quanto maior o valor dessa subtração, mais acertivo esta sendo o modelo.





A seguir as funções resposta impulso (utilizando um intervalo de confiança de 10% para não poluir o gráfico).



Como podemos reparar as funções reposta impulso fazem sentido, se o mercado começar a fazer um Equity muito maior do que o Book Value de uma empresa, significa que um mercado em sí esta otimista com o futuro dessa empresa, ou seja, não faz sentido pensar que o preço de um CDS subiria.

Ao mesmo tempo se o índice de alavancagem do banco sobe, significa que começa a surgir dúvidas referentes a sua solidez e isso pode fazer com que papéis de CDS se valorizem nesse cenário.

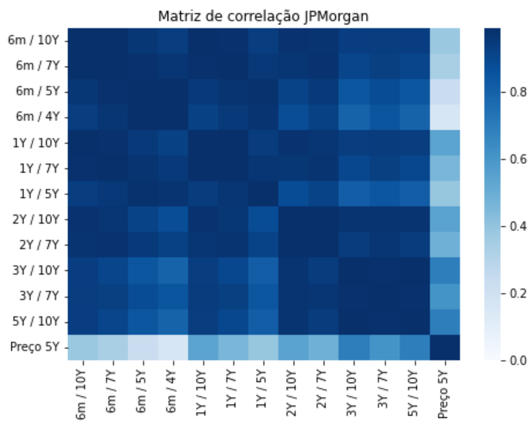
Em anexo, segue um arquivo Jupyter Notebook com esse modelo replicado para outros bancos.

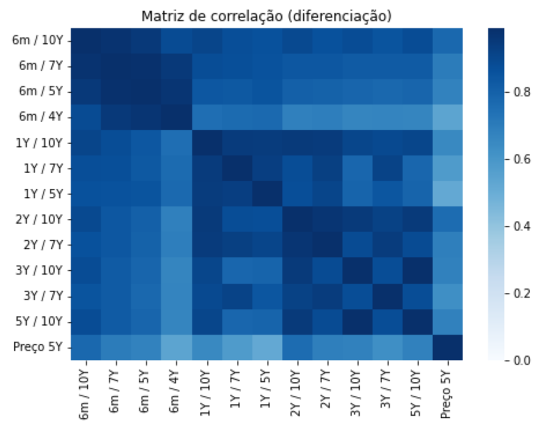
6. Incrementando a variável “spread” no modelo VAR

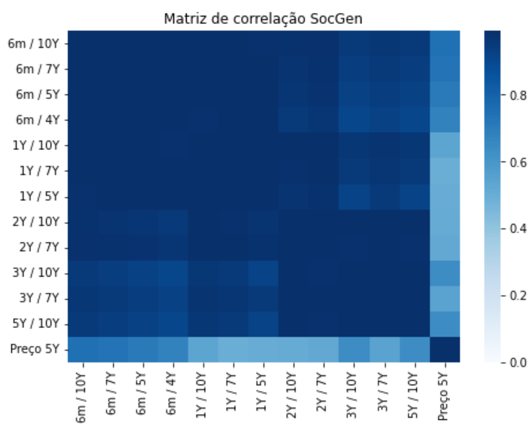
Dado que as unicas coisas que diferenciam um CDS para outro é a duração do contrato e o risco que está sendo protegido, uma hipótese que pode ser levantada é, em momentos de crise de confiança na liquidez de um banco, o preço dos CDS deveriam se aproximar já que todos eles cobrem esse periodo conturbado.

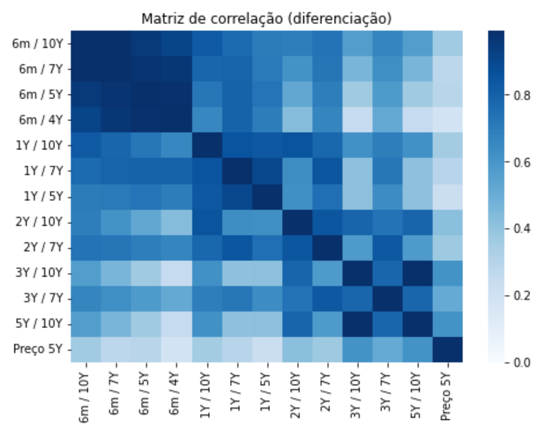
Caso essa dinâmica não ocorresse, individuos conseguiriam comprar CDS 1Y e vender CDS 5Y de modo que estariam assegurados da crise sem ter necessariamente saido com saldo negativo, ou seja, seria possivel realizar arbitragem.

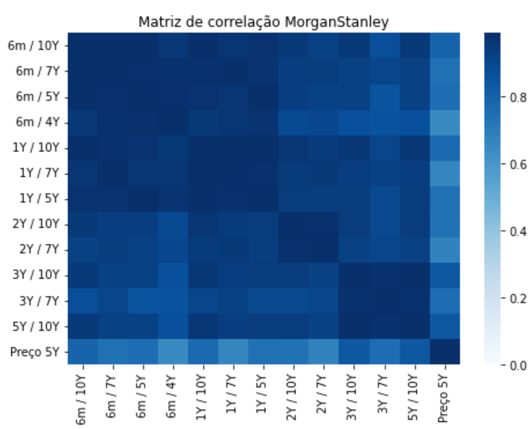
Abaixo vamos fazer uma matriz de correlação dentre os possiveis spreads com pelo menos 4 anos de diff e o preço de um CDS 5Y, uma matriz de correlação da diferenciação e isto para cada banco.

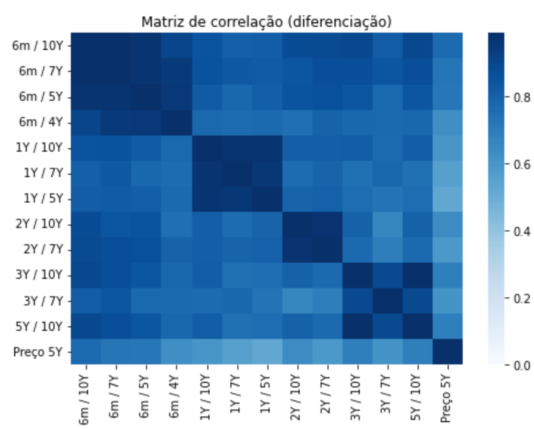




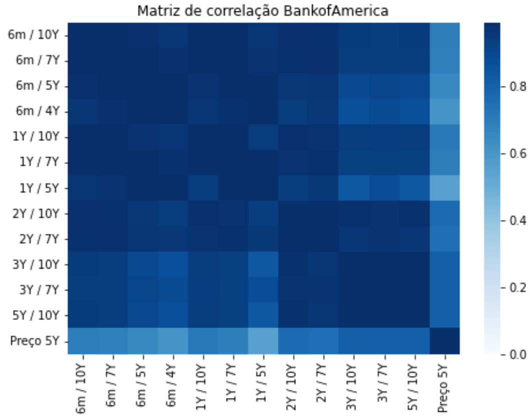




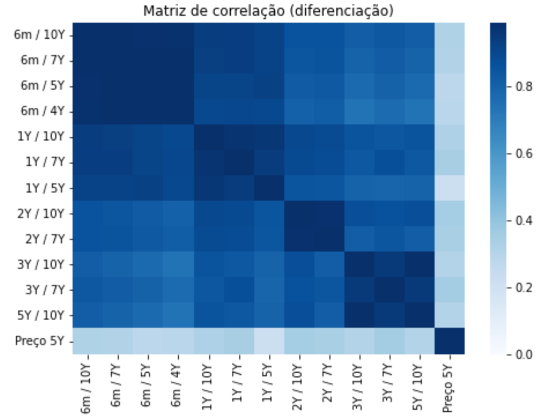




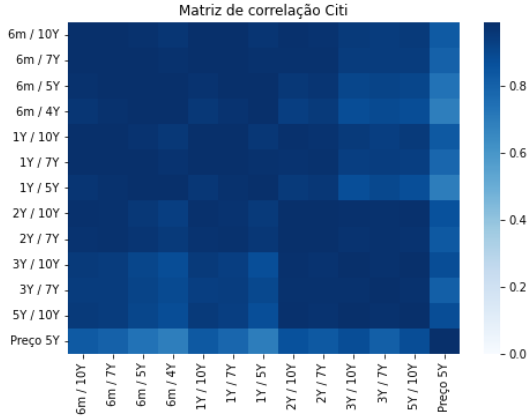
Matriz de correlação BankofAmerica



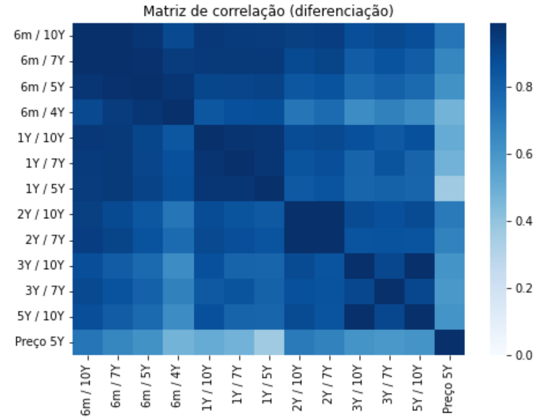
Matriz de correlação (diferenciação)



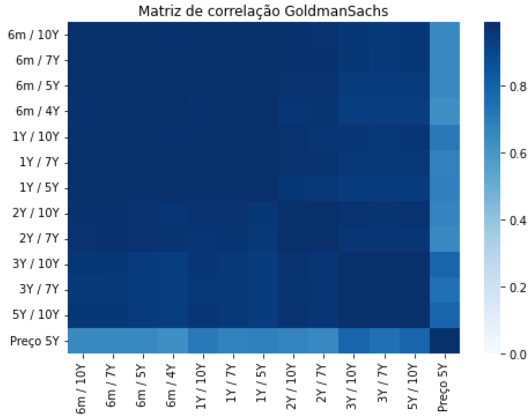
Matriz de correlação Citi



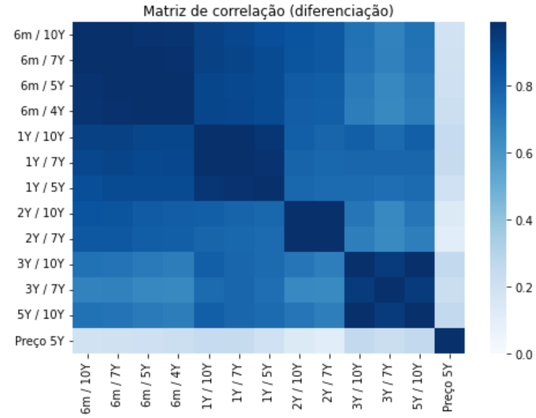
Matriz de correlação (diferenciação)



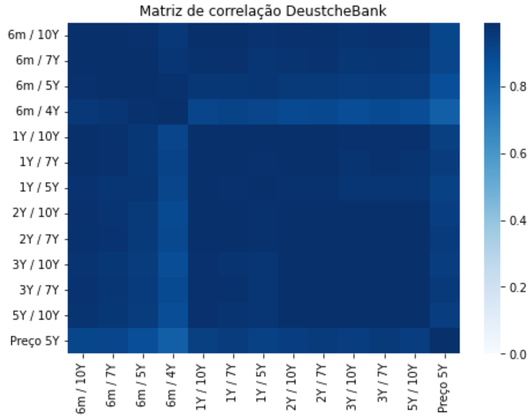
Matriz de correlação GoldmanSachs



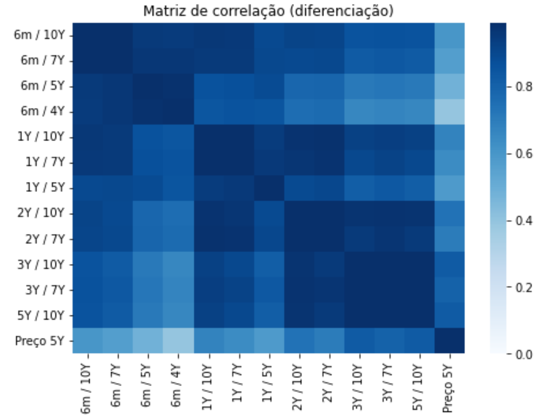
Matriz de correlação (diferenciação)



Matriz de correlação DeustcheBank

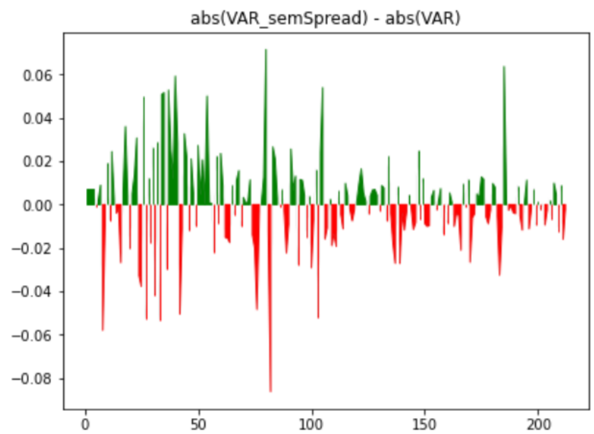


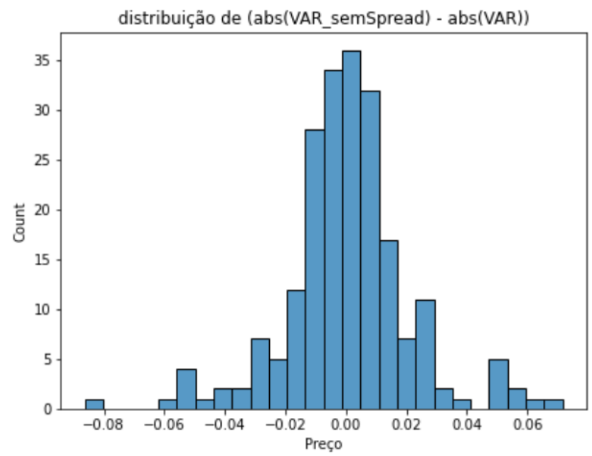
Matriz de correlação (diferenciação)



Como os outros contrato, exceto 5Y, possuem baixa liquidez, não há dados o suficiente para utilizar o spread com maior correlação, por isso vamos utilizar o spread com maior número de dados para cada banco e comparar os critérios de informação do modelo com spread / sem spread.

A seguir a compração de erros entre os modelos com/sem spread e sua comparação de Critério Information, o spread utilizado foi a razão de 1y / 5y e para exemplificar aqui utilizamos o *JP Morgan* em anexo segue um jupyter com os demais bancos.





	S_spread	C_spread
aic	-32.242	-36.259
bic	-30.419	-33.796
hqic	-31.505	-35.263
fpe	0.000	0.000

7. Conclusão

Com o trabalho foi concluido que é possivel utilizar o conhecimento de econometria, para entender melhor o funcionamento de preços de diferentes tipos de derivativos no mercado financeiro. O foco do trabalho não foi apenas fazer previsão, mas foi tentar acompanhar com ferramentas estátisticas e de econometria os movimentos e flutuações nos preços dos CDS, entender o que se mostra estatisticamente relevante no tocante ao preço da categoria de CDS estudado, como também busquei enter qual o "perfil" do movimento dessas flutuações.

8. Bibliografia

- Bloomberg: Dados fundamentalistas e histórico de preços.

- Python: Ferramenta utilizada para os modelos e plotagem dos gráficos. (Caso queira auditar o código, recomendo a seguinte sequencia de atalhos Ctrl +A , CTRL+K+0 para resumir todos os códigos.)