

MESTRADO
ECONOMETRIA APLICADA E PREVISÃO

TRABALHO FINAL DE MESTRADO
DISSERTAÇÃO

REGRESSÃO DO ÍNDICE DE CAUDA: UMA APLICAÇÃO EMPÍRICA

FERNANDO MIGUEL LAIRES CASÇÃO

OUTUBRO - 2018

MESTRADO
ECONOMETRIA APLICADA E PREVISÃO

TRABALHO FINAL DE MESTRADO
DISSERTAÇÃO

REGRESSÃO DO ÍNDICE DE CAUDA: UMA APLICAÇÃO EMPÍRICA

FERNANDO MIGUEL LAIRES CASÇÃO

ORIENTAÇÃO:
JOÃO NICOLAU

OUTUBRO - 2018

RESUMO

Para efeitos de enquadramento desta dissertação, de referir que na explicação de valores extremos, um parâmetro importante para medir o comportamento de distribuições com caudas pesadas é o índice de cauda. Uma das classes de distribuições mais relevantes é a de Pareto.

No presente trabalho é apresentada uma metodologia de estimação do índice de cauda, que assenta numa regressão exponencial do parâmetro função de variáveis explicativas. O método de estimação é o de Quase Máxima Verosimilhança baseada na função log-verosimilhança de Pareto de tipo I. A metodologia em estudo é aplicada às observações do prémio de risco do mercado acionista. Neste sentido, pretende-se explicar os valores extremos da aba esquerda da distribuição dos dados, com recurso a um conjunto de variáveis estudadas na literatura, no contexto do mercado de ações.

Os resultados sugerem que as variáveis mais relevantes para explicar a variável de interesse são regressores que representam situações de crise e incerteza social, política e económica, para cada momento de tempo. Os resultados finais indicam que o prémio de risco tem uma massa de probabilidade considerável associada a valores extremos da série.

PALAVRAS-CHAVE: Valor extremo; Regressão do índice de cauda; Distribuição de Pareto; Prémio de risco.

JEL CODES: C58; G01; G11.

ÍNDICE DE CONTEÚDOS

Resumo	i
Índice de Conteúdos	ii
Índice de Figuras	iv
Índice de Tabelas	v
Agradecimentos	vi
1. Introdução	1
2. Revisão de Literatura	3
2.1. Distribuições com Caudas Pesadas	3
2.2. Distribuições com Caudas de Pareto	4
2.3. Estimação do Índice de Cauda	4
2.3.1. Estimador de Hill	6
2.3.2. Regressão do Índice de Cauda por Nicolau e Rodrigues	6
2.3.3. Regressão Exponencial do Índice de Cauda	7
3. Aplicação Empírica	10
3.1. Prémio de risco e o retorno do mercado de ações	10
3.2. Estimação	14
3.2.1. Dados e Metodologia de Estimação	14
3.2.2. Análise Descritiva	15
3.2.3. Resultados de Estimação	17
3.2.4. Discussão de Resultados	18
3.2.5. Limitações e Considerações Finais	23
4. Conclusão	25
Referências	27
Anexos	28

Anexo A.....	28
Anexo B.....	29
Anexo C.....	30
Anexo C1.....	30
Anexo C2.....	31
Anexo C3.....	32

ÍNDICE DE FIGURAS

FIGURA 1: Histograma das observações do PRMA.	16
FIGURA 2: Gráfico de Quantis.	17
FIGURA 3: Variáveis explicativas ao longo do histórico de análise.	19
FIGURA 4: Valor do índice de cauda em função das variáveis explicativas.	22

ÍNDICE DE TABELAS

TABELA I: Estatísticas Descritivas	16
TABELA II: Estimação para o Quantil de 5% (n = 55).....	18
TABELA III: Estimação para o Quantil de 10% (n = 110)	18
TABELA IV: Estimação para o Quantil de 15% (n = 164)	18
TABELA V: Efeitos Parciais e Índice de Cauda Médio	21
TABELA VI: Log-Verosimilhança Média por Quantil.....	23

AGRADECIMENTOS

Gostaria de começar por endereçar uma forte palavra de agradecimento ao meu orientador, Professor Doutor João Nicolau, pela sua elevada disponibilidade e profundos conhecimentos que me foi transmitindo ao longo deste projeto.

Agradeço também ao corpo docente da parte curricular do mestrado de Econometria Aplicada e Previsão, pelos ensinamentos relevantes que me foram transmitidos e que me permitiram consolidar vários conhecimentos que concorreram para a minha formação atual nesta área.

Um reconhecimento também ao ambiente positivo do ISEG, o qual me permitiu uma maior facilidade no meu percurso e sucesso académicos.

Por último e não menos importante, gostava de dar uma palavra de agradecimento ao meu pai, pelo apoio e estímulo constantes e ao meu amigo António Relvas, pela preciosa ajuda na parte de programação.

1. INTRODUÇÃO

A importância da análise de valores extremos tem vindo a crescer nos mais diversos campos da sociedade. O seu estudo é efetuado através da cauda da distribuição subjacente aos dados em análise. As distribuições apropriadas para representar dados que apresentam valores extremos são chamadas de distribuições de caudas pesadas. É neste contexto que se justifica o desenvolvimento deste trabalho, em que a distribuição usada pertence a uma classe de distribuições de Pareto, a saber: Pareto de tipo I. O parâmetro mais importante que a caracteriza é o índice de cauda. Este parâmetro é de extrema importância para a medição do risco de determinada variável - quanto mais baixo for o seu valor, maior o risco associado à variável respetiva.

Neste trabalho, pretende-se explicar os valores extremos do prémio de risco do mercado acionista (de agora em diante designado por PRMA). Conceptualmente, o PRMA é a compensação que os investidores exigem pelo risco adicional de possuírem um portfólio com ações de mercado em relação a um ativo sem risco, ou seja, é um indicador do que se ganha por se assumir o risco. O risco é materializado em situações em que o prémio de risco e, por construção, o retorno de mercado de ações, é negativo, situações que todos os investidores pretendem evitar. É neste seguimento que o interesse deste trabalho se desenvolveu, no sentido de se perceber quais os fatores que levam à ocorrência de valores extremos negativos no PRMA.

Para o efeito, estima-se uma regressão do índice de cauda da distribuição de Pareto função de variáveis explicativas. Deste modo, assume-se como hipótese inicial que as observações negativas mais extremas do PRMA seguem uma distribuição com caudas de Pareto. O conjunto de variáveis explicativas testadas são as estudadas por Goyal e Welch (2008), que resumem o largo conjunto de variáveis utilizadas ao longo do tempo na explicação e previsão do PRMA. O método de estimação é o de Quase Máxima Verosimilhança (de agora em diante designado por QMV) baseada na função log-verosimilhança da Pareto. A principal conclusão é a de que as variáveis que melhor explicam os valores extremos (negativos) do PRMA são: *Default Yield Spread*, definido como sendo a diferença entre o retorno de obrigações classificadas com BAA e AAA e a volatilidade do mercado acionista. Estas variáveis funcionam como indicadores da situação social, política e económica atravessada em cada momento de tempo.

Os resultados parecem indicar que a cauda da distribuição do PRMA apresenta uma massa de probabilidade considerável associada a valores extremos na amostra, face ao baixo valor estimado para o índice de cauda ($\alpha=2,1965$). Os resultados sugerem também que a cauda da distribuição começa num quantil mais baixo relativamente a outros estimadores e variáveis estudadas. De facto, existe evidência de que o ponto de corte da distribuição de 5% é preferível ao de 10% e 15%.

De referir o carácter inovador da abordagem aqui descrita a qual sustentou o desenvolvimento deste trabalho. Parece ainda mais relevante a sua utilização, face à escassez da sua aplicação e divulgação académica. Na realidade, não se encontra na literatura a aplicação desta metodologia no PRMA, podendo emergir aqui o carácter pioneiro de uma nova linha de investigação.

O trabalho apresenta-se estruturado da seguinte forma: na secção 2 é realizada uma revisão de literatura ao estudo dos valores extremos. Neste sentido é apresentada a distribuição estudada no trabalho, a distribuição de Pareto e um conjunto de estimadores que se baseiam nesta mesma distribuição. São eles: o estimador de Hill (1975); a regressão do índice de cauda por Nicolau e Rodrigues (2018); e a metodologia aplicada no presente trabalho, a regressão exponencial do índice de cauda em função de variáveis explicativas. Na secção 3 é apresentada a aplicação empírica em estudo, *i.e.*, estudam-se os valores extremos negativos com potencial de emergirem no mercado accionista, através de uma análise da literatura, seguida da obtenção de resultados e discussão dos mesmos. Na secção 4 expõem-se as principais conclusões do trabalho.

2. REVISÃO DE LITERATURA

Nesta secção do trabalho é feita uma revisão de literatura à estimação de um parâmetro cada vez mais importante na medição do risco em diversos campos da sociedade: o índice de cauda. Neste sentido começa-se por apresentar o conceito de distribuição de caudas pesadas, seguindo-se o exemplo da distribuição de Pareto e, por último, faz-se uma descrição das características essenciais de três estimadores do índice de cauda da Pareto.

2.1. Distribuições com Caudas Pesadas

As distribuições que descrevem amostras com um elevado número de valores extremos são designadas de distribuições com caudas pesadas (ou que são leptocúrticas). Por valor extremo entenda-se um valor que se distancia substancialmente dos restantes valores da amostra e, como tal, se encontra muito afastado da respetiva média. Estas distribuições são assim designadas dado que apresentam um decaimento polinomial para zero, significativamente mais lento relativamente à velocidade da exponencial. Uma definição mais informal envolve o parâmetro de *kurtosis* (k), quarto momento estandardizado em relação à média da amostra, que mede o achatamento da distribuição: quando $k > 3$ diz-se que a distribuição dos dados tem caudas pesadas. O valor de referência 3 é o que resulta da distribuição Normal. Desta forma, quanto maior o valor do k maior a espessura da cauda da distribuição, o que faz com que os eventos extremos tendam a acontecer com maior frequência, comparativamente a uma distribuição de caudas mais finas (como é exemplo a distribuição Normal).

Um indicador de enorme importância no estudo dos valores extremos é o índice de cauda (α). Este parâmetro controla a velocidade do decaimento da cauda da distribuição. Quanto mais baixo for o seu valor, mais lento será o decaimento da cauda, e como tal, maior será a probabilidade de ocorrência de eventos extremos e *outliers*.

Uma outra característica importante do índice de cauda tem a ver com o facto de caracterizar a ordem máxima de momentos finitos de uma variável, isto é, se $h \leq \alpha$, então existem todos os momentos (finitos) até à ordem h , *i.e.* $E[|X|^h] < \infty$, se $h > \alpha$, então não existem os momentos (finitos) de ordem h , *i.e.* $E[|X|^h] = \infty$. Por exemplo, para existir o parâmetro de *kurtosis* é condição necessária que $\alpha \geq 4$.

2.2. Distribuições com Caudas de Pareto

Para caracterizar probabilisticamente os eventos extremos, uma das distribuições com caudas pesadas mais comumente usadas na literatura é a distribuição de Pareto. O nome provém do economista italiano Vilfredo Pareto que usou pela primeira vez a distribuição para explicar a distribuição de riqueza numa sociedade, uma vez que se ajustava bem ao facto de uma larga proporção da riqueza ser detida por uma pequena percentagem da população.

Uma variável aleatória Y segue uma distribuição de Pareto, se a sua função densidade de probabilidade, $f(y)$, for dada por:

$$y \sim \text{Pareto}(y_0, \alpha)$$

$$(1) \quad f(y) = \begin{cases} \alpha \frac{y_0^\alpha}{y^{\alpha+1}} & , y \geq y_0, \quad \alpha > 0, \quad y_0 > 0 \\ 0 & , y < y_0 \end{cases}, y_0 = \min_i\{y_i\}$$

A função de sobrevivência é dada por:

$$(2) \quad \bar{F}(y) = P(Y > y) = \left(\frac{y_0}{y}\right)^\alpha, \text{ com } y > y_0 > 0, \quad y_0 = \min_i\{y_i\}$$

O α nas Equações (1) e (2) representa o índice de cauda, que assume apenas valores positivos (*i.e.*, $\alpha > 0$). A distribuição está apenas definida para $y > y_0$, ou seja, apenas para a aba direita da distribuição (para a aba esquerda faz-se a transformação $-y$), com y_0 a representar o mínimo valor da variável y a que se atribui probabilidades.

Note-se que a distribuição de Pareto de tipo I não é a única utilizada na literatura para modelar séries com um decaimento lento para zero. De facto, existem outras distribuições do tipo de Pareto, como por exemplo Pareto Generalizada, Burr, t-Student e F (Beirlant et al., 2004).

2.3. Estimação do Índice de Cauda

Dada a importância da explicação de valores extremos nas mais diversas áreas da sociedade (economia, finanças, engenharia, geologia, entre outras), têm sido desenvolvidos e aprimorados um conjunto de métodos de estimação do decaimento das caudas (Beirlant et al., 2004, Cap. 4).

No processo de estimação do índice de cauda de uma determinada variável aleatória y , a primeira grande decisão está relacionada com a possibilidade de se considerar que y tem distribuição exata de Pareto ou apenas caudas do tipo de Pareto, ajustando-se o número de observações a utilizar em cada um dos casos.

Considerando que q representa o quantil a partir do qual (aba direita) ou até ao qual (aba esquerda) as observações são consideradas para estimação, tem-se que o número de observações em função do q é dado para a aba direita por:

$$(3) \quad n(q) = \sum_{i=1}^n I_{\{y_i > y_0\}} \quad , \quad I_{\{y_i > y_0\}} = \begin{cases} 1 & \text{se } y_i > y_0 \\ 0 & \text{se } y_i < y_0 \end{cases}$$

O valor de y_0 depende do valor do q considerado. É o valor de q que determina a fração de observações a serem utilizadas na aba para efeitos de estimação, $n(q)$. Se y_i tiver distribuição exata de Pareto considera-se $q = 1$ (todas as observações são usadas). No caso de se admitir que y_i possa ter apenas caudas de Pareto, deve-se considerar $q < 1$.

No caso de se considerar que y_i tem apenas caudas de Pareto, qual o valor do *threshold*, q , que devemos escolher?

A determinação de onde deverá começar a cauda da distribuição é um tema crucial na estimação, dado que se apresenta um dilema de enviesamento *versus* variância. Considerando que o interesse está na aba direita da distribuição, por um lado, o ponto de corte deve ser suficientemente alto para que sejam consideradas apenas observações cujo comportamento seja bem modelado pela função densidade considerada. Se q for demasiado baixo perde-se a hipótese inicial de que a cauda da distribuição segue uma Pareto e, como consequência, o estimador $\alpha(q)$ é enviesado e mesmo inconsistente. Desta forma não interessa deslocar muito para a esquerda o valor do quantil quando se avalia a aba direita da distribuição. Por outro lado, o ponto de corte não deve ser demasiado alto dado que é necessário ter um número de observações suficiente para a estimação. Quanto mais baixo o número de observações, menor a precisão das estimativas e, como tal, não interessa deslocar muito para a direita o valor do quantil quando se avalia a aba direita da distribuição.

Não existe um método universal para determinar o ponto de corte ótimo. Neste contexto, Nicolau & Rodrigues (2018) desenvolvem um estudo de Monte Carlo para

avaliar o melhor ponto de corte. Os autores recomendam $q \leq 0,2$ para o estimador proposto e para a aba esquerda da distribuição. Já o corte ótimo no caso do estimador de Hill é mais baixo. Os estimadores em causa são apresentados de seguida.

2.3.1. Estimador de Hill

Muitos dos procedimentos estudados na literatura baseiam-se nas estatísticas de ordem $y_{(1)} \geq y_{(2)} \geq \dots \geq y_{(n)}$, obtidas a partir de uma amostra independente e identicamente distribuída (i.i.d.) $\{Y_t\}_{t=1}^n$, proveniente de uma função de distribuição $F(y)$, com n a representar o número de observações da amostra e $y_{(1)} = \min\{Y_t\}$; $y_{(n)} = \max\{Y_t\}$.

Uma das abordagens mais importantes é o estimador de máxima verosimilhança de Hill (1975). Considerando-se que $f(y)$ segue uma distribuição de Pareto de tipo I, como se apresenta na Equação (1) da página 4, o estimador de Hill (1975), para a aba direita da distribuição, é dado por:

$$(4) \quad \alpha_{Hill} = \frac{n(q)}{\sum_{t=1}^n \log(y_t/q) I_{\{y_i > y_0\}}}$$

A função indicadora $I_{\{y_i > y_0\}}$ serve para truncar a amostra apenas nas observações da cauda, ou seja, para $y_i > y_0$.

O caso em que se tem observações dependentes é tratado em Hill (2010). O estimador é o mesmo que se apresenta supra, no entanto, o desvio-padrão necessário para fazer inferência estatística é ajustado em função da dependência dos dados.

2.3.2. Regressão do Índice de Cauda por Nicolau e Rodrigues

Um estimador recentemente estudado por Nicolau & Rodrigues (2018), utiliza uma abordagem diferente, na medida em que não envolve estatísticas de ordem.

Neste sentido, os autores começam por considerar uma variável aleatória i.i.d. $\{Y_t\}_{t=1}^n$ com função de sobrevivência de Pareto de tipo I como se apresenta na Equação (2) da página 4.

Logaritimizando e reescrevendo a equação (2) obtém-se:

$$(3) \quad \log \bar{F}_n(y_i) = \vartheta + \alpha z_i + \varepsilon_i \quad i = 1, 2, \dots, n-1$$

Onde $\bar{F}_n(y_i) = \frac{1}{n} \sum_{i=1}^n I_{\{Y_i > y_i\}}$; $y_i = (1 - u_i)^{-\frac{1}{\alpha}} y_0$; $\vartheta = (\alpha - 1) \log(y_0)$; $z_i = \alpha^{-1} \log(1 - u_i)$; $u_i = \frac{i}{n}$

Dado que z_i depende de α , o procedimento consiste em dois passos: primeiro é considerada uma estimativa de α , $\bar{\alpha}$, e gera-se \bar{z}_i ; de seguida, estima-se α_{Pareto} através da aplicação do estimador *Ordinary Least Squares* à regressão da Equação (5).

Nicolau & Rodrigues (2018) argumentam que o estimador proposto é mais robusto à escolha do ponto de corte relativamente ao estimador de Hill apresentado na secção 2.3.1 anterior. Os mesmos autores tratam também o caso em que os dados apresentam dependência com forma desconhecida (Nicolau & Rodrigues, 2018, secção 3.2.3).

Os procedimentos de estimação do índice de cauda supra referidos podem ser aplicados a outros tipos de comportamento de cauda.

2.3.3. Regressão Exponencial do Índice de Cauda

Em Wang & Tsai (2009) é apresentada uma regressão do índice de cauda que permite especificar α em função de variáveis explicativas. Para garantir que o parâmetro é sempre positivo este é especificado como sendo uma exponencial.

Aplica-se, portanto, o modelo de regressão exponencial seguinte:

$$(4) \quad \alpha(x_i) = \exp(x_i' \theta) = \exp(\theta_0 + \theta_1 x_{1,i} + \dots + \theta_p x_{p,i})$$

Com vetor de variáveis explicativas $x_i' = [1 \ x_{1,i} \ x_{2,i} \ \dots \ x_{p,i}]$, observação i ($i = 1, \dots, n$).

2.3.3.1. Estimador de Máxima Verosimilhança

Para estimarem os parâmetros do modelo, os autores utilizam o método da Máxima Verosimilhança (MV). Apresenta-se de seguida o processo de obtenção do estimador.

Seja y_i uma variável aleatória i.i.d. com função densidade de probabilidade de Pareto, como se apresenta na Equação (1) (ver página 4), com a diferença de que neste caso o índice de cauda é escrito em função de variáveis explicativas, ou seja, $\alpha(x_i) = \exp(x_i' \theta)$. Apresenta-se em baixo a função descrita.

$$(5) f(y_i|x_i) = \begin{cases} \alpha(x_i) \frac{y_0^{\alpha(x_i)}}{y_i^{\alpha(x_i)+1}} & , y_i \geq y_0, \quad \alpha(x_i) > 0, \quad y_0 > 0 \\ 0 & , y_i < y_0 \end{cases}, y_0 = \min_i\{y_i\}$$

Aplicando o logaritmo do produtório a $f(y_i|x_i)$, obtém-se a função log-verosimilhança seguinte:

$$\begin{aligned} \ln \prod_{i=1}^n f(y_i|x_i, \theta) &= \sum_{i=1}^n \ln[f(y_i|x_i, \theta)] = l(\theta|y_i, x_i) \\ l(\theta|y_i, x_i) &= \sum_{i=1}^n \ln \left[\alpha(x_i) \frac{y_0^{\alpha(x_i)}}{y_i^{\alpha(x_i)+1}} \right] \\ &= \sum_{i=1}^n \{ \ln[\alpha(x_i)] + \alpha(x_i) \ln(y_0) - [\alpha(x_i) + 1] \ln(y_i) \} \end{aligned}$$

$$(6) \Leftrightarrow l(\theta|y_i, x_i) = \sum_{i=1}^n \{ x_i' \theta + \exp(x_i' \theta) \ln(y_0) - [\exp(x_i' \theta) + 1] \ln(y_i) \}$$

Seguidamente, prova-se que este estimador é uma generalização do estimador de Hill (1975).

Assumindo que $p = 0$ e $x_i = 1$, então a Equação (8) fica:

$$\begin{aligned} l(\theta|y_i) &= \sum_{i=1}^n \{ \theta + \exp(\theta) \ln(y_0) - [\exp(\theta) + 1] \ln(y_i) \} \\ \Leftrightarrow l(\theta|y_i) &= \sum_{i=1}^n \{ \theta + \exp(\theta) [\ln(y_0) - \ln(y_i)] - \ln(y_i) \} \end{aligned}$$

$$(7) \quad \Leftrightarrow l(\theta|y_i) = n\theta + \exp(\theta) \sum_{i=1}^n \ln\left(\frac{y_0}{y_i}\right) - \sum_{i=1}^n \ln(y_i)$$

Diferenciando a função (9) em ordem ao coeficiente θ , obtém-se:

$$(8) \quad l'(\theta|y_i) = \frac{\partial l(\theta|y_i)}{\partial \theta} = n + \exp(\theta) \sum_{i=1}^n \ln\left(\frac{y_0}{y_i}\right)$$

Por último, igualando a zero a derivada da função calculada em (10), têm-se:

$$l'(\theta|y_i) = 0 \Leftrightarrow n + \exp(\theta) \sum_{i=1}^n \ln\left(\frac{y_0}{y_i}\right) = 0$$

$$(9) \Leftrightarrow \exp(\theta) = \alpha = \frac{n}{\sum_{i=1}^n \ln\left(\frac{y_0}{y_i}\right)}$$

Desta forma, chega-se ao estimador de Hill, tal como proposto por Hill (1975).

Relativamente à matriz de variâncias-covariâncias, esta é dada pelo inverso da matriz de informação de Fisher, $I(\theta)^{-1}$, com $I(\theta) = A(\theta) = B(\theta)$.

$$A(\theta) = -E\left(\frac{\partial^2 l(\theta|y_i, x_i)}{\partial \theta \partial \theta'}\right), \quad B(\theta) = E\left(\frac{\partial l(\theta|y_i, x_i)}{\partial \theta} \frac{\partial l(\theta|y_i, x_i)}{\partial \theta'}\right)$$

2.3.3.2. Estimador de Quase Máxima Verosimilhança

Como não se pode garantir que a cauda da distribuição segue, de facto, uma densidade de Pareto, uma opção de estimação mais robusta é considerar o estimador de Quase Máxima Verosimilhança (QMV). O estimador de QMV é o de MV com o cuidado adicional de se tomar como matriz de variâncias-covariâncias do estimador a expressão $A(\theta)^{-1}B(\theta)A(\theta)^{-1}$ e não $I(\theta)^{-1}$.

Sob certas condições prova-se que este estimador é consistente, assintoticamente normal e, para certos casos, é mesmo um estimador eficiente (Gourieroux & Monfort, 1995). As condições são:

- A distribuição pertence à família exponencial (em baixo prova-se que a distribuição de Pareto pertence à família exponencial);
- A média condicional está bem especificada;
- A variância condicional está bem especificada.

Demonstra-se que $\{f(y|\theta): \theta \in \Theta\}$ é uma família exponencial se existirem funções $\eta(\theta)$, $c(\theta)$, $t(y)$ e $h(y)$ tal que:

$$(10) \quad f(y|\theta) = h(y)c(\theta)\exp\left\{\sum_{j=1}^k \eta_j(\theta)t_j(y)\right\}$$

onde o suporte da distribuição não depende de θ , $h(y) \geq 0$, $t_j(y)$ são funções reais dependentes apenas de y (não podem depender de θ), $c(\theta) \geq 0$ e $\eta_j(\theta)$ são funções reais dependentes apenas de θ (não podem depender de y).

De seguida apresenta-se a prova para o caso da distribuição de Pareto.

$$\begin{aligned}
f(y|\alpha; y_0) &= \frac{\alpha y_0^\alpha}{y^{\alpha+1}} = \exp\left\{\log\left(\frac{\alpha y_0^\alpha}{y^{\alpha+1}}\right)\right\} = \exp\{\log(\alpha y_0^\alpha) - \log(y^{\alpha+1})\} \\
&= \exp\{\log(\alpha y_0^\alpha)\} * \exp\{-\log(y^{\alpha+1})\} \\
&= \alpha y_0^\alpha \exp\{(\alpha + 1) \log(y^{-1})\} \\
(11) \qquad &= \alpha y_0^\alpha \exp\left\{(\alpha + 1) \log\left(\frac{1}{y}\right)\right\}
\end{aligned}$$

Com este resultado, demonstra-se que a distribuição de Pareto pertence à classe da família exponencial com $h(y) = 1$, $c(\alpha, y_0) = \alpha y_0^\alpha$, $\eta_j(\alpha, y_0) = (\alpha + 1)$ e $t(y) = \log\left(\frac{1}{y}\right)$.

Note-se que esta prova exige que y_0 seja uma constante fixa. De facto, y_0 é o valor que resulta do *threshold* (q) definido previamente à estimação e representa o menor valor (em módulo) de y considerado na estimação de α , *i.e.*, representa o início preestabelecido para a cauda da distribuição de y .

3. APLICAÇÃO EMPÍRICA

Depois de analisados os estimadores do índice de cauda em causa, efetua-se neste ponto uma análise empírica que pretende explicar os valores extremos negativos com potencial de emergirem no mercado acionista. No seguimento, estuda-se a literatura de estimação das variáveis de interesse, sendo posteriormente analisada e discutida a estimação e resultados objeto deste estudo.

3.1. Prémio de risco e o retorno do mercado de ações

O PRMA é uma medida importante para a análise de risco do mercado financeiro. Nesta secção é descrita a sua importância na análise de mercado e os seus determinantes.

Todo o investidor é avesso ao risco (uns mais que outros) e como tal está disposto a pagar um menor preço para investimentos de maior risco, relativamente a investimentos mais seguros, para um mesmo valor esperado de rendimento. O PRMA é o retorno exigido para se assumir o risco. Rozeff (1984) fornece no seu trabalho uma analogia de fácil compreensão: quando o consumidor compra, por exemplo, uma *pizza*, sabe o seu preço e tem uma perceção mais ou menos clara do retorno (não monetário, ou seja, o

prazer) que lhe irá proporcionar o seu consumo. Quando um investidor compra ações, sabe o seu preço, mas não tem conhecimento de antemão do retorno associado. É aqui que começa o interesse pelo PRMA. Uma estimativa do seu valor fornece uma indicação do valor da recompensa por se deterem ações em detrimento de um ativo sem risco, ou pelo menos, de risco próximo de zero.

Ao longo dos anos o interesse por esta variável tem crescido nas áreas de Economia e Finanças. É vista como uma importante medida de risco e um determinante do custo de financiamento das empresas, das decisões de poupança das famílias e na elaboração de orçamentos de Estado (Duarte & Rosa, 2015). Por consequência, esta tem sido uma variável de grande interesse prático e académico. Duarte & Rosa (2015) apresentam de forma concisa um conjunto de vinte modelos de previsão e explicação do PRMA, estudados ao longo dos anos. De facto, existe uma vasta literatura na modelação da variável, que difere na sua forma de cálculo, na metodologia usada e nas variáveis explicativas consideradas na estimação.

Um primeiro desafio na estimação do PRMA prende-se com a própria definição da variável. Não é claro o que realmente constitui o retorno de mercado e a taxa de mercado sem risco. Na prática, o que é mais usado como medida do preço do mercado de ações são os índices compostos, como é o caso do S&P 500¹. A taxa de mercado sem risco mais comumente usada é a das obrigações do Tesouro (títulos de dívida soberana). No entanto, não são completamente isentas de risco, uma vez que a taxa nominal está exposta à inflação, a riscos de liquidez e ao próprio risco soberano.

A larga literatura na explicação e previsão do PRMA e dos retornos de mercado tem sugerido uma série de variáveis explicativas, nomeadamente estudadas por Goyal & Welch (2008), as quais se apresentam na tabela do Anexo A.

O prémio de risco relaciona-se diretamente com o tipo de perfil do investidor. Um investidor demasiado otimista acerca do seu valor vai poupar menos e investir mais em ativos de risco, contrariamente ao investidor mais conservador. A grande questão está em saber quais os *timings* de investimento mais acertados. Por um lado o investidor não quererá possuir ações de empresas com retornos e prémios de risco negativos. Por outro

¹ *Standard & Poor's 500* (S&P 500) é um índice ponderado pela capitalização de mercado das 500 maiores empresas norte-americanas cotadas nas bolsas de NYSE ou NASDAQ.

lado, um prémio de risco negativo é também um método rápido de avaliar o valor das ações. Quanto mais negativo for o seu valor, maior terá de ser o crescimento dos ganhos futuros para compensar o investidor pelo risco assumido.

De acordo com dados publicados pelo Financial Times em 2011², um aspeto interessante no prémio de risco é o facto das transições entre valores positivos e negativos estarem correlacionadas com os pontos de viragem do ciclo económico. As observações mais extremas negativas no PRMA e retorno do mercado são, indubitavelmente, resposta a períodos em que o mercado financeiro apresenta dificuldades. Interessa em primeira instância perceber quais as variáveis que melhor poderão representar os períodos de crise.

Bloom (2009) refere que as recessões aparecem em períodos de maior incerteza, sugerindo a medição desta variável (incerteza) para modelar os ciclos económicos. O autor explica que a incerteza numa economia de mercado aumenta substancialmente depois de grandes choques políticos e económicos como a crise dos mísseis de Cuba, o assassinato de John F. Kennedy (35º presidente dos Estados Unidos), a crise nos preços do petróleo da OPEP (Organização dos Países Exportadores de Petróleo) ou os ataques terroristas de 11 de setembro de 2001. O autor constrói uma variável que serve de *proxy* para os períodos de maior incerteza económica observada ao longo dos anos. Este indicador assume o valor 1 nos 17 períodos representados na tabela do Anexo B. Estes eventos são escolhidos porque são os períodos para os quais a volatilidade no mercado de ações apresenta valores acima da média (com o filtro de Hodrick–Prescott, popularizado nos anos 90 por este autor).

Nesta linha de investigação, Fama & French (1989) já tinham efetuado trabalhos de medição do retorno esperado em ações e obrigações a partir de 3 variáveis, as quais consideram representativas do ciclo económico e da incerteza latente no mercado financeiro. Essas variáveis são:

1. *Default Spread (DS)*, *spread* entre rendimentos de obrigações com uma classificação inferior (variável análoga aos retornos de obrigações classificadas com

² Disponível em: <https://www.ft.com/content/30eec02e-03d4-11e1-bbc5-00144feabdc0> [Acesso em: 20/10/2018]

Baa³) e superior (obrigações classificadas com Aaa). A variável DS estudada assume os seus maiores valores no histórico considerado pelo artigo (1926-1987), para os anos da Grande Depressão de 1929 e durante as recessões ainda assim menos severas depois da segunda Guerra Mundial, por exemplo 1957-1958, 1974-1975 e 1980-1982. Efetivamente, a variável reflete alguma da variação no ciclo económico, no entanto, as suas oscilações vão para além disso. De facto, a variável assume valores altos em períodos de incerteza económica, como é o caso dos anos posteriores à Grande Depressão e os anos anteriores ao início da Segunda Guerra Mundial. Durante o período de 1953-1973 apresenta valores consistentemente baixos, um período de maior estabilidade económica que, no entanto, apresenta quatro recessões medidas pelo NBER (*National Bureau of Economic Research*);

2. *Dividend Yield*, calculado como sendo o rácio entre os dividendos pagos e o valor do portfólio de ações ou obrigações considerado. É uma variável com uma elevada correlação com a *Default Spread* (0.61 para 1927-1987 e 0.75 para 1941-1987);

3. *Term Spread*, diferença entre o retorno de obrigações classificadas com Aaa e o retorno das obrigações do Tesouro. Segundo os autores, a variável identifica um padrão associado ao ciclo económico, uma vez que as duas variáveis que o compõem têm flutuações diferentes ao longo do tempo. O retorno das obrigações classificadas com Aaa é uma série mais alisada, enquanto que a série do retorno das obrigações do Tesouro tem maior variabilidade. Isto é, o retorno de obrigações Aaa tende a aumentar menos durante expansões da economia e a cair menos durante as contrações da Economia, relativamente às obrigações do Tesouro.

Fama & French (1989) apresentam evidência de que as 3 variáveis têm poder explicativo dos retornos em obrigações e ações e, conseqüentemente, a variação nos retornos esperados, pelo menos em parte, deve ser explicada através dos ciclos e das situações de incerteza no âmbito da vida política, económica e social.

³ Fama & French (1989) utilizam o retorno das obrigações do mercado, medido com um portefólio de 100 obrigações emitidas por empresas.

3.2. Estimação

Nesta secção do trabalho é apresentado o processo, resultados e principais conclusões da estimação do índice de cauda da distribuição do PRMA, função de variáveis explicativas.

3.2.1. Dados e Metodologia de Estimação

De forma a calcular o PRMA segue-se a prática habitual de usar o índice S&P 500 (*index*) como medida do preço do mercado acionista e as obrigações do Tesouro como medida da taxa de retorno sem risco.

As variáveis são definidas da seguinte forma, com t a representar o período correspondente à observação:

- Retorno contínuo do mercado acionista ajustado aos dividendos (*div*):

$$(12) \quad y_t^r = \log\left(\frac{\text{index}_t + \text{div}_t}{\text{index}_{t-1}}\right)$$

- Retorno contínuo das obrigações do Tesouro dos Estados Unidos (*rfr* representa o retorno discreto das obrigações do Tesouro):

$$(13) \quad y_t^{\text{rfr}} = \log(\text{rfr}_t + 1)$$

- Prémio de risco do mercado acionista:

$$(14) \quad y_t^{\text{prm}} = y_t^r - y_t^{\text{rfr}}$$

A hipótese inicial é a de que o PRMA tem caudas do tipo de Pareto, com função densidade, como se apresenta na Equação (7) da página 8.

A especificação a aplicar na estimação do índice de cauda da distribuição do PRMA é a regressão exponencial, apresentada de seguida.

$$(15) \quad \alpha(x_t)_t = \exp(x_t' \theta) = \exp(\theta_0 + \theta_1 x_{1,t} + \dots + \theta_p x_{p,t})$$

A estimação dos parâmetros θ é efetuada por QMV baseada na função log-verosimilhança de Pareto (obtida na Equação (8) da página 8), $x_t' = [1 \ x_{1,t} \ x_{2,t} \ \dots \ x_{p,t}]$ define o vetor de variáveis explicativas consideradas na estimação, com t a representar o período a que se refere a observação ($t = 1, \dots, n$). As variáveis explicativas consideradas são as estudadas por Goyal & Welch (2008). A base

de dados está disponível no *website* de Amit Goyal⁴. O conjunto de 14 variáveis testadas pode ser encontrado na tabela do Anexo A.

O histórico considerado na estimação vai desde dezembro de 1926 a dezembro de 2017, com periodicidade mensal. O elevado número de observações é crucial para a obtenção das estimativas, uma vez que os eventos extremos da cauda, por definição, acontecem com pouca frequência. Desta forma, apenas uma pequena porção dos dados (indicada pelo quantil considerado) é de facto utilizada.

No presente trabalho, considera-se a condição comumente usada na literatura de que $q \leq 0.2$, mais concretamente, o estimador é avaliado para os quantis da aba esquerda correspondentes a 5%, 10% e 15% das observações.

De notar ainda que apesar do espaço de resultados da variável y^{prm} ser \mathbb{R} (variável pode assumir qualquer valor em \mathbb{R}), a escolha da distribuição de Pareto como função densidade do processo gerador de dados não caracteriza uma limitação na metodologia. De facto, a distribuição de Pareto apenas está definida para valores superiores a uma qualquer constante positiva (Equação (7) da página 8); no entanto, apenas se pretende caracterizar probabilisticamente os eventos extremos. Entre estas observações, pretende-se estudar as da aba esquerda da distribuição, ou seja, os valores mais baixos assumidos pela variável. Neste sentido e para que a estimação por QMV baseada na função log-verosimilhança da Pareto seja possível, é efetuado o seguinte processo:

1. Selecionam-se as observações correspondentes a cada um dos quantis considerados;
2. Efetua-se a estimação para a aba direita da distribuição para a seguinte retransformação⁵ da variável dependente: $-y^{prm}$.

3.2.2. Análise Descritiva

Na Tabela I apresenta-se um conjunto de estatísticas descritivas relativamente às três variáveis supra descritas para todo o histórico de análise.

⁴ Disponível em: <http://www.hec.unil.ch/agoyal/> [Acesso em: 23/09/2018]

⁵ Retransformação é válida uma vez que todos os valores de y^{prm} são negativos para o maior quantil considerado.

TABELA I: Estatísticas Descritivas

Variável	Média	Desvio-padrão	Mínimo	Máximo	<i>Kurtosis</i>	% de valores negativos
y^r	0,0420	0,0544	-0,2710	0,4187	10,7035	16,65
y_t^{rfr}	0,0028	0,0025	0,0000	0,0135	4,2433	0
y^{prm}	0,0392	0,0545	-0,2714	0,4184	10,7173	18,39

Algumas conclusões positivamente expectáveis podem ser retiradas da tabela supra:

- Variáveis y^{prm} e y^r são muito semelhantes entre si face ao baixo valor de y_t^{rfr} . De facto, o retorno das obrigações do Tesouro apenas acentua ligeiramente a magnitude e o número de valores negativos assumidos pelo prémio de risco;
- A média das variáveis é próxima de zero para y^{prm} e y^r face ao facto dos valores positivos e negativos se compensarem entre si. A média de y_t^{rfr} vem ainda mais baixa, indicando o fraco retorno (médio) das obrigações do Tesouro ao longo do histórico considerado;
- Todas as variáveis têm um valor para o coeficiente de achatamento (*kurtosis*) superior a três (valor de referência da Normal), indicador de que o processo gerador de dados subjacente deverá ser uma distribuição leptocúrtica.

De forma a ilustrar o formato da distribuição de y^{prm} , apresenta-se de seguida o histograma das observações.

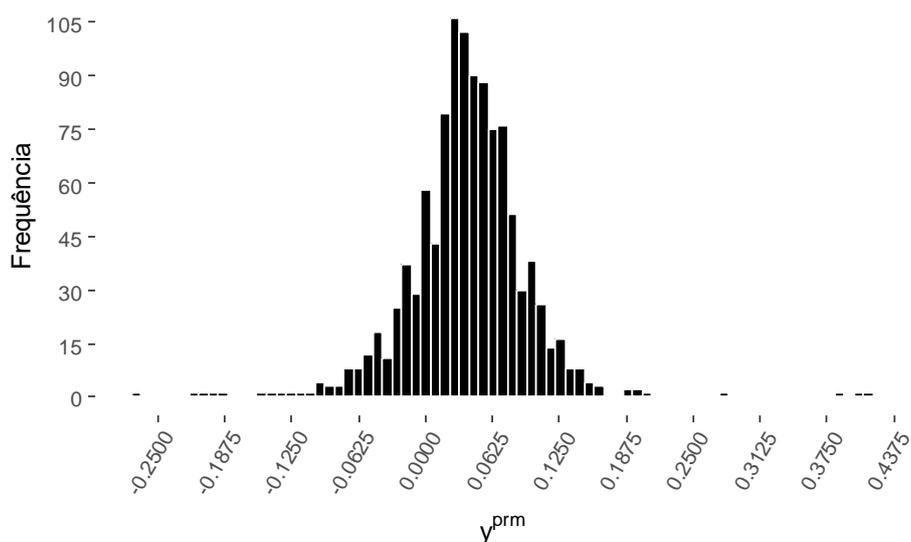


FIGURA 1: Histograma das observações do PRMA.

Através do histograma da Figura 1 é perceptível a existência de observações extremas em ambas as abas da distribuição. No entanto, à primeira vista até parece ser possível que a distribuição Normal se ajuste bem a y^{prm} . Uma forma intuitiva de contrariar esta ideia é a análise do gráfico de Quantis (QQ Plot), que se apresenta na Figura 2. De facto, a figura sugere, tal como o histograma, que a distribuição subjacente ao PRMA tem caudas pesadas em ambas as abas, dado que o gráfico tem a forma de um “S” em torno da reta representada.

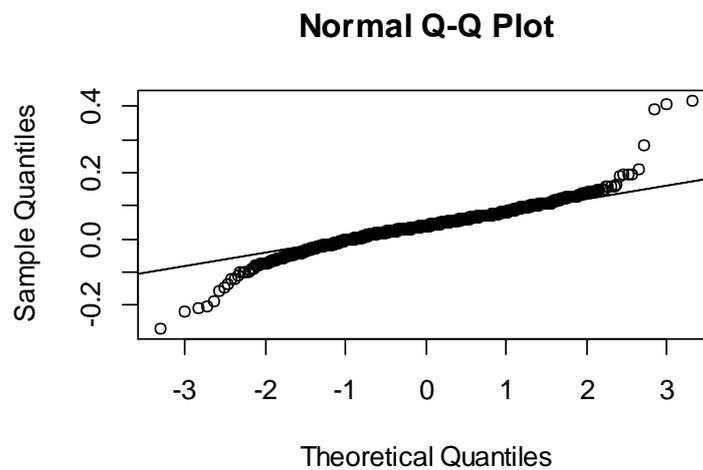


FIGURA 2: Gráfico de Quantis.

Em suma, as duas grandes conclusões da análise descritiva efetuada são:

- 1) Distribuição Normal não é a mais adequada na representação dos dados de y^{prm} ;
- 2) Parece ser razoável assumir a hipótese de que os valores extremos negativos de y^{prm} têm como distribuição implícita uma Pareto de tipo I.

3.2.3. Resultados de Estimação

Efetuada a estimação⁶ por QMV, o modelo final para cada um dos pontos de corte é apresentado nas tabelas infra. Os modelos finais considerados tiveram em conta a significância económica e estatística individual dos coeficientes estimados.

⁶ Todo o processo de estimação foi desenvolvido recorrendo ao *software* R. Os *packages* utilizados foram: readxl, bbmle, numDeriv, ggplot2, reshape2, scales, lubridate, moments e VGAM.

TABELA II: Estimação para o Quantil de 5% (n = 55)

Regressor ⁷	Coefficiente	Desvio-padrão ⁸	Estatística-Teste ⁹	Valor-p
Termo constante	1,2746	0,1900	6,7069	0,0000
DFY	-15,777	7,2275	-2,1829	0,0290
SVAR	-20,0821	4,3116	-4,6577	0,0000

TABELA III: Estimação para o Quantil de 10% (n = 110)

Regressor	Coefficiente	Desvio-padrão	Estatística-Teste	Valor-p
Termo constante	0,5181	0,0983	5,2705	0,0000
DFY	-9,8040	5,0073	-1,9579	0,0502
SVAR	-22,3919	3,7967	-5,8977	0,0000

TABELA IV: Estimação para o Quantil de 15% (n = 164)

Regressor	Coefficiente	Desvio-padrão	Estatística-Teste	Valor-p
Termo constante	-0,3350	0,0512	-6,5474	0,0000
DFY	-5,8957	3,0069	-1,9608	0,0499
SVAR	-19,7878	3,5945	-5,5051	0,0000

Foram ainda considerados e testados outros modelos com mais variáveis explicativas. No entanto, nenhum deles obteve variáveis adicionais estatisticamente significativas, relativamente aos modelos estimados supra. Algumas das estimações adicionais efetuadas, por quantil, podem ser encontradas nas tabelas do Anexo C.

3.2.4. Discussão de Resultados

Um primeiro resultado evidenciado é o de que a escolha do ponto de corte não altera as variáveis explicativas consideradas para o modelo final. De facto, as variáveis DFY e SVAR são estatisticamente significativas (e.s.) a 5% para os casos em que $q = 5\%$ e $q = 15\%$. Quando $q = 10\%$ a variável DFY é e.s. no limite de 5%.

⁷ DFY representa o *Default Yield Spread*, calculado como sendo a diferença entre o retorno de obrigações classificadas com BAA e AAA; SVAR representa a volatilidade do retorno de mercado de ações.

⁸ Os desvios-padrão são obtidos a partir da matriz de variâncias-covariâncias $A(\theta)^{-1}B(\theta)A(\theta)^{-1}$. Para mais detalhe consultar secção 2.3.3 do presente trabalho.

⁹ Estatística teste: $t = \frac{\hat{\theta}_j}{se(\hat{\theta}_j)} \xrightarrow{a} Normal(0,1)$ sob H_0 ($H_0: \theta_j = 0$)

Interessa perceber o significado económico dos resultados. Para o efeito, segue infra o gráfico das variáveis ao longo do histórico de análise.

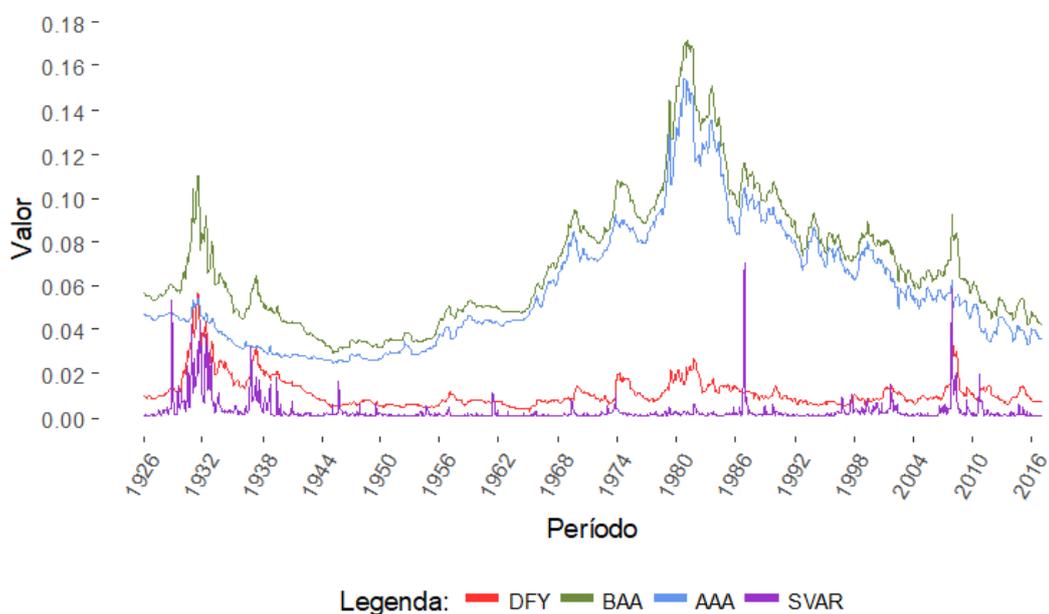


FIGURA 3: Variáveis explicativas ao longo do histórico de análise.

A variável DFY, apresentada a vermelho neste gráfico, assume valores mais elevados em períodos de maior crise e incerteza. De facto, os 15% maiores valores correspondem aos seguintes períodos:

- Pós Grande Depressão (década de 1930). A Grande Depressão foi um período de crise que teve início em 1929 e que persistiu na década de 1930. É considerado o pior e mais extenso período de recessão económica do século XX.
- Os anos que antecederam a Segunda Guerra Mundial até ao período de 1940, cujo conflito militar global durou de 1939 a 1945.
- Finais de 1974 até inícios de 1976. Foi um período de recessão que pôs fim à expansão da Economia pós Segunda Guerra Mundial. Durante este período são os seguintes dois eventos de maior volatilidade de mercado identificados por Bloom (2009) (tabela do Anexo B):
 1. Crise nos preços do petróleo face à política seguida pela OPEC (Organização dos Países Exportadores de Petróleo) em resposta ao conflito Israelo-Árabe que se inicia em outubro de 1973;

2. Colapso de um dos maiores bancos americanos, *Franklin National Bank* em outubro de 1974, sendo à época a primeira vez que um grande banco americano ia à falência na história do país. O primeiro mês de 1974 a aparecer nas 15% observações mais elevadas de DFY é novembro, mês posterior ao colapso do Banco.
- Recessão no início dos anos 80 (1980 a 1984). Bloom (2009) identifica dois eventos de maior volatilidade de mercado neste intervalo de tempo (tabela do Anexo B):
 1. Mudança na política monetária levada a cabo pelo Sistema de Reserva Federal dos Estados Unidos para combater a elevada inflação e os efeitos da crise energética (1982);
 2. Crise dos reféns americanos (diplomatas no Irão) entre 1979 e 1981. O acontecimento teve um grande aparato dos *media*.
 - Crise económica e financeira internacional: meses finais de 2008 e ano de 2009. O primeiro mês de 2008 a aparecer nas 15% observações mais elevadas de DFY é outubro, mês identificado por Bloom (2009), como sendo o de máxima volatilidade para o período de crise.

Todos os períodos de crise descritos supra são também referenciados por Fama & French (1989).

Em relação à variável explicativa SVAR, representada a roxo no gráfico da Figura 3, percebe-se que tende a ter um comportamento semelhante a DFY. A variável assume os dois valores mais extremos em dois períodos de grande declínio dos mercados financeiros: novembro de 1987, um mês após a ocorrência da Segunda-feira Negra (19 de outubro de 1987). Esta data é marcada pela queda abrupta no mercado de ações por todo o mundo (também identificado por Bloom (2009) – tabela do Anexo B); e novembro de 2008, dois meses após o pedido de falência do banco de investimentos Lehman Brothers, em plena crise financeira.

Por outro lado, como pode ser observado no gráfico, o período de 1953 a 1973 apresenta valores consistentemente baixos, um período de maior estabilidade económica pós Segunda Guerra Mundial.

Esta análise permite identificar um padrão nas variáveis escolhidas: os períodos em que as variáveis tomam valores mais extremos correspondem a momentos de crise e

incerteza na vida política, social e económica. Nesse sentido, o sinal negativo das estimativas obtidas parece corresponder à realidade. Aumentos no valor das variáveis DFY e SVAR diminuem o valor do índice de cauda da distribuição de y^{prm} . Isto significa que, aumentos no valor de DFY ou SVAR, aumentam a probabilidade de ocorrência de valores extremos negativos no PRMA.

De seguida apresentam-se os efeitos parciais¹⁰ calculados para cada quantil considerado, bem como o índice de cauda médio ($\bar{\alpha}$) estimado.

TABELA V: Efeitos Parciais e Índice de Cauda Médio

Quantil	Efeito Parcial DFY	Efeito Parcial SVAR	$\bar{\alpha}$
5%	-32,2496	-41,0497	2,1965
10%	-11,4867	-26,2351	1,2249
15%	-3,3871	-11,3682	0,5882

Como se pode observar na Tabela V, à medida que o quantil considerado aumenta, a magnitude absoluta dos efeitos parciais e o valor médio do índice de cauda estimado diminuem. Na verdade, o valor estimado para $\bar{\alpha}$ é demasiado baixo para os pontos de corte de 10% e 15%. Estes valores parecem indicar a inconsistência das estimativas obtidas para estes casos, potenciadas pela consideração de observações que se afastam da hipótese inicial de que o PRMA segue uma distribuição com caudas de Pareto. No entanto, de notar que a estimação por QMV é robusta à escolha das variáveis explicativas para cada quantil considerado.

Face à estimativa mais plausível para o α médio, considera-se que o quantil mais adequado é o de 5%. Neste sentido, estima-se que quando a diferença entre o retorno de obrigações classificadas com BAA e AAA aumenta 1 ponto percentual, face à sua média, o índice de cauda diminui aproximadamente 0,3225 pontos, mantendo-se constante a volatilidade do mercado de ações. Já o aumento de 1 ponto percentual na média da volatilidade do mercado de ações, estima-se que diminua o índice de cauda aproximadamente em 0.4105 pontos, mantendo-se constante a variável DFY

¹⁰ Efeito parcial médio (EPM), calculado para a média das observações das variáveis explicativas, *i.e.*, $EPM_j = \theta_j * \exp(\theta_0 + \theta_1 DFY + \theta_2 SVAR)$, $j = 1, 2$.

O valor de α em função de cada uma das variáveis explicativas consideradas no modelo final é apresentado em cada um dos gráficos da Figura 4. Os gráficos refletem a relação inversa existente entre cada uma das variáveis explicativas e o índice de cauda α .

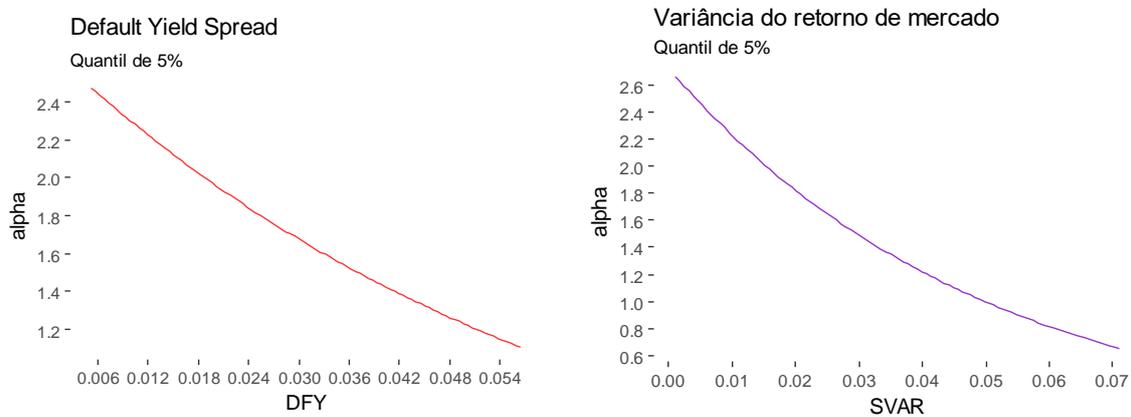


FIGURA 4: Valor do índice de cauda em função das variáveis explicativas.

Pelos gráficos, percebe-se que não é razoável admitir que os momentos de ordem 4 possam existir (valor de $\bar{\alpha}$ está bastante distante de 4).

Em suma, as conclusões mais importantes a retirar da estimação efetuada são as seguintes:

1. DFY e SVAR são duas variáveis que tendem a seguir de perto o ciclo económico e os períodos de incerteza económica, política e social que se atravessa. Ao longo do histórico de análise, o seu valor é substancialmente maior em períodos identificados como períodos de crise;
2. Os fatores mais relevantes para explicar os valores extremos negativos do PRMA identificam de alguma maneira os períodos de crise e incerteza. Esses fatores têm uma relação inversa com o índice de cauda α , ou seja, a ocorrência de valores extremos no PRMA, tende a ocorrer em períodos em que DFY e SVAR apresentam valores mais elevados;
3. O ponto de corte mais apropriado para o caso em concreto é o menor considerado (5%);
4. O índice de cauda estimado é relativamente baixo o que sugere que existe uma massa de probabilidade considerável associada a valores extremos na amostra.

3.2.5. Limitações e Considerações Finais

Tal como indicado na secção 2.3 do presente trabalho, o ponto de corte é essencial no processo de estimação do índice de cauda, já que estabelece o número de observações a considerar.

Os valores obtidos para a média dos α 's dos quantis de 10% e 15%, parecem indicar que à medida que o quantil da aba esquerda aumenta, a distribuição dos valores extremos de y^{prm} começa a afastar-se da hipótese inicial considerada: y^{prm} segue uma cauda de Pareto de tipo I. Uma forma *naïve* de avaliar este pressuposto inicial é calcular o valor médio da função log-Verosimilhança, para cada um dos quantis considerados. Os valores são apresentados na Tabela VI.

TABELA VI: Log-Verosimilhança Média por Quantil

Quantil (q)	Valor médio da função log-verosimilhança
5%	2,2688
10%	2,1786
15%	1,9022

Os valores mostrados supra indicam que as observações referentes aos 5% de valores mais baixos de y^{prm} são as que estão mais próximas do comportamento esperado da função densidade de Pareto, dado o maior valor (médio) obtido para a função log-verosimilhança. Este resultado vai ao encontro das conclusões anteriormente referidas. No entanto, reforça-se que se trata apenas de uma análise *naïve*, uma vez que os diferentes pontos de corte alteram o poder explicativo das variáveis.

Uma das limitações que pode ser apontada à metodologia aplicada é o facto de não ter em conta formas desconhecidas de dependência dos dados. Na verdade, assume-se neste trabalho que todo o tipo de dependências é capturado pelo modelo final. Esta hipótese é essencial para que as estimativas de QMV sejam consistentes e assintoticamente eficientes, dado que a média e variâncias condicionais têm de ser corretamente especificadas. Face à periodicidade mensal dos dados os efeitos ARCH podem estar relativamente atenuados (Nicolau, 2012).

Neste sentido, considera-se que a investigação futura acerca dos valores extremos deve prosseguir para os seguintes tópicos:

- 1) Desenvolvimento de uma metodologia de selecção do ponto de corte ótimo com justificação teórica;
- 2) Avaliação do impacto da estimação por QMV na escolha do ponto de corte;
- 3) Desenvolvimento de uma metodologia de regressão, função de variáveis explicativas, que tenha em conta potenciais dependências temporais dos dados;
- 4) Extensão da regressão do índice de cauda, função de variáveis explicativas, a outros tipos de distribuições de caudas pesadas;
- 5) Verificar em que medida as variáveis explicativas consideradas no modelo final podem antecipar valores extremos do PRMA e mesmo de crises financeiras.

4. CONCLUSÃO

Neste trabalho estudaram-se os valores extremos observados para o PRMA. Para o efeito é aplicada uma metodologia que consiste na regressão do índice de cauda da distribuição de Pareto, função de variáveis explicativas. Esta abordagem permite verificar se determinada variável está correlacionada com a ocorrência de valores extremos da variável dependente.

Neste sentido, conclui-se que as variáveis de maior relevância na explicação dos valores extremos observados no prémio de risco estão associadas a situações de crise e incerteza de teor social, político e económico do momento em análise. As referidas variáveis são o *Default Yield Spread*, calculado como sendo a diferença entre o retorno de obrigações classificadas com BAA e AAA e a volatilidade do mercado de ações. Assim, quando as variáveis descritas assumem valores mais elevados, estima-se que haja uma diminuição do valor do índice de cauda, que por sua vez implica uma maior probabilidade da ocorrência de eventos extremos. A forte probabilidade de ocorrência de valores extremos é evidente face ao baixo valor estimado para o α ($\alpha = 2,1965$). Este valor considera o ponto de corte de 5% das observações da aba esquerda da distribuição, dado que no caso em concreto existem evidências de que o ponto de corte mais apropriado é o menor considerado (5%, relativamente aos usados, 10% e 15%). Esta conclusão é retirada dos valores estimados para o α . De facto, o valor das estimativas para os quantis correspondentes a 10% e 15% das observações vêm demasiado baixos, indicador de que as estimativas são inconsistentes face à utilização na estimação de observações que se afastam das caudas. Para além disso, a análise *naïve* do valor médio da função log-verosimilhança, reforça a ideia de que o quantil de 5% é o mais apropriado.

Em síntese, quando se investe em ações, o risco assumido é real, principalmente se incluir momentos de crise ou catástrofe, *i.e.*, eventos que ocorram com pouca frequência (considerando um grande histórico em análise) mas que causem quedas abruptas no rendimento e no mercado de capitais. Desta forma, para que um investidor naturalmente avesso ao risco, o possa evitar ao máximo na seleção das suas carteiras de investimento, um dos comportamentos a ter em consideração, será o de avaliar o ciclo económico do

momento e tentar utilizar este conhecimento como vantagem relativamente ao restante mercado.

A metodologia considerada deixa espaço ao desenvolvimento de futura investigação. Esta deve visar a determinação do valor do ponto de corte ótimo e avaliar o impacto da estimação por QMV nessa escolha. No âmbito dos modelos de regressão do índice de cauda, estes devem ser desenvolvidos por forma a que se possam incorporar eventuais dependências temporais dos dados, e ainda considerar diferentes distribuições de caudas pesadas. De referir ainda que poderá ser interessante verificar o potencial previsionial das variáveis explicativas nos valores extremos do PRMA e/ou de crises financeiras.

REFERÊNCIAS

- Beirlant, J., Goegebeur, Y., Segers, J. & Teugels, J. (2004). *Statistics of extremes: theory and applications*, Ed. Wiley Series in Probability and Statistics.
- Bloom, N. (2009). The Impact of Uncertainty Shocks. *Econometrica* 77 (3), 623-85.
- Duarte, F. & Rosa, C. (2015). The Equity Risk Premium: A Review of Models. *Economic Policy Review* 2, 39-57.
- Fama, E. & French, K. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics* 25 (1), 23–49.
- Gourieroux, C. & Monfort, A. (1995). *Statistics and Econometric Models*, Ed. Cambridge University Press.
- Goyal, I & Welch, A. (2008). A Comprehensive Look at The Empirical Performance of Equity Premium Prediction. *The Review of Financial Studies* 21 (4), 1455–1508.
- Hill, B. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics* 3, 1163-1174.
- Hill, J. (2010). On tail index estimation for dependent, heterogeneous data. *Econometric Theory* 26, 1398–1436.
- Nicolau, J. (2012). *Modelação de Séries Temporais Financeiras*, Ed. Almedina.
- Nicolau, J. & Rodrigues, P. (2018). A New Regression-Based Tail Index Estimator. *The Review of Economics and Statistics*, to appear (2018).
- Rozeff, M. (1984). Dividend yields are equity risk premiums. *Journal of Portfolio Management* 11 (1), 68–75.
- Wang, H. & Tsai, Chih-Ling (2009). Tail Index Regression. *Journal of the American Statistical Association* 104 (487), 1233-1240.

ANEXOS

Anexo A

VARIÁVEIS EXPLICATIVAS TESTADAS

Dividend Price Ratio (DP), diferença entre o log dos dividendos e o log dos preços (índice S&P 500);

Dividend Yield (DY), diferença entre o log dos dividendos e o log dos preços desfasados (índice S&P 500);

Earnings Price Ratio (EPR), diferença entre log dos ganhos e o log dos preços (índice S&P 500);

Dividend Payout Ratio (DE), diferença entre log dos dividendos e o log dos ganhos (índice S&P 500);

Book-to-Market Ratio (BM), rácio entre valor contabilístico e valor de mercado (índice *Dow Jones*);

Net Equity Expansion (NTIS), rácio entre a soma de emissões líquidas em ações listadas na NYSE (*New York Stock Exchange*) e o valor total de capitalização de mercado no final do ano das ações da NYSE;

Treasury-bill rate (TBL);

Long-Term Government Bond Yield (LTY);

Long-Term Government Bond Rate of Returns (LTR);

Term Spread (TMS), diferença entre LTY e TBL;

Default Yield Spread (DFY), diferença entre o retorno de obrigações classificadas com BAA e AAA;

Default Return Spread (DFR), diferença entre obrigações de longo-prazo de empresas e LTR;

Inflação (INFL), Índice de Preços do Consumidor;

Variância do retorno do mercado acionista (SVAR), medida pela variância do índice S&P 500.

Para mais detalhe consultar Goyal & Welch (2008).

Anexo B

PERÍODOS DE MAIOR INCERTEZA ECONÓMICA POR BLOOM (2009)			
Evento	Volatilidade Máxima	Primeira Volatilidade	Tipo
Crise dos mísseis cubanos	Outubro 1962	Outubro 1962	Terror
Assassinato de JFK	Novembro 1963	Novembro 1963	Terror
Guerra do Vietname	Agosto 1966	Agosto 1966	Guerra
Estado de Kent, Cambodja	Mai 1970	Mai 1970	Guerra
OPEP I, Guerra Israelo-Árabe	Dezembro 1973	Dezembro 1973	Petróleo
Colapso do <i>Franklin National Bank</i>	Outubro 1974	Setembro 1974	Económico
OPEP II	Novembro 1978	Novembro 1978	Petróleo
Crise dos reféns americanos (diplomatas no Irão)	Março 1980	Março 1980	Guerra
Alteração na política monetária	Outubro 1982	Agosto 1982	Económico
Segunda-feira Negra	Novembro 1987	Outubro 1987	Económico
Guerra do Golfo I	Outubro 1990	Setembro 1990	Guerra
Crise Asiática	Novembro 1997	Novembro 1997	Económico
Crise Financeira Russa; Falência de LTCM	Setembro 1998	Setembro 1998	Económico
Ataques terroristas	Setembro 2001	Setembro 2001	Terror
Escândalos empresariais (Worldcom e Enron)	Setembro 2002	Julho 2002	Económico
Guerra do Golfo II	Fevereiro 2003	Fevereiro 2003	Guerra
Crise Financeira	Outubro 2008	Agosto 2007	Económico

Tabela apresentada em Bloom (2009).

Anexo C

Anexo C1

QUANTIL 5% - EXEMPLOS DE MODELOS TESTADOS					
Regressor	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo Final
Termo constante	-1,59740 (1,4543)	1,0416*** (0,3755)	1,1004*** (0,3313)	1,2334*** (0,2063)	1,2746*** (0,1900)
BM	0,7873* (0,4248)	–	–	–	–
DFR	1,4405 (5,4126)	0,3576 (5,3439)	–	–	–
DFY	-10,8474 (20,3997)	-13,4537 (13,5909)	-8,6801 (12,7549)	-18,9211** (7,6937)	-15,777** (7,2275)
DY	-0,6193** (0,3155)	–	–	–	–
INFL	26,9568 (32,1229)	23,9327 (30,6863)	22,2890 (29,7205)	–	–
LTR	-2,0375 (5,8168)	-0,9299 (6,1213)	–	–	–
SVAR	-13,1554** (6,1762)	-18,1483*** (5,2625)	-18,3057*** (5,5014)	-20,8010*** (4,0622)	-20,0821*** (4,3116)
TMS	8,0616 (6,6279)	7,6020 (5,8880)	–	5,6190 (5,6550)	–

Significância estatística: 0,01 ‘***’ 0,05 ‘**’ 0,1 ‘*’

Anexo C2

QUANTIL 10% - EXEMPLOS DE MODELOS TESTADOS					
Regressor	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo Final
Termo constante	0,8368 (0,8723)	0,59019*** (0,1388)	0,56355*** (0,1207)	0,51413*** (0,1064)	0,5180*** (0,0983)
BM	-0,0843 (0,2985)	–	–	–	–
DFR	-2,7343 (3,0032)	-2,6697 (2,9529)	–	–	–
DFY	-13,2669 (8,9255)	-13,8446** (6,4036)	-11,6750** (5,6841)	-10,1505* (5,5607)	-9,8040* (5,0073)
DY	0,0589 (0,2084)	–	–	–	–
INFL	-8,6693 (11,0759)	-9,5943 (11,1339)	-6,80601 (9,7274)	–	–
LTR	-2,3826 (2,3258)	-2,5485 (2,3451)	–	–	–
SVAR	-23,6645*** (5,1695)	-22,7318*** (4,3770)	-22,97230*** (4,0394)	-22,4916*** (3,7758)	-22,3919*** (3,7967)
TMS	0,2209 (4,9547)	0,3095 (4,2516)	–	0,5720 (4,0593)	–

Significância estatística: 0,01 ‘****’ 0,05 ‘***’ 0,1 ‘**’

Anexo C3

QUANTIL 15% - EXEMPLOS DE MODELOS TESTADOS					
Regressor	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo Final
Termo constante	-0,2373 (0,4943)	-0,2949*** (0,0759)	-0,3042*** (0,0671)	-0,3326*** (0,0577)	-0,3350*** (0,0512)
BM	-0,0795 (0,1812)	–	–	–	–
DFR	0,9039 (1,7232)	0,7290 (1,6683)	–	–	–
DFY	-4,4367 (5,4345)	-6,8695* (3,9016)	-7,2012** (3,5682)	-5,7912* (3,1875)	-5,8957** (3,0069)
DY	0,0112 (0,1190)	–	–	–	–
INFL	-3,3794 (6,3252)	-5,4938 (6,4171)	-5,3595 (6,1571)	–	–
LTR	0,5174 (1,1822)	0,4440 (1,1897)	–	–	–
SVAR	-21,1133*** (4,2599)	-20,0713*** (3,7287)	-20,0843*** (3,6900)	-19,6634*** (3,5508)	-19,7878*** (3,5945)
TMS	-0,9722 (2,7484)	-0,7376 (2,5905)	–	-0,2899 (2,4474)	–

Significância estatística: 0,01 ‘****’ 0,05 ‘***’ 0,1 ‘**’