Online Appendix: Implied Dividend Volatility and Expected Growth

Niels J. Gormsen Ralph S.J. Koijen Ian W.R. Martin*

January 9, 2021

I. Online appendix

A. *True correlation equals risk-neutral correlation under lognormality*

If $M_{t+1} = e^{-r_{f,t}-\frac{1}{2}\sigma^2_{1,t}-\sigma_{1,t}Z_{1,t+1}}$ and $R_{d,t+1} = e^{\mu_{d,t}-\frac{1}{2}\sigma^2_{2,t}+\sigma_{2,t}Z_{2,t+1}}$ and $R_{t+1} = e^{\mu_t-\frac{1}{2}\sigma^2_{3,t}+\sigma_{3,t}Z_{3,t+1}}$ then we must have $\mu_{d,t}-r_{f,t} = \rho_{12,t}\sigma_{1,t}\sigma_{2,t}$ and $\mu_t-r_{f,t} = \rho_{13,t}\sigma_{1,t}\sigma_{3,t}$ so that $\mathbb{E}\, MR = 1$ holds; we write $\rho_{ij,t}$ for $\text{corr}_t(Z_{i,t+1}, Z_{j,t+1})$. Standard calculations show that $\text{Cov}_t(R_{d,t+1}, R_{t+1}) = e^{\mu_{d,t}+\mu_t}(e^{\rho_{23,t}\sigma_{2,t}\sigma_{3,t}}-1)$ and $\text{Cov}^*_t(R_{d,t+1}, R_{t+1}) = e^{2r_{f,t}}(e^{\rho_{23,t}\sigma_{2,t}\sigma_{3,t}}-1)$; similarly, $\text{Var}_t R_{d,t+1} = e^{2\mu_{d,t}}(e^{\sigma^2_{2,t}}-1)$ and $\text{Var}_t R_{t+1} = e^{2\mu_t}(e^{\sigma^2_{3,t}}-1)$, while risk-neutral variances are $\text{Var}^*_t R_{d,t+1} = e^{2r_{f,t}}(e^{\sigma^2_{2,t}}-1)$ and $\text{Var}^*_t R_{t+1} = e^{2r_{f,t}}(e^{\sigma^2_{3,t}}-1)$. It follows that the true and risk-neutral correlations are equal:

$$\text{corr}_t\left(R_{d,t+1}, R_{t+1}\right) = \text{corr}^*_t\left(R_{d,t+1}, R_{t+1}\right) = \frac{e^{\rho_{23,t}\sigma_{2,t}\sigma_{3,t}}-1}{\sqrt{\left(e^{\sigma^2_{2,t}}-1\right)\left(e^{\sigma^2_{3,t}}-1\right)}}.$$

B. *Conditions under which* $\text{Cov}_t(M_{t:t+\tau}R^M_{t:t+\tau}, R^\tau_{t:t+\tau}) \le 0$

Suppose that

1. The SDF takes the form

$$M_{t:t+\tau} = \beta\frac{V_W(W_{t+\tau}, z_{1,t+\tau},\ldots,z_{N,t+\tau})}{V_W(W_t, z_{1,t},\ldots,z_{N,T})}$$

*Gormsen: University of Chicago, Booth School of Business, 5807 South Woodlawn Avenue, 60637 Chicago, Illinois, USA, niels.gormsen@chicagobooth.edu. Koijen: University of Chicago, Booth School of Business, 5807 South Woodlawn Avenue, 60637 Chicago, Illinois, USA, ralph.koijen@chicagobooth.edu. Martin: London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom, i.w.martin@lse.ac.uk. We thank Lasse Pedersen for comments and suggestions.

where $z_1, \ldots, z_N$ are state variables (and we allow for the possibility that there are no state variables).

2. Wealth is invested in the market and in some other asset or portfolio of assets with return $\widetilde{R}_{t:t+\tau}$:

$$W_{t+\tau} = \underbrace{\alpha_t (W_t - C_t) R^M_{t:t+\tau}}_{\text{market wealth, } W^M_{t+\tau}} + \underbrace{(1-\alpha_t)(W_t - C_t)\widetilde{R}_{t:t+\tau}}_{\text{non-market wealth}}, \quad \text{where } 0 \leq \alpha_t \leq 1.$$

3. Risk aversion is sufficiently high:

$$-\frac{W V_{WW}}{V_W} \geq \frac{W_{t+\tau}}{W^M_{t+\tau}}.$$

4. The market return, non-market return, dividend return, and the state variables are associated random variables (when the signs of the state variables are chosen so that $V_{Wz_n} \leq 0$ for $n = 1, \ldots, N$, so that the marginal value of wealth is decreasing in each state variable, just as it is decreasing in wealth).

Then we have

$$\mathrm{Cov}_t(M_{t:t+\tau} R^M_{t:t+\tau}, R^\tau_{t:t+\tau}) \leq 0.$$

Assumption 3 says that if, say, at least a third of wealth is in the market, then the condition holds so long as risk aversion is at least three. We have made Assumption 4 as general as possible by using the concept of associated random variables (Esary et al., 1967), which generalizes (weak) positive correlation to collections of more than two random variables. If, say, the random variables are Normal random variables—or increasing functions of Normal random variables—then they are associated if and only if the pairwise correlations between the Normal random variables are nonnegative (Pitt, 1982).

Proof. We follow a similar strategy to the proof given for Example 3b in Martin (2017). We must show that

$$\mathrm{Cov}_t(-R^M_{t:t+\tau} V_W(W_{t+\tau}, z_{1,t+\tau}, \ldots, z_{N,t+\tau}), R^\tau_{t:t+\tau}) \geq 0.$$

So we must prove that the covariance of two functions of $R^M_{t:t+\tau}$, $R^\tau_{t:t+\tau}$, $\widetilde{R}_{t:t+\tau}$, $z_{1,t+\tau}$, $\ldots$, $z_{N,t+\tau}$ is positive. The two functions are

$$f(R^M_{t:t+\tau}, R^\tau_{t:t+\tau}, \widetilde{R}_{t:t+\tau}, z_{1,t+\tau}, \ldots, z_{N,t+\tau}) = -R^M_{t:t+\tau} V_W \left(\alpha_t (W_t - C_t) R^M_{t:t+\tau} + (1-\alpha_t)(W_t - C_t)\widetilde{R}_{t:t+\tau}, z_{1,t+\tau}, \ldots,\right.$$

and

$$g(R^M_{t:t+\tau}, R^\tau_{t:t+\tau}, \widetilde{R}_{t:t+\tau}, z_{1,t+\tau}, \ldots, z_{N,t+\tau}) = R^\tau_{t:t+\tau}.$$

(As the covariance is conditional on time t information, we can treat α_t and $W_t - C_t$ as known constants.) As the random variables are associated, the result follows if f and g are each weakly increasing functions of their arguments. But this follows from the assumptions above. For example, differentiating f with respect to $R^M_{t:t+\tau}$, we need

$$-V_W(W_{t+\tau}, z_{1,t+\tau}, \ldots, z_{N,t+\tau}) - \alpha_t(W_t - C_t)R^M_{t:t+\tau}V_{WW}(W_{t+\tau}, z_{1,t+\tau}, \ldots, z_{N,t+\tau}) \geq 0\,.$$

Rearranging, this reduces to the constraint on risk aversion provided in assumption 3 above. The other necessary conditions on f and on g are trivially satisfied.

C. Implied volatility in equilibrium models

Although we have focussed on dividend volatility, consumption-based models also have difficulty matching the time series behavior of price volatility. Martin (2017, Table IV) reports time series of various statistics (mean, median, standard deviation, min, max, skewness, kurtosis, and autocorrelation) of sample paths of the VIX and SVIX indices generated in the model economies of Campbell and Cochrane (1999), Bansal and Yaron (2004), Bansal et al. (2012), Bollerslev et al. (2009), Drechsler and Yaron (2011), and Wachter (2013). None of these is able to generate sample paths that resemble those observed empirically. All the models apart from Drechsler and Yaron (2011) generate volatility series that are more persistent than the data. The empirically observed mean gap between VIX and SVIX—a measure of the average importance of extreme left-tail events—is outside the support of the million sample paths in every model: Wachter (2013) overstates the importance of such events, and all other models understate it. All the models apart from Wachter (2013) fail, on 99% of sample paths, to generate the maximum levels of VIX that have been observed in reality. All models apart from Drechsler and Yaron (2011) fail, on 99% of sample paths, to match the kurtosis of VIX and SVIX.

We also refer to Dew-Becker et al. (2017) for the a set of related challenges of models to match the term structure of variance swaps and variance risk premia. Dew-Becker et al. show that the monthly risk premium on short-term variance (≤ 3 months) is negative and large whereas the risk premium on longer-term variance (> 3 months) is essentially zero. The fact that very-short run variance is priced but longer-term variance is not implies the existence of a transitory element in realized volatility that investors are highly averse towards. The results also imply that investors are not averse towards changes in long-term expected volatility – something that is hard to reconcile with long-run risk models (Drechsler and Yaron, 2011) and disaster models with Epstein-Zin preferences (Wachter, 2013) as investors in such models are averse towards increases in expected volatility, which is modeled to be persistent, meaning that claims on longer-term variance should be priced.

D. Campbell and Cochrane (1999)

In this subsection, we use the notation from Campbell and Cochrane (1999) without further comment. The price of a claim to the first dividend is

$$\begin{aligned}
\mathbb{E}_t\left(M_{t+1}D_{t+1}\right) &= D_t\,\mathbb{E}_t\left(M_{t+1}\frac{D_{t+1}}{D_t}\right)\\
&= \delta D_t\,\mathbb{E}_t\left(\exp\left\{-\gamma g-\gamma\left\{(\phi-1)(s_t-\overline{s})+[1+\lambda(s_t)]v_{t+1}\right\}+g+w_{t+1}\right\}\right)\\
&= \delta D_t\exp\left\{g(1-\gamma)+\gamma(1-\phi)(s_t-\overline{s})+\frac{\gamma^2\sigma^2}{2\overline{S}^2}(1-2(s_t-\overline{s}))+\frac{1}{2}\sigma_w^2-\frac{\gamma\rho\sigma\sigma_w}{\overline{S}}\sqrt{1-2(s_t-\overline{s})}\right\}
\end{aligned}$$

So the return on this dividend strip is

$$R_{d,t+1}=\frac{D_{t+1}}{\mathbb{E}_t\left(M_{t+1}D_{t+1}\right)}=\frac{D_{t+1}}{D_t}\delta^{-1}\exp\left\{g(\gamma-1)-\gamma(1-\phi)(s_t-\overline{s})-\frac{\gamma^2\sigma^2}{2\overline{S}^2}(1-2(s_t-\overline{s}))-\frac{1}{2}\sigma_w^2+\frac{\gamma\rho\sigma\sigma_w}{\overline{S}}\sqrt{1-\right.$$

Hence the expected return is

$$\begin{aligned}
\mathbb{E}_t\,R_{d,t+1} &= \delta^{-1}\exp\left\{g\gamma-\gamma(1-\phi)(s_t-\overline{s})-\frac{\gamma^2\sigma^2}{2\overline{S}^2}(1-2(s_t-\overline{s}))+\frac{\gamma\rho\sigma\sigma_w}{\overline{S}}\sqrt{1-2(s_t-\overline{s})}\right\}\\
&= R_f\exp\left\{\frac{\gamma\rho\sigma\sigma_w}{\overline{S}}\sqrt{1-2(s_t-\overline{s})}\right\},
\end{aligned}$$

and

$$\mathrm{Var}_t\log R_{d,t+1}=\sigma_w^2\,.$$

As M_{t+1} and $R_{d,t+1}$ are conditionally lognormal, $\mathrm{Cov}_t(M_{t+1}R_{d,t+1},R_{d,t+1})\leq 0$ if and only if

$$\frac{\log\mathbb{E}_t\,R_{d,t+1}-\log R_f}{\sigma_t(\log R_{d,t+1})}\geq\sigma_t(\log R_{d,t+1})\,.$$

The quantity on the left-hand side of the inequality is, essentially, the Sharpe ratio of the dividend strip. The preceding results therefore imply that $\mathrm{Cov}_t(M_{t+1}R_{d,t+1},R_{d,t+1})\leq 0$ if and only if

$$\frac{\gamma\rho\sigma}{\overline{S}}\sqrt{1-2(s_t-\overline{s})}\geq\sigma_w.$$

Figure **??** shows that the condition holds in sufficiently bad states of the world, but not in good states of the world and not at the steady state level of habit, $\overline{S}$.

E. Bansal and Yaron (2004)

In this subsection, we use the notation from Bansal and Yaron (2004) without further comment. We focus on the Case II calibration that features stochastic volatility.

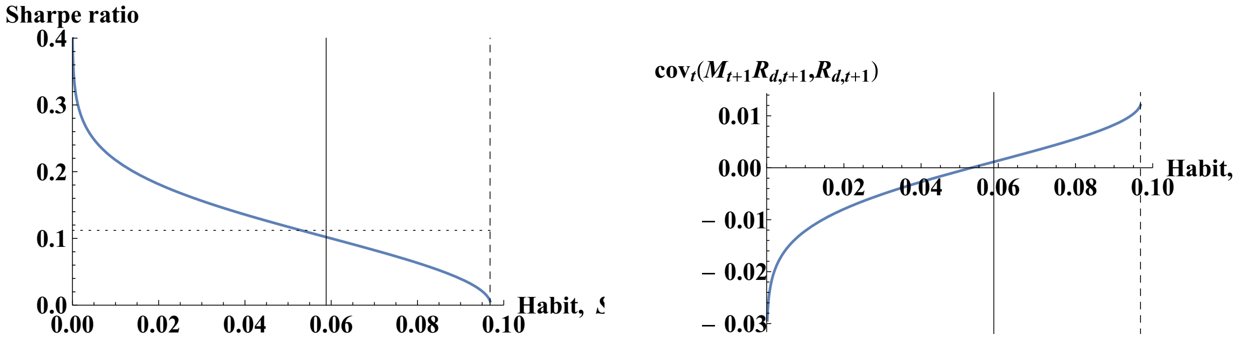

Figure A.1. The Sharpe ratio, and the covariance term, for the one-period dividend strip in Campbell and Cochrane (1999). The dotted line in the left panel indicates the volatility of the dividend strip return, σ_w. The solid vertical line indicates the steady state level of habit. The dashed vertical line indicates the maximum attainable level of habit.

Single-period calculations. A claim to the first dividend earns zero risk premium because dividend growth is conditionally uncorrelated with the log stochastic discount factor, $\text{Cov}_t(g_{d,t+1}, m_{t+1}) = 0$. (We are following Bansal and Yaron by treating the model as conditionally lognormal, relying on loglinearizations.) Hence $\mathbb{E}_t\, R_{d,t+1} = R_{f,t+1}$. Again exploiting lognormality, the risk-neutral variance takes the form above,

$$\text{Var}_t^* R_{d,t+1} = R_{f,t+1}^2 \left(e^{\text{Var}_t \log g_{d,t+1}} - 1\right) = R_{f,t+1}^2 \left(e^{\varphi_d^2 \sigma_t^2} - 1\right),$$

and

$$\text{Cov}_t\left(M_{t+1}R_{d,t+1}, R_{d,t+1}\right) = \frac{1}{R_{f,t+1}} \text{Var}_t^* R_{d,t+1} - \left(\mathbb{E}_t\, R_{d,t+1} - R_{f,t+1}\right) = R_{f,t+1}\left(e^{\varphi_d^2 \sigma_t^2} - 1\right).$$

Hence the conditional covariance is positive in Bansal and Yaron (2004).

Bansal et al. (2012) make dividends and consumption growth correlated in the short run. As a result, $\text{Cov}_t(g_{d,t+1}, m_{t+1}) = -\gamma\pi\sigma_t^2$, and hence $\mathbb{E}_t\, R_{d,t+1} = R_{f,t+1} e^{\gamma\pi\sigma_t^2}$, where π is a new parameter introduced in equation (3) of Bansal et al. (2012). Risk-neutral variance changes slightly:

$$\text{Var}_t^* R_{d,t+1} = R_{f,t+1}^2 \left(e^{\left(\pi^2+\varphi^2\right)\sigma_t^2} - 1\right)$$

and

$$\text{Cov}_t\left(M_{t+1}R_{d,t+1}, R_{d,t+1}\right) = R_{f,t+1}\left(e^{\left(\pi^2+\varphi^2\right)\sigma_t^2} - e^{\gamma\pi\sigma_t^2}\right).$$

This is positive in their calibration, in which $\pi^2 + \varphi^2 = 42.3 > \gamma\pi = 26$.

Multi-period calculations. Bansal and Yaron (2004) calibrate the model to a monthly frequency, while we use 1- to 3-year dividend claims. We therefore compute the risk-neutral variance and the covariance for longer horizons for completeness. The model implies that

$$P_t(\tau) = D_t \exp(A_\tau + B_\tau x_t + C_\tau \sigma_t^2),$$

where the coefficients follow from

$$\begin{aligned} P_t(\tau) &= D_t \,\mathbb{E}_t \left(M_{t+1} \exp\left\{ g_{d,t+1} + A_{\tau-1} + B_{\tau-1} x_{t+1} + C_{\tau-1} \sigma_{t+1}^2 \right\} \right) \\ &= D_t \exp\left\{ \mu_d + \phi x_t + A_{\tau-1} + B_{\tau-1} \rho x_t + C_{\tau-1} \sigma^2 (1-\nu_1) + C_{\tau-1} \nu_1 \sigma_t^2 \right\} \times \\ &\quad R_{ft}^{-1} \exp\left\{ \frac{1}{2} \left(\varphi_d^2 + B_{\tau-1}^2 \varphi_e^2 \right) \sigma_t^2 + \frac{1}{2} C_{\tau-1}^2 \sigma_w^2 - \lambda_{m,e} B_{\tau-1} \varphi_e \sigma_t^2 - \lambda_{m,w} C_{\tau-1} \sigma_w^2 \right\} \end{aligned}$$

where $R_{ft} = \exp\{s_0 + s_1 x_t + s_2 \sigma_t^2\}$ and thus

$$\begin{aligned} A_\tau &= -s_0 + \mu_d + A_{\tau-1} + C_{\tau-1} \sigma^2 (1-\nu_1) + \frac{1}{2} C_{\tau-1}^2 \sigma_w^2 - \lambda_{m,w} C_{\tau-1} \sigma_w^2, \\ B_\tau &= -s_1 + \phi + B_{\tau-1} \rho, \\ C_\tau &= -s_2 + C_{\tau-1} \nu_1 + \frac{1}{2} \left(\varphi_d^2 + B_{\tau-1}^2 \varphi_e^2 \right) - \lambda_{m,e} B_{\tau-1} \varphi_e. \end{aligned}$$

Returns are given by

$$\begin{aligned} R_{d,t+1}(\tau) &= \exp\left\{ g_{d,t+1} + A_{\tau-1} - A_\tau + B_{\tau-1} x_{t+1} - B_\tau x_t + C_{\tau-1} \sigma_{t+1}^2 - C_\tau \sigma_t^2 \right\} \\ &= f_t(\tau, x_t, \sigma_t^2) \exp\left\{ \varphi_d \sigma_t u_{t+1} + B_{\tau-1} \varphi_e \sigma_t e_{t+1} + C_{\tau-1} \sigma_w w_{t+1} \right\}, \end{aligned}$$

where

$$\begin{aligned} f_t(\tau, x_t, \sigma_t^2) &= \exp\left\{ \mu_d + \phi x_t + A_{\tau-1} - A_\tau + \left(B_{\tau-1} \rho - B_\tau \right) x_t + C_{\tau-1} \sigma^2 (1-\nu_1) + \left(C_{\tau-1} \nu_1 - C_\tau \right) \sigma_t^2 \right\} \\ &= R_{ft} \exp\left\{ -\frac{1}{2} C_{\tau-1}^2 \sigma_w^2 - \frac{1}{2} \left(\varphi_d^2 + B_{\tau-1}^2 \varphi_e^2 \right) \sigma_t^2 + \lambda_{m,w} C_{\tau-1} \sigma_w^2 + \lambda_{m,e} B_{\tau-1} \varphi_e \sigma_t^2 \right\}. \end{aligned}$$

Exploiting lognormality,

$$\mathrm{Var}_t^* R_{d,t+1} = R_{ft}^2 \left(\exp\left\{ \mathrm{Var}_t \log R_{d,t+1}(\tau) \right\} - 1 \right) = R_{ft}^2 \left(\exp\left\{ \left(\varphi_d^2 + B_{\tau-1}^2 \varphi_e^2 \right) \sigma_t^2 + C_{\tau-1}^2 \sigma_w^2 \right\} - 1 \right).$$

For the risk premium, we have

$$\mathbb{E}_t \left(R_{d,t+1}(\tau) \right) - R_{ft} = R_{ft} \left(\exp\{ \lambda_{m,w} C_{\tau-1} \sigma_w^2 + \lambda_{m,e} B_{\tau-1} \varphi_e \sigma_t^2 \} - 1 \right).$$

This implies for the covariance

$$\begin{aligned}\mathrm{Cov}_t\left(M_{t+1}R_{d,t+1}(\tau),R_{d,t+1}(\tau)\right) &= R_{ft}(\exp\{\left(\varphi_d^2+B_{\tau-1}^2\varphi_e^2\right)\sigma_t^2+C_{\tau-1}^2\sigma_w^2\} \\ &\quad -\exp\{\lambda_{m,w}C_{\tau-1}\sigma_w^2+\lambda_{m,e}B_{\tau-1}\varphi_e\sigma_t^2\}),\end{aligned}$$

which we linearize (and approximating $R_{ft}=1$ as it does not affect the sign and is the relevant empirical case)

$$\mathrm{Cov}_t\left(M_{t+1}R_{d,t+1}(\tau),R_{d,t+1}(\tau)\right) \simeq \varphi_d^2\sigma_t^2+B_{\tau-1}\varphi_e\left(B_{\tau-1}\varphi_e-\lambda_{m,e}\right)\sigma_t^2+C_{\tau-1}\left(C_{\tau-1}-\lambda_{m,w}\right)\sigma_w^2.$$

Note that $B_\tau, B'_\tau > 0$, $C_\tau, C'_\tau < 0$, for $\tau > 0$, and $\lambda_{m,e} > 0$ and $\lambda_{m,w} < 0$. At longer horizons, the NCC will be satisfied, but the coefficients (B_τ, C_τ) change only slowly with maturity due to the persistence of the processes. We therefore conclude that the NCC condition is likely not satisfied in Bansal and Yaron (2004) when calibrated to our sample period.

II. Additional figures

We summarize the broad patterns in the data as opposed to high-frequency event studies.[1] We therefore sample the data at three moments in time: the pre-pandemic peak of the market, the bottom of the market, and at the end of our sample. For both Europe and the US, we determine the peak of the index level before the start of the pandemic and compute the average prices and volatilities of each asset during the three week period before the peak. Similarly, we determine the bottom of the market indexes and average the prices and implied volatilities in the three weeks surrounding the bottom. We also average the prices during the last three weeks of our sample. To succinctly present the results, we present the returns averaged for the 2021, 2022, and 2023 dividend futures prices. The dividend prices are indicated by "ST" in the legend of the figure.

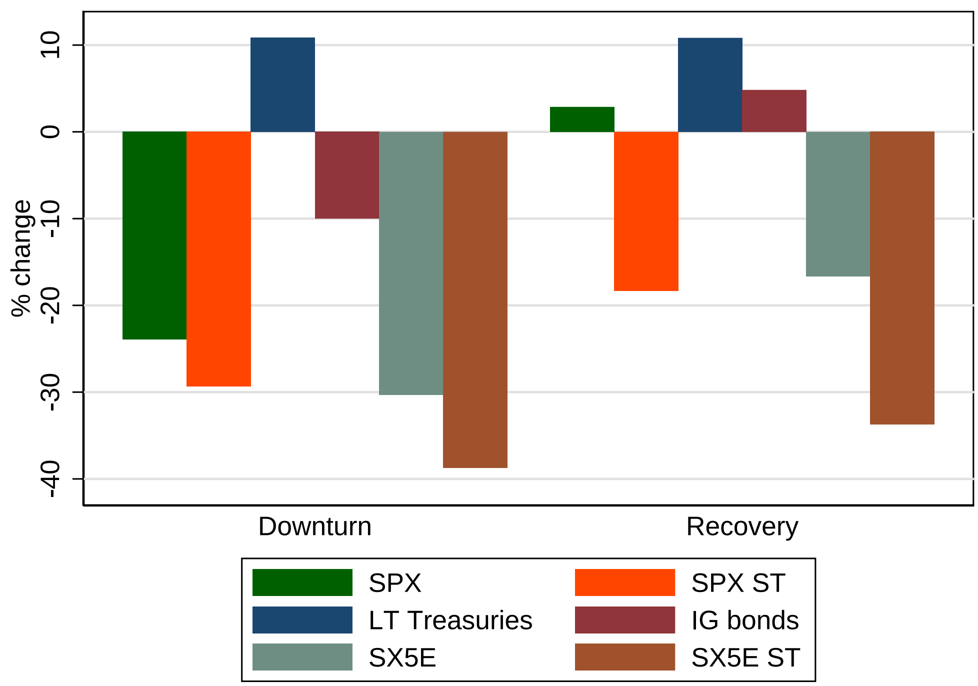

Figure A.2. The dynamics of asset prices during the COVID-19 crisis.

[1] Gormsen and Koijen (2020) analyze the dynamics of the index and dividend futures around some of the key events during the crisis for the European, Japanese, and US market.

REFERENCES

Ravi Bansal and Amir Yaron. Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance*, 59(4):1481–1509, 2004.

Ravi Bansal, Dana Kiku, and Amir Yaron. An empirical evaluation of the long-run risks model for asset prices. *Critical Finance Review*, 1:183–221, 2012.

Tim Bollerslev, George Tauchen, and Hao Zhou. Expected stock returns and variance risk premia. *Review of Financial Studies*, 22:4463–4493, 2009.

John Y. Campbell and John H. Cochrane. By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy*, 107(2):205–251, 1999.

Ian Dew-Becker, Stefano Giglio, Anh Le, and Marius Rodriguez. The price of variance risk. *Journal of Financial Economics*, 123(2):225 – 250, 2017. ISSN 0304-405X. doi: https://doi.org/10.1016/j.jfineco.2016.04.003.

Itamar Drechsler and Amir Yaron. What's vol got to do with it. *Review of Financial Studies*, 24:1–45, 2011.

J. D. Esary, F. Proschan, and D. W. Walkup. Association of random variables, with applications. *Annals of Mathematical Statistics*, 38(5):1466–1474, 1967.

Niels Joachim Gormsen and Ralph S J Koijen. Coronavirus: Impact on Stock Prices and Growth Expectations. *The Review of Asset Pricing Studies*, 10(4):574–597, 2020.

Ian Martin. What is the Expected Return on the Market? *The Quarterly Journal of Economics*, 132(1):367–433, 2017.

Loren D. Pitt. Positively correlated normal variables are associated. *Annals of Probability*, 10(2):496–499, 1982.

Jessica A. Wachter. Can time-varying risk of rare disasters explain aggregate stock market volatility? *Journal of Finance*, 68:987–1035, 2013.