South African Reserve Bank Working Paper Series WP/22/14

Can information on the distribution of ZAR returns be used to improve SARB's ZAR forecasts?

Matthew Greenwood-Nimmo, Daan Steenkamp and Rossouw van Jaarsveld

Authorised for distribution by Konstantin Makrelov

27 September 2022

South African Reserve Bank

South African Reserve Bank Working Papers are written by staff members of the South African Reserve Bank and, on occasion, by consultants under the auspices of the South African Reserve Bank. The papers deal with topical issues and describe preliminary research findings, and develop new analytical or empirical approaches in their analyses. They are solely intended to elicit comments and stimulate debate.

South African Reserve Bank Working Papers are externally refereed.

Information on South African Reserve Bank Working Papers can be found at
https://www.resbank.co.za/en/home/publications/Papers/working-papers

Enquiries relating to the Working Paper Series can be addressed to:

Head: Economic Research Department
South African Reserve Bank
P O Box 427
Pretoria 0001

Tel. +27 12 313 3911

Can information on the distribution of ZAR returns be used to improve SARB's ZAR forecasts?

Matthew Greenwood-Nimmo*, Daan Steenkamp† and Rossouw van Jaarsveld‡

27 September 2022

Abstract

We characterise the distribution of the rand/US dollar exchange rate returns in two ways: (i) using realised moments estimated from intra-day data; and (ii) using implied market expectations of the rand exchange rate extracted from options prices. We begin by showing that these distributions can be used to contextualise rand movements and summarise the extent of rand uncertainty and tail risk. We then conduct a simple exercise to evaluate whether information on the variance of the rand returns distributions can be used to improve the South African Reserve Bank's rand forecasts. We find that, while these rand forecasts show little bias, the magnitude of the forecast errors can be reduced and the directional accuracy improved substantially by incorporating historical and implied estimates of rand variance in the forecasting model. This effect is particularly pronounced at longer forecast horizons.

JEL classification: F31, G01, G15

Keywords: Foreign exchange markets, realised and options-implied return distributions, variance risk premium, forecast evaluation

*Faculty of Business and Economics, University of Melbourne and Centre for Applied Macroeconomic Analysis, Australian National University. Email: matthew.greenwood@unimelb.edu.au.

†Corresponding author. Codera Analytics and Research Fellow, Department of Economics, Stellenbosch University. Email: daan@codera.co.za.

‡Codera Analytics and Research Fellow, Department of Economics, Stellenbosch University. Email: rossouw@codera.co.za.

1 Introduction[1]

Exchange rate forecasting is widely regarded as a difficult task, because exchange rates typically exhibit the characteristics of a random walk (i.e. they often follow no clear trend and appear to move at random). Nevertheless, central banks typically publish exchange rate forecasts and they represent an important input into monetary policy decision-making. For an inflation-targeting central bank in a small open economy, like the South African Reserve Bank (SARB), variations in the exchange rate may impact on inflation and inflation expectations via their effect on the prices of imported goods and services. Exchange rate fluctuations may also impact on the terms of trade, the distribution of economic activity among the tradeable and non-tradeable sectors and the expected rate of economic growth. Regular evaluation and continuous development of central bank exchange rate forecasting procedures is therefore an important area of activity.

The SARB produces quarterly forecasts of the rand/US dollar exchange rate by creating an implied rand forecast from its model of the nominal effective exchange rate of the rand. In this paper, we investigate whether information on the shape of the rand return distribution can be used to refine the SARB's rand forecast. We characterise the distribution of rand returns in two ways. First, we estimate the realised distribution of rand returns by combining information on intra-day rand bid and ask prices. In this way, we obtain a backward-looking historical record of the behaviour of rand returns. Second, we compute forward-looking estimates of the rand return distribution from options data following the computational routines described by Malz (1997, 2014). Options are financial derivatives that confer the right to the buyers of such instruments to buy or sell the underlying currency at a specific price within a specific period. The options-implied distributions summarise market expectations of rand exchange rate movements that are embedded in options prices, which reflect traders' beliefs about the range of possible outcomes and their expected probabilities. Many central banks use options prices to characterise the distribution of market expectations of future currency movements, as well as to gauge expected exchange rate uncertainty (examples include Eitrheim 1999, Castren 2005, Gereben and Pinter 2005, Lewis 2012, Blake and Rule 2015 or Korkmaz et al. 2019).[2]

As a precursor to our forecast analysis, we demonstrate the utility of the realised and options-implied distributions and their moments. We show that they can be used to contextualise rand movements, especially around key events like the 2017 downgrade of South African sovereign debt, and to summarise the extent of rand uncertainty and tail risk. The options-implied distribution can also be used to estimate the range of values within which market participants believe that the rand will trade over a given horizon.

[1] We are grateful to Pieter Pienaar for assistance with data, Sechaba Mokobane for help with earlier versions of the options-implied volatility estimation, Michelle Lewis for the provision of code and Allan Malz and Nick Prisk for helpful suggestions.

[2] For further discussion of the usefulness of information from options for forecasting see Christoffersen et al. (2013).

Next, we test whether information on the variance of the realised and options-implied distributions can be used to refine the SARB's rand forecasts.[3] There is considerable evidence that the difference between implied variance and expected realised variance evaluated over the same horizon may help to predict foreign exchange (FX) returns in certain settings. This quantity is known as the variance risk premium (VRP). It measures the cost of hedging a unit of variance risk, and tends to rise in volatile states of the world when the willingness of investors to pay for volatility protection rises. The VRP has been shown to have predictive power for spot exchange rates and is a popular measure of the risk appetite of FX investors (see Egbers and Swinkels 2015 or Della Corte et al. 2016). It may be particularly relevant in countries such as South Africa, where investors are concerned about sovereign credit risk.

A number of existing studies have demonstrated the usefulness of the VRP for FX forecasting. For example, Rombouts et al. (2020) show that the VRP can predict market returns, Londono and Zhou (2017) show that currency and stock VRPs can predict currency movements, and Della Corte et al. (2016) show that FX volatility insurance and exchange rate predictability are related.[4] However, there has been little research into emerging market VRPs, and we are not aware of any studies of the VRP for the rand. Despite receiving a lot of attention in South African policy debates, there has also been very little research into the relative volatility of the rand and the factors contributing to it. Notable exceptions are Mavee et al. (2016), who use the Johannesburg Stock Exchange measure of options-implied rand volatility as well as generalised autoregressive conditional heteroscedascity models of daily rand returns in simple ordinary least squares (OLS) models to show that commodity price volatility and global market volatility are the main drivers of rand volatility and that South African political uncertainty can explain rand volatility, and May et al. (2018), who suggest that monetary policy surprises affect returns and volatility in the rand/US dollar exchange rate. In a related vein, Greenwood-Nimmo et al. (2021) document realised volatility spillovers between emerging and developed FX markets, and identify the rand as a 'bellwether' emerging market currency.

We show that a simple strategy to augment the SARB's forecast with the VRP leads to significant gains in forecast performance, most notably at the 4-quarters ahead horizon, where it reduces the proportion of forecast errors attributable to bias by 57% and increases the proportion of unsystematic forecast errors by 128%. Interestingly, we find that a subtly different specification in which the constituent parts of the VRP (the implied variance and a forecast of realised variance) enter the model separately performs better still, which suggests that it is not only the VRP but also the level of rand variance that helps predict exchange rate movements. The biggest gain that we obtain is in terms of directional accuracy of the forecasts, or the accuracy of the model to correctly forecast appreciation or depreciation of the currency. Our preferred specification correctly predicts the directional change approximately 90% and

[3] We limit our attention to the second moment owing to the small sample available for out-of-sample forecast evaluation.

[4] Other papers that demonstrate the usefulness of FX moments for FX predictability include Szakmary et al. (2003) and Ornelas and Mauad (2019). Jurek and Xu (2013) identify FX risk premia by controlling for cross-sectional FX moments.

88% of the time over a 1- to 2-quarters horizon and 3- to 4-quarters horizon respectively. By comparison, the SARB forecast matches this performance at the 1- to 2-quarters horizon but performs worse over the 3- to 4-quarters horizon, predicting the directional change correctly only about 60% of the time. We show that there is therefore considerable scope to enrich the SARB's forecasting model with information on the realised and options-implied distributions of rand returns.

2 Options-implied and realised moment estimation

We consider two different approaches to characterising the distribution of currency returns: options-implied and realised distributions. Thereafter, we document the historical behaviour of the distributions of rand returns and assess whether the SARB can use distributional information to improve its rand forecasts.

2.1 The options-implied distribution of exchange rate returns

FX derivatives incorporate market expectations of future movements in exchange rates. Options prices are particularly useful, because they can be used not just to estimate the expected level of the currency but also to characterise the extent of expected uncertainty and tail risk. We use options data to estimate the moments of the implied distribution of the exchange rate that is embedded in the actions of market participants.

FX options are not quoted in currency units. Instead, the convention is for them to be quoted in implied volatility terms and often as a combination of options (i.e. strategies like 'straddles', defined later).

The models we estimate use market prices for three options strategies to characterise the market's expected distribution of each currency's values in future. A straddle strategy provides the expected variance (i.e. the market's view of future currency volatility), a risk reversal strategy provides a measure of expected skewness (i.e. the implied risk that the quote currency could move in a particular direction), and a strangle strategy provides the expected kurtosis (i.e. the implied probability of large shifts in the quote currency). The strikes (i.e. exercise prices) of currency options are quoted as 'deltas', which measure how much option prices change for a small change in the currency. Delta approximates the 'moneyness' of an option: an at-the-money option will have a delta of approximately 50, implying that there is a 50:50 chance that it will expire in the money. To extract information from options prices, a range of exercise prices for options with sufficient trading liquidity is required. For this reason, most approaches typically exclude illiquid options, such as 'deep-out-of-the-money' options (with deltas approaching 100).

Implied volatilities can be estimated using the formula of Black and Scholes (1973) as the parameter value that makes options prices (derived from their model) consistent with market prices. However, the Black-Scholes approach assumes that all options on an underlying asset have the same implied volatility. In practice, currency options markets are often characterised by a 'volatility smile': the implied volatilities of options with the same time to expiration tends to vary across different strikes. The existence of a volatility smile implies that the options-implied probability distribution of future exchange rate values may not be normal and requires that the higher moments be estimated since the distributions often have long tails (indicative of skewness) and/or fat tails (indicative of excess kurtosis).

We use the techniques proposed by Malz (1997) and Malz (2014) to estimate the risk-neutral probability density from the options data. Whereas the approach of Malz (1997) uses a parametric estimation methodology to interpolate over observed volatilities (using at-the-money and three options contracts at delta of 25), the Malz (2014) framework is based on a non-parametric estimation approach (using at-the-money and deltas of 10, 25 and 35) and has the advantage of not requiring strong assumptions about the shape of the volatility smile to be estimated. Both techniques have been widely adopted among other central banks to estimate risk-neutral probability densities.

The approach of Malz (1997) is summarised concisely in Gereben (2002), Lewis (2012) and Blake and Rule (2015). Given the definition of delta as the first derivative of call prices with respect to the spot rate:

$$\delta = \frac{\partial C(S_t)}{\partial S_t}, \tag{1}$$

Malz (1997) proposes a simplification of the call function as:

$$C_t = N\left[-\frac{log(Q_t) - \tau(\frac{\sigma^2}{2})}{\sigma\sqrt{\tau}}\right] - Q_t N\left[-\frac{log(Q_t) - \tau(\frac{\sigma^2}{2})}{\sigma\sqrt{\tau}}\right], \tag{2}$$

where N denotes the cumulative probability distribution function of the standard normal distribution, τ is an option's time to maturity (i.e. $\tau = T - t$, where T is the settlement date and t the current date) and σ denotes the volatility of the currency pair. Meanwhile, Q_t is the implied return calculated as follows:

$$Q_t = \frac{X_t}{F_{t,T}}, \tag{3}$$

where X denotes the strike price and the forward rate, F, is the spot rate, S, adjusted for the cost of carry to the specific future period:

$$F_{t,T} = S_t e^{(r-r^*)\tau}, \tag{4}$$

where r is the domestic risk free rate and r^* the foreign risk free rate. Malz (1997) suggests the following simplification for δ:

$$\delta = e^{-r^*\tau}\left[-\frac{log(Q_t) - \tau(\frac{\sigma^2}{2})}{\sigma\sqrt{\tau}}\right]. \tag{5}$$

As a continuum of options prices across strikes is needed, Malz (1997) proposes that one interpolates between available delta quotes to obtain a volatility smile, $\sigma(\delta)$, as follows:

$$\sigma(\delta) = b_0 atm_t + b_1 rr_t(\delta - 0.5) + b_2 str_t(\delta - 0.5)^2, \tag{6}$$

where σ denotes the implied volatility, atm denotes at-the-money volatility, rr denotes the quoted risk reversal and str the strangle volatility.[5] The location of the smile is described by *atm*, *rr* summarises its skewness and *str* its kurtosis. To calculate the implied volatility and δ at every strike price, equations (5) and (6) are solved simultaneously. Thereafter, prices at every potential strike are generated by substituting the implied volatilities into the call function. Finally, the options-implied probability density is obtained by twice differentiating the call price function (see Blake and Rule, 2015, for further details).

The approach proposed by Malz (2014) instead interpolates between the observed values of the volatility smile using a clamped cubic spline,[6] using in- and out-of-the-money deltas. Following Blake and Rule (2015), the delta for at-the-money volatility can be calculated by equating the strike price and that day's spot price in the equation for delta. The next step is to use the clamped cubic spline to obtain the volatility smile and therefore volatilities at each delta value. To obtain the necessary volatility-strike pairs needed to construct the call price function, numerical iteration is used to find a value of implied volatility for which the corresponding delta at that specific strike price is equal to the delta from the cubic spline, that is an implied volatility that satisfies: $\delta_{s+1} = \delta_s$ where:

$$\hat{\delta} = e^{-r^*\tau} N\left[-\frac{log(\frac{S}{X}) - (r - r^* - \frac{\sigma^2}{2})\tau}{\sigma_s\sqrt{\tau}} \right], \tag{7}$$

and:

$$\delta_{s+1} = a_k + b_k(\hat{\delta} - \delta_k) + c_k(\hat{\delta} - \delta_k)^2 + d_k(\hat{\delta} - \delta_k)^3, \tag{8}$$

with spline coefficients a_k, b_k, c_k and d_k that minimise the squared errors between actual values of implied volatilities, δ_k, and the constructed knot points, $\hat{\delta}$. Finally, option prices are calculated for every strike price. As before, the risk-neutral density is estimated numerically using the second derivative of the call price with respect to each strike price.

We express our estimates in foreign exchange rate return form, as the difference between the logarithm of each value of the spot rate from the risk-neutral density and the logarithm of current spot rates. We estimate options-implied moments using both Malz (1997) and Malz (2014), but use the Malz (2014)-based estimates in our forecasting exercise after demonstrating that the estimates obtained are similar. All of the data used to estimate our options-implied moments is sourced from Bloomberg. Apart from daily quotes for the level of the rand, this

[5] Malz (1997) uses $b_0 = 1$, $b_1 = 2$ and $b_2 = 16$, noting that this matches the shape of the volatility smile for most currency pairs.

[6] A clamped spline has the advantage of accurately approximating the shape of the curve, provided it is assumed that the curve flattens before the first and after the last knot points. See Malz (2014) for a discussion.

includes the quoted rand forward rate, the at-the-money implied volatility, and risk reversal and butterfly spreads for each maturity we consider (i.e. 1-, 2-, 3-, 6-, 9- and 12-month horizons).[7] These quotes are used to calculate volatilities for each delta (e.g. $d = 10, 25, 35$ when applying the approach of Malz (2014)) as:

$$Volatility_{d,h,t} = \frac{Strangle_{d,h,t} + RiskReversal_{d,h,t}}{2} \quad (9)$$

where market strangle volatility is calculated as $str_{d,h,t} = atm_{h,t} + bf_{d,h,t}$. and d is delta, h is the horizon (i.e. 1-, 2-, 3-, 6-, 9- and 12-month) and t is the value for a specific day.

2.2 The realised distribution of exchange rate returns

Unlike the options-implied distribution of exchange rate returns, which is inherently forward-looking owing to the nature of options contracts, the realised distribution is computed from historical returns. To construct the realised distribution, we use data on intra-day mid-rates, which we construct as the mid-point between intra-day bid and offer rates sourced from Refinitiv. We adopt the convention of Andersen et al. (2001) and consider a trading day to start at 21:05GMT and to end at 21:00GMT (roughly between when markets open in the East and close in the United States). Furthermore, we exclude weekends and fixed and floating US public holidays, which are associated with thin rand trading. We also eliminate several days where we observe problems with the Refinitiv datafeed, reflected in strings of zero returns.

The simplest estimator of the realised volatility involves summing the squared intra-day returns, as follows:

$$RV_t = \sum_{i=1}^{n} r_{t,i}^2,$$

where $r_{t,i}$ is the ith period-to-period intra-day log-return on day t and n is the number of intra-day observations. The convention is to use five-minute intervals ($n = 288$), which balances asymptotic considerations against microstructure noise.

Interpretation of the realised variance is complicated by the fact that it is expressed in variance units (squared per cent where the underlying returns are expressed in per cent). Consequently, the realised volatility is often provided as an alternative measure of the scale of the realised distribution that can be expressed in common units. For example, the realised volatility expressed on an annualised basis in per cent is given by:

$$RVol_t = 100\sqrt{N(RV_t)},$$

where N is the number of trading days per year.

[7] Butterfly describes the 'wings' (i.e. curvature or kurtosis) of the volatility smile.

The realised skewness of the exchange rate is defined as follows:

$$RSkew_t = \sqrt{n}\sum_{i=1}^{n} r_{t,i}^3 \Big/ (RV_t)^{3/2},$$

where, using the rand as an example, $RSkew_t > 0$ indicates that the right tail of the distribution of rand returns is fatter than the left tail and vice versa for $RSkew_t < 0$. $RSkew_t > 0$ ($RSkew_t < 0$) implies that there is a tendency toward depreciation (appreciation) of the rand against the US dollar.

Finally, realised kurtosis measures the mass in the tails of the distribution, with larger values of $RKurt_t$ implying a higher probability of extreme rand returns. It is estimated as follows:

$$RKurt_t = n\sum_{i=1}^{n} r_{t,i}^4 \Big/ (RV_t)^2 .$$

3 Estimated rand return distributions

3.1 Options-implied distribution

Options-implied probability densities can be used to measure both the direction of future currency movements and the balance of risk that the market is pricing in. The evolution of the rand during April 2020 is a good example. On 20 April 2020, options prices suggested that market participants expected the rand to continue to trade around its prevailing level with a higher risk of depreciation than of appreciation (Figure 1). The light blue area shows the range of values that the rand was expected to be within, with 80% probability, over the next 12 months. Apart from having a probability mass that is skewed towards higher values of the rand (i.e. rand depreciation), the density widens over the horizon, consistent with increased uncertainty over the range of possible future values of the currency at longer horizons.[8]

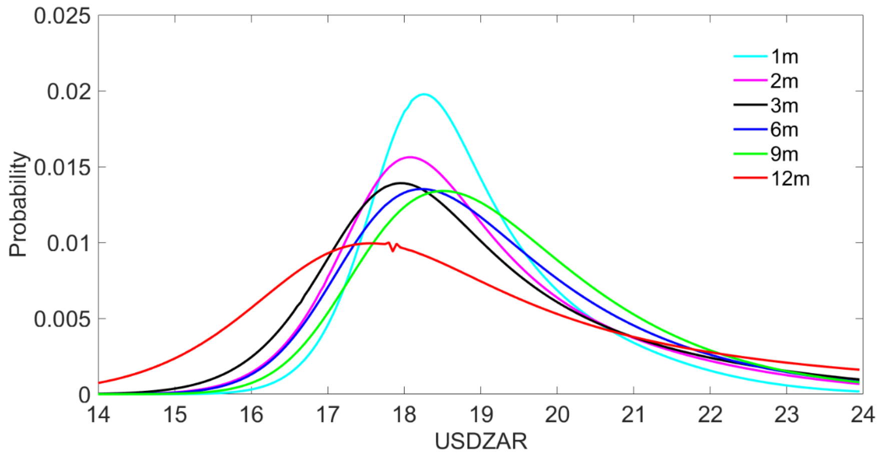
Figure 2 plots the options-implied probability densities for the rand across option maturities as of 20 April 2020. It shows that the distribution was relatively peaked at shorter maturities, but much wider at longer maturities, owing to the greater uncertainty associated with longer investment horizons. Interestingly, market quotes on 20 April implied a sharp shift of the mode of the distribution toward appreciation between maturities of 9 and 12 months.

[8] The accuracy of options-implied distributions depends crucially on the liquidity of the options market. Low liquidity, particularly for options that are far out-of-the-money, can create volatility in the tails of the implied distributions over time. The options market for the rand is relatively liquid on average, implying that it should capture market expectations reasonably accurately. According to the Bank for International Settlements (2019) survey on currency market liquidity, South Africa's options market is the 13th largest globally, with daily turnover of USD268 million.

Figure 1: Options-implied fan chart for rand over the 12 months from 20 April 2020

Source: Authors' calculations, estimated using Malz (1997). Additional results based on Malz (2014) available on request.

Figure 2: Options-implied rand probability density functions (as at 20 April 2020)



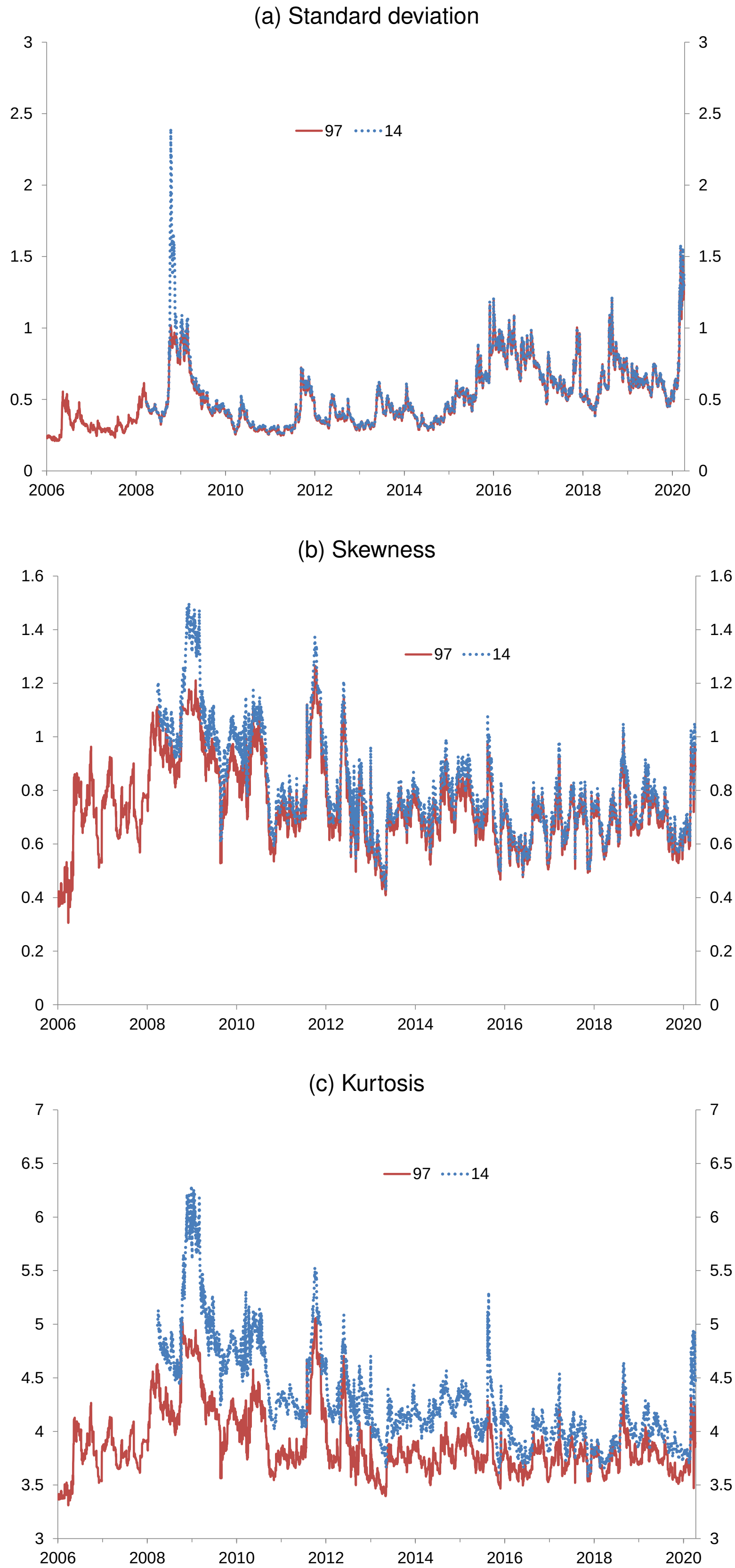
Source: Authors' calculations, estimated using Malz (1997). Additional results based on Malz (2014) available on request.

Having estimated a sequence of daily options-implied probability densities, their moments can be calculated to provide a time series representation summarising the time variation in market expectations of the rand. Figure 3 plots the estimated moments of the distributions over time based on Malz (1997) and Malz (2014) at the 1-month ahead horizon by way of example. While the standard deviations and skewness of the distributions based on Malz (1997) are very similar to those based on Malz (2014), the estimates of kurtosis are generally slightly higher based on the former approach, which suggests that the difference between the two procedures is primarily confined to the tails of the resulting densities.

The implied standard deviation of the rand has historically spiked during periods of heightened financial and political instability, such as following the global financial crisis, the firing of Finance Minister Nene in December 2015, the sovereign downgrades in April 2017 and the recent COVID-19 crisis (panel (a)). During some of these periods (especially the COVID-19 crisis), the market also priced in a higher risk of rand depreciation than the average over the sample (panel (b)). Kurtosis of greater than 3 implies that the rand distribution has fat tails relative to a normal distribution, implying a higher probability of large exchange rate movements than is normal. The tails of the distribution did not, however, fatten significantly during most of these volatile periods (panel (c)).

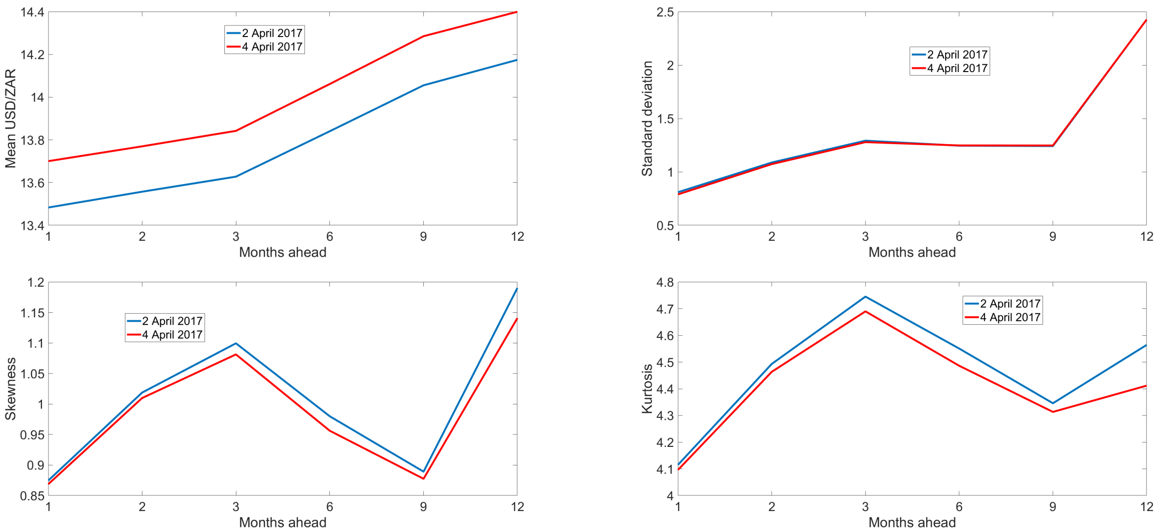
The options-implied distribution can also be used to assess how events affect currency expectations. For example, Figure 4 shows that the sovereign downgrade by S&P Global Ratings on 3 April 2017 was associated with shifts in the mean of the options-implied rand distribution, indicative of a higher implied level of the rand at all maturities (i.e. a sustained rand depreciation). The standard deviation of the distribution was unaffected, indicating that the downgrade was perceived to be neutral with respect to rand volatility. By contrast, skewness increased slightly at longer maturities, suggesting that market participants expected that values of the rand above its mean were subsequently more likely than values below it. The kurtosis also increased at maturities of three months and beyond, implying market expectations of more extreme changes in the rand following the downgrade.

Figure 3: Moments of 1-month-ahead options-implied densities over time



Source: Authors' calculations. '97' and '14' refer to the approaches of Malz (1997, 2014).

Figure 4: Term structure of implied moments (before and after downgrade)



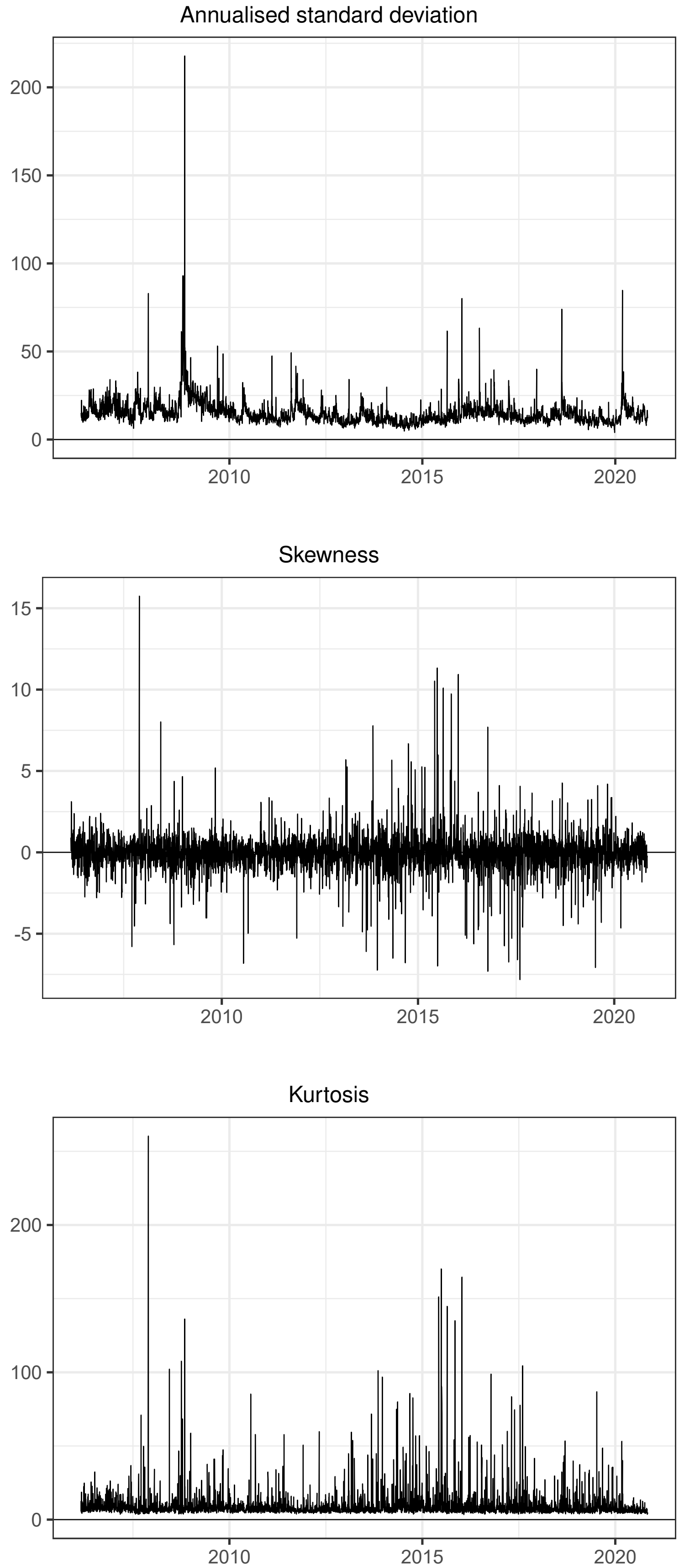
Source: Authors' calculations

3.2 Realised moments for the rand

Figure 5 plots daily estimates of the moments of the realised probability density for the rand.[9] Consistent with the widespread evidence of volatility clustering documented in the academic finance literature, the rand exhibits extended periods of high or low realised standard deviation, with strong persistence. As was the case for the options-implied distribution, the realised distribution has a marked right-skew and fat tails.

[9] Note that we plot daily realised moments in Figure 5, as it is the daily realised variance that is used as an input in our forecasting exercise below. It is straightforward to construct estimates at other frequencies (e.g. monthly) if desired.

Figure 5: Moments of daily realised probability density



Source: Authors' calculations

4 Can rand variance be used to improve SARB's rand forecasts?

The SARB uses a model of the nominal effective exchange rate to create an implied rand forecast for each forecasting quarter.[10] These forecasts do not exploit information on the shape of either the realised or options-implied distributions. Yet the moments of these distributions convey a great deal of information on realised and expected currency movements. The remainder of this paper tests whether this distributional information can enhance the SARB's forecasting framework. Because the SARB's forecasts are only available quarterly, there is a relatively small sample size available for forecast evaluation. Therefore, we limit our attention to the use of second-order moments to keep our models as small as possible. We leave the development of forecasting models that exploit the higher-order moments of the realised and/or options-implied distributions of rand returns for future work.

4.1 Implied variance, realised variance and the VRP

The options-implied and realised densities described above are related on a conceptual level. To see this, note that the options-implied variance for the rand represents a market-based forecast of the evolution of the realised variance over a horizon corresponding to the options maturity (and likewise for other moments). In general, the options-implied variance will exceed the realised variance, because investors typically overestimate the probability of adverse events. The difference between the implied variance at a given horizon and the expected realised variance at the same horizon is a widely studied quantity known as the VRP.

The VRP measures the cost to hedge variance risk and also describes the returns from speculative trades involving the writing of insurance options against FX variance. Such a VRP trading strategy could provide attractive returns over long periods of time for currencies subject to downside risk which has not materialised, while a spike in the VRP creates a potential opportunity for trades that are short variance (i.e. that bear the risk that the VRP narrows in future). This phenomenon may be able to explain a significant portion of the excess returns associated with FX carry strategies with long positions in the currencies of high interest rate economies (see Egbers and Swinkels, 2015; Della Corte et al., 2016).[11]

Given the international evidence of the forecasting power of the VRP, we begin by testing whether the VRP can be used to improve the SARB's forecasts. In practice, expected realised variance is unknown and must be estimated. Early empirical research into the VRP

[10] The nominal effective rate is modeled as an identity that incorporates the real effective exchange rate, international producer price index and domestic producer price index. It is estimated by OLS. The real effective exchange rate is modelled as a function of interest rate differentials, a risk premium, the ratio of public sector borrowing requirement to GDP and the VIX.

[11] Della Corte et al. (2016), for example, demonstrate that FX VRP helps to forecast cross-sectional FX returns. Their findings suggest that using VRP in an FX strategy involving buying currencies with 'low insurance costs' (i.e. low implied variance relative to realised variance) may provide excess return opportunities, as such currencies tend to appreciate. They show that the predictive power of VRP is not related to risk factors, implying that VRP developments do not represent compensation for systemic risk and that their usefulness for forecasting is not based on interest rate differentials but spot returns.

(e.g. Bollerslev et al., 2009) typically invoked the simplifying assumption that the realised variance would remain unchanged over the forecast horizon. This assumption has been criticised by Bekaert and Hoerova (2014), who note that a realised variance forecasting model should be employed instead. To this end, we use a modified version of the heterogeneous autoregressive-realised variance (HAR-RV) model as set out in Corsi (2009). The model uses additive volatility combinations to reduce a realised variance model to a simple autoregressive type model that incorporates information over different interval sizes.

In light of the variance forecasting model comparison conducted by Bekaert and Hoerova (2014), we augment the HAR-RV model with information on the options-implied variance. The model has an expanding specification for each h-quarter ahead forecast, where the lag structure of the model is reflective of the forecast horizon, h. The augmented HAR-RV model for each forecast horizon can be written as follows:

(i) 1 quarter ahead

$$RV_t = \beta_0 + \beta_1 RV_{t_5} + \beta_2 RV_{t_{22}} + \beta_3 RV_{t_{66}} + \beta_4 IV_t^{3-month} + \epsilon_t \tag{10}$$

(ii) 2 quarters ahead

$$\begin{aligned} RV_t = &\beta_0 + \beta_1 RV_{t_5} + \beta_2 RV_{t_{22}} + \beta_3 RV_{t_{66}} + \beta_4 IV_t^{3-month} + \\ &\beta_5 RV_{t_{132}} + \beta_6 IV_t^{6-month} + \epsilon_t \end{aligned} \tag{11}$$

(iii) 3 quarters ahead

$$\begin{aligned} RV_t = &\beta_0 + \beta_1 RV_{t_5} + \beta_2 RV_{t_{22}} + \beta_3 RV_{t_{66}} + \beta_4 IV_t^{3-month} + \\ &\beta_5 RV_{t_{132}} + \beta_6 IV_t^{6-month} + \\ &\beta_7 RV_{t_{196}} + \beta_8 IV_t^{9-month} + \epsilon_t \end{aligned} \tag{12}$$

(iv) 4 quarters ahead

$$\begin{aligned} RV_t = &\beta_0 + \beta_1 RV_{t_5} + \beta_2 RV_{t_{22}} + \beta_3 RV_{t_{66}} + \beta_4 IV_t^{3-month} + \\ &\beta_5 RV_{t_{132}} + \beta_6 IV_t^{6-month} + \\ &\beta_7 RV_{t_{196}} + \beta_8 IV_t^{9-month} + \\ &\beta_9 RV_{t_{264}} + \beta_{10} IV_t^{12-month} + \epsilon_t \end{aligned} \tag{13}$$

where $RV_{t_m} = \frac{1}{m}\sum_{i=1}^{m} RV_{t-i}$ is the average realised variance over the period $t-1$ to $t-m$, IV_t^h is the implied variance at horizon h estimated at time t and ϵ_t is the error term.

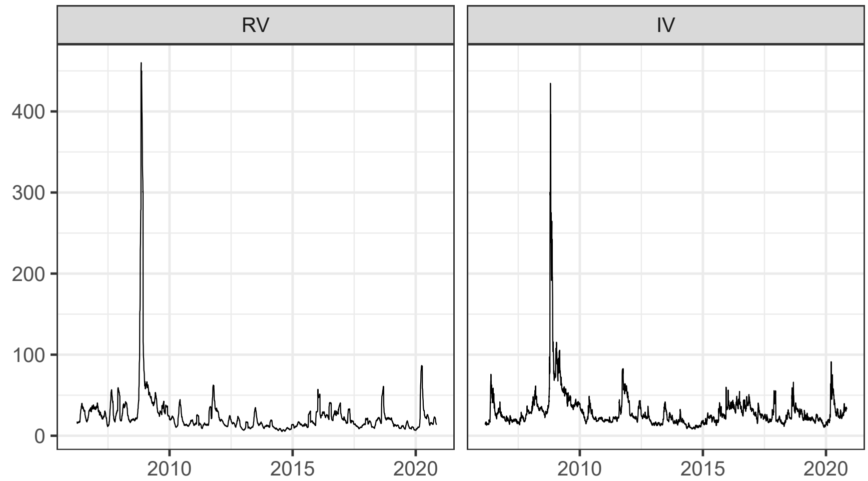
The VRP at horizon h, VRP_t^h, is defined as the difference between the implied variance at horizon h, IV_t^h, and the h-step-ahead forecast of realised variance:

$$VRP_t^h = IV_t^h - E_{t+h,t}(RV_t), \tag{14}$$

where $E_{t+h,t}()$ denotes the forecast for period $t+h$ computed using information available at time t.

The convention in the equity VRP literature is to focus on a 1-month horizon, in no small part because the VIX, as the most well-known example of an options-implied volatility statistic, is defined over that horizon. Adopting the same convention, Figure 6 plots both the RV and the implied variance (IV) for the rand at the 1-month horizon. As is the case for major currencies during normal times, implied variance for the rand tends to be higher than realised variance (Figure 6). This implies that the VRP is positive on average (Figure 7).

Figure 6: 1-month RV and IV for the rand

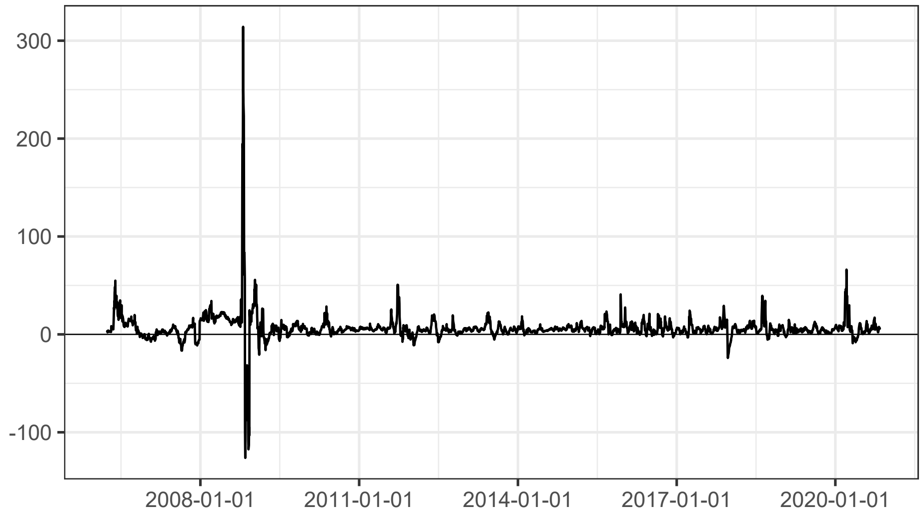


Source: Authors' calculations

In general, one would expect the VRP to fluctuate over time, taking high values during volatile periods of high risk aversion and lower values or falling in periods of higher risk appetite. In practice, during volatile periods, RV may rise more rapidly than IV, leading to brief periods where the VRP is negative. A good example of this can be seen in 2008. At these times, unhedged VRP harvesting strategies (i.e. those involving the writing of options that provide buyers with insurance against downside FX risk) will suffer arbitrarily large losses. However, over our sample period, the VRP for the rand is relatively stable, taking an average value of approximately 25 variance units. The largest spike in the rand VRP occurred during the global financial crisis, while the second largest spike in our sample occurred at the onset of the COVID-19 pandemic. Figure 8 shows that the average level of the rand VRP over our sample period is relatively high by international standards.[12]

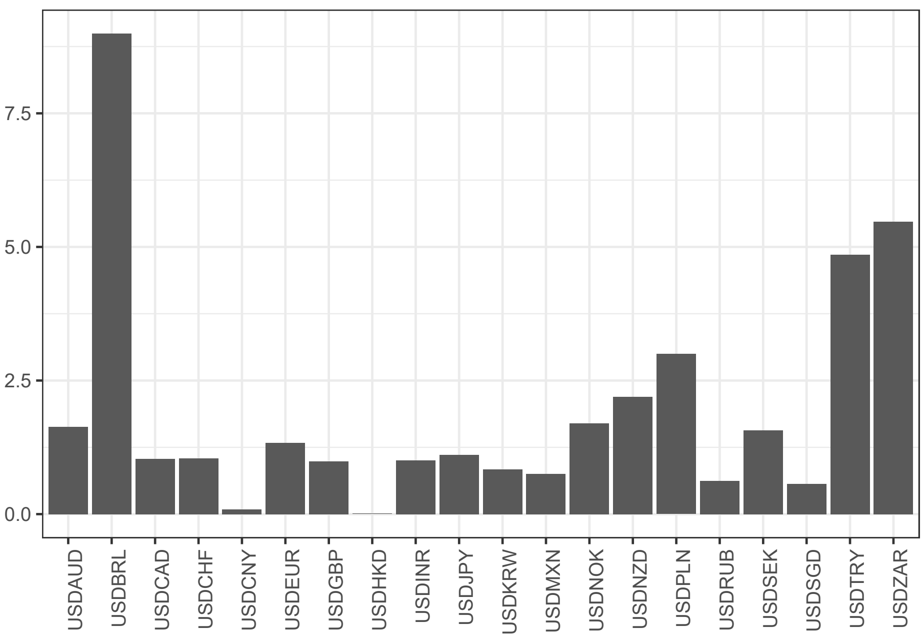
[12] We compute the VRP for each currency using the same procedure described above. The only exception is for the US dollar/Brazilian real, where we find it necessary to augment the HAR-RV model with two shifting intercepts during the periods of January 2008 to November 2009 and January 2015 to December 2016. Both of these periods are characterised by dramatic swings in the Brazilian exchange rate. As a result of the sharp

Figure 7: Rand VRP



Source: Authors' calculations

Figure 8: VRP comparison across major currencies



Note: Median estimate over the sample 27 February 2006 to 30 July 2021.

Source: Authors' calculations

An interesting question is the extent to which the high relative level of the rand VRP reflects differences in liquidity relative to other FX markets. Olds et al. (2021) use the same currency

currency movements, the HAR-RV model with a single intercept term produces uncharacteristically poor out-of-sample estimates. This behaviour is improved significantly by the introduction of shifting intercepts. The 2008-2009 period can be attributed to the aftermath of the global financial crisis. The 2015-2016 period was characterised by large economic and political shocks. Iron ore and soy beans, which account for 25% of Brazil's export basket, saw a dramatic decrease in prices. Furthermore, a combination of political crisis, prolonged economic recession and an aggressive interest rate cutting cycle contributed to the rapid depreciation of the currency.

sample as used in Figure 8 and find that FX liquidity (as measured using FX bid-ask spreads) has a meaningful impact on the volatility of most currency pairs, including the rand. Olds et al. (2021) show that periods of higher liquidity are associated with lower FX RV for the individual currency pairs and that common variation in these currencies tends to be associated with common variation in FX market liquidity. In the case of the rand, its realised volatility is strongly affected by global FX liquidity. Interestingly, in the case of the US dollar/Brazilian real, Olds et al. (2021) find that, although global FX liquidity has a statistically significant impact on liquidity for the exchange rate, the relationship between liquidity and volatility is relatively weak in this case.

In addition to testing whether the VRP can be used to enhance the SARB's rand forecasts, we also consider the use of its constituent parts separately. Including the VRP linearly in a forecasting model is equivalent to including the IV and the expected RV separately subject to the restriction that the parameters on the two regressors share the same magnitude but opposite signs. Consequently, a model where the IV and the expected RV are included separately (rather than being combined via the VRP) relaxes this restriction and opens the possibility that the level of realised and/or implied variance, and not just the difference between them, may have predictive power for the exchange rate.

4.2 Forecast evaluation

Our goal is not to propose a new SARB forecasting model but rather to assess whether information on the variance of the rand could be fruitfully incorporated into the SARB's forecasting model. We compare three alternatives: (i) the existing SARB forecasts; (ii) the SARB forecasts augmented with the VRP via an auxiliary regression, denoted SARB+VRP; and (iii) the SARB forecasts augmented with IV and RV separately via an auxiliary regression, denoted SARB+RV+IV.

We construct the augmented forecasts as follows:

$$y_t = \alpha + \beta_1 y^{SARB}_{t,t-h} + \beta_2 \mathbf{x_{t,t-h}} + \epsilon_t, \qquad (15)$$

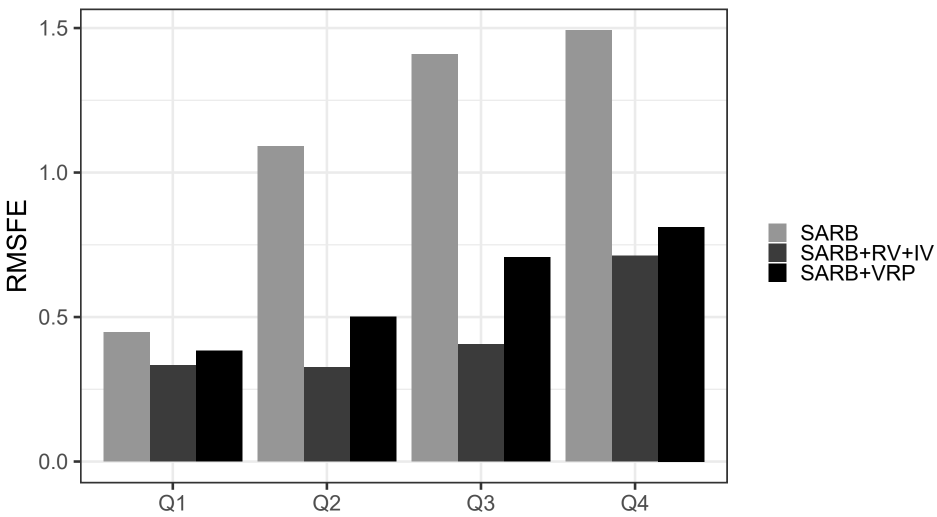
where y_t is the actual value of the rand at time t, $y^{SARB}_{t,t-h}$ is the SARB's h-step ahead forecast of y_t that is available at time $t-h$ and $\mathbf{x_{t,t-h}}$ is a vector of information for time t that is available at time $t-h$ that contains either: (i) the VRP; or (ii) the IV and RV (defined in equations 10 to 13). We exploit the fact that the VRP and its components are updated at a higher frequency than the SARB forecasts (daily as opposed to quarterly). Specifically, we assume that the SARB forecasts are produced on the Friday before each Monetary Policy Committee (MPC) meeting and use VRP, IV and RV estimates as of this date.[13] Nevertheless, our approach relies only on information available at period t and therefore provides a true forecast. We estimate the

[13] Although the SARB forecasts may be constructed prior to the last Friday before the MPC meeting in practice, that Friday represents the last opportunity for them to be updated to reflect revisions to the SARB's view of the likely path of the rand prior to the meeting. Consequently, the SARB forecast may be treated as representative of the SARB's view as of the Friday prior to each meeting.

parameters of equation (15) by OLS using data from 2006Q2 to 2017Q4 and perform out-of-sample forecast evaluation over the period 2018Q1 to 2021Q2 (see Table 3 in the Appendix).[14]

Figure 9 compares the root mean squared forecast error (RMSFE) of each forecast at horizons of 1-, 2-, 3- and 4-quarters ahead. Forecast errors tend to be larger at longer horizons owing to the increase in implicit uncertainty over time. At all horizons, the SARB forecast has the largest RMSFE, indicating that it has the largest average forecast error of the three models. Both the SARB+VRP and SARB+RV+IV forecasts yield considerably lower RMSFEs, with the SARB+RV+IV model marginally preferred at all horizons. This suggests that the implicit parameter restriction introduced through the use of the VRP in place of the IV and expected RV is detrimental for forecast performance.

Figure 9: Root mean squared forecast error for each forecasting model



Source: Authors' calculations

Table 1 presents the results of a simple decomposition proposed by Theil (1971), whereby the mean squared forecast error is decomposed into three proportions. The bias proportion captures the proportion of forecast error that can be attributed to systematic bias in the level of the forecast, i.e. systematic under- or over-prediction. The regression proportion measures whether the forecasts are systematically more or less variable than the actual values of the exchange rate. Lastly, the disturbance proportion represents the part of the forecast error that is random. The first two quantities reflect systematic sources of error that can be addressed through refinements to the forecasting model, while the last is unsystematic. Consequently, for an optimal forecast, the theoretical values of the three proportions are (0,0,1). See Ahlburg (1984) for further detail.

[14] It is important to note that the SARB's forecasting approach underwent significant changes in 2017 with the introduction of the Quarterly Projection Model (QPM) (Botha et al. 2017). The SARB exchange rate forecasts are a consolidation of views based on the SARB's core model (Smal et al. 2007), QPM and a set of assumptions about the current state of the economy.

Table 1: Sources of forecast error

	SARB				SARB + VRP				SARB + RV + IV			
Horizon	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Bias	0.125	0.177	0.167	0.510	0.115	0.121	0.156	0.219	0.109	0.113	0.131	0.208
Regression	0.281	0.201	0.118	0.197	0.244	0.226	0.136	0.112	0.228	0.239	0.147	0.077
Disturbance	0.594	0.622	0.645	0.293	0.641	0.653	0.708	0.669	0.663	0.648	0.722	0.715

Source: Authors' calculations

First consider the composition of the SARB's forecast errors. Over horizons 1 to 3, the majority of the SARB's forecast errors are unsystematic, although the regression proportion is relatively high for the 1-quarter ahead forecast. However, bias becomes a major contributor to the SARB's forecast errors at the 4-quarters ahead horizon, accounting for 51% in the decomposition. At this horizon, only 29% of the SARB forecast error is unsystematic.

Both the SARB+VRP and SARB+RV+IV models improve on the SARB forecasts in several respects. First, both deliver a reduction in the bias proportion at all horizons. The gain at the 4-quarters ahead horizon is very large – the bias proportion of the SARB+VRP and SARB+RV+IV models are 57% and 59% lower than that of the SARB model, respectively. Both the SARB+VRP and SARB+RV+IV models also deliver a reduction in the regression proportion at the 1-quarter and 4-quarters ahead horizons, although both are beaten slightly by the SARB model at the 2- and 3-quarters ahead horizons. Overall, however, the disturbance proportions for the SARB+VRP and SARV+RV+IV models are higher than for the SARB model at all horizons, indicating that they are less prone to systematic errors than the SARB approach regardless of the horizon. This effect is most pronounced at the 4-quarters ahead horizon, where the disturbance proportions for the SARB+VRP and SARV+RV+IV models are 128% and 144% higher than that of the SARB model, respectively. The implication is that augmenting the SARB forecast with market data reduces the prevalence of systematic forecasting errors.

Lastly, in Table 2, we reported estimated hit rates for each model across the four forecast horizons. As our forecasts are specified in level terms, the hit rate test amounts to testing whether the forecasts gets the subsequent direction of change in the currency right. The SARB forecast has a very high hit rate at the 1-quarter ahead horizon, correctly predicting the direction of the exchange rate change more than 90% of the time. The hit rate drops over the 2- and 3-quarters ahead horizons before reaching 50% at the 1-year horizon, which indicates that the SARB forecasts perform no better at predicting the direction of the rand one year ahead than tossing a fairly weighted coin. The SARB+VRP model is beaten by the SARB forecast at horizon 1, matches it at horizon 2 and outperforms it substantially at horizons 3 and 4. Meanwhile, the SARB+RV+IV model matches or beats the directional accuracy of the SARB forecast at all horizons. Its performance is remarkably consistent across horizons, with the hit rate remaining in the range 0.875 to 0.909 throughout. Consequently, the gain relative to the SARB forecast is particularly striking at the 1-year horizon, where the SARB+RV+IV model correctly the predicts the direction of change almost 90% of the time compared to just

50% of the time for the SARB forecast. The difference in hit rates between the SARB+VRP and SARB+RV+IV models indicates that removing the restriction implicit in the definition of the VRP has a much larger effect on directional forecasting accuracy than on other aspects of forecast performance.

Table 2: Hit rates across forecast model specifications

	Q1	Q2	Q3	Q4
SARB	0.909	0.800	0.667	0.500
SARB + VRP	0.818	0.800	0.778	0.625
SARB + RV + IV	0.909	0.900	0.888	0.875

Source: Authors' calculations

5 Conclusion

In this paper, we begin by characterising the distribution of rand returns in both an historical and a forward-looking manner. We show that the realised and options-implied probability densities of the rand can be used to construct narratives contextualising currency movements and that the options-implied probability density can be used to identify the range within which market participants expect the rand to trade in the future.

Next, we undertake a forensic study of the SARB's historical rand forecast errors and show that information on the realised and implied variance of the rand can be used to improve the SARB's forecasts in several important dimensions. Not only can this information be used to reduce the magnitude of the SARB's forecast errors at all horizons, but it can also be used to mitigate systematic errors in the SARB's long-horizon forecasts. Arguably the biggest gain, however, lies in improving the directional accuracy of the SARB's forecasts. While the SARB's 4-quarter ahead forecast is equivalent to a fair coin toss in terms of directional accuracy, a model augmented with realised and implied variance measures yields directional accuracy of 87.5% at the same horizon. Given that exchange rate forecasts feed into a central bank's assessment of inflationary pressure arising from imported goods and services, optimising their directional accuracy is essential for an inflation-targeting central bank.

We close by highlighting several avenues for continuing research. First, while we do not propose a new forecasting model in this paper per se, our results motivate for the development of an updated SARB forecasting framework that uses information on market expectations. Second, although our analysis is limited to the use of second-order moments owing to the small sample of SARB forecasts available for evaluation, there may be gains from using higher-order moments and risk measures for forecasting purposes. Skewness measures may be particularly relevant, as they explicitly convey information on expected directional changes in the currency. Lastly, it would be useful to benchmark the SARB's augmented forecasts against historical market analyst forecasts.[15]

[15] At present, the only available source of historical market analyst forecasts of which we are aware, Consensus

References

Ahlburg, D A. 1984. 'Forecast evaluation and improvement using Theil's decomposition'. *Journal of Forecasting* 3: 345–351.

Andersen, T, Bollerslev, T, Diebold, F and Ebens, H. 2001. 'The distribution of realized stock return volatility'. *Journal of Financial Economics* 61: 43–76.

Bank for International Settlements. 2019. 'Triennial central bank report on global foreign exchange market activity in 2019'.

Bekaert, G and Hoerova, M. 2014. 'The VIX, the variance premium and stock market volatility'. *Journal of Econometrics* 183: 181–192.

Black, F and Scholes, M. 1973. 'The pricing of options and corporate liabilities'. *Journal of Political Economy* 81: 637–654.

Blake, A and Rule, G. 2015. 'Deriving option-implied probability densities for foreign exchange markets' . Centre for Central Banking Studies, Bank of England.

Bollerslev, T, Tauchen, G and Zhou, H. 2009. 'Expected stock returns and variance risk premia'. *Review of Financial Studies* 22: 4463–4492.

Botha, B, de Jager, S, Ruch, F and Steinbach, R. 2017. 'The quarterly projection model of the SARB', *Working Papers WP/17/01*, South African Reserve Bank.

Castren, O. 2005. 'Estimating and analysing currency options implied risk-neutral density functions for the largest new EU member states', *Working Paper Series 440*, European Central Bank.

Christoffersen, P, Jacobs, K and Chang, B Y. 2013. 'Forecasting with option-implied information', Elsevier, volume 2, chapter 10, 581–656.

Corsi, F. 2009. 'A simple approximate long-memory model of realized volatility'. *Journal of Financial Econometrics* 7: 174–196.

Della Corte, P, Ramadorai, T and Sarno, L. 2016. 'Volatility risk premia and exchange rate predictability'. *Journal of Financial Economics* 120: 21–40.

Egbers, T and Swinkels, L. 2015. 'Can implied volatility predict returns on the currency carry trade?' *Journal of Banking & Finance* 59: 14–26.

Eitrheim, O. 1999. 'Can the price of currency options provide an indication of market perception of the uncertainty attached to the krone exchange rate?' *Norges Bank Economic Bulletin* 99-03.

Economics' Foreign Exchange Consensus Forecasts, are only available at the 1-month horizon.

Gereben, A. 2002. 'Extracting market expectations from option prices: an application to over-the-counter New Zealand dollar options', *Reserve Bank of New Zealand Discussion Paper Series DP2002/04*, Reserve Bank of New Zealand.

Gereben, A and Pinter, K. 2005. 'Implied volatility of foreign exchange options: is it worth tracking?' *Central Bank of Hungary Working Paper* 05-39.

Greenwood-Nimmo, M, Steenkamp, D and van Jaarsveld, R. 2021. 'Risk and return spillovers in a global model of the foreign exchange network', *Working Papers WP/21/15*, South African Reserve Bank.

Jurek, J and Xu, Z. 2013. 'Option-implied currency risk premia'. *SSRN Electronic Journal*.

Korkmaz, H I, Kucuksarac, D, Onay, Y and Senol, A. 2019. 'Estimation of FX option implied density functions: nonparametric-Malz approach'. *Central Bank of Turkey Working Paper* 19-03.

Lewis, M. 2012. 'Market perceptions of exchange rate risk', *Reserve Bank of New Zealand Analytical Notes series AN2012/12*, Reserve Bank of New Zealand.

Londono, J M and Zhou, H. 2017. 'Variance risk premiums and the forward premium puzzle'. *Journal of Financial Economics* 124: 415–440.

Malz, A M. 1997. 'Option-implied probability distributions and currency excess returns', *Staff reports no. 32*, Federal Reserve Bank of New York.

Malz, A M. 2014. 'Simple and reliable way to compute option-based risk-neutral distributions', *Staff Reports 677*, Federal Reserve Bank of New York.

Mavee, N, Perrelli, R and Schimmelpfennig, A. 2016. 'Surprise, surprise; what drives the rand / U.S. dollar exchange rate volatility?', *IMF Working Papers 16/205*, International Monetary Fund.

May, C, Farrell, G and Rossouw, J. 2018. 'Do monetary policy announcements affect exchange rate returns and volatility of returns? Some evidence from high-frequency intra-day South African data'. *South African Journal of Economics* 86: 308–338.

Olds, T, Steenkamp, D and van Jaarsveld, R. 2021. 'The impact of global FX liquidity on the rand'. *Economic Research Southern Africa Policy Bulletin* 21-05.

Ornelas, J R H and Mauad, R B. 2019. 'Implied volatility term structure and exchange rate predictability'. *International Journal of Forecasting* 35: 1800–1813.

Rombouts, J V, Stentoft, L and Violante, F. 2020. 'Dynamics of variance risk premia: A new model for disentangling the price of risk'. *Journal of Econometrics* 217: 312–334.

Smal, D, Pretorius, C and Ehlers, N. 2007. 'The core forecasting model of the South African Reserve Bank', *Working Papers WP/07/02*, South African Reserve Bank.

Szakmary, A, Ors, E, Kyoung Kim, J and Davidson, W. 2003. 'The predictive power of implied volatility: evidence from 35 futures markets'. *Journal of Banking and Finance* 27: 2151–2175.

Theil, H. 1971. *Applied Economic Forecasting*, Amsterdam: North Holland.

Appendices

Table 3: Forecast combination with SARB forecast and RV and IV

	Dependent variable:			
	USDZAR			
	(1-Q ahead)	(2-Q ahead)	(3-Q ahead)	(4-Q ahead)
β_1	0.997*** (0.025)			
β_1		0.989*** (0.040)		
β_1			0.998*** (0.055)	
β_1				1.012*** (0.072)
β_{21}	5.699*** (0.544)	5.686*** (0.251)	7.054*** (0.792)	9.571*** (0.809)
β_{22}	−9.173*** (1.585)	−2.753*** (0.970)	−5.768*** (01.018)	−4.505 (0.855)
Constant	0.027 (0.257)	0.310 (0.408)	0.309 (0.553)	0.089 (0.721)
Observations	46	46	46	46
R^2	0.976	0.938	0.891	0.831
Adjusted R^2	0.974	0.933	0.883	0.819
Residual Std. Error (df = 41)	0.419	0.671	0.890	1.106
F Statistic (df = 3; 41)	552.315***	206.586***	111.516***	67.423***

Note: * implies statistical significance at 10%; ** at 5%; *** at 1% significance.*