Supplementary material

On the role of interest rate differentials in the dynamic asymmetry of exchange rates

J. Hambuckers[1,†] and M. Ulm[1]

[1] University of Liège - HEC Liège

† jhambuckers@uliege.be

Preamble

In this document, we provide additional information to the paper "On the role of interest rate differentials in the dynamic asymmetry of exchange rates". The material is structured as follows: in Section 1, we study the finite sample properties of our econometric approach by means of a simulation study. Section 2 includes additional results on coefficient estimates, model selection and model performance for EUR and CHF currencies. In Section 3, we conduct a robustness check of our main results using three additional currency pairs. Section 4 provides technical details on the CUSUM tests, and Section 5 contains additional information on the sinh-arcsinh distribution.

1 Simulation study

In this section, we study the finite sample properties of our econometric approach. We start by investigating the quality of the proposed maximum likelihood estimation procedure in 6 data generating processes (DGP), namely DGP1 to DGP6. Then, we study the size and power of the hypothesis test based on Wald statistics, focusing on the regression parameters in the skewness equation (DGP7). We consider 3 sample sizes: $T \in \{500, 1500, 3000\}$. The parameters of the DGP are displayed in Table 1. In DGP1 to DGP3, we assume λ and c to equal zero, such that $\mathbb{E}(R_t) = 0$. We introduce a mean structure in DGP4 to DGP7. If $a_3 \neq 0$, we assume that $x_{t-1} \overset{iid}{\sim} N(0,1)$. For each DGP, we simulate $1,000$ time series.

Results in term of root-mean squared error (RMSE) are presented in Tables 2 and 3. Overall, we observe a decrease in the RMSE of the estimated parameters when the sample size increases, and no differences across DGPs. As for most time-series models, the simulations highlight the need for large samples (i.e. several thousands observations) to attain a high level of precision.

Then, using DGP7, we investigate the size and power of the suggested Wald-type tests. We consider two explanatory variables $x_{t-1,1}$ and $x_{t-1,2}$ in eq. (2.1)[1]. We generate $x_{t,1}$ and $x_{t,2}$ from two AR(1) processes, where the AR parameters are equal to 0.9 and the error terms follow a bivariate normal distribution with a correlation parameter of -0.4 (a value observed in our data). Baseline values of the parameters are presented in Table 1 (DGP7). To compute full power curves, we sequentially replace the value of the parameters reported in Table 1 of values (including 0), keeping the others at their baseline values. In line with our empirical study, we set $n = 1,500$. Figure 1 summarizes our results, indicating good powers under the various scenarios. For a_3 and a_4, we also obtain excellent sizes. For λ, a_1 and a_2, we reject a little too often.

Table 1: Values of the parameters considered in the different simulation setups.

Data generating processes											
	ω	α	β	c	λ	δ	a_0	a_1	a_2	a_3	a_4
DGP1	10^{-4}	0.05	0.88	-	-	.8	-0.1	0.4	-0.9	-	-
DGP2	10^{-4}	0.05	0.88	-	-	.8	-0.1	0.4	-0.9	0.85	-
DGP3	10^{-4}	0.1	0.8	-	-	.85	-0.1	0.5	0.3	-0.5	-
DGP4	10^{-4}	0.05	0.88	10^{-4}	0.1	.8	-0.1	0.4	-0.9	-	-
DGP5	10^{-4}	0.05	0.88	10^{-4}	0.1	.8	-0.1	0.4	-0.9	0.85	-
DGP6	10^{-4}	0.1	0.8	10^{-4}	0.1	.85	-0.1	0.5	0.3	-0.5	-
DGP7	10^{-4}	0.05	0.88	10^{-4}	-0.1	.8	-0.1	0.5	0.1	0.15	0.3

[1]We do not look at inference for the GARCH parameters, since these parameters are estimated under positivity constraints.

Table 2: RMSE divided by the value of the corresponding parameter, for DGPs without mean structure

RMSE									
No mean	T	ω	α	β	a_0	a_1	a_2	a_3	δ
DGP1	500	1.07	0.32	0.1	0.52	0.15	0.77	-	0.23
	1500	0.26	0.16	0.03	0.23	0.08	0.38	-	0.08
	3000	0.17	0.11	0.02	0.16	0.06	0.27	-	0.06
DGP2	500	0.53	0.27	0.05	0.64	0.12	0.17	0.18	0.6
	1500	0.21	0.13	0.02	0.3	0.06	0.08	0.09	0.33
	3000	0.13	0.09	0.01	0.23	0.04	0.06	0.06	0.23
DGP3	500	0.43	0.23	0.07	0.61	0.15	0.27	0.22	0.44
°	1500	0.18	0.12	0.03	0.29	0.07	0.13	0.11	0.22
	3000	0.11	0.08	0.02	0.2	0.05	0.09	0.07	0.16

Table 3: RMSE divided by the value of the parameters, for the DGPs with a GARCH-in-Mean structure.

RMSE											
GARCH-in-Mean	T	ω	α	β	c	λ	a_0	a_1	a_2	a_3	δ
DGP4	500	1.02	0.31	0.1	275.55	7.82	0.61	0.16	0.71	-	0.21
	1500	0.25	0.17	0.03	42.19	1.13	0.34	0.09	0.43	-	0.1
	3000	0.16	0.11	0.02	26.76	0.71	0.21	0.062	0.28	-	0.07
DGP5	500	0.76	0.28	0.07	88.54	2.79	0.77	0.13	0.2	0.21	0.65
	1500	0.22	0.14	0.02	33.31	1.39	0.37	0.07	0.09	0.09	0.32
	3000	0.13	0.09	0.01	22.89	1.17	0.24	0.04	0.05	0.06	0.23
DGP6	500	0.41	0.3	0.07	97.2	3.21	0.63	0.14	0.3	0.23	0.43
	1500	0.19	0.15	0.03	35.06	1.38	0.31	0.07	0.15	0.11	0.24
	3000	0.13	0.12	0.02	23.99	1.19	0.20	0.05	0.09	0.07	0.17

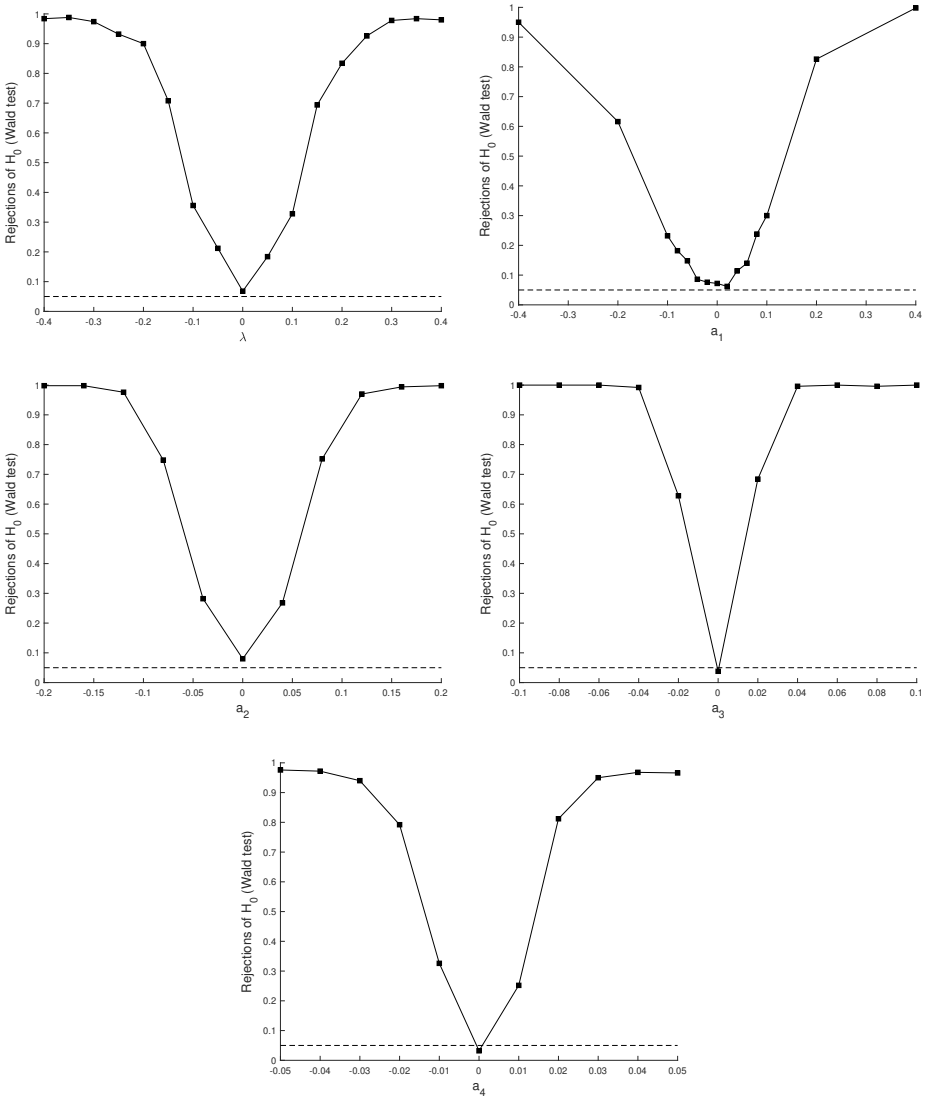


Figure 1: Rejection rates of Wald tests at the 5% test level, for various values of the parameters. Dashed: 5% threshold.

2 Additional empirical results for EUR and CHF

2.1 Coefficient estimates

In Tables 4 and 5, we report the full information on the estimated coefficients for skewness specifications that rely on covariate information (specifications 7 to 12) for the equation

$$\epsilon_t = g(\mathcal{I}_{t-1}) = a_0 + a_1\epsilon_{t-1} + a_2 z_{t-1} + a_3 x_{t-1}.$$

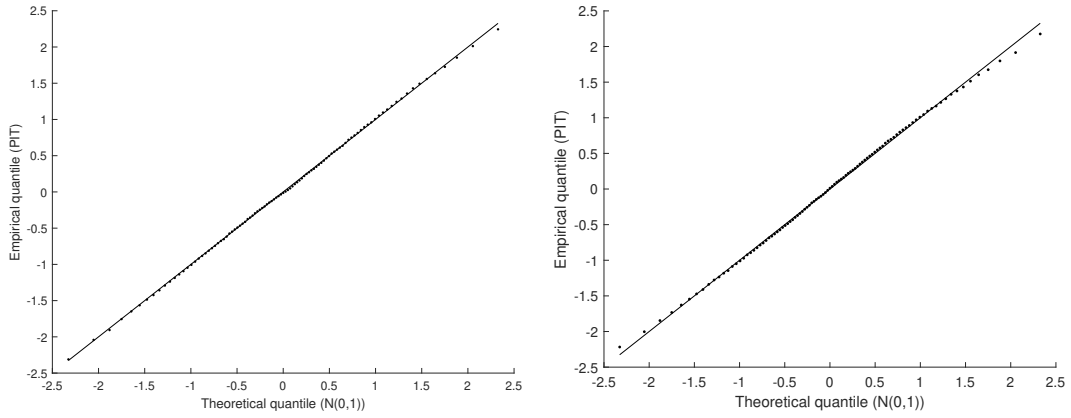
A detailed discussion of this equation is given in Section 2.1 in the main text.

2.2 Model specification and selection

Table 6 contains detailed information on model performance criteria and test statistics for skewness specifications that rely on covariate information (specifications 7 to 12).

2.3 QQ-plots of the reference models

On Figure 2, we display the QQ-plots of the residuals for the two currencies, using ARMAX(2) for EUR (left panel) and MAX(IRD) for CHF (right panel).



(i) (ii)

Figure 2: QQ-plot of the residuals for (i) ARMAX(2) (EUR), and (ii) MAX(IRD) (CHF), fitted on the complete period.

Table 4: Estimated coefficients for selected skewness models

USD/EUR (1999M1 - 2019M3)											
Model	ω	α	β	c	λ	δ	a_0	a_1	a_2	IRD (a_3)	VIX (a_4)
ARMAX(2)	0.000	0.029	0.968	0.000	-0.032	0.772	-0.043 *	0.631 ***	0.042 ***	1.520 *	0.171 *
	(0.699)	(0.177)	(0.007)	(0.000)	(0.052)	(0.022)	(0.025)	(0.184)	(0.014)	(0.813)	(0.101)
ARMAX(IRD)	0.000	0.029	0.968	0.000	-0.054	0.769	-0.007	0.589***	0.042***	1.205*	-
	(0.615)	(0.166)	(0.006)	(0.000)	(0.052)	(0.022)	(0.007)	(0.199)	(0.014)	(0.679)	-
ARMAX(VIX)	0.000	0.03	0.968	0.000	-0.025	0.771	-0.025	0.543***	0.045***	-	0.10
	(0.628)	(0.164)	(0.006)	(0.000)	(0.052)	(0.022)	(0.019)	(0.198)	(0.014)	-	(0.079)
MAX(2)	0.000	0.029	0.968	0.000	-0.027	0.772	-0.103***	-	0.045***	3.821***	0.401**
	(0.435)	(0.132)	(0.006)	(0.000)	(0.054)	(0.022)	(0.038)	-	(0.014)	(1.058)	(0.160)
MAX(IRD)	0.000	0.029	0.968	0.000	-0.047	0.770	-0.016	-	0.043***	2.805***	-
	(0.435)	(0.132)	(0.004)	(0.000)	(0.054)	(0.022)	(0.015)	-	(0.014)	(0.978)	-
MAX(VIX)	.000	0.029	0.968	0.000	-0.023	0.772	-0.045	-	0.046***	-	0.184
	(0.435)	(0.132)	(0.004)	(0.000)	(.053)	(.022)	(0.034)	-	(0.014)	-	(0.147)

a_0, a_1, a_2 are the constant, AR and MA parameters in the skewness equation (2.1), respectively. IRD and VIX refer to the estimated parameters of the corresponding predictors. Standard errors are put in parentheses. *, ** and *** denote Wald tests significant at the 10%, 5% and 1% test level.

Table 5: Estimated coefficients for selected skewness models

USD/CHF (1999M1 - 2019M3)											
Model	ω	α	β	c	λ	δ	a_0	a_1	a_2	IRD (a_3)	VIX (a_4)
ARMAX(2)	0.000	0.037	0.959	0.000	-0.045	0.687	-0.021	-0.227	0.041***	4.304***	0.094
	(0.253)	(0.098)	(0.004)	(0.000)	(0.048)	(0.017)	(0.043)	(0.218)	(0.013)	(1.341)	(0.161)
ARMAX(IRD)	0.000	0.037	0.960	0.000	-0.049	0.687	0.000	-0.220	0.041***	4.180***	-
	(0.252)	(0.098)	(0.04)	(0.000)	(0.048)	(0.017)	(0.022)	(0.216)	(0.013)	(1.304)	-
ARMAX(VIX)	0.000	0.036	0.960	0.000	-0.032	0.700	0.061	-0.235	0.040***	-	-0.008
	(0.252)	(0.098)	(0.04)	(0.000)	(0.047)	(0.017)	(0.039)	(0.211)	(0.013)	-	(0.158)
MAX(2)	0.000	0.037	0.959	0.000	-0.046	0.686	-0.017	-	0.040***	3.524***	0.072
	(0.253)	(0.098)	(0.004)	(0.000)	(0.048)	(0.017)	(0.035)	-	(0.013)	(0.901)	(0.131)
MAX(IRD)	0.000	0.037	0.959	0.000	-0.050	0.686	-0.000	-	0.040***	3.448***	-
	(0.253)	(0.098)	(0.004)	(0.000)	(0.047)	(0.017)	(0.018)	-	(0.013)	(0.890)	-
MAX(VIX)	0.000	0.037	0.959	0.000	-0.033	0.686	0.050	-	0.039***	-	-0.012
	(0.253)	(0.098)	(0.004)	(0.000)	(0.047)	(0.017)	(0.031)	-	(0.013)	-	(0.129)

a_0, a_1, a_2 are the constant, AR and MA parameters in the skewness equation (2.1), respectively. IRD and VIX refer to the estimated parameters of the corresponding predictors. Standard errors are put in parentheses. *, ** and *** denote Wald tests significant at the 10%, 5% and 1% test level.

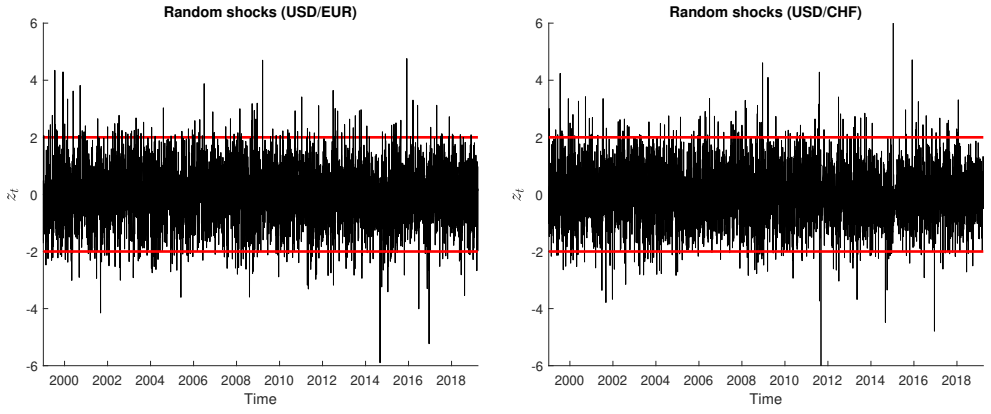
Table 6: Model selection and specification criteria.

EUR	ARMAX(2)	ARMAX(IRD)	ARMAX(VIX)	MAX(2)	MAX(IRD)	MAX(VIX)
LLF	-19,026.94	-19,022.82	-19,019.48	-19,024.26	-19,020.47	-19,017.16
AIC	-38,031.88	-38,025.64	-38,018.96	-38,028.52	-38,022.94	-38,016.32
BIC	-37,960.02	-37,960.32	-37,953.64	-37,963.20	-37,964.14	-37,957.53
LR	-	8.24***	14.92***	5.36**	12.94***	19.56***
BK	0.702	0.758	0.840	0.252	0.735	0.770
DH	2.885	3.030	3.217	2.758	2.895	3.019
AD	0.353	0.371	0.328	0.369	0.339	0.357
CHF	ARMAX(2)	ARMAX(IRD)	ARMAX(VIX)	MAX(2)	MAX(IRD)	MAX(VIX)
LLF	-18,670.06	-18,669.88	-18,662.59	-18,669.52	-18,669.38	-18,661.98
AIC	-37,318.11	-37,319.77	-37,305.18	-37,319.05	-37,320.75	-37,305.97
BIC	-37,246.26	-37,254.44	-37,239.86	-37,253.73	-37,261.96	-37,247.18
LR	-	0.35	14.94***	1.07	1.37	16.15***
BK	3.32	3.36	3.06	3.39	3.41	3.13
DH	18.63***	18.74***	16.57***	18.95***	19.04***	16.90***
AD	1.08	1.09	1.07	1.12	1.12	1.11

LLF denotes the value of the negative log-likelihood function. The line *LR* displays the likelihood ratio test statistics between ARMAX(2) and the competing models. The lines labelled BK, DH and AD report the test statistics for Berkowitz [2001], Doornik and Hansen [2008] and Anderson-Darling tests, respectively. *** denote tests significant at the 1% level.

2.4 Crash risk measures

We provide several additional insights regarding our crash risk measures. In essence, we simply compute the probability that random shocks (z_t) are large (resp. small). In Figure 3, we display the time series of unexpected shocks $\hat{z}_t$ for EUR (left) and CHF (right). The red lines are the -2/2 bounds used in our crash risk definition. Crash risk is thus the probability, at a given point in time, that z_t takes values above or below these bounds. Data points above 2 indicate therefore a large realized depreciation shock to USD (resp. appreciation shock for EUR or CHF, or ρ_t^+), while data points below -2 indicate large USD appreciation shocks (reps. large depreciation shocks for EUR or CHF, or ρ_t^-). This is what we could call "currency crashes", although what matter more to us is the actual probability that the event takes place. In Figure 4, we display our estimates of the crash risk measures (upper panel: depreciation of USD, middle panel: appreciation of USD). The red lines indicate the crash risk under the hypothesis of a symmetric Gaussian distribution for z_t, which amounts to 0.0228 and can be used as a reference point. In panels (v) and (vi) we display smooth estimates of the crash risk over time. The global financial crisis is clearly apparent in the crash risk measure. We also see long periods of decline and increase in crash risk, reflecting the dynamic identified in the paper.



(i) (ii)

Figure 3: Time series of $\hat{z}_t$, obtained with ARMAX(2) for EUR (left) and MAX(IRD) for CHF (left). Red: -2/2 bounds used for crash risk calculations. Rem.: for CHF (panel (ii)), the large peaks in 2011 and 2015 take place the days of the introduction and removal of the CHF cap against EUR. Their actual values are around -8 and 24, but for readability and comparison purposes with EUR, we kept the y-axis between -6 and 6.

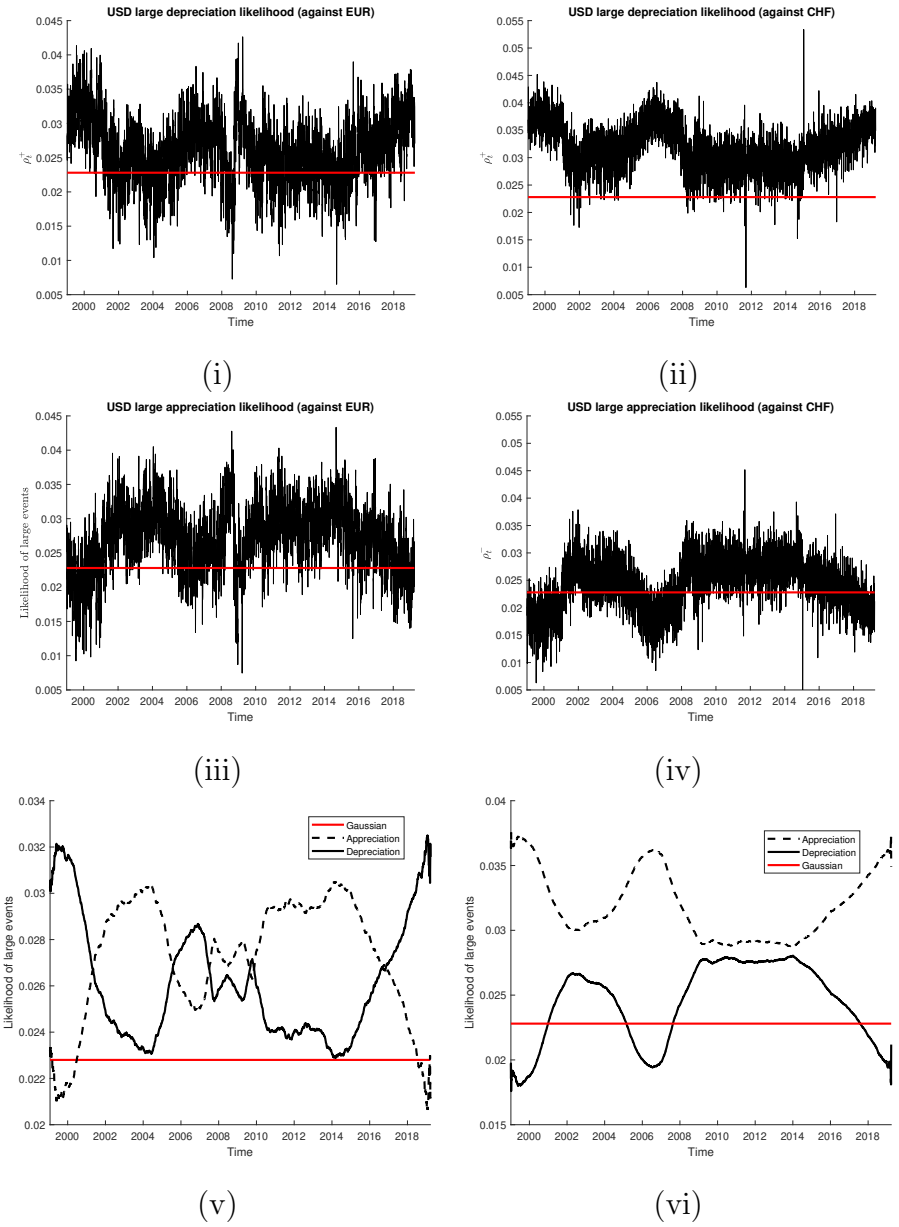
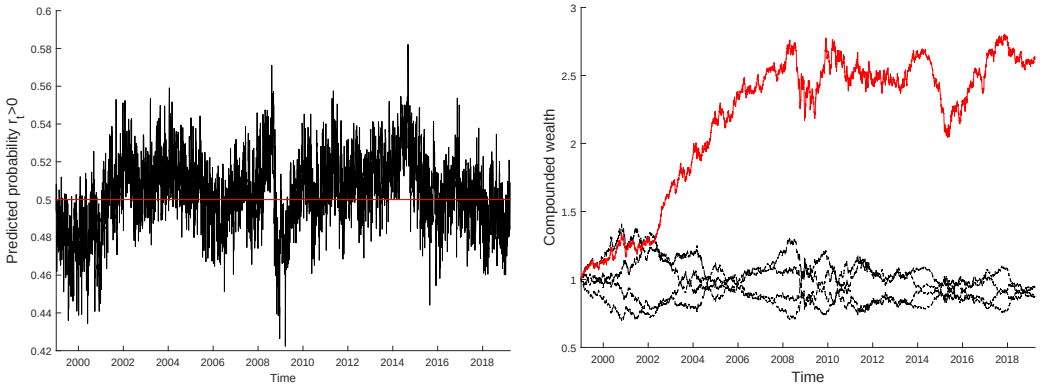


Figure 4: Upper panels: large depreciation likelihood of USD against the foreign currency. Middle panel: large appreciation likelihood. Bottom panel: smooth estimates. Left: USD/EUR, obtained with ARMAX(2). Right: USD/CHF, obtained with MAX(IRD).

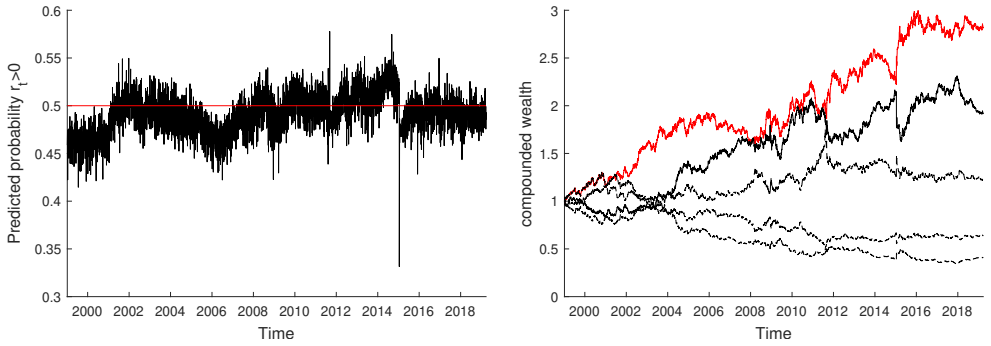
2.5 In-sample and subsample performance

Figures 5 and 6 display for EUR and CHF, respectively, i) the predicted probability of an appreciation and ii) the cumulative returns over time of an initial investment of 1 USD with reinvestment.



(i) (ii)

Figure 5: (i) Predicted probability of an appreciation of EUR based on the ARMAX(2) specification and (ii) Evolution over time of an initial investment of 1 USD with reinvestment of the proceed. ARMAX(2) in red. Solid black: RW^-. Dashed black: BH, AS, RW^+.



(i) (ii)

Figure 6: (i) One-step ahead predicted probability of an appreciation for CHF. (ii) Evolution over time of an initial investment of 1 USD with reinvestment of the proceed. Red: MAX(IRD). Solid black: RW^-. Dashed black: BH, AS, RW^+.

Figures 7 and 8 display the evolution over time of an initial investment of 1 USD in EUR and CHF, respectively for a subsample analysis. For EUR, the analysis is conducted for the periods (i) 6/01/1999 - 1/12/2008, (ii) 2/12/2008 - 14/12/2015 and (iii) 15/12/2015 - 29/03/2019. For CHF, the displayed period covers 6/01/1999 - 6/09/2011 (before the ceiling). Similarly, Tables 7 and 8 display details on the forecasting performance of all model specifications over these time periods for EUR and CHF, respectively.

Table 7: Forecasting performance for all models in subsamples

USD/EUR (1999M1 - 2019M3)												
	1999M1 - 2008M10 ($n = 2,470$)				2008M10 - 2015M12 ($n = 1,792$)				2015M12 - 2019M3 ($n = 814$)			
Spec.	CR	PT09	$\hat{m}$	*Full SSPA*	CR	PT09	$\hat{m}$	*Full SSPA*	CR	PT09	$\hat{m}$	*Full SSPA*
ARMAX(2)	54.98%	12.31***	14.50%	0.000	51.25%	0.56	0.64%	0.989	52.09%	0.10	1.14%	0.977
ARMAX(IRD)	54.17%	5.75**	12.36%	0.000	50.87%	0.00	0.98%	0.984	53.07%	0.02	2.61%	0.903
ARMAX(VIX)	52.72%	3.28*	6.14%	0.121	50.48%	0.00	-1.08%	1.000	52.21%	0.33	2.55%	0.908
MAX(2)	54.66%	10.89***	14.67%	0.000	51.31%	0.27	0.89%	0.986	48.40%	0.00	-0.35%	0.999
MAX(IRD)	53.44%	3.12*	9.88%	0.007	52.15%	2.48	8.54%	0.180	48.89%	0.66	1.25%	0.975
MAX(VIX)	52.51%	1.96	4.78%	0.269	51.31%	1.65	0.51%	0.992	49.51%	0.00	1.85%	0.951
ARMA	53.04%	12.84***	5.89%	0.142	50.92%	0.00	0.95%	0.985	52.58%	0.09	2.94%	0.874
MA	52.55%	2.15	4.98%	0.246	52.09%	2.05	7.95%	0.235	49.51%	0.15	1.42%	0.972
ARX(2)	54.17%	4.85**	11.30%	0.001	48.80%	0.20	-1.94%	1.000	48.77%	0.15	2.48%	0.910
ARX(IRD)	53.32%	3.92**	8.97%	0.015	50.53%	0.88	4.71%	0.670	47.91%	1.02	-2.21%	1.000
ARX(VIX)	52.31%	0.90	8.11%	0.034	49.81%	0.00	1.04%	0.983	50.25%	0.58	6.00%	0.449
CST	52.19%	0.50	7.74%	0.041	49.08%	0.74	3.30%	0.842	49.51%	0.15	-0.13%	0.998
RW^+/RW^-	50.65%	0.00	0.01%	0.945	49.97%	0.00	0.16%	0.928	49.88%	0.00	1.84%	0.969
BH/AS	49.19%	-	0.74%	0.886	51.14%	-	2.31%	0.744	48.77%	-	0.83%	0.988

Forecasting performance for all models during the 3 subsamples defined by the following dates: 6/01/1999, 23/10/2008, 15/12/2015 and 25/03/2019. PT09 refers to the independence test of Pesaran and Timmermann [2009].

Table 8: Forecasting performance for all models in subsamples

USD/CHF (1999M1 - 2019M3)												
	1999M1 - 2011M09 ($n = 3,188$)				2011M09 - 2015M01 ($n = 840$)				2015M01 - 2019M3 ($n = 1,045$)			
Spec.	CR	PT09	$\hat{m}$	*Full SSPA*	CR	PT09	$\hat{m}$	*Full SSPA*	CR	PT09	$\hat{m}$	*Full SSPA*
ARMAX(2)	53.98%	7.61***	11.20%	0.002	51.73%	1.13	10.73%	0.145	51.39%	0.00	-0.88%	0.923
ARMAX(IRD)	53.48%	2.71*	9.93%	0.008	51.85%	0.21	5.91%	0.609	51.96%	0.64	2.94%	0.597
ARMAX(VIX)	51.41%	0.32	4.49%	0.312	52.80%	1.97	12.63%	0.056	52.82%	0.01	2.99%	0.593
MAX(2)	53.83%	5.45***	11.18%	0.002	50.54%	0.00	3.69%	0.824	52.34%	0.26	3.22%	0.562
MAX(IRD)	53.76%	6.95***	10.63%	0.003	51.85%	0.49	4.87%	0.720	52.34%	0.10	3.15%	0.574
MAX(VIX)	51.35%	0.22	4.50%	0.311	52.32%	0.94	8.67%	0.308	52.73%	0.55	2.62%	0.631
ARMA	51.51%	0.66	5.43%	0.197	51.97%	0.89	9.01%	0.275	51.87%	1.54	2.08%	0.685
MA	51.57%	0.73	5.48%	0.191	51.49%	0.26	3.33%	0.847	52.92%	0.15	2.54%	0.640
ARX(2)	53.26%	0.02	9.84%	0.008	52.56%	2.39	5.12%	0.693	53.59%	1.67	6.34%	0.225
ARX(IRD)	53.42%	0.02	10.52%	0.004	51.49%	1.16	7.85%	0.384	52.54%	5.36**	3.22%	0.562
ARX(VIX)	50.60%	0.01	0.51%	0.899	51.73%	0.69	7.90%	0.382	53.40%	0.19	4.60%	0.410
CST	50.53%	0.08	0.83%	0.867	50.77%	0.53	1.76%	0.951	53.59%	0.58	4.58%	0.411
RW^+/RW^-	51.19%	0.00	5.34%	0.193	51.01%	0.00	4.84%	0.696	49.86%	0.00	1.10%	0.788
BH/AS	49.75%	-	3.85%	0.434	49.70%	-	4.92%	0.687	51.58%	-	2.50%	0.626

Forecasting performance for all models during the 3 subperiods defined by the following dates: 6/01/1999, 6/09/2011, 15/01/2015.

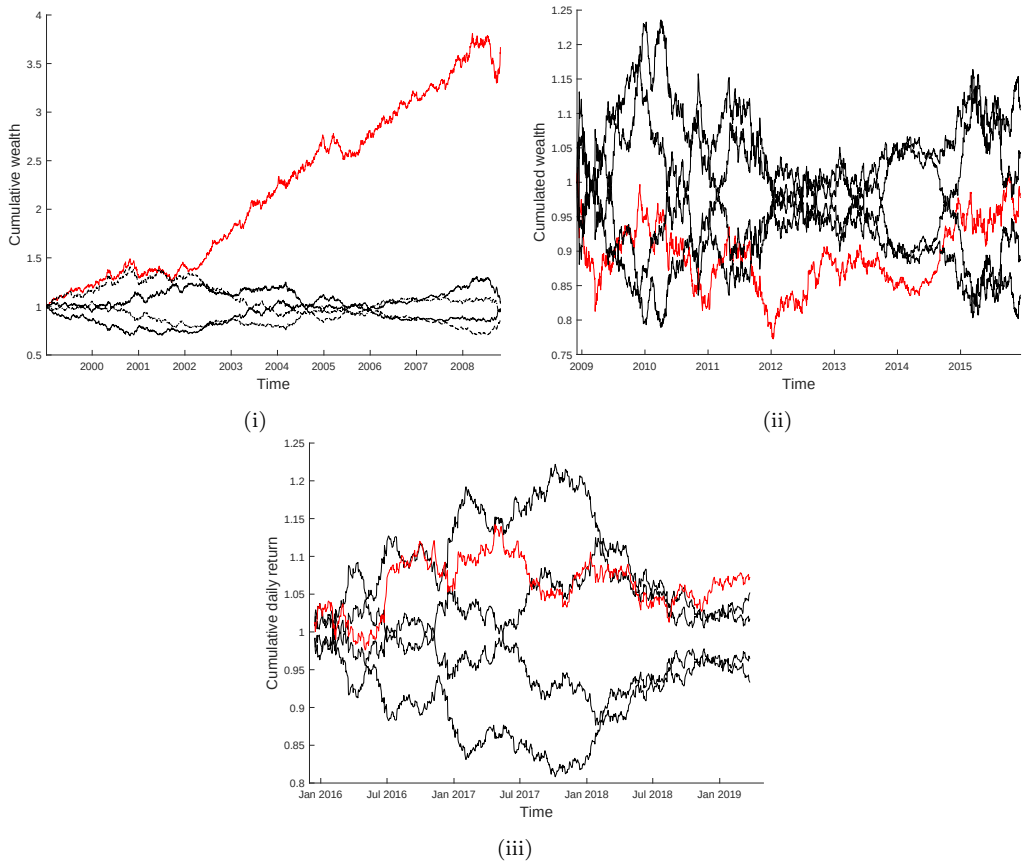


Figure 7: Evolution of an initial investment of 1 USD in EUR, with reinvestment of the proceed, for the periods (i) 6/01/1999 - 1/12/2008, (ii) 2/12/2008 - 14/12/2015 and (iii) 15/12/2015 - 29/03/2019. Red: ARMAX(2) specification. Solid black: RW^-. Dashed black: BH, AS, RW^+.

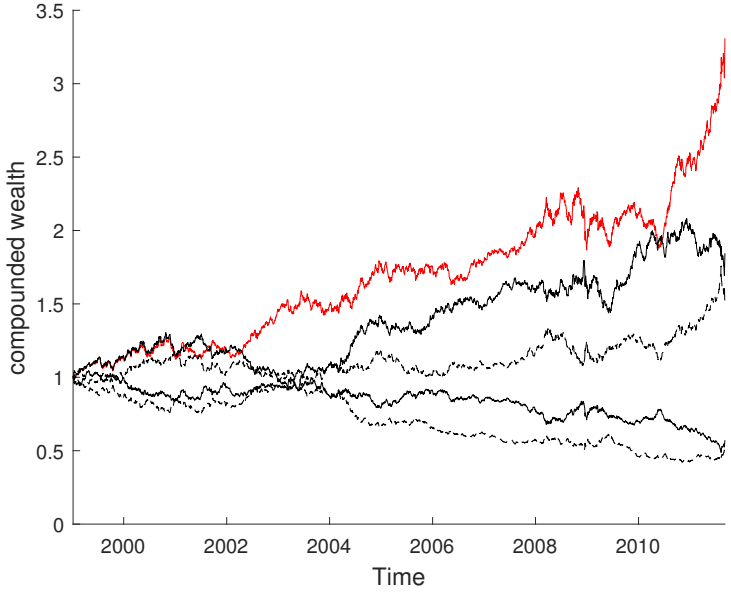


Figure 8: Evolution, over time, of an initial investment of 1 USD in CHF (red), with reinvestment of the proceed, for the period 6/01/1999 - 6/09/2011 (before ceiling). Black: benchmark strategies (solid: BH and RW^+, dashed: AS and RW^-).

2.6 Out-of-sample performance

Figure 9 displays the cumulative profit of a trading rule derived from the different model specifications that earns 1 if the sign is correctly forecast, and -1 otherwise for EUR.

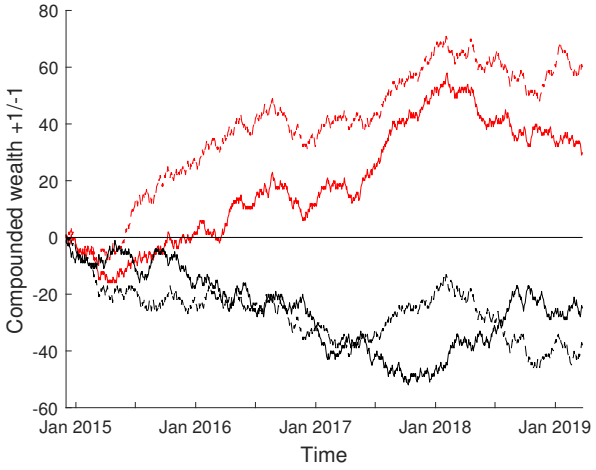
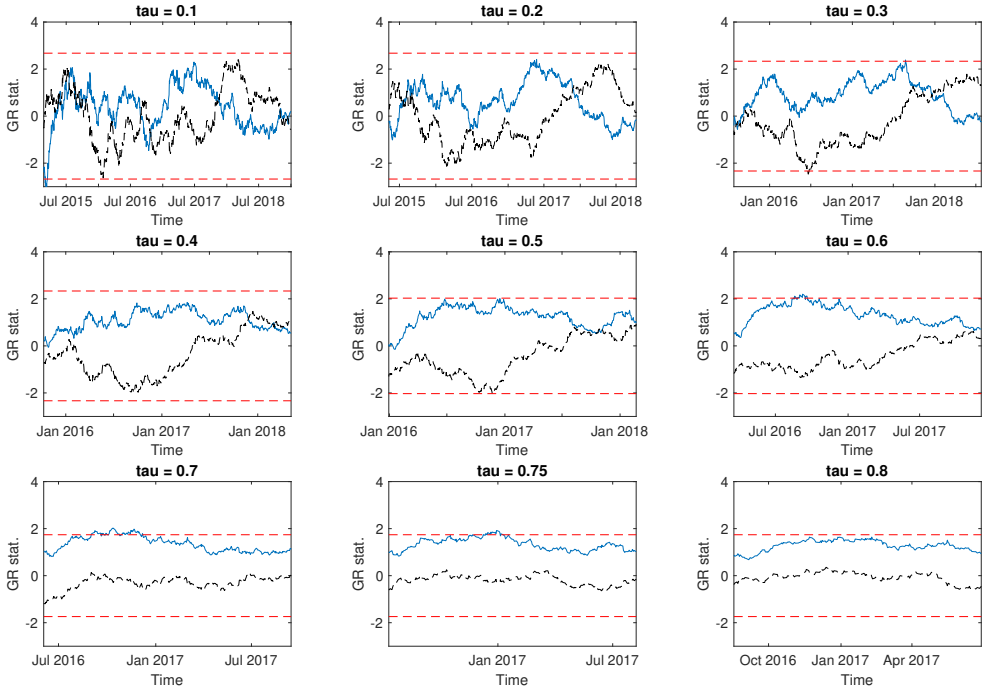


Figure 9: Cumulative profit over time of a trading rule derived from ARMAX(2) (solid red), ARMAX(IRD) (dashed red), random walk (dashed black) and momentum (solid black) that earns 1 if the sign is correctly forecast, -1 otherwise; for EUR.

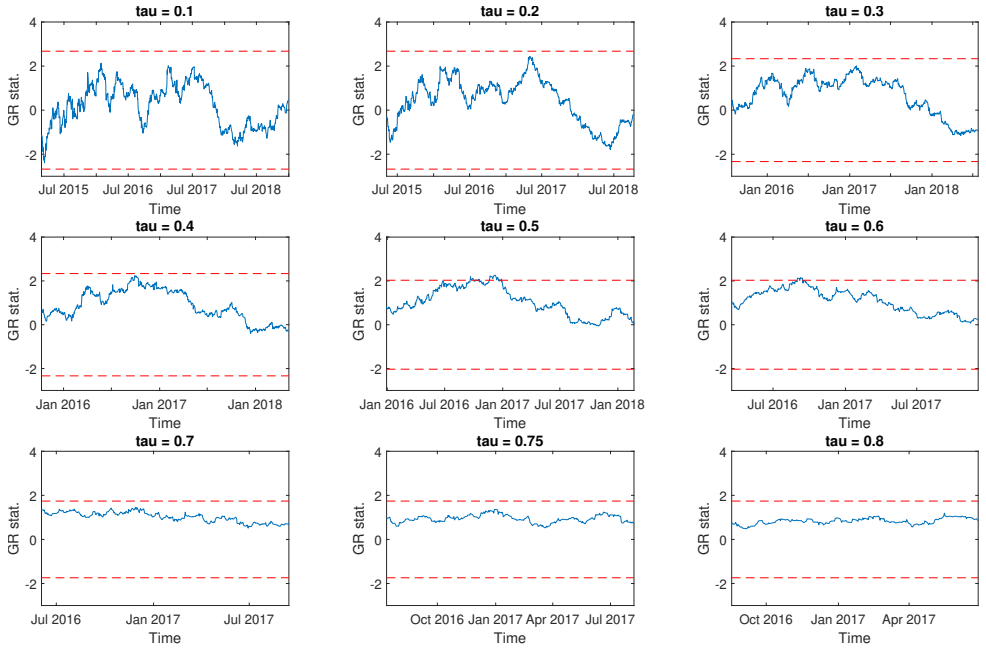
Figure 11 displays the results of the GR tests using $L_{\Delta}^{(2)}$ as loss function, defined by

$$L_{\Delta}^{(2)} = \frac{1}{h} \sum_{j=t+1}^{t+h} (\hat{p}_j^* - \hat{p}_j^{*,RW}) \text{sign}(R_j), \tag{1}$$

where $\hat{p}_j^{*,RW}$ is the sign forecast obtained from the random walk benchmark. Figure 12 displays the results of the GR test using the best ex-post benchmark (i.e. AS with a classification rate of 51.38%). We also report some results using $L_{\Delta}^{(3)} = \frac{1}{h} \sum_{j=t+1}^{t+h} (\hat{p}_j^* - \hat{p}_j^{*,AS}) \text{sign}(R_j)$ as loss function in Figure 12, where AS is the benchmark model.

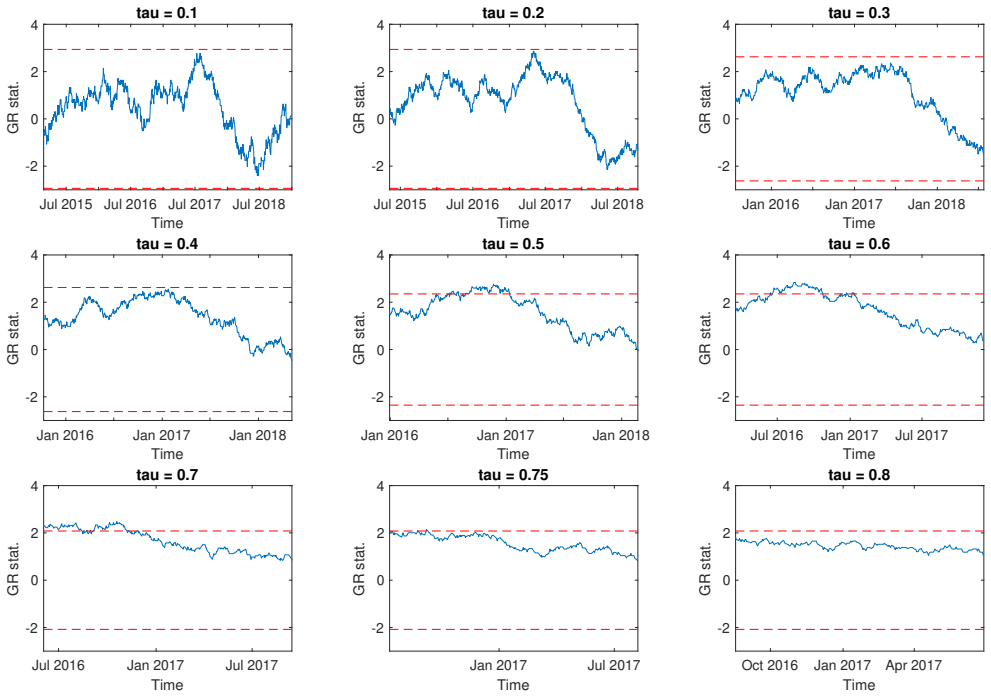


(i) GR test for ARMAX(IRD) ($H_0 : |\hat{m}| \leq 0$)

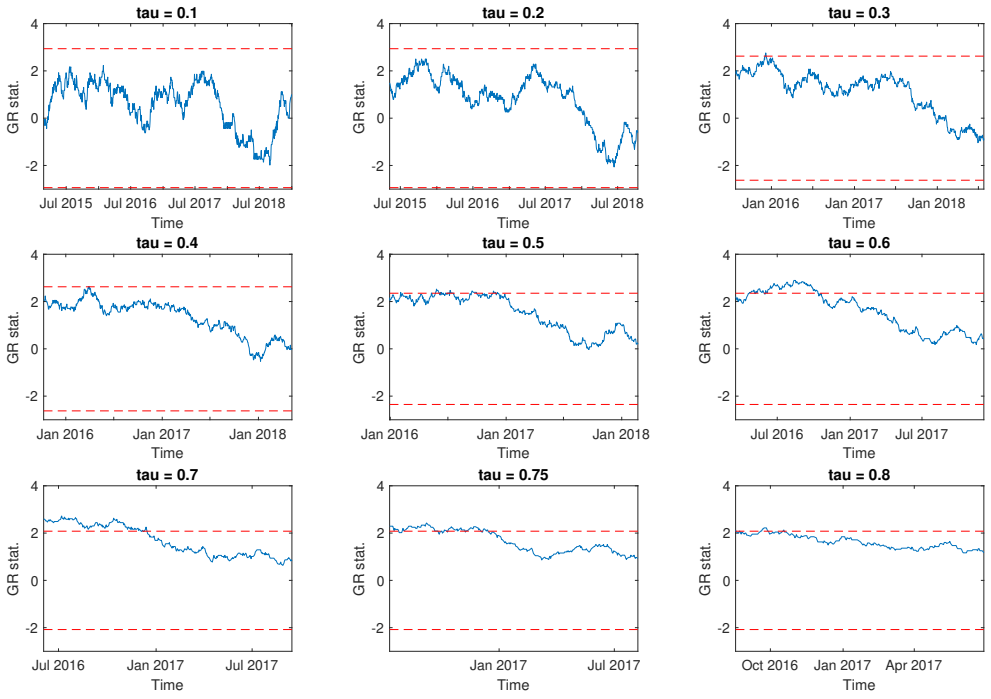


(ii) GR test for ARMAX(IRD) ($H_0 : |\Delta\hat{m}| \leq 0$)

Figure 10: For EUR, GR test statistic with $\tau \in [.1, .85]$ using (i) $\hat{m}$ and (ii) $L_{\Delta}^{(1)}$ as loss functions, for ARMAX(IRD). If the statistic is above the rejection threshold (dashed red), we reject the null hypothesis (i) $H_0 : |\hat{m}| \leq 0$ (upper panel) or (ii) $H_0 : |\Delta\hat{m}| \leq 0$ (lower panel). Dashed black: test statistic for the RW strategy. Dashed red: rejection threshold at the 10% test level.

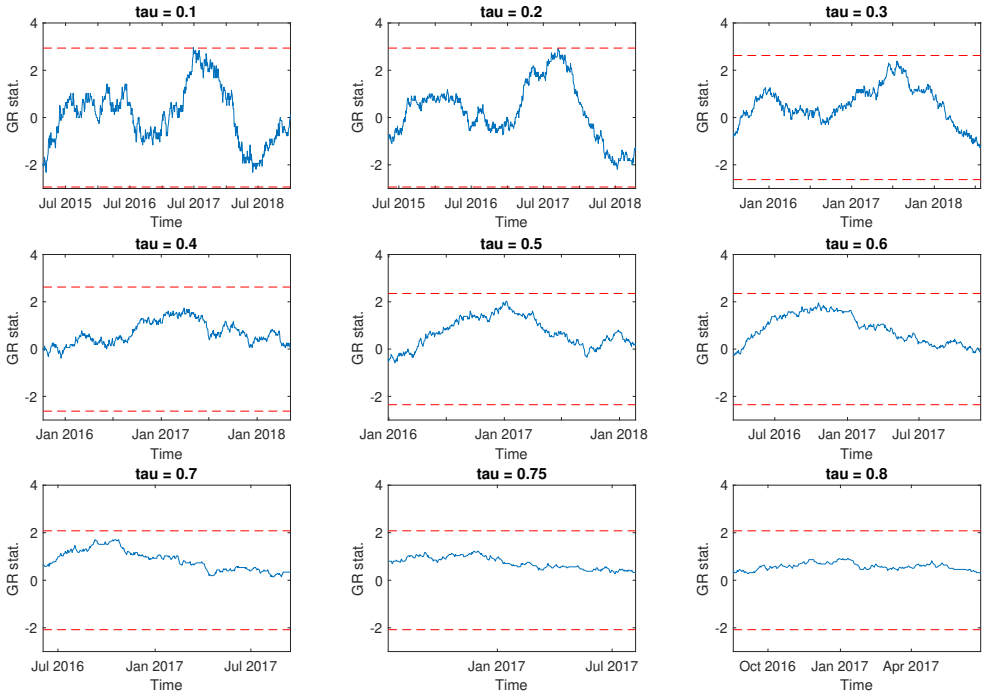


(i) GR test for ARMAX(2) ($H_0 : |L_\Delta^{(2)}| \leq 0$)

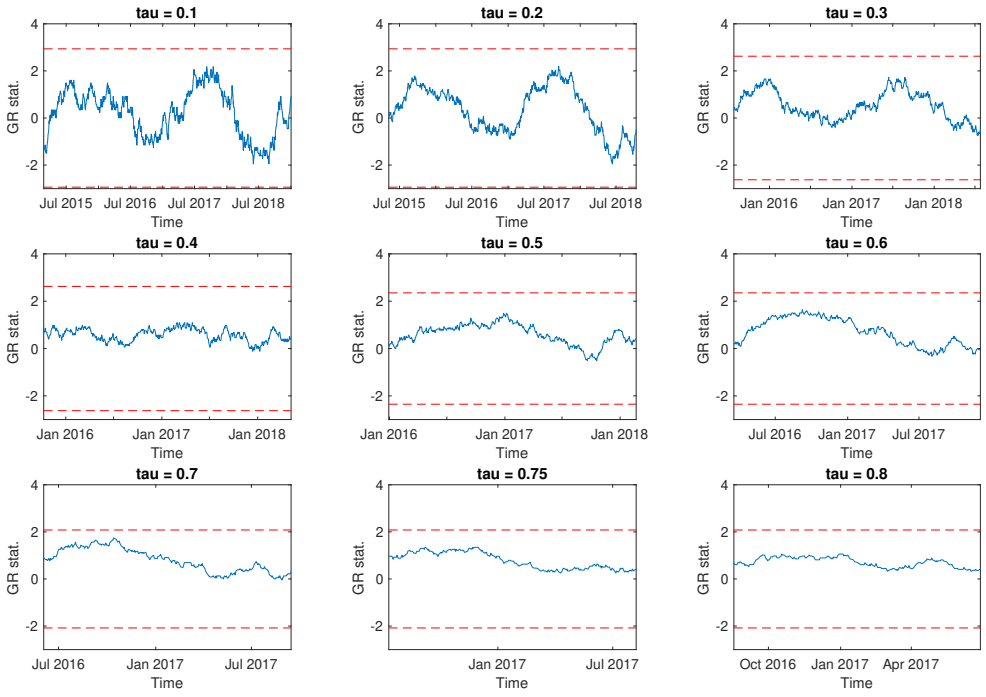


(ii) GR test for ARMAX(IRD) ($H_0 : |L_\Delta^{(2)}| \leq 0$)

Figure 11: For EUR, GR test statistic (blue) with $\tau \in [.1, .85]$ using $L_\Delta^{(2)}$ as loss function. If the statistic is above the rejection threshold (dashed red), we reject the null hypothesis $H_0 : |L_\Delta^{(2)}| \leq 0$. (i) ARMAX(2) and (ii) ARMAX(IRD). Both rejection thresholds are at the 5% test level.

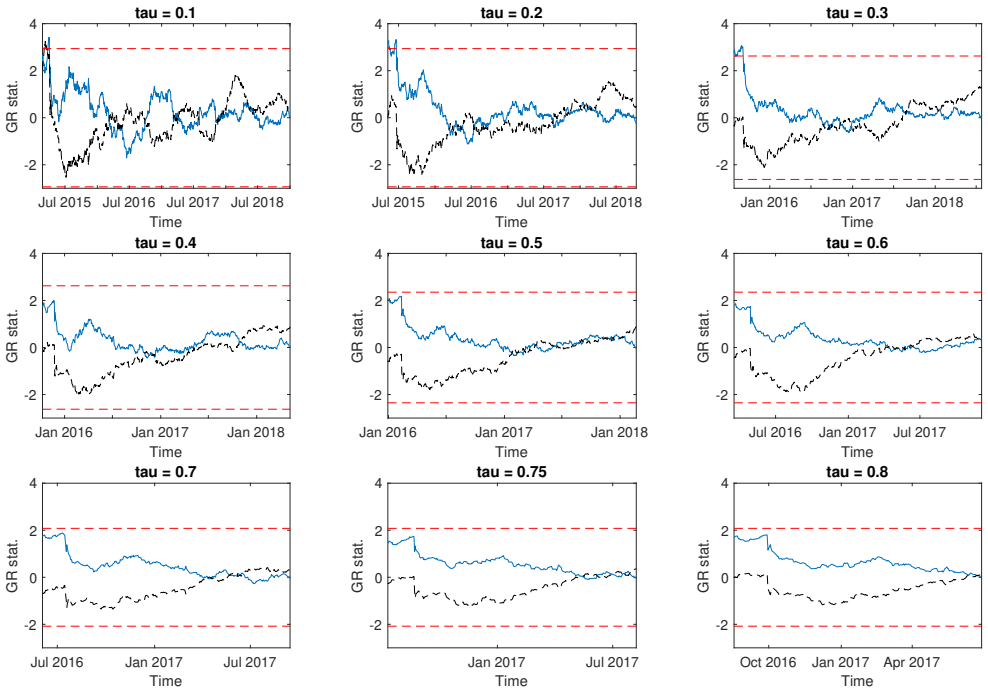


(i) GR test for ARMAX(2) ($H_0 : |L_\Delta^{(3)}| \leq 0$)

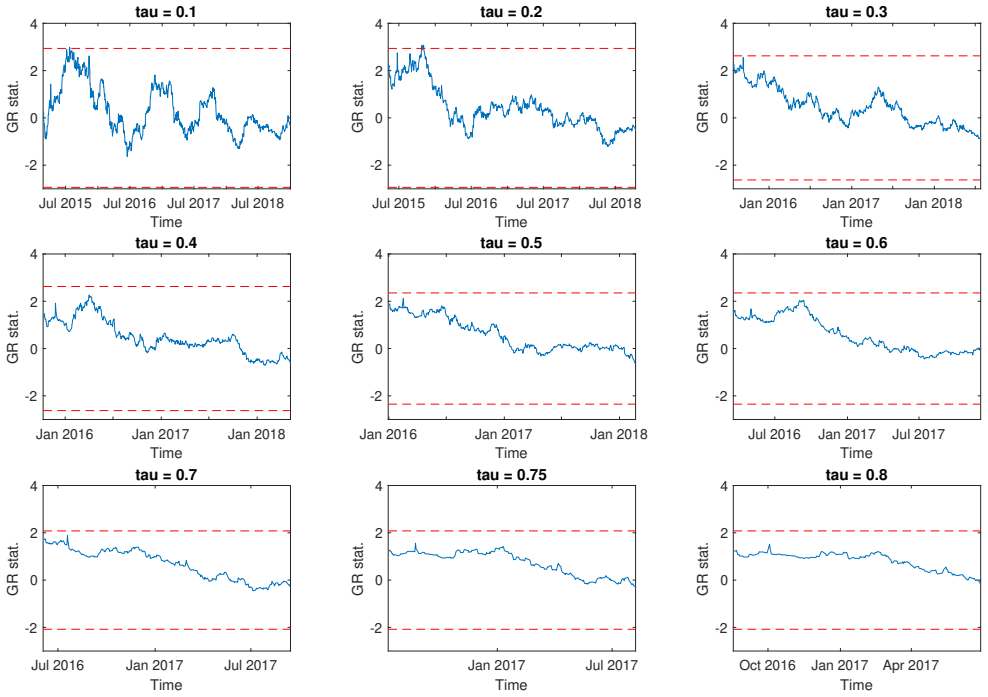


(ii) GR test for ARMAX(IRD) ($H_0 : |L_\Delta^{(3)}| \leq 0$)

Figure 12: For EUR, GR test statistic (blue) with $\tau \in [.1, .85]$ using $L_\Delta^{(3)}$ as loss function. If the statistic is above the rejection threshold (dashed red), we reject the null hypothesis $H_0 : |L_\Delta^{(3)}| \leq 0$. (i) ARMAX(2) and (ii) ARMAX(IRD). Both rejection thresholds are at the 5% test level.

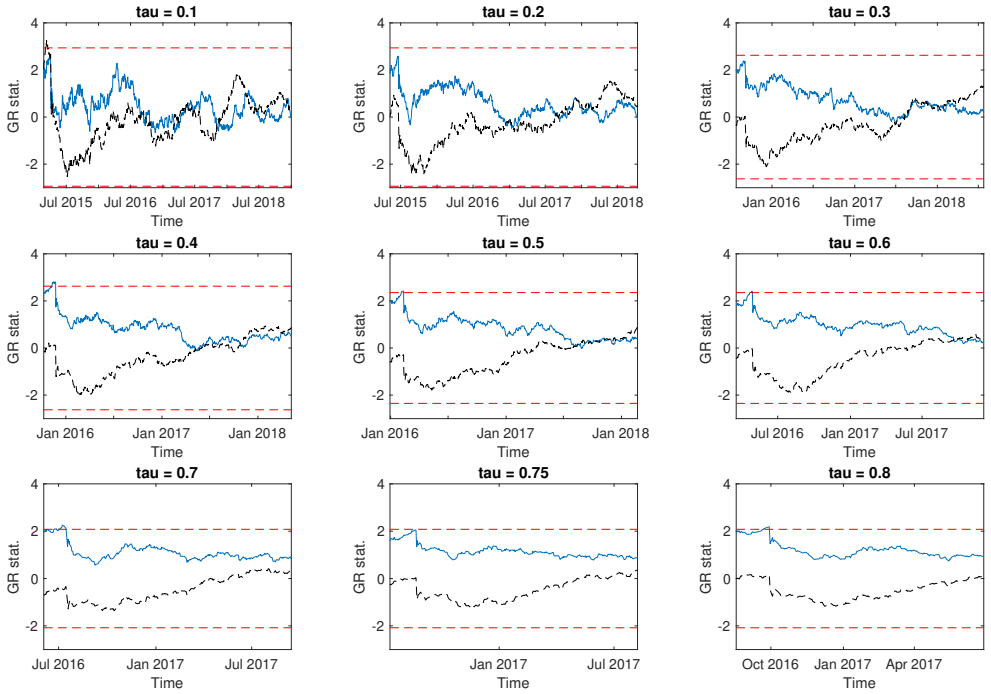


(i) GR test for MAX(2) ($H_0 : |\hat{m}| \leq 0$)

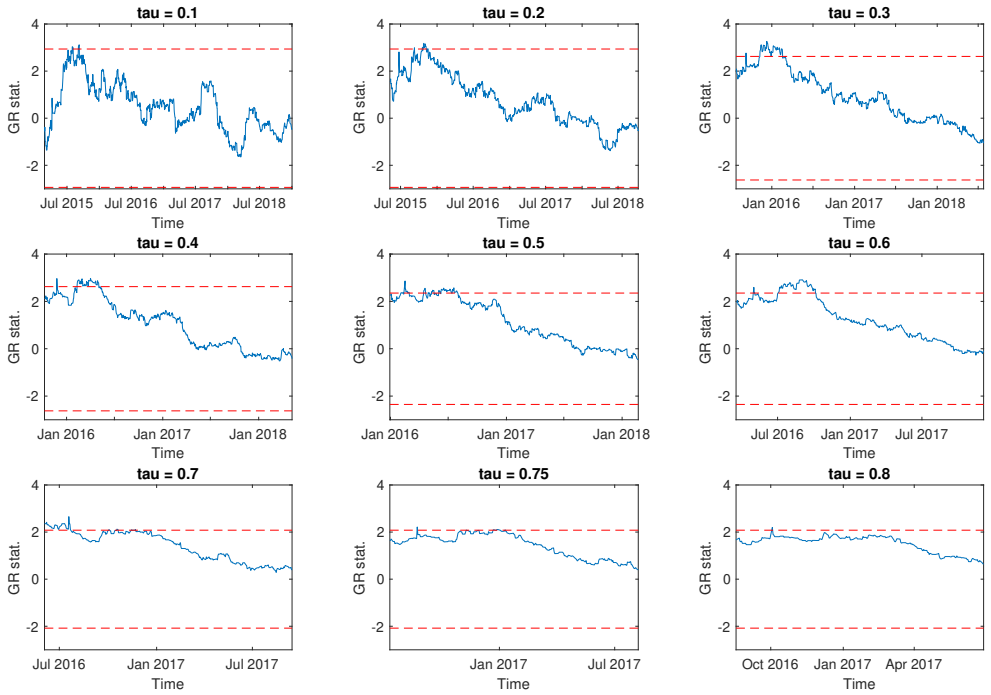


(ii) GR test for MAX(2) ($H_0 : |\Delta\hat{m}| \leq 0$)

Figure 13: For CHF, GR test statistic (blue) with $\tau \in [.1, .85]$ using (i) $\hat{m}$ and (ii) $L_{\Delta}^{(1)}$ as loss functions. If the statistic is above the rejection threshold (dashed red), we reject the null hypothesis at the 5% test level. Dashed black: test statistics for the RW benchmark.



(i) GR test for ARX(IRD) ($H_0 : |\hat{m}| \leq 0$)



(ii) GR test for ARX(IRD) ($H_0 : |\Delta\hat{m}| \leq 0$)

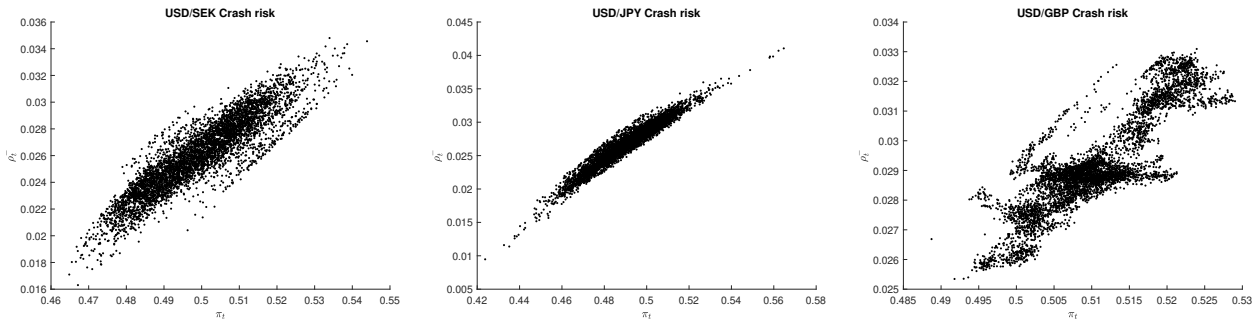
Figure 14: For CHF, GR test statistic (blue) with $\tau \in [.1, .85]$ using (i) $\hat{m}$ and (ii) $L_{\Delta}^{(1)}$ as loss functions. If the statistic is above the rejection threshold (dashed red), we reject the null hypothesis at the 5% test level. Dashed black: test statistics for the RW benchmark.

3 Additional currencies

In our main analysis, we study the dynamics of EUR and CHF currencies vis-a-vis USD. To investigate if our results hold also for other currencies, we reproduce some of the main analysis for the British Pound (GBP), the Japanese Yen (JPY) and the Swedish Kronor (SEK). The use the same time period as for EUR and CHF. As interest rates for JPY and GBP, we use the 3-month LIBOR denominated in these currencies. For SEK, we use the Stockholm Interbank Offered Rate (STIBOR) with a 3-month tenor provided by the Sveriges Riksbank.

In Table 9, we report the BIC and the results of the LR tests for ARMAX(2), ARMAX(IRD), MAX(IRD) and ARMA specifications, for each currency. The information criterion favors models including the IRD, in particular MAX(IRD).

In Table 10 we report coefficient estimates for the selected specifications. The coefficient related to IRD is found to be highly significant and positive for the three currencies, a result similar to the one presented in the main manuscript for EUR and CHF. In Figure 15, we display the historical crash risk trade-off, indicating the likelihood of depreciation crash of the foreign currency as a function of the probability of an appreciation. As for EUR and CHF, we observe a positive association: an increase in the probability of an appreciation is associated to a higher likelihood of a depreciation crash. Finally, we compare the empirical conditional skewness of the residuals of the MAX(IRD) specification to the one obtained for a simple GARCH(1,1) model (Figures 16 to 18). As for EUR and CHF, conditional skewness (with respect to IRD) is present in the raw data (solid blue line) and appears to be correctly filtered out with MAX(IRD) (dashed blue line).



(i) USD/SEK (ii) USD/JPY (iii) USD/GBP

Figure 15: Crash-risk trade-off between the probability of an appreciation (π_t) and the probability of an extreme depreciation (ρ_t^-) for the foreign currency, inferred from MAX(IRD).

Table 9: Model selection and specification criteria.

SEK	ARMAX(2)	ARMAX(IRD)	MAX(IRD)	ARMA	BH/AS	RW
BIC	-36,196	-36,204	-36,212	-36,207	-	-
LLF	-18,145.1	-18,144.9	-18,144.4	-18,141.9	-	-
LR	-	0.40	1.4	6.4**	-	-
AD	0.49	0.51	0.48	0.53	-	-
JPY	ARMAX(2)	ARMAX(IRD)	MAX(IRD)	ARMA	BH/AS	RW
BIC	-37,539.3	-37,547.3	-37,550	-37,543	-	-
LLF	-18,816.6	-18,816.3	-18,813.3	-18,810	-	-
LR	-	0.60	6.6**	13.2***	-	-
AD	0.78**	0.78**	0.79**	0.93**	-	-
CR	50.94%	51.17%	51.15%	49.83%	50.62%	51.92%
GBP	ARMAX(2)	ARMAX(IRD)	MAX(IRD)	ARMA	BH/AS	RW
BIC	-38,687	-38,695	-38,703	-38,699	-	-
LLF	-19.390.3	-19,390	-19,390	-19,388.1	-	-
LR	-	0.6	0.6	4.4**	-	-
AD	0.26	0.26	0.26	0.27	-	-
CR	50.07%	50.58%	50.66%	49.85%*	50.19%	51.51%

Notes: The lines *LR* displays the likelihood ratio test statistics between ARMAX(2) and the competing models. *** and ** denotes tests significant at the 1% and 5% levels, respectively.

Table 10: Estimated coefficients for MAX(IRD) skewness model, for SEK, JPY and GBP

1999M1 - 2019M3										
MAX(IRD)	ω	α	β	c	λ	δ	a_0	a_2	IRD (a_3)	
SEK	0.000	0.031	0.963	0.000	0.061	0.804	-0.002	0.028	1.965***	
	(0.340)	(0.155)	(0.005)	(0.000)	(0.055)	(0.025)	(0.038)	(0.065)	(0.431)	
JPY	0.000	0.040	0.958	0.000	-0.021	0.687	-0.010	0.034***	1.432***	-
	(0.346)	(0.088)	(0.004)	(0.000)	(0.126)	(0.027)	(0.055)	(0.012)	(0.272)	
GBP	0.000	0.041	0.951	0.000	-0.035***	0.785	-0.028***	-0.003	2.164***	
	(0.107)	(0.014)	(0.002)	(0.022)	(0.01)	(0.015)	(0.01)	(0.014)	(0.065)	

a_0 and a_2 are the constant and MA parameters in the skewness equation (2.1) (AR component is omitted in MAX(IRD) specification). IRD refers to the estimated parameters of the corresponding predictor. Standard errors are put in parentheses. *** denote Wald tests significant at the 1% test level.

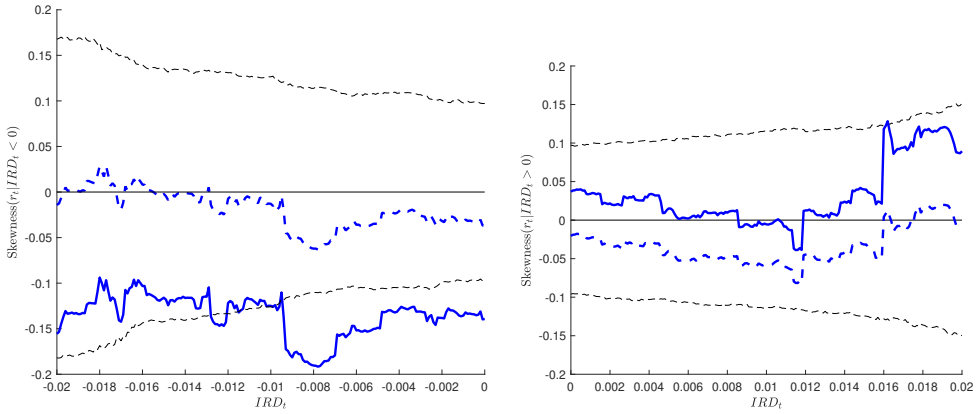


Figure 16: Empirical skewness of USD/SEK residuals, conditional on observing IRD smaller (resp. larger) than zero. Solid blue: residuals of a GARCH(1,1) model. Dashed blue: residuals of MAX(IRD).

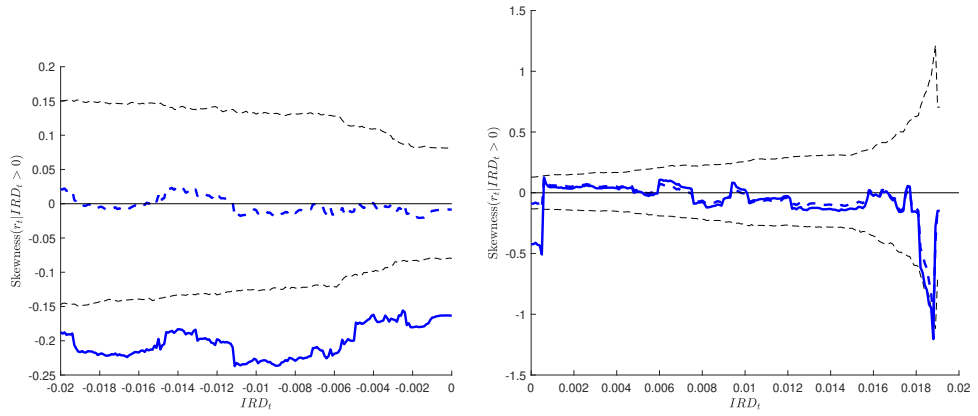


Figure 17: Empirical skewness of USD/GBP residuals, conditional on observing IRD smaller (resp. larger) than zero. Solid blue: residuals of a GARCH(1,1) model. Dashed blue: residuals of MAX(IRD).

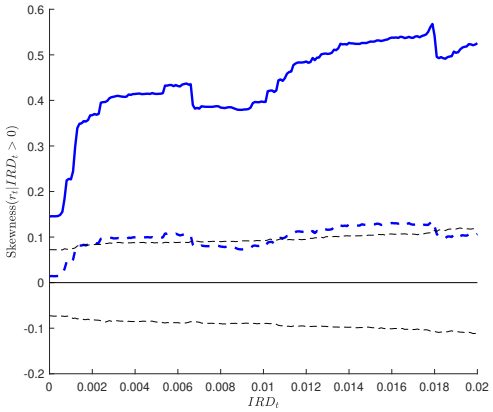


Figure 18: Empirical skewness of USD/JPY residuals, conditional on observing IRD larger than zero. Solid blue: residuals of a GARCH(1,1) model. Dashed blue: residuals of MAX(IRD).

Looking at predictive ability (lines $\hat{m}$ and CR for in-sample results, and $\hat{m}^{oos}$ and CR^{oos} for out-of-sample results, in Table 11), evidence are more mixed. The economic performance of our model is always positive in-sample for MAX(IRD) and largely superior to the ARMA model, as well as the RW and BH benchmarks. For SEK, $\hat{m}$ is found significantly larger than zero at the 10% test level with MAX(IRD), and at the 1% test level with ARMAX(2). For GBP and JPY, however, it is not found to deliver a significantly positive profit, although for JPY it is rather positive. Similarly, the out-of-sample performance is not found to be significantly different from zero, although it is strongly positive for SEK with MAX(IRD), reaching 4.31%. Comparisons with the benchmarks lead to similar conclusions. These results suggest that the predictive ability of IRD, in our setting, is not sufficiently strong for these currencies to generate an economic profit despite the existence of a significant in-sample link with asymmetry. These discrepancies are in accordance with the currency typology highlighted in Hossfeld and MacDonald [2015]: while CHF is a safe heaven currency, and EUR a hedge currency, SEK is categorized as a speculative currency, while neither JPY nor GBP were found to have any specific pattern.

Table 11: In-sample and out-of-sample performance of the trading rules derived from the different skewness models and benchmarks.

SEK	ARMAX(2)	ARMAX(IRD)	MAX(IRD)	ARMA	BH/AS	RW
CR	52.49%***	50.76%	51.78%**	50.25%	50.27%	50.68%
$\hat{m}$	6.80%***	1.42%	4.33%*	-0.43%	0.73%	0.38%
CR^{oos}	50.51%	49.86%	51.53%	49.30%	50.33%	48.65%
$\hat{m}^{oos}$	2.78%	0.70%	4.31%	1.28%	-3.39%	-5.17%
JPY	ARMAX(2)	ARMAX(IRD)	MAX(IRD)	ARMA	BH/AS	RW
CR	50.94%	51.17%	51.15%	49.83%	50.62%	51.92%
$\hat{m}$	1.82%	1.19%	2.63%	-0.49%	0.05%	-0.20%
CR^{oos}	50.23%	51.26%	50.33%	51.91%	48.47%	49.30%
$\hat{m}^{oos}$	1.83%	1.92%	-1.00%	2.36%	1.84%	0.75 %
GBP	ARMAX(2)	ARMAX(IRD)	MAX(IRD)	ARMA	BH/AS	RW
CR	50.07%	50.58%	50.66%	49.85%*	50.19%	51.51%
$\hat{m}$	-0.19%	-1.13%	0.42%	-1.73%	1.13%	1.07%
CR^{oos}	49.02%	49.67%	49.67%	49.30%	49.40%	49.12%
$\hat{m}^{oos}$	-5.20%	-3.79%	-2.00%	-4.91%	-3.94%	1.15%

Notes: *** and ** denotes tests significant at the 1% and 5% levels, respectively. The lines $\hat{m}$ and CR display the in-sample performance of the corresponding model, while the lines $\hat{m}^{oos}$ and CR^{oos} report the out-of-sample performance.

4 Testing for structural breaks via CUSUM tests

In this Section, we provide technical details regarding the CUSUM tests mentioned in Section 3.2 in the main text. In particular, we discuss the necessary modifications to be made, in order to account for the specifics of our model. We then report the results of this test for EUR and CHF selected models.

In Kulperger and Yu [2005], the authors derived the asymptotic properties of partial sum processes constructed on k^{th} power of GARCH residuals, showing that it converges toward a Brownian process plus a correction term. Such CUSUM statistics can be used to test for a change in conditional (potentially high-order) moments over time. As implied by their Theorem 1.1 and 1.2, the partial sum process behaves as if the residuals $\hat{z}_t = r_t/\hat{\sigma}_t$ were asymptotically the same as the innovations z_t. However, in usual GARCH models, z_t are assumed (unconditionally) i.i.d, whereas in our GARCH-SH model, it is not the case under the null hypothesis of no breaks. To circumvent this issue, we suggest working instead with Gaussian pseudo-residuals $\hat{u}_t$, based on the inverse of the sinh-arcsinh transform given by (3). Thus, we defined these pseudo-residuals as

$$\hat{u}_t = \sinh\left(\hat{\delta}_t \sinh^{-1}(\hat{z}_t) - \hat{\epsilon}_t\right),$$

where $\hat{z}_t = (R_t - \hat{c} - \hat{\lambda}\hat{\sigma}_t)/\hat{\sigma}_t$. Under a correct specification of the sinh-arcsinh distribution, $\hat{u}_t$ is asymptotically N(0,1) distributed and fulfills the main assumptions of Kulperger and Yu [2005]. It also fulfills the assumptions of zero-mean and unit-variance. Two additional requirements are the finiteness of the k^{th} moment of z_t and that u_0 is a non-degenerate random variable. These conditions are fulfilled when we assume that ϵ_t and δ_t are finite. Then, we suggest using the following test statistic, similar to the one proposed in Kulperger and Yu [2005]:

$$CUSUM^{(k)} = \max_{1 \leq i \leq T} \frac{|\sum_{t=1}^{i} \hat{u}_t^k - i\hat{\mu}_k|}{\hat{s}_k \sqrt{T}}, \tag{2}$$

where $\hat{\mu}_k$ is the empirical moment of order k of the residuals, and $\hat{s}_k$ an estimate of $E(u_0^k - \mu_k)^2$. A formal proof of the asymptotic properties of (2) is beyond the scope of the paper. On the basis of the theoretical arguments enumerated previously, we use the (approximated) results that (2) converges to the supremum of a Brownian bridge:

$$CUSUM^{(k)} \xrightarrow{a.s.} \sup_{0 \leq u \leq 1} |B_0(u)|.$$

Additionally, since we might face several breaks in the time series, we need an algorithm to sequentially identify the dates of the breaks. We simply apply the procedure detailed in Inclan

and Tiao [1994], consisting in repeatedly partitioning our time series, until no more breaks are found.

Relying on the simulation set-up described in Appendix B in the main text, we briefly study the size of this test. Results are displayed in Table 12. CUSUM tests appear slightly under-sized.

Rejections of H_0: no breaks

No mean					**GARCH-in-Mean**				
DGP	T	k=2	k=3	k=4	DGP	T	k=2	k=3	k=4
DGP1	500	0.024	0.04	0.012	DGP4	500	0.04	0.048	0.018
	1500	0.032	0.032	0.026		1500	0.044	0.038	0.026
	3000	0.044	0.052	0.05		3000	0.042	0.052	0.04
DGP2	500	0.024	0.042	0.034	DGP5	500	0.028	0.036	0.028
	1500	0.03	0.046	0.032		1500	0.042	0.058	0.038
	3000	0.046	0.04	0.036		3000	0.042	0.034	0.036
DGP3	500	0.042	0.038	0.03	DGP6	500	0.04	0.036	0.042
	1500	0.036	0.038	0.028		1500	0.07	0.044	0.05
	3000	0.044	0.042	0.032		3000	0.024	0.042	0.022

Table 12: Type-I error for testing the null hypothesis of no structural breaks, for DGP either with no mean structure (left) or with a GARCH-in-Mean structure (right).

Finally, we display the value over time of the CUSUM test statistic used to investigate potential structural breaks (Figure 19).

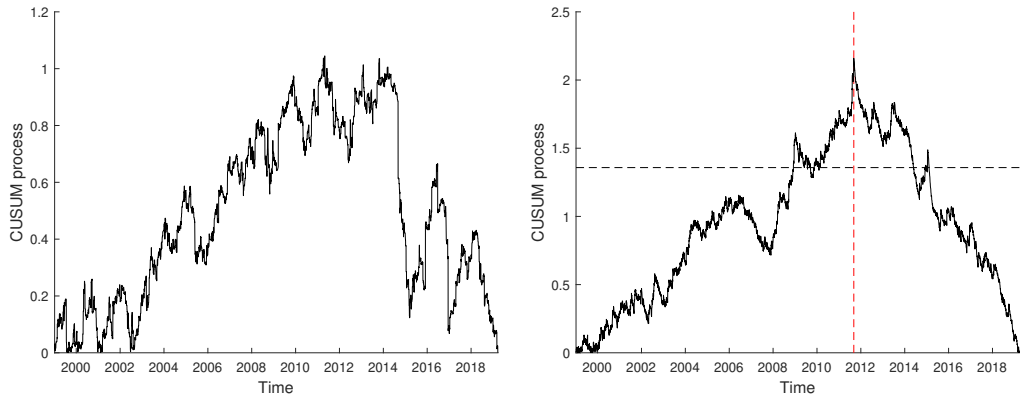


Figure 19: CUSUM process over time. Left: for EUR, with $k = 3$. Right: CHF, with $k = 2$ (solid). Dashed red: break date (6 September 2011) identified with the algorithm of Inclan and Tiao [1994]. Dashed black: rejection threshold of the test.

5 Additional information on the SH distribution

In this Section, we provide additional information regarding the sinh-arcsinh (SH) distribution used in our econometric approach.

5.1 The sinh-arcsinh distribution

The probability density function of the sinh-arcsinh distribution is given by

$$f(z;\epsilon,\delta) = \eta^{-1} Z_{\xi,\eta}(x)^{-1/2} \delta C_{\epsilon,\delta}((x-\xi)\eta) \exp(-S^2_{\epsilon,\delta}((x-\xi)/\eta)/2),$$

where

$$\begin{aligned}
Z_{\xi,\eta}(x) &= (2\pi(1+((x-\xi)/\eta)^2)), \\
C_{\epsilon,\delta}(x) &= \cosh(\delta \sinh^{-1}(x) - \epsilon) = (1+S^2_{\epsilon,\delta}(x))^{1/2}, \\
S_{\epsilon,\delta}(x) &= \sinh(\delta \sinh^{-1}(x) - \epsilon), \\
\xi &= -\eta \sinh(\epsilon_t/\delta) P_{1/\delta} \\
\eta &= \sqrt{1/(0.5(\cosh(2\epsilon_t/\delta_t)P_{2/\delta} - 1) - \sinh(\epsilon_t/\delta)P_{1/\delta})^2}, \\
P_q &= \frac{\exp(1/4)}{8\pi^{1/2}} \left(K_{(q+1)/2}(1/4) + K_{(q-1)/2}(1/4)\right).
\end{aligned}$$

with K being the modified Bessel function of the second kind. ξ and η are the location and the scale parameters, respectively, whose values are fixed to ensure zero mean and unit variance. The cumulative distribution function $F(z;\epsilon,\delta)$ is obtained from the transformation given in (3) and is simply:

$$F(z;\epsilon,\delta) = \Phi(\sinh(\delta \sinh^{-1}((z-\xi)/\eta) - \epsilon),$$

where $\Phi(\cdot)$ is the CDF of the standardized Gaussian distribution. The quantile function F^{-1} is easily derived in the same way; and is given by

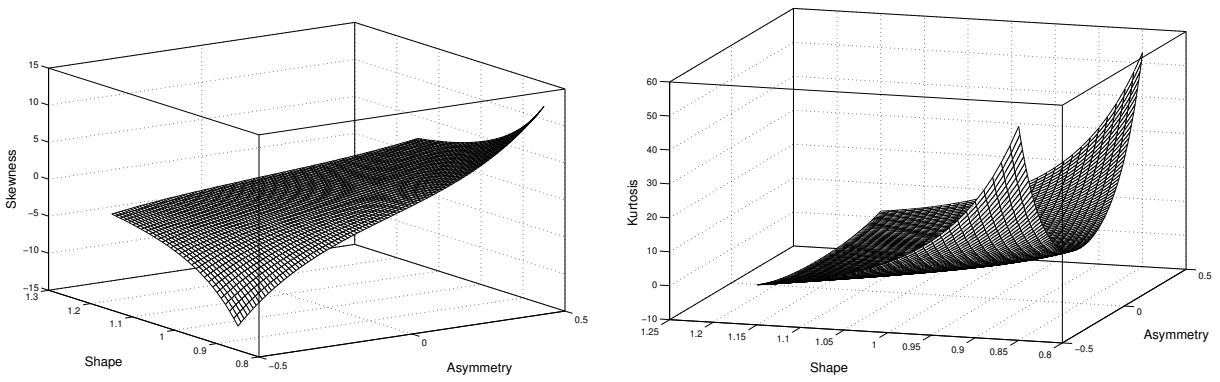
$$F^{-1}(p;\epsilon,\delta) = \sinh\left((1/\delta) * \sinh^{-1}(\Phi^{-1}(p)) + (\epsilon/\delta)\right)\eta + \xi,$$

where Φ^{-1} is the quantile function of the Gaussian distribution.

The skewness and kurtosis of z_t are given by

$$\begin{aligned} SK_t &= \frac{1}{4}\{\sinh(3\epsilon_t/\delta)P_{3/\delta_t} - 3\sinh(\epsilon_t/\delta_t)P_{1/\delta}\}, \\ KU_t &= \frac{1}{8}\{\cosh(4\epsilon_t/\delta)P_{4/\delta} - 4\cosh(2\epsilon_t/\delta_t)P_{2/\delta} + 3\}, \end{aligned}$$

where P_q is defined by equation (3). Hence, one observes that both quantities depend on both parameters. The response surfaces for skewness and kurtosis are displayed in Figure 20 for various values of ϵ and δ.



(a) (b)

Figure 20: (a) Skewness and (b) Kurtosis

As depicted in Figure 20 (left panel), δ has a limited impact on the skewness (the response surface is quite flat along this dimension).

5.2 Characteristics of SH distribution

As explained in Jones and Pewsey [2009], the SH distribution is conveniently built around the Gaussian distribution such that, assuming a random variable $Y \sim N(0,1)$, we can define $f(z;\epsilon,\delta)$ by the sinh-arcsinh transformation:

$$Z = \sinh\left(\frac{\sinh^{-1}(Y) + \epsilon}{\delta}\right). \tag{3}$$

Skewness increases with increasing ϵ for $\epsilon \in]-\infty, +\infty[$, where $\epsilon > 0$ corresponds to positive skewness. Notice that positive (negative) skewness, for a standardized random variable, implies

that there is more probability mass below (above) zero. The kurtosis decreases with increasing δ, $0 < \delta < +\infty$, $\delta < 1$ yielding heavier tails than the normal distribution. Thus, the Gaussian distribution has a central position in the SH distribution. This is an advantage compared to other distributions, for which the Gaussian distribution is usually a limiting case. Another advantage of the SH distribution is the existence of all its moments for finite values of the parameters. This is particularly useful for inference and residuals-based tests.

For instance, assuming that all parameters in (2.1) take the value zero, we are back to the symmetric case. Setting $a_1 = a_2 = a_3 = 0$ leads to a model without dynamics but including asymmetry, whereas assuming a_1 and a_3 to be zero leads to a model where only past innovations impact on the asymmetry. Furthermore, as explained in Jondeau and Rockinger [2003], a model in which $a_2 = a_3 = 0$ and $a_1 \neq 0$ is not properly identified: for t sufficiently far from zero, ϵ_t equals its stationary value $\epsilon^* = a_0/(1 - a_1)$. Since we do not set an additional restriction linking a_0 and a_1, an infinity of pairs exists (a_0, a_1) to solve this equation. In practice, the estimation will converge at random, depending on the starting value chosen for ϵ_0. Thus, in our application, we assume that at least one of the other coefficients is always different from zero. In addition, the stability of the process is fulfilled when $|a_1| < 1$.

References

J. Berkowitz. Testing density forecasts, with applications to risk management. *Journal of Business & Economic Statistics*, 19(4):465–474, 2001.

J.A. Doornik and H. Hansen. An omnibus test for univariate and multivariate normality. *Oxford Bulletin of Economics and Statistics*, 70(s1):927–939, 2008.

O. Hossfeld and R. MacDonald. Carry funding and safe haven currencies: A threshold regression approach. *Journal of International Money and Finance*, 59:185–202, 2015.

C. Inclan and G. Tiao. Use of cumulative sums of squares for retrospective detection of changes of variance. *Journal of American Statistical Association*, 89(427):913–923, 1994.

E. Jondeau and M. Rockinger. Conditional volatility, skewness, and kurtosis: Existence, persistence, and comovements. *Journal of Economic Dynamics and Control*, 27(10):1699–1737, 2003.

M.C. Jones and A. Pewsey. Sinh-arcsinh distributions. *Biometrika*, 96(4):761–780, 2009.

R. Kulperger and H. Yu. High moment partial sum processes of residuals in garch models and their applications. *The Annals of Statistics*, 33(5):2395–2422, 2005.

H. Pesaran and A. Timmermann. Testing dependence among serially correlated multicategory variables. *Journal of the American Statistical Association*, 104(485):325–337, 2009.