What events matter for exchange rate volatility?

Igor Martins*and Hedibert Freitas Lopes †

October 2023

Abstract

This paper expands stochastic volatility models by proposing a data-driven method to selected the macroeconomic events more likely to impact volatility. The paper identifies and quantifies the effect of macroeconomics events of multiple countries on exchange rate volatility using high frequency currency returns while accounting for persistent stochastic volatility effects and seasonal components capturing time of the day patterns. Due to the hundreds of macroeconomic announcements and its lags, we rely on sparsity based methods to selected relevant events for the model. We contribute to the exchange rate literature in four ways: First, we identify the macroeconomic events that drives currency volatility, estimate their effect, connect them to macroeconomic fundamentals and show how they can be linked to lower frequency currency returns using a model averaging argument. Second, we find a connection between intraday seasonality, trading volume and opening hours of majors markets across the globe and provide a simple labor-based argument for the pattern found. Third, we show that inclusion of macroeconomic events and seasonal components are key for forecasting exchange rate volatility. Fourth, applying our proposed model for multiple currencies alongside a dynamic copula yields a Sharpe ratio 3.5 times higher than using standard SV and GARCH models.

Key words: Stochastic Volatility; Macroeconomic Announcements; Sparsity; Seasonality.

1 Introduction

Exchange rate volatility is a central topic in macroeconomic and finance. Accurate forecasting and the understanding of the mechanisms behind it have been crucial for policymakers and investors. Blanchard et al. [2015] and Fratzscher et al. [2019] indicate the willingness of central bankers to intervene in foreign exchange (FX) markets to smooth exchange rate fluctuations and, therefore, limit FX volatility supported by recent theories of welfare gains

*Insper. igorfbm@al.insper.edu.br
†Insper. hedibertfl@insper.edu.br

due to interventions such as Gabaix and Maggiori [2015]. Bhansali [2007] connected the popular carry trading, a strategy based on buying currencies from countries with a high interest rate, denominated investment currencies, and selling the exchange rate of countries with a low interest rate, known as funding currencies, is effectively a form of short volatility trade. This paper tackles the volatility forecasting problem and its determinants by modeling the volatility of high frequency FX returns using a new stochastic volatility model capable of capturing the announcement effect of hundred of macroeconomic variables from multiple countries via spike and slab priors while also taking into account seasonal components capturing time of the day patterns.

Our model captures three main features of intraday returns: volatility persistence, time of the day effects and macroeconomic announcements. Volatility persistence is ubiquitous in financial markets in high and low frequency settings. Currency returns are no exceptions e.g. Andersen and Bollerslev [1998] and Bauwens et al. [2005]. We opt to capture volatility persistence via SV models due to their record of outperforming GARCH in volatility forecasting and trading for intraday index returns shown in Stroud and Johannes [2014].

The model accounts for time of the day patterns by considering dummy variables in each 5-minute window while also accounting for persistence and the effect of macroeconomic announcements. Several papers employ time of the day effects for intraday FX volatility modeling with mixed conclusions. Ito and Hashimoto [2006] observes a U-shaped pattern for both the Japanese Yen and the Euro quoted in US Dollars starting at 8:00 GMT going up to 15:00 GMT. Ederington and Lee [2001], however, points to the U-shaped pattern on FX markets being due to macroeconomic announcements on specific days of the week, and after accounting for this feature the U-shaped seasonal effect vanishes. Thus, our proposed model provides a reasonable setting to test the claims of Ito and Hashimoto [2006], Ederington and Lee [2001] as well as other possible patterns. While being common in the intraday literature, seasonal effects may also play a relevant role in lower frequency returns and in other assets classes e.g. Sørensen [2002] and our model can be easily be adapted to such settings.

Our methodological contribution comes when modeling announcements. Previous papers, such as Andersen and Bollerslev [1998], Bauwens et al. [2005] and Andersen et al. [2007], select a small number of events based solely on his experience, at most 25, and estimate their effect. This approach may lead to several problems. First, there is a clear possibility of cherry-picking the announcements. Second, by neglecting the inclusion of relevant announcements, estimates of the seasonal or even the persistence component may be contaminated. Third, it may hinder the relevant macroeconomic channels. Fourth, if the researcher select irrelevant events, the model is over-parameterized with potential increases in the uncertainty of other parameters. Selecting relevant events solely based on experience is equivalent to assigning probability one to its inclusion and zero otherwise. We contribute by allowing the probability of inclusion to be determined not only by the researcher's prior knowledge but also from the data. Specifically, we allow for the effect of the announcement on volatility to come from a mixture of two distributions. One component of the mixture comes from a Dirac's delta on zero, reflecting no effect, while the other is represented by a Gaussian with large variance in order to accommodate a large range of effects on volatility.

We model 5-minutes returns of the Australian Dollar, AUDUSD, a currency used as an investment currency on carry trade strategies as shown in Lustig et al. [2011]. We allow for all macroeconomic announcements from the US and Australia available at Bloomberg's Economic Calendar to possibly affect FX volatility via announcement dummies. Among hundreds of macroeconomic announcements, the model identifies two groups of variables as having more than 95% posterior probability of inclusion: Taylor rule variables and external imbalances measures. Both groups are tied to macroeconomic fundamentals. Additionally, we link the high frequency volatility induced by news of those variables to lower frequency predictability of exchange rate using a model average argument. Similarly to Ito and Hashimoto [2006], we obtain a U-shaped pattern starting at 8:00 and going up to 14:30 GMT. However, in contrast to Ederington and Lee [2001] and Ito and Hashimoto [2006], we find an additional U starting at 1:30 GMT and going up to 8:00 GMT, leading to a W pattern. The new W-shaped pattern may reflect the growth experienced by the Chinese and Japanese markets since the early 2000s. Additionally, the estimated seasonal component is informative about traded volume. Finally, we connect the spikes in our estimated seasonal component to the opening of major markets and propose a simple labor economics explanation for the observed pattern. Our proposed model outperform SV and GARCH competitors in terms of out-of-sample R2 and MAE, and also improves upon the model with all events but without spike and slab priors.

We also investigate the performance of our model on a portfolio allocation problem. We allow an investor to choose the amount of money to allocate between the Swiss Franc, CHFUSD, and Australian Dollar. FX traders commonly use them as funding and investment currencies on carry strategies. By combining the volatilities from our proposal with the dependence structure obtained via Dynamic copula of Hafner and Manner [2012], an investor would achieve a Sharpe ratio of at least 3.5 times higher than when he using traditional SV or GARCH models.

The paper is organized as follows. It starts by describing the FX returns, all macroeconomic announcements, our model and estimation approach on Section 2. Section 3 shows our estimates for the macroeconomic announcements, time of the day effects, and volatility persistence. Section 4 presents the volatility forecasting and portfolio allocation applications. Section 5 concludes.

2 Data and proposed model

2.1 Data

The paper models 5 minutes returns of Australian Dollar across 24 hours from January 2, 2018 to December 31, 2022 a period with 1825 days. We use the dataset from January 2, 2018 to June 30, 2022 for estimation and the remaining observations for out-of-sample analysis. The AUDUSD is traded 24h a day from Sunday 22h (GMT) to Friday 22h (GMT) and comes from Dukas Copy Swiss Banking Group's Historical data feed.

We consider 119 macroeconomic events from Australia and the US as possible sources of volatility for the Australian Dollar. By considering six lags, we obtain 714 event-related sources of volatility. Our approach is flexible enough to include events from other countries as well. However, since the US and Australia already provide hundreds of events, we consider just the events of those two countries in this analysis. All timestamps for the macroeconomic announcements are obtained via Bloomberg's economic calendar.

For our portfolio allocation, we also consider returns of the Swiss Franc during the same time span and using the same splits for estimation and forecasting as used for the AUDUSD. Similarly, we include macroeconomic announcements for the US and Switzerland when modeling CHFUSD. The whole list of events considered is available in Appendix A.

Our choice of modeling the AUDUSD, and latter inclusion of CHFUSD in our portfolio application, is due to their critical role in carry trade strategies. AUDUSD and CHFUSD are commonly used as investment and funding currencies, respectively, as shown in Lustig et al. [2011].

2.2 Model, priors and estimation

We model 5-min log-returns, y_t, allowing for time-varying volatility v_t as shown in Equation (1).

$$y_t = v_t \varepsilon_t \text{ with } \varepsilon_t \sim N(0,1) \tag{1}$$

We follow Stroud and Johannes [2014] in using total volatility with a multiplicative specification as shown by Equation (2) which also implies a linear specification for the log variance h_t represented by Equation (3).

$$v_t = \sigma X_t S_t E_t \tag{2}$$

$$h_t = log(v_t^2) = log(\sigma^2) + log(X_t^2) + log(S_t^2) + log(E_t^2) = \mu_h + x_t + s_t + e_t \tag{3}$$

σ represents the volatility level i.e. v_t when $X_t = S_t = E_t = 1$. X_t represents a latent persistent stochastic volatility component represented by Equation (4)

$$x_t = \phi x_{t-1} + \sigma_x \eta_{x,t} \text{ with } \eta_{x,t} \sim N(0,1) \tag{4}$$

S_t and E_t represent the seasonal and event components, respectively. We model the seasonal effect via $s_t = \sum_{k=1}^{288} H_{tk}\beta_k$, where H_{tk} is an indicator variable that assumes the value 1 if time t corresponds to the 5 minute window k and β_k denotes the coefficient associated with the seasonal effect of period k.

We model the event effects similarly. $e_t = \sum_{i=1}^{n} I_{itk}\alpha_{ik}$. I_{itk} is an indicator variable assuming the value 1 if an event i occurred at time t which lies into window k and α_{ik} is the coefficient that captures the effect of such event. Our model assumes that each event may impact volatility up to 30 minutes after it's release. Since our modeling approach relies on five minutes windows, we include six lags on our model. For a small number of events, the announcement component could be recovered via linear regressions shown in Shumway et al. [2000] or via smoothing splines exemplified in Stroud and Johannes [2014] without compromising predictive performance. Since we consider hundreds of events, a sparsity inducing approach may also improve forecast performance. Additionally, due to our interest in selecting which announcements may affect volatility, we want to disentangle inclusion probability and the associated effect after inclusion. Thus, we consider spike and slab priors for α_{ik}. This approach is not only sparsity inducing but also disentangles inclusion probability and the effect associated with the event.

The spike and slab prior was introduced by Mitchell and Beauchamp [1988], George and McCulloch [1993] and Geweke [1996] being recently expanded by Ishwaran and Rao [2005]. Its core idea, as presented in Equation (5), is to allow each α_{ik} from the announcement component to be modeled as having come either from a distribution $p_{spike}(\alpha_{ik}|\theta)$ with mass concentrated around zero or from a distribution $p_{slab}(\alpha_{ik}|\theta)$ with mass covering a long range of values.

$$\alpha_{ik}|\pi \sim (1-\pi)p_{spike}(\alpha_{ik}|\theta) + \pi p_{slab}(\alpha_{ik}|\theta) \tag{5}$$

For example, George and McCulloch [1993] consider both $p_{spike}(\cdot)$ and $p_{slab}(\cdot)$ to be Gaussian distributions with different variances. Here, we consider the spike to be a Dirac delta at zero and the slab to be a Normal with mean 0 and variance $\sigma_a^2\tau^2$. The probability of α_{ik} coming from either the spike or the slab is modeled via $\pi \sim Bern(\theta)$. Since θ must be between 0 and 1, we assume a beta prior. We use IG priors σ_a and τ^2.

For the remainder of the parameters, we assume natural conjugate priors. For both the seasonality coefficients and for the persistence of the AR(1) in Equation (4), we assume normal priors. For all variance parameters we consider IG priors. We assume no effect due to the seasonality and announcement components and $x_t = log(y_t^2)$ as initialization values for the MCMC procedure. Appendix B presents a full description of the priors, hyperpriors, and a summary of the model.

We employ a Bayesian approach and use MCMC methods to simulate from the posteriors distribution, i.e., the joint distribution of parameters and latent states conditional on the observed returns $\{\mu_h, \phi, \sigma_x^2, \{\beta_k\}_{k=1}^{288}, \{\alpha_{ik}\}_{i=1}^{n}$, $\pi, \theta, \tau^2, \sigma_\alpha^2, \{x_t\}_{t=1}^{T}\}|y_t$ a full description of the MCMC strategy employed to sampled from the posterior is discussed in Appendix B. Our choice of priors allow for great simplification of the sampling scheme. In summary, we can sample from the posterior of (conditional) linear regression coefficients and variances via Normal and IG distributions. $p(\theta|\cdot)$ falls into the Beta - Bernoulli conjugate case leading to a Bernoulli posterior, we can avoid problems due to the Dirac's delta when sampling from $p(\pi|\cdot)$ using the approach proposed by Geweke [1996] and, finally, we can recover $\{x_t\}_{t=1}^{T}|\cdot$ using Kim et al. [1998] seven Gaussian components approach. We use 6000 draws with 2000 as burn-in.

3 Results

3.1 Macroeconomic announcements

From the 119 events considered as possible sources of volatility, only six have at least 95% posterior inclusion probability of its first lag: RBA Cash target rate, FOMC rate decision, US non - farm payroll, Australian GDP, American CPI and Australian Trade Balance. The effects on volatility produce by the six events up to 30 minutes, i.e. up to six lags, after their release are presented in Figure (1). The posterior inclusion probability for the fist lag of all events is present in Appendix C.

Figure 1: Heatmap containing the announcement effect,the posterior mean of $exp\left(\frac{a}{2}\right)$, for events with average posterior probability of inclusion higher than 95% for the first lag. x-axis contain the number of the lag corresponding to the five minute window after the announcement occurred capturing the effect on volatility up to 30 minutes after the announcement. Exchange rate volatility is connected to news about two groups of macroeconomic fundamentals: Taylor rule and external imbalance measures.

The six events are informative about two groups of variables used in fundamental-based macroeconomic models: Taylor rule and external imbalance measures. Taylor [1993] indicates that a central bank sets the interest rates as a function of the difference between current inflation and its target level, and also as a function of the output gap which is the difference between actual and potential output of an economy. The decision about the nominal interest rate are represented by FOMC rate decision for the USA and RBA cash target rate for Australia. American CPI is used to determine current inflation in the US. Australian GDP releases are used to construct the output gap. Due to the FED dual mandate of price stability and maximum sustainable employment, US non-farm payroll influences the central bank decision of interest rates and is useful for forecasting output gap as shown in Chan and Grant [2017]. Therefore, the first five variables are selected as Taylor rule related variables. The link between current account and trade balance is straightforward.

We can connect currency returns with our findings based on Campbell and Clarida [1987] and Gourinchas and Rey [2007]. Starting with the definition of log excess currencies returns, Campbell and Clarida [1987] obtain Equation (6) connecting nominal exchange rate, s_t, via differences of expected nominal interest rates between countries, $E_t \sum_{\tau=0}^{\infty} i_{t+\tau} - i^*_{t+\tau}$, expectations about currency risk premium, $E_t \sum_{\tau=0}^{\infty} r^e_t$, and the expected long-run exchange rate $E_t \lim_{\tau\to\infty} s_{t+\tau}$. Therefore, exchange rates are connected to interest rate differentials. Since central banks decide interest rates based on Taylor rule variables, news about such variables influence interest rates which in turn influence the exchange rate.

$$s_t = E_t \sum_{\tau=0}^{\infty} (i_{t+\tau} - i^*_{t+\tau}) + E_t \sum_{\tau=0}^{\infty} r^e_t + E_t \lim_{\tau\to\infty} s_{t+\tau} \tag{6}$$

Gourinchas and Rey [2007] shows that when a country experiences a current account imbalance, it must run future trade surpluses or receive a wealth transfer from countries via currency depreciation. Based on this observation, they propose a measure of cyclical external imbalance, nxa_t, representing deviations from a trend that combines exports, imports, assets, and liabilities which incorporates information from both the trade balance and the foreign asset position. Since information about trade balance can be useful to forecast currency returns, it is reasonable to consider that investors react to the new information leading to increase in exchange rate volatility.

While one may claim that such are events are obvious inclusions, they are often neglected. For example, Bauwens et al. [2005] ruled out the possibility of trading balance or non-US central bank meetings affecting currency volatility. Marshall et al. [2012] ignores the possibility of trading balance, inflation and non-US interest rate decisions affecting volatility. Chen and Gau [2010] rules out all non-US events.

The same two group of variables, Taylor rule related and external balance measures, are shown by Rossi [2013] to be the more likely to predict currency returns on monthly and quarterly frequencies. The connection between currency returns predictability on lower frequencies and volatility right after macroeconomic announcements can be understood using a model averaging argument. Suppose that an investor aims to forecast currency returns using models $M = \{M_1, M_2, \dots\}$. The posterior model probability of model M_j with parameters Θ_j is given by Equation (7). One of its components, $f(y|M_j)$, is the predictive density of model j, and the other, $Pr(M_j)$ is model j prior probability. Therefore, assuming same prior probabilities $Pr(M_j)$ for all j, models with higher predictive likelihood will have higher probabilities.

$$Pr(M_j|y) \propto f(y|M_j)Pr(M_j) = \int f(y|\Theta_j, M_j)P(\Theta_j|M_j)p(M_j)d\Theta_j \tag{7}$$

When news on a macroeconomic variable is released, the investor will look if the variable is part of a model with non-zero weight. If the weight is non-zero, he will adjust his portfolio to reflect the news causing volatility. Otherwise, he won't adjust and there will be no impact on volatility.

Since, according to Rossi [2013], Taylor rule variables and external balance measures have higher predictive likelihood, a investor will given more weight for those models when accessing currency returns and news about such variables will be reflect as currency volatility. Additionally, as show in Meese and Rogoff [1983], several models have no predictive power for forecasting exchange rate, so it is reasonable to assume that several models will have 0 or close to 0 posterior probability and therefore news about fundamentals of those models won't impact or will have a small effect on currency volatility.

3.2 Seasonality

There are three main results connected to the seasonality component. First, we obtain a W-shaped curve for the seasonality effects, i.e. for the posterior mean of $exp\left(\frac{s}{2}\right)$, shown in Figure(2). Our estimate presents a clear U shape that starts on the Shanghai and Hong Kong opening hours (01h30 GMT) and finishes on the opening of the London market (08:00 GMT) and a second U pattern starts at the opening hours of the London market and finishes on the opening hours of the New York stock exchange (14h30 GMT) with peaks of 1.465, 1.477 and 1.579 representing increases of 46.5%, 47.7% and 57.9%, respectively, when compared to the baseline volatility level of 0.031 (8.3% on an annualized scale) represented by σ. The opening of the three mentioned markets are represented by vertical red, blue and green dashed lines on the figure. This W pattern differs from the typical U-shaped seasonal effects on volatility. For example, Harvey and Huang [1991], Hautsch [2011] andStroud and Johannes [2014] obtain a single U for the Japanese Yen quoted in dollars, individual assets and for the SP500, respectively, with higher effects during the opening and closing minutes of the US Market.

Second, the estimated seasonality effects peaks when a major exchange opens. While the W-shaped pattern is connected to the opening hours of the Chinese, London and New York markets, those are not the only markets that affect the estimated seasonal component. For example, at the Tokyo (00:00 GMT) and Frankfurt (07:00 GMT) openings, the seasonal component increases on more than 40% from the baseline volatility level. For the Australian opening (23:00 GMT) we have a more modest increase of 4.2%, however this value is high when compared to the hours after its opening and prior to the Tokyo opening.

Third, while we only use return information in our estimation, our seasonality component is informative about traded dollar volume. Figure(3) plots the average dollar amount traded on each 5 minutes window yielding a similar shape as the one present in Figure(2). The similarity between the plots is confirmed by the scatter plot in Figure(4) and a simple linear regression of average traded volume on the posterior mean of the seasonal effects. The regression implies that an increase in one standard deviation of the baseline volatility is associated with a traded volume of 726 millions of dollars and also leads to an R2 of 0.84.

The connection between volatility of returns and trading volume have been pointed out before by Abanto-Valle et al. [2010] on a daily framework. Our finding expands the connection between volatility and trading volume by showing that their connection is not limited to daily stock returns but is also present on other frequencies and asset

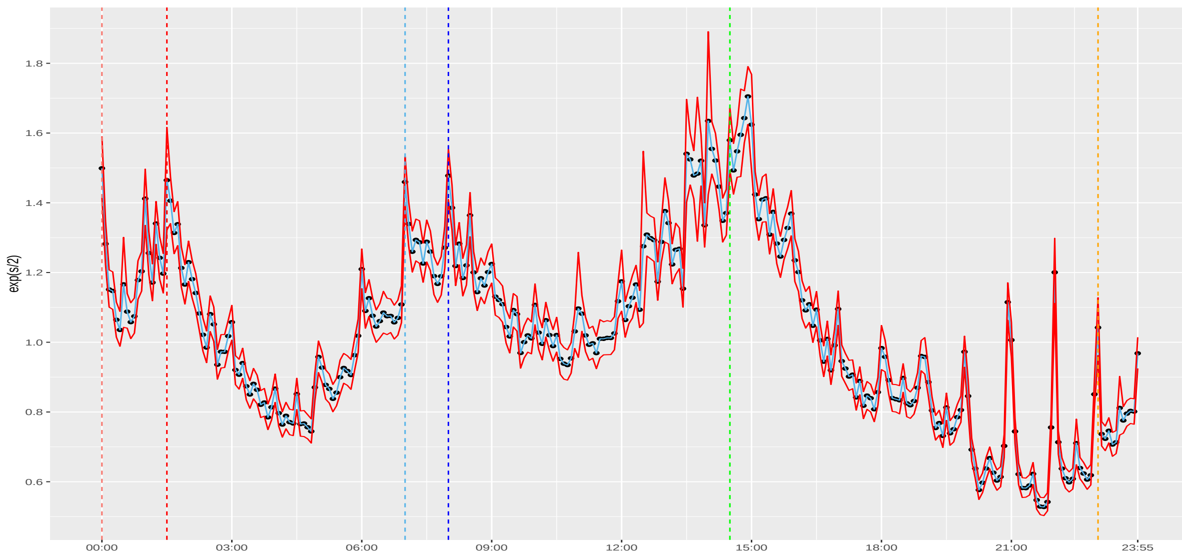

Figure 2: Estimated seasonal effect, posterior mean of $exp\left(\frac{s}{2}\right)$, for the Australian Dollar. An y-axis value of 1.5 indicates that volatility is 50% higher than its baseline value throughout the day. x-axis represents the time of the day split into five - minute windows. Black dots represents the posterior mean of each 5 minute interval connected by the solid blue line. Solid red lines indicate a 90% credible interval for the seasonal effect. Spikes on the seasonality component are linked to opening hours of major exchanges around the globe. For example, Tokyo (00:00 GMT), Hong Kong and Shanghai (01:30 GMT), Frankfurt (07:00 GMT), London (08:00 GMT) and New York (14h30 GMT) show an increase of over 40% the baseline level while Sidney (23h GMT) has a more modest increase of 4.2%. The opening hours are represented as dashed vertical lines in pink, red, light and dark blue, green and orange, respectively. The seasonality effect decreases within 90 minutes after the opening for each exchange. There is a notable W pattern peaking on the Chinese, European and American market openings.

classes, and that trading volume is largely associated with a specific portion of volatility: the seasonal component.

The spikes on seasonal volatility and trading volume on the opening of a market can be partly due to a simple labor-leisure. Traders begin their working day when a market opens and higher outputs have been largely associated with initial working hours, e.g. Pencavel [2015]. If we consider that a trader knows the dollar amount that he should have invested at the end of the day, he may benefit from trading on the beginning of the working day and using part of the remainder hours for leisure purposes.

3.3 SV component

Volatility persistence is a common characteristic in asset returns and its also present in our model. Figure(5) plots the combination of the level effect with the SV term, i.e. $\sigma X_t = exp\left(\frac{\mu_h}{2}\right)exp\left(\frac{x_t}{2}\right)$ when considering the posterior mean for both μ_h and x_t. The estimated posterior mean for the baseline volatility level $\sigma = exp\left(\frac{\mu_h}{2}\right)$ is 0.031 (8.3% on an annualized scale). By interacting with SV component, we get increases of 10 times the baseline value during the beginning of the COVID19 outbreak in the first trimester of 2020 persisting for months until getting close to its baseline level. This persistence, posterior mean of 0.990 for ϕ, is in line with values founded in other asset returns e.g. Stroud and Johannes [2014].

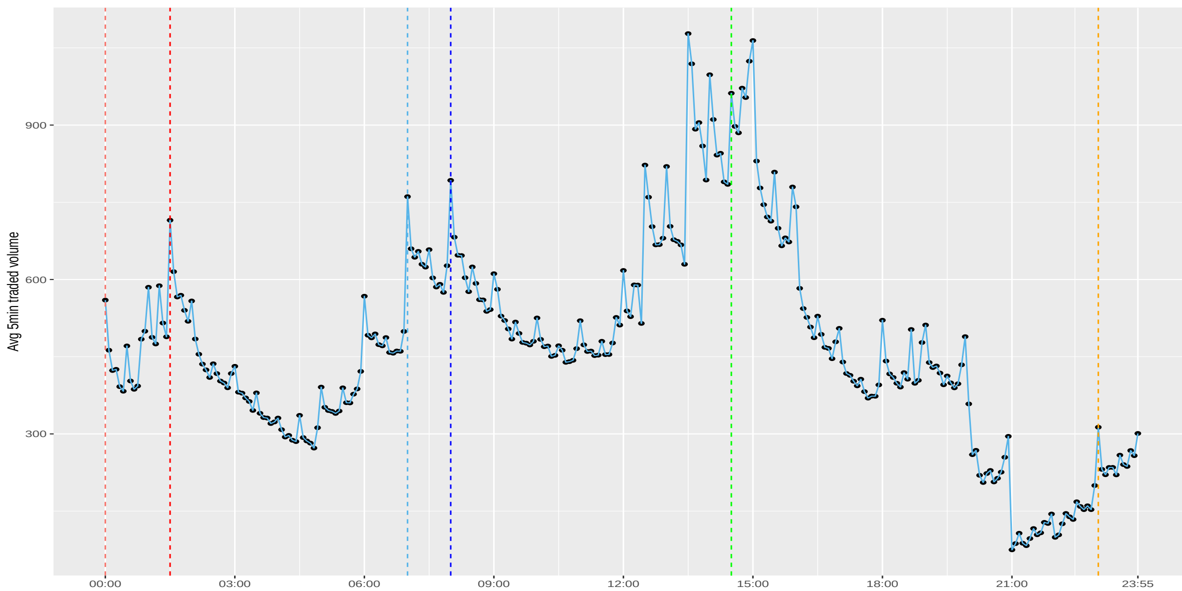

Figure 3: Average traded volume for the Australian Dollar, in millions, for each 5 minutes interval within a day. x-axis shows the time of the day split into five - minute windows. Black dots represents the average value connected via solid blue line. As in the estimated seasonality, we observe spikes on the opening of the major exchange and a W pattern with peaks on the Chinese, European and American market openings. As in Figure(2), the opening hours for Tokyo (00:00 GMT), Hong Kong and Shanghai (01:30 GMT), Frankfurt (07:00 GMT), London (08:00 GMT), New York (14h30) and Sidney (23h GMT) are represented as dashed vertical lines in pink, red, light and dark blue, green and orange, respectively. Both the W-shaped pattern and peaks when major markets open are also presented on the average traded volume

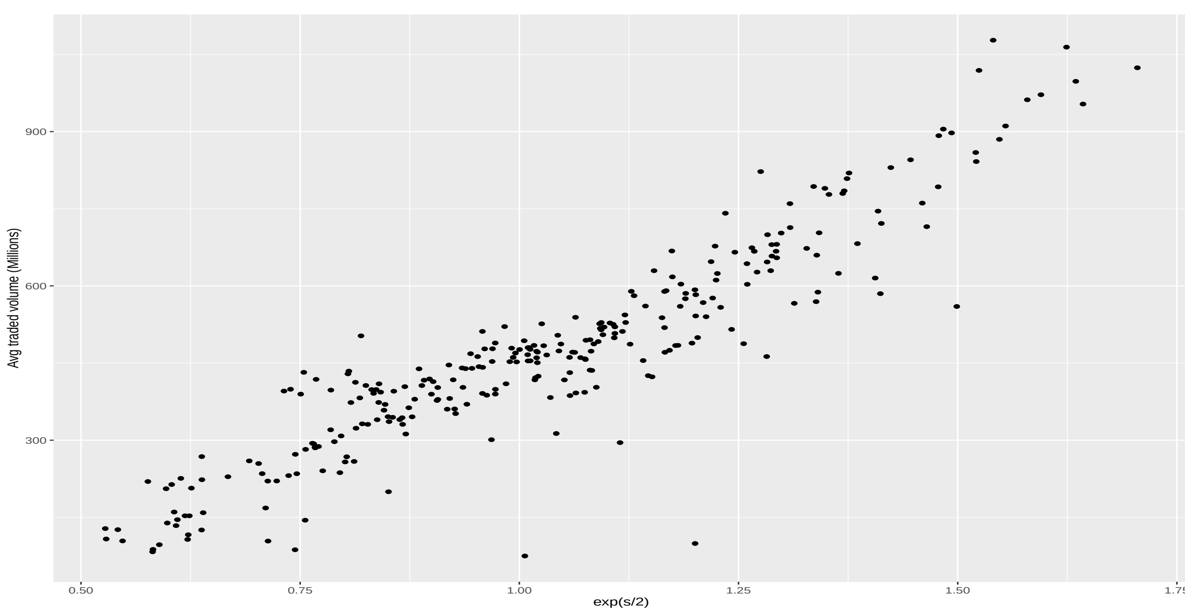

Figure 4: Scatter plot with Australian Dollar average traded volume, in millions, for each 5 minutes interval within a day on the y-axis and posterior mean of $exp\left(\frac{s}{2}\right)$ on the x-axis. The plot indicates a clear relationship between estimated seasonality and trading volume presented in Figures (2) and (3), respectively. A simple linear regression indicates that an increase in one standard deviation of the baseline volatility is associated with 726 millions of dollars. Such linear regression has an R2 of 0.84

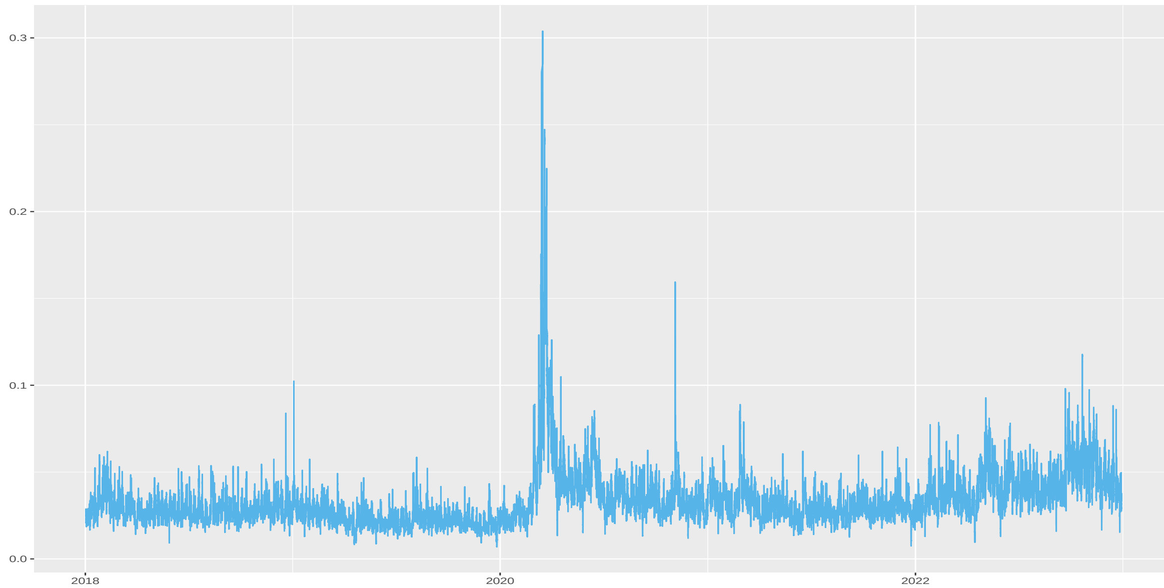

Figure 5: Effect of the level and stochastic volatility components represented by the posterior mean of $exp\left(\frac{\mu_h}{2}\right)exp\left(\frac{x_t}{2}\right)$. The rise of volatility on 2020 can be associated with the COVID-19 outbreak.

4 Forecast and portfolio applications

4.1 Forecasting

As discussed in Andersen and Benzoni [2008] and Stroud and Johannes [2014], volatility forecasting is required for nearly every financial application being considered the gold-standard for evaluating intraday models. We must make a comprise about how to evaluate such forecasts since volatility is unobserved. Our approach is to consider 5 minute RV for the AUDUSD as our forecasting target. As described in Section 2, we use return data after June 30 23:55 GMT to perform such forecasts.

Following Stroud and Johannes [2014], we compare the performance of our model by looking at out-of-sample R2, MAE and by running horse-race regressions. Table (1) indicates that our model produces the highest R2 out-of-sample and the smallest mean absolute errors for predicting the realized log-volatility 5 minutes ahead for the AUDUSD when compared to standard SV and GARCH models and even models that use realized volatility as part of its estimation. Additionally, it out performs more complex GARCH based models such as GJR-GARCH and apARCH. The results for the horse-race regressions presented on Table (2) provides more favorable evidence for our model by indicating that competitors provide little to no additional information for forecasting future volatility.

By restricting attention to the data points up to 30 minutes after an macroeconomic announcement, the importance of an announcement component is make evident when comparing our proposal to traditional methods that neglect this characteristic such as SV, GARCH, GJR-GARCH, apARCH and logRVAR1. Changing from our proposal to one of the traditional methods implies declines of R2 of at least 85% and increases of MAE of at least 37%. A natural question is if the variable selection approach compromises forecasting performance in order to improve

	Proposal	SV	GARCH(1,1)	logRVAR1	GJR-GARCH	apARCH
b0	-0.01	-0.01	0.02	-0.66	0.02	0.01
	(-18.22)	(-18.21)	(45.46)	(-79.91)	(45.58)	(25.65)
b1	1.11	1.29	0.59	3.65	0.58	0.74
	(147.82)	(86.72)	(95.12)	(85.48)	(94.45)	(99.20)
R2	0.37	0.17	0.20	0.16	0.19	0.21
$\Delta R2$	-	-54.51%	-47.08%	-55.59%	-47.68%	-43.42%
ΔMAE	-	14.25%	7.73%	12.20%	7.91%	6.22%

Table 1: Coefficients and statistics of $logRV_t = b_0 + b_1 log\widehat{Vol_{i,t|t-1}} + \varepsilon_t$. $logRV_t$ represents the out-of-sample 5-minutes log-realized volatility based on 1 minute returns. $log\widehat{Vol_{i,t|t-1}}$ is the log-volatility forecasted by the model i for time t using information up to t-1. Values in parenthesis indicates t-statistics associated with the regression coefficients. $\Delta R2$ and ΔMAE represent the percentage change of models R2 and MAE when compared to the proposed model. The proposed model produces the highest t-statistics for b1, the highest R2, and the lowest MAE.

	SV	GARCH(1,1)	logRVAR1	GJR-GARCH	apARCH
b1	1.00	0.87	1.00	0.88	0.89
	(95.69)	(127.32)	(123.57)	(128.04)	(104.55)
1 - b1	0.00	0.13	0.00	0.12	0.11

Table 2: Horse-race regression $logRV_t = b_0 + b_1 Proposal\_log\widehat{Vol_{t|t-1}} + (1-b_1)Competitor\_log\widehat{Vol_{t|t-1}} + \varepsilon_t$. $logRV_t$ represents the out-of-sample 5-minutes log-realized volatility based on 1 minute returns. $log\widehat{Vol_{t|t-1}}$ is the log-volatility forecasted by a model for time t using information up to t-1. b1 coefficients close to 1 indicate that competitor models provide small additional contribution to forecasting realized volatility when compared to the proposed model.

interpretability. Table (3) show that our variable selection method enhances the out of sample perform in terms of both R2 and MAE when compared to the SSVA model which consider the same SV and seasonality components as our proposal but estimates the full vector of announcements effects without performing variable selection. Table (4) indicates that while the SSVA model has some information about volatility forecasting after macroeconomic announcements, our proposal still accounts for the majority of the informational share. When compared to model without an announcement component, our model dominates all of its competitors.

	Proposal	SSVA	SSV	SV	GARCH	logRV	GJR-GARCH	apGARCH
b0	-0.05	-0.04	-0.04	-0.10	0.10	-0.68	0.10	0.10
	(-5.67)	(-4.81)	(-1.49)	(-2.96)	(4.72)	(-1.22)	(4.77)	(4.26)
b1	1.86	1.47	2.29	4.33	0.11	4.08	0.06	0.20
	(24.24)	(23.63)	(5.85)	(6.24)	(0.28)	(1.41)	(0.17)	(0.48)
R2	0.61	0.59	0.08	0.09	-0.00	0.00	-0.00	-0.00
Δ R2	-	-2.02%	-86.81%	-85.12%	-100.04%	-99.57%	100.42%	-100.33%
Δ MAE	-	2.35%	36.07%	37.53%	51.03%	48.75%	52.23%	50.62%

Table 3: Coefficients and statistics of $logRV_t = b_0 + b_1 log\widehat{Vol_{i,t|t-1}} + \varepsilon_t$ when restricting to periods in which an event occurred no more than 30 minutes ago. The restriction emphasizes the role of the announcement component on the proposed model. $logRV_t$ represents the out-of-sample 5-minutes log-realized volatility based on 1 minute returns. $log\widehat{Vol_{i,t|t-1}}$ is the log-volatility forecasted by the model i for time t using information up to t-1. Values in parenthesis indicates t-statistics associated with the regression coefficients. $\Delta R2$ and ΔMAE represent the percentage change of models R2 and MAE when compared to the proposed model. The proposed model produces the highest t-statistics for b1, the highest R2, and the lowest MAE even when compared to the SSVA model which includes all events in the database.

	SSVA	SSV	SV	GARCH(1,1)	logRV	GJR - GARCH	apARCH
b1	0.76	1.00	1.00	1.00	1.00	1.00	1.00
	(2.73)	(12.79)	(13.17)	(16.74)	(16.33)	(16.78)	(16.58)
1 - b1	0.24	0.00	0.00	0.00	0.00	0.00	0.00

Table 4: Horse-race regression $logRV_t = b_0 + b_1 Proposal\_log\widehat{Vol}_{t|t-1} + (1-b_1)Competitor\_log\widehat{Vol}_{t|t-1} + \varepsilon_t$ when restricting to periods in which an event occurred no more than 30 minutes ago. $logRV_t$ represents the out-of-sample 5-minutes log-realized volatility based on 1 minute returns. $log\widehat{Vol}_{t|t-1}$ is the log-volatility forecasted by a model for time t using information up to t-1. As in Table (2) coefficients close to 1 indicate that competitor models provide small additional contribution to forecasting realized volatility when compared to the proposed model. SSVA provide some additional contribution to forecasting log RV. However, our proposal remains with the bulk of the contribution while also reducing the number of events considered from hundreds to only six events.

4.2 Portfolio application

One of the main applications of volatility forecasting is to serve as an input for portfolio allocation problems. We consider a global minimum variance portfolio (GMVP) in which a mean-variance investor can choose to take long or short position in either the AUDUD or CHFUSD. The choice of the AUDUSD and CHFUSD is due to their common use on carry trade, a strategy based on accruing returns due to interest rate differentials between the countries, which is largely affected by volatility as discussed in Bhansali [2007]. Therefore, our portfolio application can also be related to which side of the carry trade the investor should take on each period of time.

The GMVP problem is a reasonable tool to evaluate variance-covariance matrices since the weights of each asset, $w_i, i = \{1, 2\}$ don't depend on a model for the average return but only on the volatility of each asset, vol_i and their correlation, cor_{12}, as presented in Equation(8). Campbell [2017] provides a detailed textbook derivation of Equation (8).

$$w_{1,t} = \frac{vol_{2,t}^2 - cor_{12,t}}{vol_{1,t}^2 + vol_{2,t}^2 - 2cor_{12,t}} \text{ and } w_{2,t} = 1 - w_{1,t} \tag{8}$$

Our proposed model and its competitors produce information about the volatility of individual exchange rates. However, we must also model the correlation between asset return. A reasonable approach is to model the dependence structure between asset returns after accounting for the individual volatilities via copula models as show in Ausin and Lopes [2010], Tsay [2013] and Nguyen and Virbickaite [2022].

To account for time-varying correlation, we model the dependence between the marginals via dynamic stochastic Gaussian copula proposed by Hafner and Manner [2012] represented by Equations (9) to (11). We could also use other copula models such as Gumbel and Student's t and recover implicit correlations as shown, for example, in Patton [2013]. However, the dynamic Gaussian copula is simpler and suffices the needs of a GMVP allocation.

Equation (9) indicates that the bivariate time series $(u_{1,t}, u_{2,t})$ follows a time-varying copula with parameter Ξ_t which is driven by a latent stationary Gaussian AR(1), Equation (10), process of a transformation $\Xi_t = \psi(\lambda_t)$ of the copula-specific parameter shown in Equation (11).

$$(u_{1,t}, u_{2,t}) \sim C(u_1, u_2|\Xi_t) \tag{9}$$

$$\lambda_t = \alpha_c + \beta_c \lambda_{t-1} + \nu \xi_t \text{ with } \xi_t \sim N(0,1) \tag{10}$$

$$\Xi_t = \psi(\lambda_t) = \frac{exp(2\lambda_t) - 1}{exp(2\lambda_t) + 1} \tag{11}$$

Due to the non-linear non-Gaussian state space formed by Equations (9) to (11), we can't use Kalman filter based approaches to recover the latent state $\{\lambda\}$ we rely on Hafner and Manner [2012] Effective Importance Sampling approach to recover the path of the copula parameter. As before, we estimate using information from January 2, 2018 up to June 30, 2022 and perform forecast for the remaining observation up to December 30 2022. Table(4.2) shows that an investor would achieve a Sharpe ratio of at least 3.5 times higher by using our proposed model than when using standard SV or GARCH(1,1). Of course, rebalacing every 5-min is unlikely to be profit maximizing when considering transaction costs. However, this approach is useful to illustrate the performance implied by the different volatility models.

	Proposal	SV	GARCH(1,1)
Mean	4.82×10^{-4}	0.49×10^{-4}	0.62×10^{-4}
Sd	8.37×10^{-2}	4.51×10^{-2}	3.90×10^{-2}
5-min Sharpe	0.57×10^{-2}	0.11×10^{-2}	0.16×10^{-2}
"Annualized" Sharpe	1.55	0.30	0.42

Table 5: Summary statistics of minimum variance portfolios when considering our proposal, SV and GARCH(1,1) models for Australian Dollar and Swiss Franc returns for the out-of-sample period. In all cases, the dependence between the marginals is modeled via dynamic stochastic Gaussian copula, proposed by Hafner and Manner [2012], yielding time-varying correlations with average close to the sample correlation. "Annualized" Sharpe ratio is the Sharpe Ratio $\times\sqrt{252 \times 288}$ the factor $\sqrt{252 \times 288}$ aims to reflect the 288 5-minutes windows within the 24h day and 252 represent an average number of trading days. The portfolio returns based on our proposed model yields a Sharpe Ratio more than 3.5 times higher than any of its competitors.

5 Conclusion

This paper develops a stochastic volatility model of 5 minute FX returns accounting for hundred of macroeconomic events and seasonal components capturing time of the day effects. Of the possible hundreds of events, only announcements related to Taylor rule variables and external imbalance measures have more than 95% inclusion probability for their first lag. We reconcile why news about those macroeconomic fundamental may affect exchange rates via macro models proposed by Campbell and Clarida [1987] and Gourinchas and Rey [2007]. The estimated seasonality effect shows a W-shaped pattern peaking on China - London - New York openings while also reflecting volatility due to the opening of other markets such as Tokyo and Frankfurt. In addition, our seasonal volatility component is informative about average traded dollar volume. We also show that the same increases in the seasonal volatility during major markets opening hours also appear on the average traded volume. On a forecasting application, our model not only enhances interpretability but also increases out of sample R2 and decreases out of sample MAE. Finally, by combining the proposed method with the dynamic copula by Hafner and Manner [2012], we obtain a global minimum variance portfolio with a Sharpe ratio at least 3.5 times higher than SV and GARCH(1,1) models.

References

C. A. Abanto-Valle, H. S. Migon, and H. F. Lopes. Bayesian modeling of financial returns: A relationship between volatility and trading volume. *Applied Stochastic Models in Business and Industry*, 26(2):172–193, 2010.

T. G. Andersen and L. Benzoni. Realized volatility, working paper 2008-14. 2008.

T. G. Andersen and T. Bollerslev. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International economic review*, pages 885–905, 1998.

T. G. Andersen, T. Bollerslev, F. X. Diebold, and C. Vega. Real-time price discovery in global stock, bond and foreign exchange markets. *Journal of international Economics*, 73(2):251–277, 2007.

M. C. Ausin and H. F. Lopes. Time-varying joint distribution through copulas. *Computational Statistics & Data Analysis*, 54(11):2383–2399, 2010.

L. Bauwens, W. B. Omrane, and P. Giot. News announcements, market activity and volatility in the euro/dollar foreign exchange market. *Journal of International Money and Finance*, 24(7):1108–1125, 2005.

V. Bhansali. Volatility and the carry trade. *The Journal of Fixed Income*, 17(3):72–84, 2007.

O. Blanchard, G. Adler, and I. de Carvalho Filho. Can foreign exchange intervention stem exchange rate pressures from global capital flow shocks? Technical report, National Bureau of Economic Research, 2015.

J. Y. Campbell. *Financial decisions and markets: a course in asset pricing*. Princeton University Press, 2017.

J. Y. Campbell and R. H. Clarida. The dollar and real interest rates. In *Carnegie-Rochester Conference Series on Public Policy*, volume 27, pages 103–139. Elsevier, 1987.

J. C. Chan and A. Grant. Measuring the output gap using stochastic model specification search. 2017.

Y.-L. Chen and Y.-F. Gau. News announcements and price discovery in foreign exchange spot and futures markets. *Journal of Banking & Finance*, 34(7):1628–1636, 2010.

L. Ederington and J. H. Lee. Intraday volatility in interest-rate and foreign-exchange markets: Arch, announcement, and seasonality effects. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, 21(6):517–552, 2001.

M. Fratzscher, O. Gloede, L. Menkhoff, L. Sarno, and T. Stöhr. When is foreign exchange intervention effective? evidence from 33 countries. *American Economic Journal: Macroeconomics*, 11(1):132–156, 2019.

X. Gabaix and M. Maggiori. International liquidity and exchange rate dynamics. *The Quarterly Journal of Economics*, 130(3):1369–1420, 2015.

E. I. George and R. E. McCulloch. Variable selection via gibbs sampling. *Journal of the American Statistical Association*, 88(423):881–889, 1993.

J. Geweke. Variable selection and model comparison in regression. *In Bayesian Statistics 5*, 1996.

P.-O. Gourinchas and H. Rey. International financial adjustment. *Journal of political economy*, 115(4):665–703, 2007.

C. M. Hafner and H. Manner. Dynamic stochastic copula models: Estimation, inference and applications. *Journal of Applied Econometrics*, 27(2):269–295, 2012.

C. R. Harvey and R. D. Huang. Volatility in the foreign currency futures market. *The Review of Financial Studies*, 4(3):543–569, 1991.

N. Hautsch. *Econometrics of financial high-frequency data*. Springer Science & Business Media, 2011.

H. Ishwaran and J. S. Rao. Spike and slab variable selection: frequentist and bayesian strategies. *The Annals of Statistics*, 33(2):730–773, 2005.

T. Ito and Y. Hashimoto. Intraday seasonality in activities of the foreign exchange markets: Evidence from the electronic broking system. *Journal of the Japanese and International Economies*, 20(4):637–664, 2006.

S. Kim, N. Shephard, and S. Chib. Stochastic volatility: likelihood inference and comparison with arch models. *The review of economic studies*, 65(3):361–393, 1998.

H. Lustig, N. Roussanov, and A. Verdelhan. Common risk factors in currency markets. *The Review of Financial Studies*, 24(11):3731–3777, 2011.

A. Marshall, T. Musayev, H. Pinto, and L. Tang. Impact of news announcements on the foreign exchange implied volatility. *Journal of International Financial Markets, Institutions and Money*, 22(4):719–737, 2012.

R. A. Meese and K. Rogoff. Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of international economics*, 14(1-2):3–24, 1983.

T. J. Mitchell and J. J. Beauchamp. Bayesian variable selection in linear regression. *Journal of the american statistical association*, 83(404):1023–1032, 1988.

H. Nguyen and A. Virbickaite. Modeling stock-oil co-dependence with dynamic stochastic midas copula models. Technical report, Working Paper, 2022.

A. Patton. Copula methods for forecasting multivariate time series. *Handbook of economic forecasting*, 2:899–960, 2013.

J. Pencavel. The productivity of working hours. *The Economic Journal*, 125(589):2052–2076, 2015.

B. Rossi. Exchange rate predictability. *Journal of economic literature*, 51(4):1063–1119, 2013.

R. H. Shumway, D. S. Stoffer, and D. S. Stoffer. *Time series analysis and its applications*, volume 3. Springer, 2000.

C. Sørensen. Modeling seasonality in agricultural commodity futures. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, 22(5):393–426, 2002.

J. R. Stroud and M. S. Johannes. Bayesian modeling and forecasting of 24-hour high-frequency volatility. *Journal of the American Statistical Association*, 109(508):1368–1384, 2014.

J. B. Taylor. Discretion versus policy rules in practice. In *Carnegie-Rochester conference series on public policy*, volume 39, pages 195–214. Elsevier, 1993.

R. S. Tsay. *Multivariate time series analysis: with R and financial applications*. John Wiley & Sons, 2013.

Appendix A: Macroeconomic Events

US_S&P Global US Manufacturing PMI	US_Total Net TIC Flows
US_MBA Mortgage Applications	US_Housing Starts
US_Construction Spending MoM	US_Philadelphia Fed Business Outlook
US_ISM Manufacturing	US_Langer Economic Expectations
US_FOMC Meeting Minutes	US_U. of Mich. Sentiment
US_Challenger Job Cuts YoY	US_Chicago Fed Nat Activity Index
US_ADP Employment Change	US_Richmond Fed Manufact. Index
US_Initial Jobless Claims	US_FHFA House Price Index MoM
US_S&P Global US Services PMI	US_Existing Home Sales
US_Langer Consumer Comfort	US_New Home Sales
US_Change in Nonfarm Payrolls	US_Leading Index
US_Trade Balance	US_Kansas City Fed Manf. Activity
US_ISM Services Index	US_Advance Goods Trade Balance
US_Factory Orders	US_GDP Annualized QoQ
US_Durable Goods Orders	US_Personal Income
US_Consumer Credit	US_Dallas Fed Manf. Activity
US_NFIB Small Business Optimism	US_S&P CoreLogic CS 20-City NSA Index
US_JOLTS Job Openings	US_S&P CoreLogic CS 20-City MoM SA
US_Import Price Index MoM	US_Conf. Board Consumer Confidence
US_Wholesale Inventories MoM	US_Employment Cost Index
US_Wholesale Trade Sales MoM	US_MNI Chicago PMI
US_PPI Final Demand MoM	US_Pending Home Sales MoM
US_Monthly Budget Statement	US_FOMC Rate Decision (Upper Bound)
US_CPI MoM	US_Nonfarm Productivity
US_Retail Sales Advance MoM	US_Mortgage Delinquencies
US_Business Inventories	US_House Price Purchase Index QoQ
US_Empire Manufacturing	US_Household Change in Net Worth
US_Industrial Production MoM	US_Current Account Balance
US_NAHB Housing Market Index	US_Interest on Reserve Balances Rate
US_U.S. Federal Reserve Releases Beige Book	

Table 6: Table with all American events considered in the analysis of both Australian Dollar and Swiss Franc. The prefix US was added to indicate that the event corresponds to the United States.

AU_Judo Bank Australia PMI Mfg	AU_CPI QoQ
AU_AiG Perf of Mfg Index	AU_CPI YoY
AU_CoreLogic House Px MoM	AU_Private Sector Credit MoM
AU_Commodity Index AUD	AU_Import Price Index QoQ
AU_Judo Bank Australia PMI Services	AU_PPI QoQ
AU_AiG Perf of Services Index	AU_Retail Sales Ex Inflation QoQ
AU_Trade Balance	AU_RBA Cash Rate Target
AU_AiG Perf of Construction Index	AU_Wage Price Index QoQ
AU_Foreign Reserves	AU_Construction Work Done
AU_ANZ Roy Morgan Weekly Consumer Confidence Index	AU_Bloomberg Feb. Australia Economic Survey
AU_Building Approvals MoM	AU_Private Capital Expenditure
AU_Building Approvals YoY	AU_Inventories SA QoQ
AU_ANZ Job Advertisements MoM	AU_BoP Current Account Balance
AU_Job Vacancies QoQ	AU_GDP SA QoQ
AU_Retail Sales MoM	AU_House Price Index QoQ
AU_Credit Card Purchases	AU_Bloomberg March Australia Economic Survey
AU_Melbourne Institute Inflation MoM	AU_Bloomberg April Australia Economic Survey
AU_Westpac Consumer Conf Index	AU_RBA Statement on Monetary Policy
AU_Home Loans MoM	AU_Bloomberg May Australia Economic Survey
AU_Investment Lending	AU_Bloomberg June Australia Economic Survey
AU_Owner-Occupier Loan Value MoM	AU_Bloomberg July Australia Economic Survey
AU_Consumer Inflation Expectation	AU_Bloomberg Aug. Australia Economic Survey
AU_Employment Change	AU_Bloomberg Oct. Australia Economic Survey
AU_Full Time Employment Change	AU_Bloomberg Nov. Australia Economic Survey
AU_RBA FX Transactions Market	AU_Bloomberg Dec. Australia Economic Survey
AU_Westpac Leading Index MoM	AU_Home Loans Value MoM
AU_Skilled Vacancies MoM	AU_Private Sector Houses MoM
AU_Bloomberg Jan. Australia Economic Survey (Table)	AU_Exports MoM
AU_NAB Business Conditions	AU_RBA 3-Yr Yield Target
AU_NAB Business Confidence	AU_CBA Household Spending YoY

Table 7: Table with all Australian events considered in the analysis of the Australian Dollar. The prefix AU was added to indicate that the event corresponds to Australia.

CH_PMI Manufacturing	CH_Credit Suisse Survey Expectations
CH_Total Sight Deposits CHF	CH_SECO Consumer Confidence
CH_CPI MoM	CH_UBS Real Estate Bubble Index
CH_CPI EU Harmonized MoM	CH_Industry & Construction Output WDA YoY
CH_Unemployment Rate	CH_GDP QoQ
CH_Foreign Currency Reserves	CH_SNB Sight Deposit Interest Rate
CH_Retail Sales Real YoY	CH_SNB 3-Month Libor Lower Target Range
CH_Producer & Import Prices MoM	CH_KOF Institute Summer Economic Forecast
CH_Money Supply M3 YoY	CH_CPI Core YoY
CH_Exports Real MoM	CH_SNB Policy Rate
CH_KOF Leading Indicator	CH_Foreign exchange transactions

Table 8: Table with all Swiss events considered in the analysis of the Swiss Franc. The prefix CH was added to indicate that the event corresponds to Switzerland.

Appendix B: Model, Priors and Estimation

The proposed model and the priors considered for both the AUDUSD and CHFUSD analysis are presented bellow.

$y_t = v_t \varepsilon_t$

$v_t = \sigma X_t S_t E_t$

$h_t = log(v_t^2) = log(\sigma^2) + log(X_t^2) + log(S_t^2) + log(E_t^2) = \mu_h + x_t + s_t + e_t$

$x_t = \phi x_{t-1} + \sigma_x \eta_{x,t}$

$s_t = H_t' \beta$ with $H_t = (H_{t1}, \ldots, H_{t288})'$

$e_t = \sum_{i=1}^{n} I_{itk} \alpha_{ik}$

$\beta \sim N(0, \tau_\beta^2)$ with $1'\beta = 0$

$\alpha_{ik} | \pi \sim (1 - \pi)\delta_0 + \pi p_{slab}(\alpha_{ik}|\theta)$

π: $\pi \sim Bern(\theta)$

$\phi | \cdot \sim N(0.97, 1)$ and $\sigma^2 | \cdot \sim IG(1, 0.1)$

$\beta | \cdot \sim N(0, 10^2)$

$\theta \sim Beta(2, 2)$

τ^2 and $\sigma_a^2 \sim IG(1/2, 1/4)$

We estimate the model via Bayes rule using MCMC methods to sample from the posterior of the model.

We must sample from $\{\mu_h, \phi, \sigma_x^2, \{\beta_k\}_{k=1}^{288}, \{\alpha_{ik}\}_{i=1}^{n}, \pi, \theta, \tau^2, \sigma_\alpha^2, \{x_t\}_{t=1}^{T}\} | y_t$. We break down the MCMC methods into the following steps:

1) $\{\mu_h, \phi, \sigma_x^2\} | \cdot$: Estimation of an AR1. Due to the Gaussian likelihood and the normal and inverse gamma priors, the posterior is conjugated.

2) $\beta|.$: Linear regression case. Due to the Gaussian likelihood and the normal prior for the coefficients, the posterior is also normal.

3) $(\alpha, \pi|.)$: The update for α is straightforward. 0 if $\pi = 0$ and sample from a conjugate normal if $\pi = 1$ due to the normal likelihood and normal distribution of the slab. Updating π is a bit trickier. The key point of Geweke [1996] strategy is to integrate over possible values of α_i to avoid problems of the sampler getting stuck on zero due to the infinite mass of the Dirac's delta when α_i is zero.

Denote all parameters with the exception of α_i and π_i as Ξ

$$\begin{aligned}
p(\pi_i = 1|\Xi, D) &= \frac{p(\alpha_i = 0, \pi_i = 1|\Xi, D)}{p(\alpha_i = 0|\pi_i = 1, \Xi, D)} \\
&= \frac{p(\Xi, D|\alpha_i = 0, \pi_i = 1)p(\alpha_i = 0, \pi_i = 1)}{p(\Xi, D)p(\alpha_i = 0|\pi_i = 1, \Xi, D)} \\
&= \frac{p(\Xi, D|\alpha_i = 0)p(\alpha_i = 0, \pi_i = 1)}{p(\Xi, D)p(\alpha_i = 0|\pi_i = 1, \Xi, D)} \\
&\propto \frac{p(\alpha_i = 0, \pi_i = 1)}{p(\alpha_i = 0|\pi_i = 1, \Xi, D)} \\
&= \frac{p(\alpha_i = 0|\pi_i = 1)p(\pi_i = 1)}{p(\alpha_i = 0|\pi_i = 1, \Xi, D)} \\
&= \frac{\theta\phi(0; 0, \tau^2)}{\phi(0; m, v)}
\end{aligned}$$

where m and v are the mean and variance of the full conditional posterior distribution for α_i and $\phi(0; a, b)$: Gaussian density at zero with mean a and variance b. Similarly, for $\pi_i = 0$:

$$p(\pi_i = 0|\Xi, D) \propto \frac{p(\alpha_i = 0|\pi_i = 0)p(\pi_i = 0)}{p(\alpha_i = 0|\pi_i = 0, \Xi, D)} = 1 - \theta$$

Therefore, π_i can be sampled from the following Bernoulli:

$$Bern\left(\frac{\frac{\theta\phi(0;0,\tau^2)}{\phi(0;m,v)}}{\frac{\theta\phi(0;0,\tau^2)}{\phi(0;m,v)} + (1-\theta)}\right)$$

4) $\theta|.$: Conjugated case of a binomial model with beta prior leading to a beta posterior.

5) $\tau^2|.$ and $\sigma^2_\alpha|.$: Conjugated case of a Gaussian model with IG prior leading to a IG posterior.

6) $\{x_t\}_{t=1}^T|.$: We have a non-linear Gaussian state space problem. We opt to linearize the problem and face a non-Gaussian problem. As inKim et al. [1998], we approximate the non-Gaussian problem with a mixture of 7 Gaussian distributions. Using data-augmentation, we can recover the latent states using the Forward Filtering Backward Sampling algorithm as shown in Kim et al. [1998].

Appendix C: Posterior inclusion probability

The table bellow show the posterior inclusion probability for the first lag of each event for the AUDUSD analysis.

Event	Inclusion Probability
US_S&P Global US Manufacturing PMI	0.04
US_MBA Mortgage Applications	0.01
US_Construction Spending MoM	0.18
US_ISM Manufacturing	0.85
US_FOMC Meeting Minutes	0.09
US_Challenger Job Cuts YoY	0.04
US_ADP Employment Change	0.79
US_Initial Jobless Claims	0.75
US_S&P Global US Services PMI	0.79
US_Langer Consumer Comfort	0.02
US_Change in Nonfarm Payrolls	1.00
US_Trade Balance	0.04
US_ISM Services Index	0.21
US_Factory Orders	0.04
US_Durable Goods Orders	0.04
US_Consumer Credit	0.03
US_NFIB Small Business Optimism	0.02
US_JOLTS Job Openings	0.54
US_Import Price Index MoM	0.09
US_Wholesale Inventories MoM	0.04
US_Wholesale Trade Sales MoM	0.04
US_PPI Final Demand MoM	0.86
US_Monthly Budget Statement	0.03
US_CPI MoM	1.00
US_Retail Sales Advance MoM	0.51
US_Business Inventories	0.03
US_Empire Manufacturing	0.07
US_Industrial Production MoM	0.20
US_NAHB Housing Market Index	0.02
US_U.S. Federal Reserve Releases Beige Book	0.03

US_Total Net TIC Flows	0.02
US_Housing Starts	0.03
US_Philadelphia Fed Business Outlook	0.05
US_Langer Economic Expectations	0.02
US_U. of Mich. Sentiment	0.08
US_Chicago Fed Nat Activity Index	0.06
US_Richmond Fed Manufact. Index	0.03
US_FHFA House Price Index MoM	0.04
US_Existing Home Sales	0.66
US_New Home Sales	0.43
US_Leading Index	0.03
US_Kansas City Fed Manf. Activity	0.02
US_Advance Goods Trade Balance	0.03
US_GDP Annualized QoQ	0.72
US_Personal Income	0.07
US_Dallas Fed Manf. Activity	0.04
US_S&P CoreLogic CS 20-City NSA Index	0.05
US_S&P CoreLogic CS 20-City MoM SA	0.03
US_Conf. Board Consumer Confidence	0.03
US_Employment Cost Index	0.29
US_MNI Chicago PMI	0.02
US_Pending Home Sales MoM	0.29
US_FOMC Rate Decision	1.00
US_Nonfarm Productivity	0.03
US_Mortgage Delinquencies	0.04
US_House Price Purchase Index QoQ	0.05
US_Household Change in Net Worth	0.05
US_Current Account Balance	0.04
US_Interest on Reserve Balances Rate	0.12
AU_Judo Bank Australia PMI Mfg	0.03
AU_AiG Perf of Mfg Index	0.03
AU_CoreLogic House Px MoM	0.16
AU_Commodity Index AUD	0.12
AU_Judo Bank Australia PMI Services	0.12

AU_AiG Perf of Services Index	0.03
AU_Trade Balance	0.97
AU_AiG Perf of Construction Index	0.15
AU_Foreign Reserves	0.04
AU_ANZ Roy Morgan Weekly Consumer Confidence Index	0.16
AU_Building Approvals MoM	0.03
AU_Building Approvals YoY	0.06
AU_ANZ Job Advertisements MoM	0.04
AU_Job Vacancies QoQ	0.04
AU_Retail Sales MoM	0.74
AU_Credit Card Purchases	0.08
AU_Melbourne Institute Inflation MoM	0.03
AU_Westpac Consumer Conf Index	0.53
AU_Home Loans MoM	0.14
AU_Investment Lending	0.11
AU_Owner-Occupier Loan Value MoM	0.78
AU_Consumer Inflation Expectation	0.17
AU_Employment Change	0.53
AU_Full Time Employment Change	0.53
AU_RBA FX Transactions Market	0.05
AU_Westpac Leading Index MoM	0.03
AU_Skilled Vacancies MoM	0.04
AU_Bloomberg Jan. Australia Economic Survey	0.14
AU_NAB Business Conditions	0.80
AU_NAB Business Confidence	0.17
AU_CPI QoQ	0.75
AU_CPI YoY	0.40
AU_Private Sector Credit MoM	0.13
AU_Import Price Index QoQ	0.18
AU_PPI QoQ	0.03
AU_Retail Sales Ex Inflation QoQ	0.07
AU_RBA Cash Rate Target	1.00
AU_Wage Price Index QoQ	0.86
AU_Construction Work Done	0.49

AU_Bloomberg Feb. Australia Economic Survey	0.06
AU_Private Capital Expenditure	0.06
AU_Inventories SA QoQ	0.07
AU_BoP Current Account Balance	0.13
AU_GDP SA QoQ	1.00
AU_House Price Index QoQ	0.05
AU_Bloomberg March Australia Economic Survey	0.12
AU_Bloomberg April Australia Economic Survey	0.26
AU_RBA Statement on Monetary Policy	0.20
AU_Bloomberg May Australia Economic Survey	0.08
AU_Bloomberg June Australia Economic Survey	0.08
AU_Bloomberg July Australia Economic Survey	0.09
AU_Bloomberg Aug. Australia Economic Survey	0.28
AU_Bloomberg Oct. Australia Economic Survey	0.09
AU_Bloomberg Nov. Australia Economic Survey	0.53
AU_Bloomberg Dec. Australia Economic Survey	0.08
AU_Home Loans Value MoM	0.10
AU_Private Sector Houses MoM	0.03
AU_Exports MoM	0.85
AU_RBA 3-Yr Yield Target	0.04
AU_CBA Household Spending YoY	0.08

Table 9: Inclusion probability of the first lag for all events considered in the AUDUSD analysis. Events with prefix US and AU correspond to American and Australian events, respectively