

3

METODOLOGIA E AMOSTRA

3.1.

Descrição da Amostra

Foram utilizados o índice da Bolsa de Valores de São Paulo (Ibovespa) como representativo da carteira de mercado e os certificados de depósitos interfinanceiros (CDIs) como uma aproximação de um título livre de risco. Como já comentado sinteticamente na introdução, a amostra é formada pela série das cotações diárias do Ibovespa e pela série das taxas praticadas no mercado interfinanceiro (taxa DI), ambas entre 1º de junho de 1994 e 31 de julho de 2012. As cotações diárias do Ibovespa foram obtidas no banco de dados do sistema Economática e as taxas do CDI no site da Cetip¹.

A doutrina há algum tempo tem debatido qual seria a melhor taxa livre de risco no mercado brasileiro. A opção mais óbvia seria considerar como tal taxa uma média da remuneração oferecida pelos títulos públicos nacionais em determinado período. Todavia, como já apontado por Neto, Lima e Araújo (2008), nem sempre os títulos públicos são aceitos como ativos sem risco pelo mercado financeiro. Além disso, esses autores apontam que a enorme dispersão apresentada pela série histórica dos retornos dos títulos públicos desaconselha seu uso como taxa livre de risco. Na verdade, mesmo a taxa do Sistema Especial de Liquidação e Custódia (Selic) apresentou forte desvio-padrão entre 1995 e 2005, como mostram Neto et al. (2008).

Barros, Famá e Silveira (2002) concluem que no caso brasileiro os “retornos da caderneta de poupança ou dos certificados de depósitos interfinanceiros são condizentes com a conceituação teórica de uma taxa livre de risco, com correlação insignificante com o mercado e desvios-padrão dos retornos também desprezíveis”.

Diante dos achados da literatura precedente e tendo em vista que os depósitos interfinanceiros são operações com grande liquidez diária desde o

¹ www.cetip.com.br

começo da série temporal (jun/1994) e refletem o que os agentes de mercado entendem ser a taxa livre de risco, este trabalho adota a rentabilidade dos CDIs como taxa livre de risco.

Optou-se por realizar a análise a partir de junho de 1994 por ser este o momento da implementação do Plano Real, que, dada a magnitude das mudanças que operou, representa um marco de descontinuidade do arcabouço econômico-financeiro que vigorava no Brasil. Em situações desse tipo se torna difícil comparar as séries financeiras anteriores e posteriores ao plano econômico em termos nominais, sendo prudente trabalhar com séries temporais um pouco mais curtas, porém consistentes.

Como o enfoque da análise é mensal, a partir dos valores diários de fechamento do índice Bovespa e dos valores médios diários da taxa DI anual foram apurados os retornos mensais dos mercados acionário e interfinanceiro, bem como o retorno em excesso de mercado. Assim, as 4497 observações diárias se transformaram em 217 observações mensais, considerados 21 dias úteis em cada mês. A tabela 1 apresenta as estatísticas descritivas da amostra já em termos mensais. Apenas os valores do último dia útil de cada mês compõem as séries mensais. A amostra se estende até julho de 2012.

Tabela 1 - Estatísticas descritivas das séries mensais

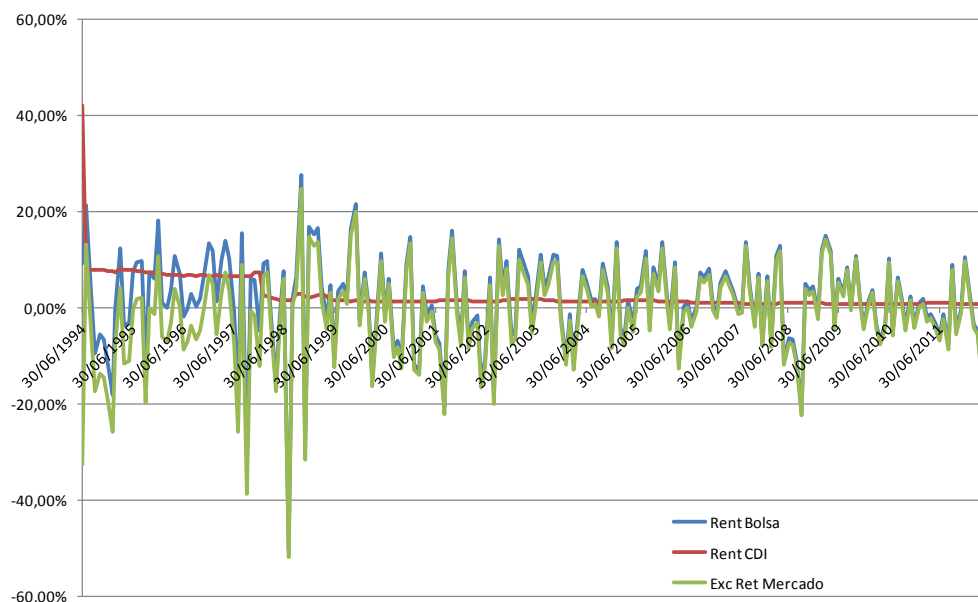
	Rentabilidade do Ibovespa	Rentabilidade do DI	Retorno de Mercado
Média	1,44%	2,78%	-1,34%
Mediana	1,77%	1,39%	-0,65%
Maior Valor	38,42%	42,09%	24,85%
Menor Valor	-50,34%	0,69%	-51,83%
Desvio Padrão	9,96%	4,38%	10,02%
Assimetria	-0,748	6,260	-1,131
Curtose	6,89	57,93	6,46
Jarque-Bera	156,74	24.420,59	154,63
Probabilidade	0,0000	0,0000	0,0000
Número de Observações	217	217	217

Apesar da média dos retornos da bolsa ser de 1,44% ao mês a série apresenta assimetria negativa, sinalizando que em muitos meses houve perdas relevantes no Ibovespa. A rentabilidade média da taxa DI foi superior à da renda variável no período.

Como consequência da composição dos retornos do Ibovespa e dos CDI, observa-se que o excesso de retorno de mercado apresenta média mensal negativa no intervalo analisado, ou seja, constata-se que no Brasil há um prêmio de risco

médio de “-1,34%” entre junho de 1994 e julho de 2012. O gráfico 1 mostra a evolução das três séries.

Gráfico 1 – Evolução dos retornos do Ibovespa, dos retornos do CDI e do excesso de retorno de mercado



Alguns pressupostos precisam ser atendidos pelas séries financeiras para que os modelos sejam livres de distorções e assim possam ser considerados confiáveis para se realizar previsões. Dentre esses pressupostos se destacam ser a série normalmente distribuída e estacionária. Para verificar se as três séries listadas na tabela 1 atendem a esses requisitos, foram realizados o teste de Jarque-Bera² (JB) e o teste *Augmented Dickey-Fuller* (ADF).

O teste de JB foi executado automaticamente pelo pacote estatístico Eviews e mostra que nenhuma das três séries apresenta distribuição próxima à normal padronizada, vide os elevados valores do teste de JB mostrados na tabela 1. A hipótese de normalidade foi rejeitada ao nível de 1% de significância nos três casos. As séries são assimétricas e leptocúrticas.

² O teste de Jarque-Bera é realizado por meio da fórmula $JB = T \left| \frac{A^2}{6} + \frac{(C-3)^2}{24} \right|$, onde A é a assimetria da amostra, C é a curtose da amostra e T é o número de observações.

3.1.1. Augmented Dickey-Fuller (ADF)

Também é desejável que as séries temporais sejam estacionárias quando utilizadas em regressões. Quando a variável dependente ou algum dos regressores não são estacionários, pode haver problemas de estimação, como a distorção dos coeficientes da regressão (isso leva ao que se chama de regressão espúria: regressões que não representam a realidade). A estacionariedade de uma série significa que ela absorve gradativamente as oscilações: numa série estacionária os choques presentes afetam muito menos os valores futuros das variáveis do que ocorreria no caso de uma série não estacionária.

Um dos testes mais aplicados para se avaliar a existência de raiz unitária numa série é o de Dickey-Fuller (DF). Partindo-se de um modelo autorregressivo simples, AR(1) sem drift:

$$y_t = \Phi y_{t-1} + u_t, \quad (6)$$

onde y_t representa a variável de interesse na data t , Φ é o coeficiente de y_t na data anterior e u_t é o termo de erro, pode-se verificar se a série possui raiz unitária, o que ocorre quando $\Phi=1$. O teste é realizado sob a hipótese nula de que a série possui raiz unitária ($\Phi=1$), contra a hipótese alternativa de que não há raiz unitária ($\Phi<1$). Sendo rejeitada H_0 conclui-se que a série é estacionária. Se não for possível rejeitar H_0 , trata-se de uma série que segue um passeio aleatório (*random walk*).

Substituindo-se o coeficiente Φ da equação 6 por “ $\rho+1$ ” chega-se à equação que é de fato utilizada pelos pacotes estatísticos para testar a existência da raiz unitária (eq.7). Como $\Phi= \rho+1$, a série terá raiz unitária e será não estacionária se $\rho=0$.

$$\begin{aligned} y_t &= (\rho+1)y_{t-1} + u_t \\ y_t &= \rho y_{t-1} + y_{t-1} + u_t \\ y_t - y_{t-1} &= \rho y_{t-1} + u_t \\ \Delta y_t &= \rho y_{t-1} + u_t \end{aligned} \quad (7)$$

Para que os resultados do teste de DF sejam confiáveis presume-se que os termos de erro (u_t) sejam um ruído branco, ou seja, que não apresentem autocorrelação. Todavia, nem sempre isto ocorre e pode haver autocorrelação nos termos de erro quando também houver correlação na variável dependente da regressão (Δy_t). Para contornar esse problema é utilizado o teste de Dickey-Fuller Aumentado (ADF), que remove os efeitos da autocorrelação pela utilização de *lags* (p) na variável dependente. O teste ADF segue um processo AR(p) e por meio de correções paramétricas nas autocorrelações, os coeficientes δ no modelo mostrado na equação 8, testa a hipótese da série ter raiz unitária. Da mesma forma que no teste padrão de Dickey Fuller, a série será não estacionária se $\rho=0$.

$$\Delta y_t = \rho y_{t-1} + \sum_{i=1}^p \delta \Delta y_{t-i} + u_t \quad (8)$$

Foi aplicado o teste ADF às séries temporais dos retornos mensais do Ibovespa, dos retornos do título sem risco e dos retornos em excesso de mercado. Como exposto na tabela 2, as probabilidades apontadas pelo ADF (p-valor) rejeitam fortemente a hipótese nula em todos os casos. Nenhuma das três séries tem raiz unitária. Portanto, elas são estacionárias e podem ser utilizadas nas regressões.

Tabela 2 - Resultados para o Augmented Dickey-Fuller Test (ADF)

	Rentab. do Ibovespa	Rentab. do CDI	Retorno de Mercado
Valor da Estatística do Teste (ADF)	-14,8057	-10,2241	-14,4586
Probabilidade (p-valor)*	0,0000	0,0000	0,0000
<u>Valores Críticos a:</u>			
1% de significância	-3,4606	-3,4606	-3,4606
5% de significância	-2,8747	-2,8747	-2,8747
10% de significância	-2,5739	-2,5739	-2,5739
Número de Observações (ajustadas)	216	216	216

*MacKinnon (1996) unicaldal p-valor

fonte: o autor

3.2. Metodologia

3.2.1. Retornos de Mercado

Foram calculados os retornos do Ibovespa e dos depósitos interfinanceiros entre o último dia útil de cada mês e o 21º dia útil precedente. No caso do Ibovespa foi usado o valor de fechamento e a seguinte fórmula:

$$RV_t = Ln \left[\frac{P_t}{P_{t-21}} \right] \quad (9)$$

onde RV_t é o retorno mensal apurado no último dia útil do mês, P_t é o valor de fechamento do Ibovespa no último dia útil do mês e P_{t-21} é o valor do fechamento do índice no 21º dia útil anterior. Para exemplificar, o retorno do Ibovespa em 29 de junho de 1994 é o logaritmo natural do valor de fechamento de 29 de junho de 1994 (P_t) sobre o valor de fechamento de 1º de junho de 1994 (P_{t-21}).

Já os retornos mensais dos depósitos interfinanceiros, nossa taxa livre de risco, foram calculados a partir da taxa DI anual vigente no 21º dia útil antecedente:

$$RF_t = Ln \left[1 + DI_{t-21} \right] * \left[\frac{21}{252} \right], \quad (10)$$

onde RF_t é a rentabilidade mensal do CDI e DI_{t-21} é a taxa anual do CDI vigente 21 dias úteis antes.

Apurados os retornos mensais da renda variável e do título sem risco, foi calculado o retorno em excesso do mercado para cada mês (R_t), que é a diferença entre as duas rentabilidades (é o prêmio de risco):

$$R_t = RV_t - RF_t \quad (11)$$

3.2.2. VaR não Paramétrico

A perda de riqueza em um ativo ou portfólio pode ser estimada pelo VaR, uma medida de risco facilmente compreendida e muito utilizada pelos agentes de mercado, como fundos de pensão e bancos de investimentos. Há diversos métodos para se estimar o VaR, que são divididos em dois grandes grupos: o paramétrico e o não paramétrico. São exemplos de métodos não paramétricos o de simulação de Monte Carlo e o histórico. Neste estudo foi escolhido o método histórico para se projetar o VaR não paramétrico.

O método histórico consiste em utilizar as séries temporais de retornos históricos dos ativos, aplicando-se às mesmas os pesos correntes que esses ativos têm no portfólio (Jorion, 1997):

$$R_{p,\tau} = \sum_{i=1}^N w_{i,t} R_{i,\tau}, \quad \tau = 1, \dots, t. \quad (12)$$

sendo $R_{p,\tau}$ o retorno do portfólio, $w_{i,t}$ o peso de cada ativo e $R_{i,\tau}$ o retorno de cada ativo. Com base nesses retornos históricos e nos pesos correntes de cada ativo, projeta-se uma distribuição hipotética de retornos e dado o nível de confiança é apontado o VaR do portfólio.

Neste estudo a metodologia para se obter o VaR não paramétrico pelo método histórico é parecida com a abordagem acima descrita, sendo no entanto mais simplificada. O primeiro passo foi apurar os retornos diários a partir dos valores de fechamento do índice Bovespa:

$$RDV_t = Ln \left[\frac{P_t}{P_{t-1}} \right] \quad (13)$$

onde RDV_t é o retorno diário do índice, P_t é valor de fechamento do dia e P_{t-1} é o valor de fechamento do dia útil anterior.

Seguindo as orientações de Bali et al. (2009), o VaR não paramétrico é simplesmente a maior perda diária do Ibovespa em determinado intervalo de meses precedentes (em termos percentuais).

Foram montadas então seis séries de VaRs não paramétricos para serem utilizadas nas regressões. Apenas os VaRs não paramétricos apurados no último dia útil de cada mês entram nas séries e são efetivamente usados como regressores.

O VaR não paramétrico estimado pelo método histórico se localiza na cauda esquerda da distribuição empírica, tendo sinal negativo. Todavia, como já explicado no capítulo 2, a teoria financeira fundamenta que uma das características esperadas das medidas de risco é que elas tenham uma relação positiva com os retornos esperados (quanto maior o risco, maior deve ser o retorno oferecido pelo ativo). Então, o VaR não paramétrico que é utilizado nas regressões é o valor encontrado na cauda esquerda da distribuição empírica multiplicado por menos um (-1).

O painel A da tabela 3 mostra as estatísticas referentes a cada uma das seis séries do VaR não paramétrico. Dependendo do número de dias úteis utilizados para se estimar o VaR de cada intervalo, os níveis de confiança ($1-\alpha$) variam entre 95,24% e 99,21%. Assim, por exemplo, o VaR estimado com base em três meses é a maior perda diária ocorrida nos 63 dias úteis anteriores à data de apuração. Essa metodologia equivale a adotar graus de significância proporcionais ao número de dias úteis utilizados para a apuração do VaR: no exemplo de três meses (63 dias úteis), a pior perda potencial será a apontada ao nível de significância de 1,59% ou “1/63” (o que representa um grau de confiança de 98,41%).

O VaR não paramétrico médio estimado com base em um mês é de 3,93%, aumentando gradativamente até atingir 6,63% quando a janela usada na apuração é de seis meses.

Tabela 3 - O painel A apresenta as estatísticas descritivas do VaR não paramétrico, apurado pelo método histórico. O painel B mostra os resultados das regressões usando modelos autorregressivos de dimensão 1, incluindo as estatísticas t ajustadas por White (entre parênteses nas segundas linhas)

Painel A - Estatísticas Descritivas						
VaR (nº meses)	VaR 1m	VaR 2m	VaR 3m	VaR 4m	VaR 5m	VaR 6m
Nível de significância (α)	4,76%	2,38%	1,59%	1,19%	0,95%	0,79%
Média	3,93%	4,80%	5,39%	5,88%	6,27%	6,63%
Mediana	3,36%	3,95%	4,44%	4,84%	5,08%	5,40%
Maior Valor	17,23%	17,23%	17,23%	17,23%	17,23%	17,23%
Menor Valor	0,86%	1,30%	1,78%	2,00%	2,12%	2,39%
Desvio Padrão	2,46%	2,80%	3,07%	3,32%	3,45%	3,55%
Assimetria	2,266	2,021	1,797	1,653	1,520	1,410
Curtose	10,034	7,913	6,473	5,552	4,948	4,515
Número de Observações	217	216	215	214	213	212

Painel B - Regressões AR(1) do VaR						
VaR (nº meses)	VaR 1m	VaR 2m	VaR 3m	VaR 4m	VaR 5m	VaR 6m
Intercepto(λ)	0,0393	0,0473	0,0521	0,0564	0,0599	0,0630
	(13,89)	(10,29)	(8,59)	(7,50)	(7,01)	(6,64)
Coefficiente AR(1), ρ^*	0,4801	0,7163	0,8112	0,8565	0,8776	0,8934
	(6,26)	(10,18)	(13,66)	(14,64)	(16,83)	(19,47)
Número de Observações	217	216	215	214	213	212
R^2 Ajustado	22,3%	51,4%	66,6%	73,9%	77,4%	80,0%

* todos os coeficientes ρ são significativos a 1%

O painel B da tabela 3 mostra os resultados de regressões executadas com modelos AR(1) para o VaR não paramétrico:

$$VaR_t = \lambda + \rho VaR_{t-1} + u_t, \quad (14)$$

sendo λ o intercepto e u_t o termo de erro. Os coeficientes (ρ) mostram que o VaR corrente em “t” é fortemente influenciado pelo VaR do momento anterior “t-1”, principalmente a partir do VaR apurado utilizando-se dois meses, com estatísticas t variando entre 6,26 e 19,27. Todas as regressões utilizadas nesta pesquisa foram executadas corrigindo-se pela abordagem de White para heterocedasticidade. O componente autorregressivo é significativo a 1% em todas as séries a que se refere o painel B da tabela 3. Diante disso, optou-se por usar o VaR_t como uma aproximação do VaR condicional ($E[VaR_{t+1}]$) nas regressões que testam a relação entre o retorno esperado e o *downside risk*, abordadas no capítulo 4.

3.2.3. VaR Paramétrico

Segundo Jorion (1997), os modelos paramétricos do VaR são construídos a partir do desvio-padrão da distribuição dos retornos do ativo/portfólio e da utilização de um fator multiplicativo que depende do nível de significância que se pretenda adotar. No caso de ativos ou portfólios cujos retornos históricos sejam normalmente distribuídos, a distribuição teórica ideal para estimar o VaR paramétrico é a normal. Além de oferecer propriedades estatísticas interessantes a normal também é a distribuição utilizada pelo *riskmetrics* (o mais difundido software de avaliação de risco na atualidade, criado pelo banco de investimentos JP Morgan).

Para estimar o VaR paramétrico por meio da normal padronizada (média igual a zero e desvio-padrão unitário) é necessário identificar a quantidade de desvios-padrão associada ao nível de significância desejado. Essa relação se dá de acordo com a função densidade de probabilidade:

$$\alpha = \int_{-\infty}^{-z} \phi(\tau) d\tau \quad (15)$$

sendo $\phi(\tau)$ a distribuição normal padronizada, $-z$ é a quantidade de desvios-padrão na cauda esquerda da distribuição e α é o grau de significância. Na prática o fator $-z$ é obtido a partir de tabelas de probabilidades acumuladas para a normal padronizada.

Identificado o valor crítico na distribuição teórica ($-z$), basta multiplicá-lo pela volatilidade do ativo/portfólio (σ) e pelo valor do ativo/portfólio (w) para se chegar ao VaR paramétrico:

$$VaR = -z\sigma w \quad (16)$$

Entretanto, a hipótese de que as séries financeiras sigam um padrão normal de distribuição é fortemente rejeitada nos testes empíricos, sendo um fato estilizado a não normalidade dessas séries. Isso sugere que distribuições que

levem em conta essa não normalidade sejam mais apropriadas aos estudos envolvendo séries financeiras.

Aplicando uma expansão de Cornish-Fisher (1937), Favre e Galeano (2002) computam o VaR analiticamente, ou seja, ajustam o VaR paramétrico tradicional com base na normal padronizada para considerar os impactos da assimetria e do excesso de curtose. O ajuste se dá no valor crítico da estatística “z”, de acordo com a equação:

$$Z_{cf} = Z_c + \frac{1}{6}(Z_c^2 - 1)S + \frac{1}{24}(Z_c^3 - 3Z_c)K - \frac{1}{36}(2Z_c^3 - 5Z_c)S^2, \quad (17)$$

onde Z_{cf} é o valor crítico modificado, Z_c é o valor crítico na distribuição normal padronizada para o nível de confiança almejado $(1-\alpha)$, S é a assimetria e K é o excesso de curtose (curtose maior que 3). Seguindo essa abordagem, neste estudo foram ajustados os valores críticos “z” obtidos da distribuição normal padronizada antes de utilizá-los para estimar o VaR paramétrico (que será tratado como VaRP para distingui-lo do VaR não paramétrico). Neste trabalho foi ignorado o valor da riqueza (w), pois o enfoque é o VaRP em percentual de perdas do Ibovespa. Assim, chega-se à equação que foi usada para achar o VaR paramétrico:

$$VaRP = -Z_{cf}\sigma \quad (18)$$

sendo σ a volatilidade diária histórica do Ibovespa, calculada por meio das janelas móveis em intervalos de um a seis meses. Os VaRPs encontrados por meio da equação 18 são multiplicados por menos um (-1) antes de serem utilizados nas regressões.

Na metodologia adotada neste trabalho, a quantidade de dias úteis referentes a cada intervalo também é a chave para se chegar aos níveis de significância utilizados para se obter os valores críticos de “z”: é o número de dias úteis que determina o fator z_c que será obtido da normal padronizada. A título de exemplo, o nível de significância utilizado no VaRP de 21 dias úteis é de 4,76% ($= 1/21$), o que leva a um fator Z_c de 1,668. Essa lógica se aplica a todas as amplitudes de estimação do VaRP. A tabela 4 mostra as estatísticas relacionadas às seis séries de VaRPs geradas utilizando-se a equação 18.

Tabela 4 - O painel A apresenta as estatísticas descritivas do VaR paramétrico (VaRP), apurado com base na distribuição normal, com ajustes para assimetria e excesso de curtose. O painel B mostra os resultados das regressões usando modelos autorregressivos de dimensão 1, incluindo as estatísticas "t" ajustadas por White (entre parênteses nas segundas linhas)

Painel A - Estatísticas Descritivas						
VaRP (nº meses)	VaRP 1m	VaRP 2m	VaRP 3m	VaRP 4m	VaRP 5m	VaRP 6m
Nível de significância (α)	4,76%	2,38%	1,59%	1,19%	0,95%	0,79%
Média	3,35%	4,00%	4,41%	4,73%	5,00%	5,24%
Mediana	2,91%	3,55%	3,89%	4,18%	4,38%	4,58%
Maior Valor	12,07%	11,94%	11,63%	11,67%	11,42%	11,59%
Menor Valor	1,18%	1,67%	1,74%	2,07%	2,28%	2,44%
Desvio Padrão	1,72%	1,83%	1,91%	1,97%	2,03%	2,13%
Assimetria	2,048	1,765	1,621	1,517	1,393	1,300
Curtose	8,443	6,640	5,737	5,097	4,463	3,962
Número de Observações	217	216	215	214	213	212

Painel B - Regressões AR(1) do VaRP						
VaRP (nº meses)	VaRP 1m	VaRP 2m	VaRP 3m	VaRP 4m	VaRP 5m	VaRP 6m
Intercepto(λ)	0,0331 (14,18)	0,0389 (12,20)	0,0421 (10,22)	0,045 (9,85)	0,0468 (9,62)	0,0492 (9,39)
Coefficiente AR(1), ρ^*	0,6122 (9,49)	0,7742 (12,29)	0,8549 (16,32)	0,8772 (14,84)	0,8915 (18,01)	0,8933 (20,48)
Número de Observações	217	216	215	214	213	212
R^2 Ajustado	37,8%	62,0%	74,3%	77,6%	80,6%	80,5%

* todos os coeficientes ρ são significativos a 1%

Os valores médios do VaRP oscilam entre 3,35% e 5,24%, sendo um pouco menores do que os do VaR não paramétrico. A volatilidade (desvio-padrão) do VaRP também é menor.

O painel B da tabela 4 traz os resultados de regressões executadas com modelos AR(1) para o VaRP:

$$VaRP_t = \lambda + \rho VaRP_{t-1} + u_t, \quad (19)$$

onde λ é o intercepto e u_t é o termo de erro. Os coeficientes ρ mostram que o VaRP corrente em "t" é fortemente influenciado pelo VaRP do momento anterior "t-1", principalmente a partir do VaRP obtido utilizando-se dois meses. Em todas as séries o componente autorregressivo é significativo a 1%. Diante disso, da mesma forma que no caso do VaR não paramétrico, optou-se por usar o $VaRP_t$ como uma aproximação do VaRP condicional ($E[VaRP_{t+1}]$) nas regressões que testam a relação entre o retorno esperado e o *downside risk* (capítulo 4).