"What's Vol got to do with it"-Replication report

Lucie Yiliu Lu[1]

1 Summary of the paper

The main research question of the current paper (Drechsler and Yaron, 2011) is how volatility risk is priced and why the volatility risk premium predicts equity returns.

The paper has an empirical section and a theoretical section. It first shows that the variance risk premium of the S&P 500 index, measured by the difference between the squared VIX index and expected realized variance, predicts future equity returns. Then the paper proposes an extension of the long-run risk model that is claimed to generate the empirical patterns.

A key feature of the theoretical model is that it introduces jump risk in the conditional mean and conditional volatility of consumption growth. This enables the model to generate the desired level of variance risk premium while matching the level and volatility of the market return and the risk free rate. In the model parameterization, the predictability is generated by the conditional volatility of consumption growth, which enters into both the jump intensity process and equity risk premia as a priced risk.

The focus of the current paper is similar to that of Bollerslev et al. (2009), in which the key model ingredient that generates the predictability of returns by the variance premium is a common volatility of volatility factor. The current paper bridges the derivatives literature and the macro-finance literature and provides economic underpinnings for the empirical findings.

This report presents the replication of results in the current paper and documents any additional assumptions and derivations that is not made explicit in the paper, and proposes explanations for inconsistencies with results in the original paper.

[1]yiliu.lu@mail.mcgill.ca

2 Empirical section

The main results in the empirical part includes (1) summary statistics of the data, (2) predictive regressions of volatility under the $\mathbb{P}$ measure, (3) summary statistics of the variance premium and (4) return predictive regressions using the variance premium.

The data source used here is the same as in the paper. Thanks Christian for sharing the TICKDATA S&P futures intraday prices. The only inconsistency in terms of data is that the paper uses realized variance measured using the cash index as the regressor to predict expected variance under $\mathbb{P}$. I did not ask for the cash index data and skipped/replaced with index futures results that require the cash index. But using realized variance measured from the futures index yields similar results.

	SP 500	VWRet	VIX^2	Fut^2	$Daily^2$
Mean	0.527	0.526	33.30	22.60	20.69
Median	0.957	1.023	25.14	15.56	13.51
Std.Dev	4.01	4.13	24.13	22.54	21.95
Skewness	-0.634	-0.837	2.00	2.58	2.68
kurtosis	4.22	4.52	8.89	10.81	11.91
AR(1)	-0.04	0.02	0.79	0.65	0.61

Table 1: Summary Statistics of Excess Returns and Variances

Table 1 provides summary statics of returns and realized variance measures. The replication is perfect except for the realized variance measured using intraday 5-minute returns of the S&P futures. The figures are very close though. This is related to how I define 5-minute bins and measure 5-minute returns within a trading day. I conform to the paper's method of treating overnight and over-weekend returns as a 5-minute return, and the current version takes the first price of a trading day as the beginning of the first 5-minute bin and takes the return between the first prices of two bins as a 5-minute return. Overnight returns are calculated as the return between the beginning price of the last bin of the previous trading day and the beginning price of the next trading day. Alternative methods, such as calculating the returns using end-of-the-bin prices and end-of-the-trading-day prices yields results that further away from the paper's result than the current choice.

Table 2 shows predictive regressions for expected variance under $\mathbb{P}$, I skipped the second regression in the original paper which involves realized variance measured

Table 2: Conditional Volatility Predictive Regressions

Dept.Var	X1	X2	intercept	β_1	β_2	R^2
$Daily^2_{t+1}$	$Daily^2_t$	MA(1)	3.84	0.81	-0.33	0.40
(t-stat)			(1.71)	(15.01)	(-4.17)	
Fut^2_t	Fut^2_t	VIX^2	-1.07	0.07	0.66	0.58
(t-stat)			(-1.03)	(0.86)	(10.13)	

from intraday returns of the cash index and replace the regressor in the third regression (realized variance measured using cash index) with realized variance measured from future returns. To estimate the first regression, which is a GARCH(1,1) model of volatility, I fit an ARMA(1,1) of realized returns using Matlab 'estimate' function. The paper does not specify which estimation method is used and the small difference in results might be due to difference in estimation method. Also, the paper does not explicitly specify the number of lags used in the Newey-West standard errors, I use 18 lags in my replication.

Overall, the results are very similar. The small differences are inevitable due to the fact that I replace cash index realized variance with futures realized variance. I think the reason that the paper chooses to use the cash index realized variance as the regressor is that the R^2 is slightly higher (0.59 vs 0.58), despite the fact that cash index is prone to the problem of stale prices.

Table 3: Properties of the Variance Premium

	VP(BTZ)	VP(daily-MA(1))	VP(Fut-forecast)
Mean	10.74	12.71	10.84
Median	8.26	8.13	8.69
Std.Dev	14.31	14.38	6.99
Minimum	-54.40	-4.16	3.70
Skewness	0.16	2.43	2.15
Kurtosis	10.82	12.44	10.25
AR(1)	0.07	0.54	0.74

Table 3 shows the summary statistics of the variance premium calculated using different methods. As in the previous table, I skipped the variance premium measured based on expected cash index realized variance and the variance premium based on expected futures variance uses futures realized variance as a predictor, and the BTZ expected variance under $\mathbb{P}$ is also measured using futures realized variance. The difference in results in the first and the third column are due to the use of futures realized variance in lieu of cash index realized variance. The variance risk premium calculated here using only futures realized variance is lower than that in the paper. This suggests that realized variance measured using the cash index is lower than that measured using futures.

Table 4: Return Predictability by the Variance Premium

	Regressors		OLS			Robust Reg.		
Dependent	X1	X2	β_1	β_2	$R^2(\%)$	β_1	β_2	$R^2(\%)$
r_{t+1}	VP_t		0.67		0.96	1.20		4.32
	(t-stat)		(1.66)			(2.72)		
r_{t+1}	VP_{t-1}		0.99		2.11	1.17		4.32
	(t-stat)		(2.38)			(2.62)		
r_{t+3}	VP_t		0.84		3.45	0.79		5.44
	(t-stat)		(1.37)			(2.80)		
r_{t+1}	VP_t	$log(P/E)_t$	1.68	-56.75	9.44	2.26	-59.51	14.51
	(t-stat)		(3.51)	(-3.62)		(4.80)	(-4.98)	
r_{t+1}	VP_t	$log(P/E)_t$	2.25	-65.77	12.90	2.32	-65.31	14.90
	(t-stat)		(4.96)	(-4.01)		(4.82)	(-5.30)	

Table 4 shows predictive regressions of returns in different horizons on the variance risk premium and the price to dividend ratio. The results are similar to those in the paper. The R^2s of OLS regressions are slightly lower than those in the paper and the R^2 in the robust regressions are slightly higher than those in the paper. The difference again, should result from the use of futures realized variance as the predictor instead of the cash index realized variance when calculating the expected variance under $\mathbb{P}$.

3 Calibration section

To provide a bit of context, the state vector is: $[\Delta c_t, x_t, \overline{\sigma}_t, \sigma_t, \Delta d_t]$. The dynamics of the system of state variables is given by:

$$\begin{aligned}
\Delta c_{t+1} &= \mu_c + x_t + z_{c,t+1} \\
x_{t+1} &= \mu_x + \rho_x x_t + z_{x,t+1} + J_{x,t+1} \\
\overline{\sigma}^2_{t+1} &= \mu_{\overline{\sigma}} + \rho_{\overline{\sigma}} \overline{\sigma}^2_t + z_{\overline{\sigma},t+1} \\
\sigma^2_{t+1} &= \mu_\sigma + (1-\rho_\sigma)\overline{\sigma}^2_t + \rho_\sigma \sigma^2_t + z_{\sigma,t+1} + J_{\sigma,t+1} \\
\Delta d_{t+1} &= \mu_d + \phi x_t + z_{d,t+1}
\end{aligned} \tag{1}$$

Table 5: Cash flow dynamics

	5%	50%	95%
$E[\Delta c]$	1.04	1.95	2.76
$\sigma(\Delta c)$	2.11	2.66	3.39
$AC1(\Delta c)$	-0.04	0.21	0.45
$E[\Delta d]$	-1.04	2.07	5.00
$\sigma(\Delta d)$	11.64	13.52	15.65
$AC1(\Delta d)$	-0.16	0.04	0.24
$\rho_{\Delta c,\Delta d}$	0.07	0.29	0.48
$kurt(\Delta c)(Q)$	3.00	4.17	8.38

Table 5 (Table 6 in the paper) gives the summary statistics of cash flow dynamics generated by the reference calibration. The figures are similar to those in the paper, one exception being the autocorrelation coefficients of consumption and dividend growth. My simulation undershoot the first-order autocorrelation of both consumption and cash flow growth. I have been unable to find out the reason for this inconsistency.

Next we solve the model numerically to obtain asset pricing implications. The solution method is to guess and verify, we assume that the equilibrium wealth to consumption ratio is a linear function of the state variables $v_t = log(P/D)_t = A_0 + \mathbf{A'Y}_t$ and solve for linear coefficients and linearization coefficients from a system of non-linear equation implied by equilibrium restrictions that include the Euler equation and parameter constraints implied by the Campbell-Shiller log-linearization. Details of the solution method is provided in Appendix A of the paper. In the appendix of this report, I provide further details including derivation of moment generating functions of the jump size distributions and the solution details for the market portfolio, which are not elaborated in the paper's appendix.

The numerical solution is very unstable, and depends critically on the choice of algorithm, starting values and whether we impose model-implied constraints. There are three model-implied constraints: (1) since the effect of Δc and Δd on asset pricing is incorporated in x_t, they have no effect on asset prices and $A_c = A_d = 0$, (2) In the reference parameterization, the representative agent prefers early resolution of uncertainty and hence the price of risk of volatility shock is negative, hence $A_{\overline{\sigma}} < 0$ and $A_\sigma < 0$,(3) the log-linearization coefficient κ_1 has to be between 0 and 1 by construction. This gives us two ways of solving the model numerically:

1. Without imposing the model implied constraints, using fsolve or vpasolve in Matlab. The advantage of this approach is that it solves the equation system almost perfectly.

2. Imposing model-implied constraints using fmincon. The advantage of this approach is that is ensures that κ_1 always falls in $(0, 1)$ and the model-implied constraints are always satisfied, however, the solution is extremely sensitive to the specification of boundaries and the equation system cannot be exactly satisfied.

I coded each of these methods for both the consumption/wealth portfolio and the market portfolio. By commenting/uncommenting these sections the user can see results using different combinations of numerical methods. For the purpose of presentation , I choose the method that yields the most stable and sensible results overall. I choose to report results obtained by solving both consumption and market systems without imposing model restrictions. This is due to the fact that fmincon is too sensitive to the specification of lower/upper-bounds and cannot solve the equation system satisfactorily. Since the authors do not disclose the intermediary parameter values, it is not possible to compare my solution to theirs. Any differences in the subsequent simulation results should be due to difference in the numerical solutions.

Table 6 (corresponds to Table 7 in the paper) shows the quantiles of asset pricing statistics in 1000 simulations of the calibrated model for 924 months. My simulation severely overshoots equity returns. It is worth to mention that due to the jump components in the model, in the absence of non-negativity restriction on the risk free rate, the risk-free rate would jump to the negative domain several times in the simulated sample path, for example Figure 1. Here I impose non-negativity constraint on the risk free rate and report the statics of the bounded non-negative

Table 6: Equity return and risk free Rate

Statistic	5%	50%	95%
$E[r_m]$	10.71	14.14	17.84
Er_f	0.54	0.88	1.21
$\sigma(r_m)$	14.95	17.75	21.61
$\sigma(r_f)$	0.48	0.63	0.81
$E[p-d]$	2.47	2.57	2.67
$\sigma(p-d)$	0.14	0.18	0.25
$skew(r_m - r_f)$(M)	-0.78	-0.20	0.16
$kurt(r_m - r_f)$(M)	3.52	5.32	10.90
$AC1(r_m - r_f)$(M)	-0.07	-0.01	0.06
$kurt(r_m - r_f)$(A)	2.41	3.22	5.96

risk free rate, which results in a similar level of mean risk free rate but a lower level of volatility of the risk free rate. The authors do not specify how they treat cases in when the risk free rate jump to insensible negative values, but I believe that they imposed such a constraint.

My simulation also has a lower level of log price to dividend ratio than that is reported in the paper. These inconsistencies are due to difference in numerical solutions.

Figure 1: Risk free rate in a sample path

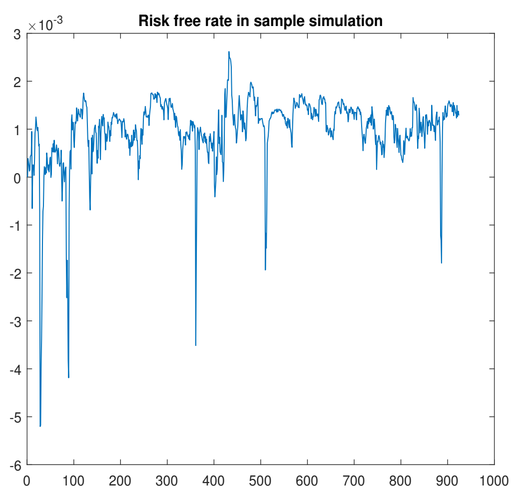


Table 7: Variance Premium

	5%	50%	95%
Variance Premium			
$\sigma(var_t(r_m))$	10.28	18.34	32.54
$AC1(var_t(r_m))$	0.77	0.85	0.91
$AC2(var_t(r_m))$	0.60	0.72	0.84
$E[VP]$	1.90	6.31	14.81
$\sigma(VP)$	2.25	8.86	24.67
$skew(VP)$	1.69	3.20	5.45
$kurt(VP)$	5.85	14.83	39.21
$kurt(\Delta VIX)$	25.04	45.55	106.95
Return Predictability (vp)			
$\beta(1m)$	-0.31	0.83	2.99
$R^2(1m)$	0.02	1.42	7.36
$\beta(3m)$	-0.65	2.09	8.02
$R^2(3m)$	0.04	3.36	15.50
Return Predictability (VP,p-d)			
$\beta_1(1m)$	-0.93	0.58	2.56
$\beta_2(1m)$	-152.27	-45.51	21.59
$R^2(1m)$	0.32	2.75	8.50
$\beta_1(3m)$	-2.11	1.40	6.14
$\beta_2(3m)$	-386.20	-134.02	39.78
$R^2(3m)$	0.73	6.82	19.43

Table 7 (corresponds to Table 8 in the paper) shows quantiles of the statistics of the variance premium in 1000 simulated samples. The numbers are different but over all qualitatively in line with the results in the paper.

Table 8 (corresponds to Table 9 in the paper) shows quantiles of long term predictability of consumption growth and excess market returns using the log price to dividend ratios in 1000 model simulations. The numbers are qualitatively similar to those in the paper. However, the evidence of predictability is weaker in my simulation. The differences result from difference in actual numerical solutions.

Table 9 (corresponds to Table 10 in the paper) shows the quantiles of predictive regressions of consumption volatility and excess market return volatility on log price-to-dividend ratio. The results are comparable to those in the paper.

Table 8: Long-horizon predictability

Consumption predictability			
	5%	50%	95%
$R^2(3y)$	0.07	6.28	26.10
$R^2(5y)$	0.08	8.00	32.77
$R^2(1y)$	0.08	6.51	32.38
Return predictability OLS			
$R^2(1y)$	0.70	4.57	11.11
$R^2(3y)$	0.68	8.25	21.21
$R^2(5y)$	0.51	9.25	25.28

Table 9: Predictability of volatility

Consumption volatility			
$R^2(1y)$	5%	50%	95%
$\beta(1y)$	-1.83	-0.67	0.75
$R^2(1y)$	0.01	1.46	9.40
$\beta(5y)$	-1.91	-0.16	1.80
$R^2(5y)$	0.02	4.37	26.83
Return Volatility			
$\beta(1y)$	-0.11	-0.05	-0.01
$R^2(3y)$	0.20	3.30	12.08
$\beta(5y)$	-0.08	-0.03	0.01
$R^2(5y)$	0.05	3.32	16.75

4 Comparative statics

4.1 Normal jump size in the conditional mean of consumption growth

In this section, I report the same set of results of model simulation under an alternative parameterization, that is, when the jump size in the conditional mean of consumption growth is normally distributed. This is provided in the paper only to show that the results do not depend on the choice of gamma distribution for jump size in the conditional mean of consumption growth.

Table 10: Cash Flows Normal Jump size

	5%	50%	95%
$E[\Delta c]$	1.01	1.90	2.83
$\sigma(\Delta c)$	2.13	2.63	3.43
$AC1(\Delta c)$	-0.04	0.20	0.44
$E[\Delta d]$	-1.40	1.79	4.85
$\sigma(\Delta d)$	11.59	13.48	15.60
$AC1(\Delta d)$	-0.17	0.04	0.24
$\rho_{\Delta c,\Delta d}$	0.09	0.28	0.48
$kurt(\Delta c)(Q)$	3.03	4.21	8.72

Table 10 (corresponds to Table 11 in the paper) reports the quantiles of the statistics of cash flow growth in 1000 simulations generated by the alternative specification with normal jump size in the conditional mean of consumption growth. Same as in the baseline case, most of the figures are in line with those reported in the paper, except for the autocorrelation coefficients, which tend to be lower.

Table 11: Equity Return and Risk Free Rate Normal Jump Size

	5%	50%	95%
$E[r_m]$	10.43	14.00	17.61
Er_f	0.10	0.34	0.69
$\sigma(r_m)$	15.09	17.73	21.55
$\sigma(r_f)$	0.22	0.51	0.78
$E[p-d]$	2.46	2.57	2.67
$\sigma(p-d)$	0.14	0.18	0.24
$skew(r_m - r_f)$(M)	-0.63	-0.11	0.25
$kurt(r_m - r_f)$(M)	3.46	5.14	10.04
$AC1(r_m - r_f)$(M)	-0.08	-0.01	0.06
$kurt(r_m - r_f)$(A)	2.39	3.23	6.37

Table 11 (corresponds to Table 12 in the paper) reports the quantiles of asset pricing statistics of 1000 samples generated by the alternative specification. The numerical solution method is chosen as the same as before, without imposing model implied restrictions and use fsolve or vpasolve to get solutions that most perfectly satisfy the equation systems. The problem here is the same as in the baseline case, I overshoot the equity return and now undershoot the risk free rate. The price-to-dividend ratio is again lower than that is reported in the paper. I think the most likely reason for the difference in results is still the numerical solution. Since I do not know which numerical algorithm is used buy the authors and there are too many degrees of freedom, a perfect replication is very difficult. The bottom line is that the results are at least qualitatively similar.

Table 12 (corresponds to Table 13 in the paper) reports the quantiles of the variance premium statistics of 1000 samples generated by the alternative specification. The results are in general similar to those in the original paper, although as before, I undershoot the average variance risk premium.

4.2 Shutting off the jump shocks

Finally, Table 13 (corresponds to Table 14 in the paper) reports comparative statics of three alternative specifications: Model A: shutting down the jump component in volatility, Model B: shutting down the jump component in the conditional mean of consumption growth, Model C: shutting down both jump components. The results are again qualitatively similar and I would attribute any differences to the possible differences in the implementation of numerical solution. It is clear that in the absence

Table 12: Variance Premium Normal Jump Size

Variance Premium			
	5%	50%	95%
$\sigma(var_t(r_m))$	9.58	17.77	33.53
$AC1(var_t(r_m))$	0.77	0.85	0.91
$AC2(var_t(r_m))$	0.59	0.72	0.83
$E[VP]$	1.56	4.53	10.93
$\sigma(VP)$	1.77	6.28	17.25
$skew(VP)$	1.55	3.21	5.32
$kurt(VP)$	5.48	15.07	38.32
$kurt(\Delta VIX)$	24.84	45.29	100.79
Return Predictability (vp)			
$\beta(1m)$	-0.34	1.12	3.77
$R^2(1m)$	0.01	1.42	7.14
$\beta(3m)$	-0.71	2.92	9.87
$R^2(3m)$	0.02	3.39	16.23
Return Predictability (VP,p-d)			
$\beta_1(1m)$	-1.24	0.81	3.26
$\beta_2(1m)$	-145.18	-44.73	21.40
$R^2(1m)$	0.28	2.93	8.38
$\beta_1(3m)$	-3.07	1.94	7.86
$\beta_2(3m)$	-397.94	-136.04	48.28
$R^2(3m)$	0.71	7.22	19.72

of jumps, the model fails to generate the desired level of variance premium.

Table 13: Comparative Statics Results

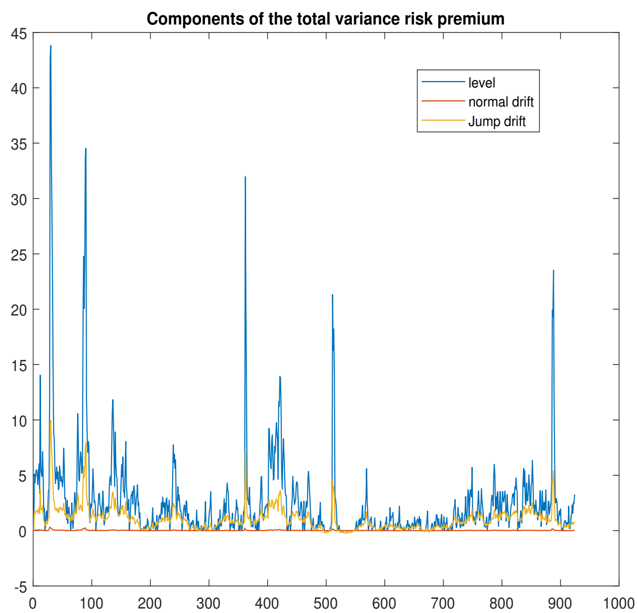
	Model 1-A			Model 1-B			Model 1-C		
Statistic	5%	50%	95%	5%	50%	95%	5%	50%	95%
Cash-flow Dynamics									
$E[\Delta c]$	0.99	1.92	2.79	1.27	1.91	2.61	1.20	1.91	2.59
$\sigma(\Delta c)$	2.20	2.66	3.29	2.02	2.50	3.16	2.08	2.49	3.00
$AC1(\Delta c)$	-0.03	0.20	0.43	-0.10	0.12	0.36	-0.09	0.13	0.34
$E[\Delta d]$	-1.12	1.81	5.05	-0.91	1.90	4.65	-1.06	1.79	4.81
$\sigma(\Delta d)$	11.54	13.45	15.39	11.37	13.27	15.27	11.35	13.28	15.18
$AC1(\Delta d)$	-0.14	0.05	0.24	-0.18	0.02	0.21	-0.18	0.02	0.21
$\rho_{\Delta c,\Delta d}$	0.09	0.29	0.47	0.04	0.24	0.43	0.06	0.25	0.41
Returns									
$E[r_m]$	3.07	5.98	9.05	2.26	5.05	7.91	-0.14	2.73	5.73
$E[r_f]$	0.67	1.05	1.40	1.31	1.57	1.84	1.27	1.58	1.86
$\sigma(r_m)$	13.73	16.02	18.61	12.45	14.74	17.29	12.32	14.52	16.66
$\sigma(r_f)$	0.54	0.68	0.83	0.42	0.58	0.76	0.41	0.57	0.73
$E[p-d]$	3.10	3.19	3.27	4.18	4.24	4.29	4.65	4.71	4.76
$\sigma(p-d)$	0.12	0.15	0.21	0.09	0.12	0.16	0.09	0.12	0.14
$skew(r_m-r_f)$(M)	-0.62	-0.13	0.10	-0.22	0.05	0.36	-0.13	0.02	0.16
$kurt(r_m-r_f)$(M)	3.05	3.76	7.08	3.08	3.79	5.82	2.85	3.14	3.51
$AC1(r_m-r_f)$(M)	-0.06	-0.00	0.05	-0.06	-0.00	0.06	-0.06	-0.00	0.05
Variance Premium									
$\sigma(var_t(r_m))$	4.29	5.39	6.77	5.31	9.81	17.17	3.16	3.99	5.05
$AC1(var_t(r_m))$	0.85	0.89	0.92	0.77	0.85	0.91	0.84	0.89	0.92
$AC2(var_t(r_m))$	0.74	0.81	0.86	0.59	0.73	0.83	0.74	0.80	0.86
$E[VP]$	1.06	2.44	4.11	0.08	0.23	0.56	0.01	0.02	0.03
$\sigma(VP)$	0.96	1.63	2.44	0.08	0.32	0.93	0.01	0.01	0.02
$skew(VP)$	0.37	0.93	1.79	1.62	3.13	5.15	0.41	0.97	1.97
$kurt(VP)$	2.35	3.47	6.92	6.06	14.80	36.22	2.36	3.56	7.47
$\beta(1m)$	-3.15	1.63	6.12	-20.82	6.76	45.16	-475.53	94.60	731.35
$R^2(1m)$	0.00	0.40	3.01	0.00	0.49	5.33	0.00	0.31	2.41
$\beta(3m)$	-7.88	4.55	16.92	-59.27	18.25	113.65	-1310.61	257.84	1873.27
$R^2(3m)$	0.01	1.09	6.87	0.01	1.30	11.14	0.01	0.85	6.17

Model A shuts down the jump component in σ_t by setting $l_{1,\sigma}=0$, Model B shuts down the jump component in x_t by setting $l_{1,x}=0$, Model C shuts down jump component in both σ_t and x_t.

5 Takeaways and additional analysis

Finally it is worth to make a few comments about how the model works. The role of jump risk is important in this model to generate the level of the variance risk premium. The paper decomposes the variance risk premium into a level component and a drift component, and the latter can be attributed to either gaussian shocks or jump shocks. Figure 2 illustrates the relative importance of these components in a sample simulation. It is immediately obvious that the level component is the dominant driver of the total variance risk premium. In the drift component, the con tribulation of the gaussian shocks is almost negligible while jump risks contribute to a significant proportion in the drift component of the variance risk premium.

Figure 2: components of the variance premium in a sample path



My biggest lesson learnt from this exercise is to never blindly trust anything that is written in a paper. It took me quite some time to realize that there is a typo in the paper's appendix (as is specified in the appendix of this report), which leads to non-existence of reasonable numerical solutions.

Another takeaway is to be careful about numerical solution methods as they are not always trustworthy.

6 Appendix

6.1 Moment generating functions (MGF) and their derivatives

The MGF of jump size distribution is used to evaluate the left-hand-side of the Euler equation in the solution of the model.

In the reference calibration, the jump size of conditional volatility σ_t follows $\Gamma(\nu_\sigma, \frac{\mu_\sigma}{\nu_\sigma})$, then the moment generating function is $MGF_{\xi_\sigma}(u) = 1 - \frac{\mu_\sigma}{\nu_\sigma}u)^{-\nu_\sigma}$ The jump in conditional mean x_t is calibrated to be $\xi_x \sim -\Gamma(\nu_x, \frac{\mu_x}{\nu_x}) + \mu_x$. Therefore the moment generating function of ξ_x is: $\psi(u) = \mathbb{E}[exp(u\xi_x)] = \mathbb{E}[exp(-\Gamma(\nu_x, \frac{\mu_x}{\nu_x}) + \mu_x)u)] = \mathbb{E}[exp(-\Gamma(\nu_x, \frac{\mu_x}{\nu_x})u]exp(\mu_x u)$. We know that Gamma distribution has MGF: $(1-\mu t)^{-\nu}$ for $t < \frac{1}{\mu}$ and to get the MGF of a negative Gamma random variable, we need to do some algebra.

$$
\begin{aligned}
MGF_{-\Gamma(\nu_x, \frac{\mu_x}{\nu_x})}(u) =& \frac{1}{\Gamma(\nu_x)(\mu_x/\nu_x)^{\nu_x}} \int_0^\infty e^{-ux} x^{\nu_x - 1} e^{-\frac{x}{(\mu_x/\nu_x)}} dx \\
=& \frac{1}{\Gamma(\nu_x)(\mu_x/\nu_x)^{\nu_x}} \int_0^\infty x^{\nu_x - 1} e^{-\frac{x}{\frac{(\mu_x/\nu_x)}{1+(\mu_x/\nu_x u}}} dx \\
=& \frac{1}{\Gamma(\nu_x)(\mu_x/\nu_x)^{\nu_x}} \Gamma(\nu_x) (\frac{\mu_x/\nu_x}{1+\mu_x/\nu_x u})^{\nu_x} \\
=& (\frac{1}{1+\mu_x/\nu_x u})^{\nu_x}
\end{aligned}
\tag{2}
$$

Then the moment generating function of ξ_x is:

$$
MGF_{\xi_x}(u) = \psi_x(u) = (\frac{1}{1+\mu_x/\nu_x u})^{\nu_x} exp(\mu_x u) \tag{3}
$$

First and second order derivatives of the MGF are also used in the model solution:

$$
\psi_x^{(1)}(u) = -\mu_x(1 + \frac{\mu_x}{\nu_x}u)^{1-\nu_x} exp(\mu_x u) + \mu_x(1 + \frac{\mu_x}{\nu_x}u)^{-\nu_x} exp(\mu_x u) \tag{4}
$$

$$
\psi_x^{(2)}(u) = \frac{\mu_x^2(1+\nu_x)(1+\frac{\mu_x}{\nu_x}u)^{-2-\nu_x} exp(\mu_x u)}{\nu_x} - 2\mu_x^2(1 + \frac{\mu_x}{\nu_x}u)^{-1-\nu_x} exp(\mu_x u) + \mu_x^2(1 + \frac{\mu_x}{\nu_x}u)^{-\nu_x} exp(\mu_x u)
$$

Likewise for ξ_σ

$$\begin{aligned} MGF_{\Gamma(\nu_\sigma,\frac{\mu_\sigma}{\nu_\sigma})}(u) = \psi_\sigma(u) &= (1-\frac{\mu_\sigma}{\nu_\sigma}u)^{-\nu_\sigma} \\ \psi_\sigma^{(1)} &= \mu_\sigma(1-\frac{\mu_\sigma}{\nu_\sigma}u)^{-1-\nu_\sigma} \\ \psi_\sigma^{(2)} &= \frac{\mu_\sigma^2(1+\nu_\sigma)(1-\frac{\mu_\sigma}{\nu_\sigma}u)^{-2-\nu_\sigma}}{\nu_\sigma} \end{aligned} \tag{5}$$

In the alternative specification in which the jump size in x is specified as a normal distribution: $\xi_X \sim N(0,\sigma_x)$, where $\sigma_x = \mu_x$, we need the moment generating function and its derivatives of a normal distribution:

$$\begin{aligned} MGF_{N(0,\sigma_x)}(u) &= exp(\sigma_x^2 u^2/2) \\ MGF^{(1)}_{N(0,\sigma_x)}(u) &= \sigma_x^2 exp(\sigma_x^2 u^2/2)u \\ MGF^{(2)}_{N(0,\sigma_x)}(u) &= \sigma_x^2 exp(\sigma_x^2 u^2/2) + \sigma_x^4 u^2 exp(\sigma_x^2 u^2/2) \end{aligned} \tag{6}$$

6.2 Solving the market portfolio system

The paper does not explicitly lay out the equation system for market return coefficients, but the structure is almost the same as that for the wealth portfolio. According to the definition of the log-linearization coefficients $\kappa_{0,m}$ and $\kappa_{1,m}$ in $r_{m,t+1} = \kappa_{0,m} + \kappa_{1,m}\nu_{m,t+1} + d_{t+1}$ (where ν is the log price-to-dividend ratio), we have a similar system of equations as in the paper's appendix (A.2.1-A.2.2):

$$\ln\kappa_{1,m} - \ln(1-\kappa_1) = \mathbb{E}(\nu_t) = A_{0,m} + \mathbf{A_m}'\mathbb{E}(\mathbf{Y}_t) \tag{7}$$

$$\kappa_{0,m} = -\kappa_{1,m}\ln\kappa_{1,m} - (1-\kappa_{1,m})\ln(1-\kappa_{1,m}) \tag{8}$$

After some manipulation, the set of equations gives us:

$$\kappa_{0,m} + (\kappa_{1,m}-1)A_{0,m} = -\ln\kappa_{1,m} + (1-\kappa_{1,m})A_m'\mathbb{E}(\mathbf{Y}) \tag{9}$$

Substitute this into the (A.3.1) to substitute out A_0, m and κ_0 we have the following equation system for $\kappa_{1,m}$ and $\mathbf{A}_{1,m}$ (the typos in the paper are highlighted in red):

$$0 = \theta ln(\delta) - (1-\theta)(\kappa_1 - 1)A_0 - (1-\theta)\kappa_0 - ln\kappa_{1,m} + (1-\kappa_{1,m})A_m'\mathbb{E}(\mathbf{Y}) + \mathbf{f}(\mathbf{e_d} + \kappa_{1,m}\mathbf{A_m} - \mathbf{\Lambda}) \tag{10}$$

$$0 = \mathbf{g}(e_d + \kappa_{1,m}\mathbf{A_m} - \mathbf{\Lambda}) + (1-\theta)\mathbf{A} - \mathbf{A_m}$$

Where $\mathbf{\Lambda} = \gamma \mathbf{e_c} + (1-\theta)\kappa_1 \mathbf{A}$ is the vector of market price of risk calculated in P.13. suppose I have the correct numerical solution for $\kappa_{0,m}$,$\kappa_{1,m}$ and $A_{0,m}$, A_m, the expression of market return is, since $\nu_{\nu,m} = A_{0,m} + \mathbf{A}'_m \mathbf{Y}_t$:

$$
\begin{aligned}
r_{m,t+1} &= \kappa_{0,m} + \kappa_{1,m}\nu_{m,t+1} - \nu_{m,t} + \Delta d_{t+1} \qquad (11)\\
&= \kappa_{0,m} + \kappa_{1,m}(A_{0.m} + \mathbf{A}'_m \mathbf{Y}_{t+1}) - A_{0,m} - \mathbf{A}'_m \mathbf{Y}_t + \mathbf{e_d} \mathbf{Y}_{t+1}\\
&= \kappa_{0,m} + \kappa_{1,m} A_{0,m} - A_{0,m} + \underbrace{(\kappa_{1,m}\mathbf{A_m}' + \mathbf{e_d})}_{\mathbf{B_r}'} \mathbf{Y}_{t+1} - \mathbf{A_m}'\mathbf{Y}_t\\
&= \kappa_{0,m} + \kappa_{1,m} A_{0,m} - A_{0,m} + B'_r(\boldsymbol{\mu} + \mathbf{F}\mathbf{Y}_t + \mathbf{G}_t \mathbf{z}_{t+1} + \mathbf{J}_{t+1}) - \mathbf{A_m}'\mathbf{Y}_t\\
&= \kappa_{0,m} + (\kappa_{1,m} - 1)A_{0,m} + \mathbf{B}'_r \mu + (\mathbf{B}'_r \mathbf{F} - \mathbf{A}'_m)\mathbf{Y}_t + \mathbf{B}'_r \mathbf{G}_t z_{t+1} + \mathbf{B}'_r \mathbf{J}_{t+1}\\
&= -ln(\kappa_{1m}) + (1-\kappa_{1m})\mathbf{A_m}'\mathbb{E}\mathbf{Y} + \mathbf{B}'_r \boldsymbol{\mu} + (\mathbf{B}'_r \mathbf{F} - \mathbf{A}'_m)\mathbf{Y}_t + \mathbf{B}'_r \mathbf{G}_t z_{t+1} + \mathbf{B}'_r \mathbf{J}_{t+1};
\end{aligned}
$$

6.3 Drift component of the variance premium due to jump risks.

Appendix C mentions the drift component of the variance premium due to jump risks, but does not specify its mathematical expression. For the purpose of simulation I need to solve for this component explicitly. Appendix C gives that the drift difference due to jump risk is:

$$
\text{drift difference} = B_r^{2'} diag(\psi^{(2)}(-\lambda))[\mathbb{E}_t^{\mathbb{Q}}(\lambda_{t+1}) - \lambda_t] - B_r^{2'} diag(\psi^{(2)}(0))[\mathbb{E}_t^{\mathbb{P}}(\lambda_{t+1}) - \lambda_t] \qquad (12)
$$

Using the model specification for jump intensity and the moment generating function, we have:

$$
\begin{aligned}
\mathbb{E}^{\mathbb{P}}[\lambda_{t+1}] &= l_1 \mathbb{E}^{\mathbb{P}}[Y_{t+1}] = l1(\boldsymbol{\mu} + \mathbf{F}\mathbf{Y}_t) + l_1 \mathbb{E}^{\mathbb{P}}[J] \qquad (13)\\
\mathbb{E}^{\mathbb{Q}}[\lambda_{t+1}] &= l_1 \mathbb{E}^{\mathbb{P}}[Y_{t+1}] = l1(\boldsymbol{\mu} + \mathbf{F}\mathbf{Y}_t) + l_1 \mathbb{E}^{\mathbb{Q}}[J]\\
\mathbb{E}^{\mathbb{P}}[J_\sigma] &= diag(\psi^{(1)}(0))\lambda_t \qquad (14)\\
\mathbb{E}^{\mathbb{Q}}[J_\sigma] &= diag(\psi^{(1)}(-\Lambda_k)\lambda_t
\end{aligned}
$$

Substitute into (12) we have explicit expression for the drift component of the variance premium due to jump risks.

References

Bollerslev, T., Tauchen, G., and Zhou, H. (2009). Expected Stock Returns and Variance Risk Premia.

Drechsler, I. and Yaron, A. (2011). What's Vol Got to Do with It. *Review of Financial Studies*, 24(1):1–45.