

Université de Montréal

Rapport de recherche

*Comparaison des modèles HAR-RV et HAR-CJ
sur les rendements de AMZN et BAC en 2010*

Rédigé par :
Mame Diarra Cissé

Dirigé par :
Ilze Kalnina

Département de sciences économiques
Faculté des arts et des sciences

30 avril 2013

Table des matières

1	Introduction	1
2	Contexte du travail	2
2.1	Objectif général	2
2.2	Objectifs spécifiques	2
3	Brève revue littérature	2
4	Création des données	2
5	Présentation des modèles	5
6	Estimation et Comparaison des modèles	5
6.1	Estimation	5
6.2	Comparaison des modèles	7
7	Conclusion	9
8	Annexes	11

Liste des tableaux

1	Estimation HAR-RV Corsi 2009 : AMZN	5
2	Estimation HAR-RV Corsi 2009 : BAC	6
3	Estimation HAR-CJ Corsi, Pirino, Renò 2010 : AMZN	6
4	Estimation HAR-CJ Corsi, Pirino, Renò 2010 : BAC	7
5	Résumé des résultats d'estimation	8

Table des figures

1	Graphique de volatilité réalisée pour AMZN	3
2	Graphique de volatilité réalisée corrigée des valeurs atypiques pour AMZN .	3
3	Graphique de volatilité réalisée pour BAC	4
4	Graphique de volatilité réalisée corrigée des valeurs atypiques pour BAC . .	4
5	Graphique de corrélogramme de la volatilité réalisée corrigée des valeurs atypiques pour AMZN	11
6	Graphique de corrélogramme de la volatilité réalisée corrigée des valeurs atypiques pour BAC	11
7	Graphique de volatilité MinRV pour AMZN	11
8	Graphique de volatilité MinRV corrigée des valeurs atypiques pour AMZN .	12
9	Graphique de volatilité MinRV pour BAC	12
10	Graphique de volatilité réalisée corrigée des valeurs atypiques pour BAC . .	13

1 Introduction

Le modèle HAR-RV (Heterogeneous Autoregressive model of Realized Volatility) est un modèle simple de type AR en volatilité réalisée prenant en considération différents composantes de la volatilité réalisée sur différents horizons de temps (jour, semaine, mois, etc.). Ce modèle capte avec succès les caractéristiques empiriques des rendements financiers (longue mémoire, queues de distribution épaisses, l'autosimilarité) de manière abordable et fournit une bonne performance dans les prévisions (Corsi 2009).

L'idée de base de ce modèle est que les agents avec différents horizons temporels perçoivent, réagissent et causent différemment les composantes de la volatilité : les agents aux fréquences journalières ou des fréquences d'échanges plus élevées, les investisseurs d'horizons moyens hebdomadaires et les agents avec des horizons de temps mensuels. Le modèle HAR-RV permet de révéler la dynamique des différents composants de marché. Il est capable de produire la dynamique des rendements et des volatilités qui reproduisent ce qui s'observe dans les données.

Ce travail estime le modèle HAR-RV sur les rendements des actifs AMZN et BAC sur l'année 2010. Il a été comparé avec une variante incluant les sauts journaliers HAR-CJ (Heterogeneous Autoregressive with Continuous volatility and Jumps) en vue de retenir celui qui permet les meilleures prévisions, la qualité de la prévision étant mesurée par deux quantités HRMSE et QLIKE.

2 Contexte du travail

2.1 Objectif général

L'objectif du travail est de comparer les modèles HAR-RV de Corsi (2009) et de Corsi Reno (2010) sur les séries de données de l'année 2010 pour les actifs AMZN et BAC.

2.2 Objectifs spécifiques

Spécifiquement, il s'agit de :

1. Créer les données de cinq minutes
2. Calculer la volatilité réalisée journalière, hebdomadaire, mensuelle.
3. Estimer les deux modèles et choisir le meilleur selon les critères HMRSE et QLIKE.

3 Brève revue littérature

Ce travail s'appuie principalement sur les trois papiers suivants :

- Corsi (2009) ;
- Corsi, Pirino et Renò (2010) ;
- Andersen Dobrislar et Shaumburg (2010) (dorénavant ADS 2010).

Le modèle 3 HAR-RV est tiré de Corsi (2009) à la page 177, le modèle 4 HAR-CJ est tiré de Corsi, Pirino et Renò (2010) à la page 282. Pour la comparaison, la partie continue n'est pas calculée de la même manière que Corsi, Pirino et Renò (2010). Nous avons utilisé l'approche de ADS (2010) à la page 7. Le saut journalier est la différence entre la volatilité réalisée et la partie continue (Voir équation 5 ci-dessous pour plus de détails).

4 Création des données

Dans un travail préliminaire, nous avons extrait les prix de cinq minutes.

$$RV_t^{(d)} = \sqrt{\sum_{j=0}^{M-1} r_{t-j.\Delta}^2} \quad (1)$$

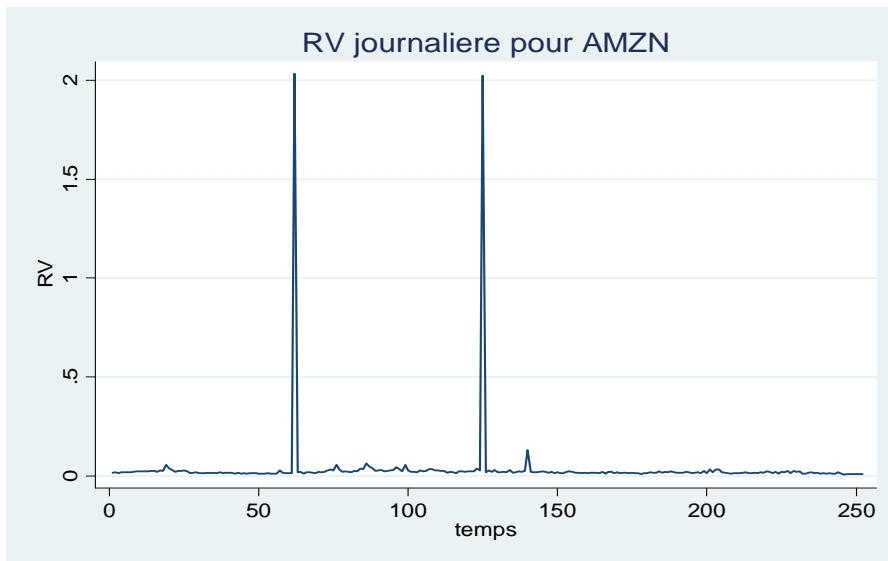
où $\Delta=5$ min ici, est la fréquence des rendements ; $r_{t-j.\Delta} = \ln P(t-j.\Delta) - \ln P(t-(j+1).\Delta)$ est le rendement de fréquence Δ . On a $M = 79$ rendements par jour.

Ensuite, nous avons calculé les volatilités réalisées journalières RV , hebdomadaires RV_w et mensuelles RV_m en utilisant les formules (3) et (4) de la page 177 de Corsi (2009).

$$RV_t^{(w)} = \frac{1}{5}(RV_t^{(d)} + RV_{t-1d}^{(d)} + \dots + RV_{t-4d}^{(d)}) \quad (2)$$

Nous disposons au total de 252 RV journalières.

FIGURE 1 – Graphique de volatilité réalisée pour AMZN



L'analyse du graphique de la série AMZN (figure 1) montre la présence de valeurs atypiques dans la RV journalière, lesquelles valeurs ont été remplacées par 0 (figure 2).

FIGURE 2 – Graphique de volatilité réalisée corrigée des valeurs atypiques pour AMZN

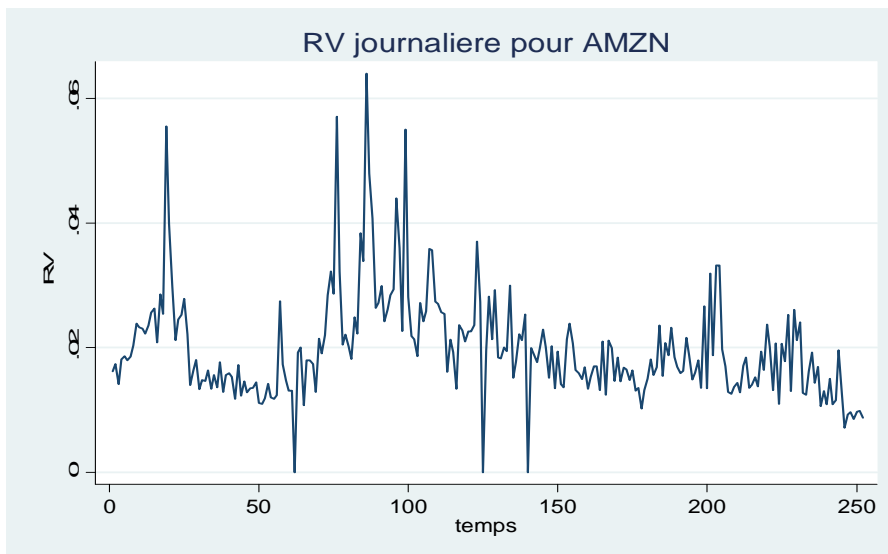
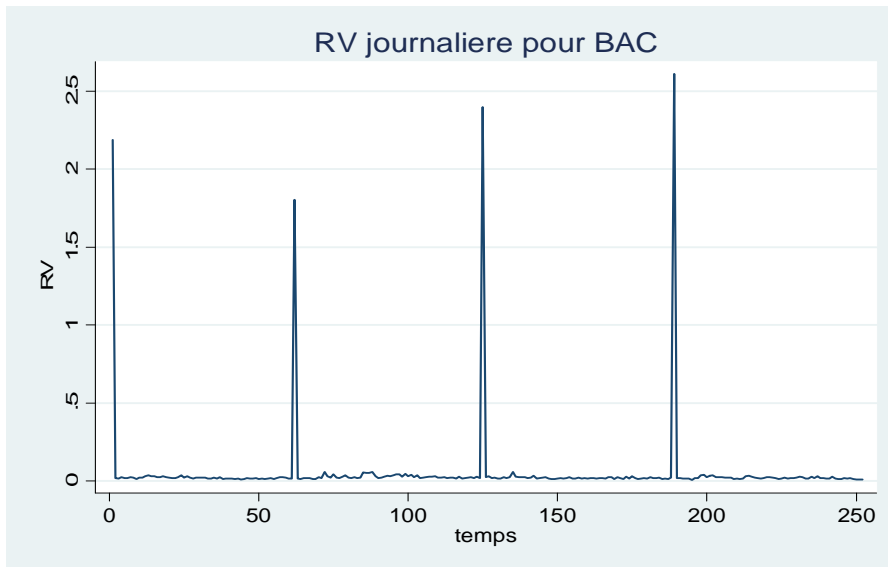
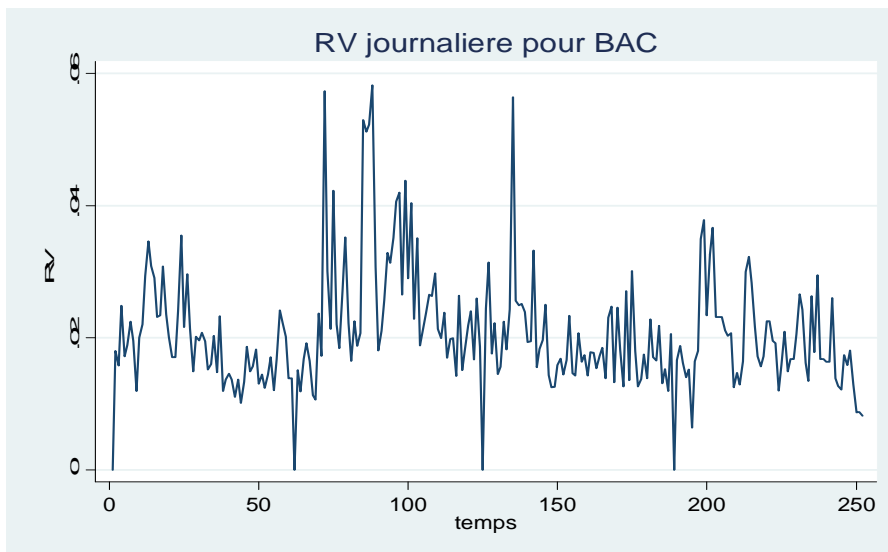


FIGURE 3 – Graphique de volatilité réalisée pour BAC



Le graphique de la série BAC (figure 3) quant à lui présente quatre valeurs atypiques, lesquelles ont également remplacées par 0 (figure 4).

FIGURE 4 – Graphique de volatilité réalisée corrigée des valeurs atypiques pour BAC



L'analyse des corrélogrammes de la volatilité réalisée pour les deux actifs révèle l'autocorrélation d'ordre 1 (voir figures 5 et 6).

5 Présentation des modèles

$$RV_{t+1d}^{(d)} = c + \beta^{(d)}RV_t^d + \beta^{(w)}RV_t^{(w)} + \beta^{(m)}RV_t^{(m)} + w_{t+1d} \quad (3)$$

$$RV_{t:t+h-1} = \beta_0 + \beta_d\widehat{C}_{t-1} + \beta_w\widehat{C}_{t-5:t-1} + \beta_m\widehat{C}_{t-22:t-1} + \beta_j\widehat{J}_{t-1} + \varepsilon_t \quad (4)$$

$$\widehat{C}_t = RV_t - \widehat{J}_t \quad (5)$$

$$\widehat{C}_t = MinRV = \frac{\pi}{\pi - 2} \left(\frac{M}{M - 1} \right) \sum_{i=1}^{M-1} \min(|r_{i\Delta}|, |r_{i+1\Delta}|)^2 \quad (6)$$

Dans notre cas M= 79 rendements par jour pour les données de cinq minutes. Le \widehat{C}_t calculé en utilisant l'équation 6 présente 4 valeurs atypiques pour chacun des deux actifs (voir figures 7 et 9), lesquelles valeurs ont été remplacées pour 0 pour donner de l'allure aux estimations (voir figures 8 et 10).

6 Estimation et Comparaison des modèles

6.1 Estimation

TABLE 1 – Estimation HAR-RV Corsi 2009 : AMZN

Source	SS	df	MS			
Model	.005961661	3	.00198722	Number of obs =	232	
Residual	.010263531	228	.000045015	F(3, 228) =	44.15	
				Prob > F =	0.0000	
				R-squared =	0.3674	
				Adj R-squared =	0.3591	
Total	.016225192	231	.000070239	Root MSE =	.00671	

RV	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
RV						
L1.	.3332978	.0767603	4.34	0.000	.1820474	.4845481
RVw						
L1.	.233139	.1278957	1.82	0.070	-.0188696	.4851477
RVm						
L1.	.253263	.1326383	1.91	0.057	-.0080905	.5146165
_cons	.0031895	.0017685	1.80	0.073	-.0002951	.0066741

TABLE 2 – Estimation HAR-RV Corsi 2009 : BAC

Source	SS	df	MS			
Model	.00503319	3	.00167773	Number of obs = 232		
Residual	.013720479	228	.000060178	F(3, 228) = 27.88		
				Prob > F = 0.0000		
				R-squared = 0.2684		
				Adj R-squared = 0.2588		
Total	.018753669	231	.000081185	Root MSE = .00776		

RV	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
RV						
L1.	.374494	.0759585	4.93	0.000	.2248236	.5241643
RVw						
L1.	.1270325	.1379427	0.92	0.358	-.1447729	.3988379
RVm						
L1.	.2419316	.1550062	1.56	0.120	-.0634962	.5473594
_cons	.0051467	.0023135	2.22	0.027	.0005882	.0097053

TABLE 3 – Estimation HAR-CJ Corsi, Pirino, Renò 2010 : AMZN

. reg RV LMinRV LMinRVw LMinRvm LJ

Source	SS	df	MS			
Model	.005805306	4	.001451327	Number of obs = 231		
Residual	.022458079	226	.000099372	F(4, 226) = 14.60		
				Prob > F = 0.0000		
				R-squared = 0.2054		
				Adj R-squared = 0.1913		
Total	.028263385	230	.000122884	Root MSE = .00997		

RV	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
LMinRV	2.221773	.7384199	3.01	0.003	.7667048	3.676842
LMinRVw	-1.76646	.9302455	-1.90	0.059	-3.599524	.066604
LMinRvm	-.3028049	.2304307	-1.31	0.190	-.7568724	.1512626
LJ	.1541845	.0716698	2.15	0.033	.0129581	.295411
_cons	.0110698	.0023982	4.62	0.000	.0063442	.0157955

TABLE 4 – Estimation HAR-CJ Corsi, Pirino, Renò 2010 : BAC

reg RV LMinRV LMinRVw LMinRVm LJ						
Source	SS	df	MS	Number of obs = 231		
Model	.005320937	4	.001330234	F(4, 226) =	22.40	
Residual	.013419657	226	.000059379	Prob > F =	0.0000	
				R-squared =	0.2839	
				Adj R-squared =	0.2713	
Total	.018740594	230	.000081481	Root MSE =	.00771	

RV	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
LMinRV	.9193858	.4318983	2.13	0.034	.0683232	1.770448
LMinRVw	-.2497487	.5539104	-0.45	0.653	-1.341238	.8417408
LMinRVm	-.3176542	.1452493	-2.19	0.030	-.6038704	-.031438
LJ	.3594717	.073297	4.90	0.000	.2150388	.5039046
_cons	.0107564	.0018589	5.79	0.000	.0070935	.0144193

6.2 Comparaison des modèles

Les résultats d'estimation sont résumés dans le tableau 5. Les coefficients $\beta^{(d)} = 0.33$, $\beta^{(w)} = 0.23$, $\beta^{(m)} = 0.25$ pour l'actif AMZN et $\beta^{(d)} = 0.33$, $\beta^{(w)} = 0.23$, $\beta^{(m)} = 0.25$ pour l'actif BAC sont proches des valeurs qu'on retrouve dans la littérature, notamment ceux utilisés par Corsi (2009) pour calibrer le modèle à savoir $\beta^{(d)} = 0.36$, $\beta^{(w)} = 0.28$, $\beta^{(m)} = 0.28$

Pour évaluer la performance des prévisions des différents modèles, on utilise (a) le R^2 des régressions de prévision de Mincer-Zarnowitz ; (b) l'hétéroscédasticité ajustée de la racine carrée de l'erreur quadratique moyenne suggérée par Bollerslev et Ghysels (1996) :

$$HMRSE = \sqrt{\frac{1}{T} \sum_{t=1}^T \left(\frac{RV_t - \widehat{RV}_t}{RV_t} \right)^2} \quad (7)$$

où RV_t est la valeur de la volatilité réalisée et \widehat{RV}_t est la valeur prédite par le modèle ; (c) la fonction de perte QLIKE définie par

$$QLIKE = \frac{1}{T} \sum_{t=1}^T \left(\log(RV_t + \frac{\widehat{RV}_t}{RV_t}) \right) \quad (8)$$

QLIKE est robuste tandis que HRMSE ne l'est pas (Patton, forthcoming). Ces deux quantités ont été calculées pour les 52 dernières observations pour chacun des actifs.

TABLE 5 – Résumé des résultats d'estimation

	AMZN		BAC	
	HAR-RV	HAR-CJ	HAR-RV	HAR-CJ
β^0	0.0032*	0.011***	0.0051**	0.011***
$\beta^{(d)}$	0.3333***	2.222***	0.3745***	0.919**
$\beta^{(w)}$	0.2331	-1.766*	0.1270	-0.2497
$\beta^{(m)}$	0.2533	-0.303	0.2419	-0.3177**
$\beta^{(j)}$	-	0.154**	-	0.359***
R^2	0.3674	0.2054	0.2684	0.2839
HMRSE	0.0488	0.0469	0.0403	0.0394
QLIKE	-2.9300	-2.9578	-2.8709	-2.8875

1.* significant at 10%, ** significant at 5%, *** significant at 1% .

Pour la série des donnée AMZN, la valeur du HMRSE vaut 0.0488 avec le modèle HAR-RV de Corsi (2009) alors quelle vaut 0.0469 avec le modèle HAR-CJ de Corsi Pirino Renò (2010). Il ressort donc que le deuxième modèle minimise l'erreur de prévision. Ce résultat est confirmé par la valeur de QLIKE qui vaut -2.9300 pour le premier modèle et -2.9578 pour le second. Donc le modèle de Corsi Pirino Renò (2010) est celui retenu pour la série de données de AMZN.

Quant aux résultats concernant la série de données de BAC, HRMSE vaut 0.0403 pour le premier modèle et 0.0394 pour le second, tandis que QLIKE vaut -2.8709 pour le premier et -2.8875 pour le second. La conclusion est la même que précédemment : Le modèle HAR-CJ de Corsi Pirino Renò minimise les erreurs de prévision. Il est donc celui retenu également pour la série de données de BAC.

7 Conclusion

L'objectif du présent travail a consisté en une comparaison des modèles HAR-RV et HAR-CJ sur les rendements des actifs AMZN et BAC sur l'année 2010. Pour ce faire, il a d'abord été question de récupérer les données de cinq minutes, ensuite calculer la volatilité réalisée RV en utilisant la formule 1, puis estimer les modèles 3 et 4 par la méthode des moindres carrés (MCO).

Quoique la qualité des régressions n'est pas très bonne (R^2 autour de 0.30), les coefficients obtenus sont proches de ceux utilisés dans la littérature. Les critères HRMSE et QLIKE calculés pour 52 prévisions permettent de retenir le modèle HAR-CJ comme celui qui fournit les meilleures prévisions, c'est-à-dire qui minimise les erreurs de prévision.

Références

- [1] **Fulvio Corsi (2009)**, « *A simple approximate long-memory model realized volatility* » journal of financial econometrics 2009, vol7, numero 2, 174-196.
- [2] **Fulvio Corsi, Davide Pirino, Roberto Reno (2010)**, « *threshold bipower variation and impact of jumps on volatility forecasting* », journal of econometrics ,159 (2010) 276-288.
- [3] **Torben G. Andersen ,Dobrislav Dobrev, Ernst Schaumburg (1974)**, « *Jump-robust volatility estimation using nearest neighbor truncation* », Federal reserve bank of new York , Staff reports numero 465, August 2010.

8 Annexes

FIGURE 5 – Graphique de corrélogramme de la volatilité réalisée corrigée des valeurs atypiques pour AMZN

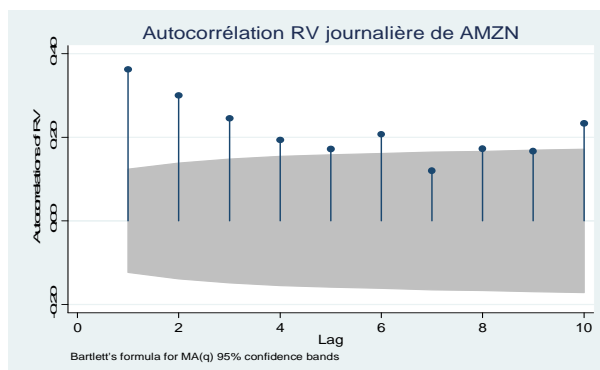


FIGURE 6 – Graphique de corrélogramme de la volatilité réalisée corrigée des valeurs atypiques pour BAC

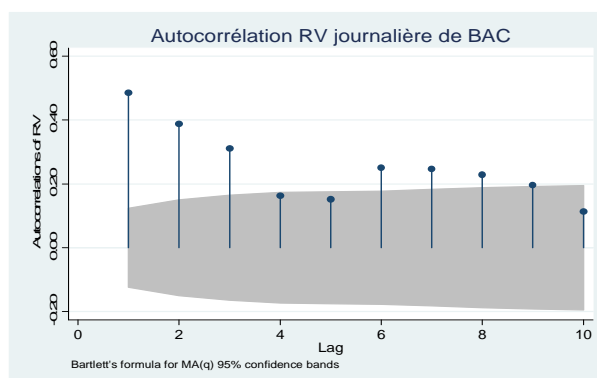


FIGURE 7 – Graphique de volatilité MinRV pour AMZN



FIGURE 8 – Graphique de volatilité MinRV corrigée des valeurs atypiques pour AMZN

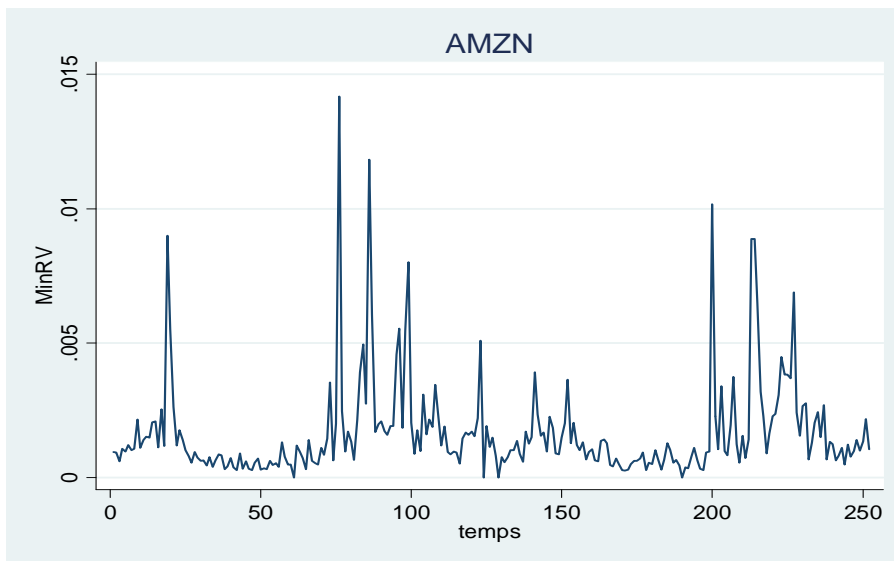


FIGURE 9 – Graphique de volatilité MinRV pour BAC

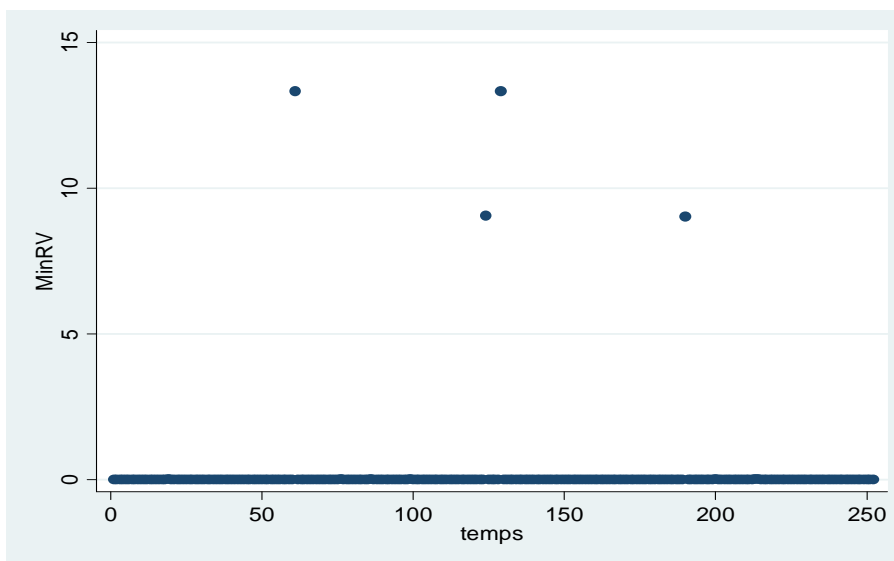


FIGURE 10 – Graphique de volatilité réalisée corrigée des valeurs atypiques pour BAC

