A Disaster Explanation for the Term Structure of Returns

Di Wu

March 29, 2018

Abstract

Disasters affect the levels and growth rates of consumption, dividend and inflation. This paper shows how the disaster framework accounts for the term structure of returns in a wide range of asset classes, including (1) for nominal bonds, the term structure of returns is upward sloping while the term structure of Sharpe ratios is downward sloping, as found by Duffee (2010), (2) the term structures of returns and Sharpe ratios are both downward sloping for dividend strips, as found by van Binsbergen, Brandt, and Koijen (2012), and (3) fixing holding period to one month, the Sharpe ratio is significantly negative for one month variance forward, and close to zero for variance forward with maturity longer than two months, as found by Dew-Becker, Giglio, Le, and Rodriguez (2017), and (4) sorting countries by either short term interest rate or negative term premia, the returns to currency carry trades decrease as the maturity of the foreign bonds increases, as found by Lustig, Stathopoulos, and Verdelhan (2017).
I. Introduction

Using data from the liquid markets on dividend strips, nominal bonds, and variance forward of different maturities, a series of recent findings present challenges to the leading asset pricing models. In this paper, I focus on the holding period return of buying an \( n \)-maturity claim and selling it in the next period as an \((n-1)\)-maturity claim. The return of dividend strips appears to be downward sloping with respect to maturity (van Binsbergen, Brandt, and Koijen (2012)), while the return of nominal bond is known to be upward sloping. The Sharpe ratios of the returns are downward sloping across many asset classes (van Binsbergen and Koijen (2017)), yet evidence suggests a very negative Sharpe ratio for the one month variance forward, and close to zero for variance forward with maturity longer than two months (Ait-Sahalia, Karaman, and Mancini (2015); Dew-Becker, Giglio, Le, and Rodriguez (2017)).

Why is the downward sloping pattern challenging? We know that the excess volatility of the stock market requires all asset pricing models to feature time varying risk premium. Closely related to excess volatility, dividend-to-price ratio is shown to predict future return. The two findings can be jointly explained by modeling the stochastic discount factor as a function of persistent and mean-reverting state variable(s). This creates a “discount rate” effect formulated in Lettau and Wachter (2007). Namely, an unexpected shock to state variable changes current discount rate, but the mean-reversion property moves all future discount rates in the opposite direction, which accumulate and affect long maturity claim much more than short maturity claim. The negative co-movement with the stochastic discount factor implies a higher required holding period return for the long maturity claim. To generate a downward sloping holding periods return curve, a model needs to balance the “discount rate” effect with a transitory “cash flow” effect, meaning that the shock needs to affect the short term cash flow more than the long term.

This paper modifies the variable disaster model of Gabaix (2012) with respect to the “discount rate” and “cash flow” perspective. In Gabaix (2012), although the disaster happens with \( i.i.d. \) probability, its impact on consumption, dividend and inflation is time varying, hence the namesake. The first change that I make fixes the impact on consumption throughout time. As a result of the constant impact on consumption, the stochastic discount factor is only affected by whether or not disaster happens. Given the \( i.i.d. \) distribution, disaster in
$t + 1$ has no bearing on the stochastic discount factor afterwards so the “discount rate” effect is completely shut down.

I then model the impact on the cash flows based on the documented empirical regularities. Specifically, (1) in addition to consumption, dividend also significantly drops, (2) based on the findings of Gourio (2008); Nakamura et al. (2013), the growth rate of dividend jumps up, reflecting the “recovery” feature that is missing in Gabaix (2012), (3) based on the findings of Barro and Ursúa (2008), inflation on average jumps up and gradually reverts to normal. As in Gabaix (2012), I model the impact on the levels of dividend and inflation to be time varying, as is evident in the historical observation of disasters. This time variation doesn’t affect the stochastic discount factor per se, but is nevertheless priced because of its impact on the cash flow in the disaster state, or more fundamentally, the state-wise correlation between the cash flow and the marginal utility.

The difference in cash flow effects across asset classes leads to the difference in the term structure. I start with the $n$-maturity dividend strips, whose payoff is the dividend in time $t + n$. The disaster in time $t + 1$ is accompanied by a large, negative impact of the level of the dividend in $t + 1$. Thanks to the recovery feature that ameliorates the “cash flow” effect, dividend in time $t + n$ is only partially affected. As a result, the holding period return of the dividend strip is downward sloping.

A by-product of the drop in dividend is the high realized variance of the stock market, which is great news for investors with a long position in variance forward. The payoff of the $n$-month variance forward is the sum of squared daily log returns of index in month $t + n$, minus a fixed payment that is called the price. Given its i.i.d. distribution, the disaster in month $t + 1$ only significantly boosts the realized variance in $t + 1$. A secondary effect of this disaster is the recovery induced jump in the dividend growth rate, which mean-reverts and slightly increases the realized variance beyond month $t + 2$. As the cash flow effect is positive and largely contemporary, the expected holding period return is negative for the one month variance forward, and close to zero for the longer maturity claims.

The nominal bond is risk free in nominal terms. Regardless of the state, the $n$-maturity zero coupon bond pays out a nominal $1$ at time $t + n$. If a disaster happens at $t + 1$, inflation jumps up on average and persists for some time. The real value of $1$ in time $t + n$ is significantly reduced because of the persistently high inflation, which has a limited impact on the real value of $1$ in time $t + 1$. As a result, the “cash flow” effect implies that the term
structure curve of the holding period return slopes upward for nominal bonds.

In this framework, the effect of a disaster to the stochastic discount factor is \textit{i.i.d.} and the explanatory power comes from the different effects of disaster to cash flow, which is contemporary for variance forward, transitory for dividend, and accumulative for nominal bond. The two key state variables are resilience of dividend and jump size of inflation in the disaster. Both are time varying and mean reversion, hence my model inherits most of the properties of \textcite{Gabaix2012} and is broadly consistent with many findings on predictability within an asset class, such as the relationship between \textit{pd} and future stock return, and the relationship between \textcite{CochranePiazzesi2005} factor and future bond excess return, as well as those across stock and variance market, such as how the unspanned jump tail risk of \textcite{Andersen2017} predicts the future stock return. Because the two state variables are persistent, they affect the return of long maturity claims more than that of short maturity claims. The movement of the two state variables is independent of the movement of the stochastic discount factor, which creates excess volatility for the long maturity claims. This is why Sharpe ratios are generally downward sloping in all asset classes.

In Section IV, I apply the knowledge on nominal bond to the international setting with a global disaster, whose impact on consumption remains fixed in time, but varies across countries based on the distribution of disaster size documented by \textcite{BarroJin2011}. In this model, if a country’s consumption drops less in the disaster (a resilient country), its short term interest rate is high due to less precautionary saving. At the same time, the term premium is low since the jump of inflation, which is positive on average, also hurts less. Assuming the market is complete, the currency of a resilient country depreciates in the global disaster as the marginal utility is lower than the rest of the world. Carry trade with long term bond equals investing in an exchange of currency, risk-free short term bonds and risky slopes. In the cross section, the exchange trade and the slope trade has opposite exposure to global disaster, so the carry trade with long term bonds is not necessarily exposed to disaster risk anymore. The model quantitatively matches the findings of \textcite{LustigStathopoulosVerdelhan2017} that the holding period return to carry trade strategy diminishes when the investor switches from short maturity bonds to long maturity bonds.

My work is a continuation of the disaster literature which starts with the seminal work of \textcite{Rietz1988, Barro2006}, who show that disaster models match the unconditional equity risk premium. The second generation disaster models of \textcite{Gabaix2012, Gourio2012};
Wachter (2013) focus on the time variation of equity premium. In the second generation, the idea of disaster is applied to other asset classes, to the international market, and to the macro economy. This paper contributes to the third generation of disaster models as pioneered by Hasler and Marfe (2016), which take a microscopic view of risk premium by examining the claims to cash flow in one certain future period, such as dividend strip, zero coupon bond and variance forward. The empirical evidence poses further restrictions on asset pricing models, which is to not only match the level of risk premium, and also distribute it properly throughout the lifetime of the asset.

Indeed, the concept of post disaster recovery is first introduced in Gourio (2008); Nakamura et al. (2013) and later incorporated into the disaster model of Wachter (2013) by Hasler and Marfe (2016). In Hasler and Marfe (2016), the recovery of the dividend produces the downward yield curve of dividend strips, and the recovery of the consumption leads to the upward yield curve of real bonds due to the less precautionary motive for the future. Both Hasler and Marfe (2016) and my model generates downward sloping holding period return of dividend strips. Different from Gabaix (2012), the time variation of Wachter (2013) and Hasler and Marfe (2016) comes from the time varying disaster probability $p_{t+1}$ that strongly affects the stochastic discount factor. The comovement between the stochastic discount factor and the payoff of variance forward of all maturities, which is the result of the varying probability of disaster, implies a too flat term structure of holding period return that is inconsistent with the evidence of Dew-Becker, Giglio, Le, and Rodriguez (2017). On the bond part, Hasler and Marfe (2016) focuses on real bond while my work is on nominal bond. A nice feature in Hasler and Marfe (2016) is the counter cyclical slope of real bond yield curve due to varying disaster probability, consistent with Buraschi and Jiltsov (2007). In my model, disaster probability and its impact on consumption are both constant which lead to a flat real bond yield. Although both Hasler and Marfe (2016) and my model match upward sloping nominal bond yield curve, Duffee (2010); van Binsbergen and Koijen (2017) point out that the Sharpe ratio of holding period return of nominal bond is decreasing in maturity, which is difficult for Hasler and Marfe (2016). In addition, modeling nominal instead of real bond allows me to address the challenging fact of term structure of carry trade in Lustig, Stathopoulos, and Verdelhan (2017). My argument relies on the jump in inflation in the disaster state, which is absent in Hasler and Marfe (2016).

This paper also adds to the vibrant literature of term structure of returns, which is an
anatomy of equity premium puzzle that reveals the distribution of risk throughout the life of the stock. Because there are more ways than one to skin a cat, it is important to distinguish between holding period return and yield. The latter assumes holding the claim to maturity, which is immune from the “discount rate” effect. The two types of returns may have similar term structure for some asset classes. For example, equity yield (van Binsbergen et al. (2013) and holding period return (van Binsbergen and Koijen (2017)) are both downward sloping with respect to maturity. However, in the case of variance forward, holding period return of long maturity claim is close to zero whereas its return to maturity is significantly negative. For carry trade with long term bond, Zviadadze (2017) shows that its return to maturity is lower than short term carry trade but still significantly positive. Since the seminal work of van Binsbergen et al. (2012) and Giglio et al. (2015), many new models have successfully generated the downward sloping dividend yield curve. However, in light of the “discount rate” effect, not all models can explain the downward sloping holding period return pattern. In addition, some of the channels proposed in the literature apply exclusively to dividend, while van Binsbergen and Koijen (2017) uncover the downward sloping pattern in a wide range of asset classes.

In what follows, I start by solving the price for stock, bond, and variance in Section II. The solution method is the Linearity Generating (LG) toolbox developed by Gabaix (2007), with a short summary provided in the Appendix for its usage in discrete time. Section III compares the predictions of the model with empirical evidence. Section IV extends the model to n countries to address the term structure of carry trade.

II. Model Setup and Asset Prices

The utility function of the agent is CRRA with \( U_0 = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \frac{C_{t+1}^{1-\gamma}}{1-\gamma} \). The coefficient of relative risk aversion is \( \gamma \), and \( \beta \) governs time preference. Absent disaster, the real consumption process \( C_t \) follows \( \frac{C_{t+1}}{C_t} = e^{\mu_c} \).\(^1\) At each future period, disaster happens with the constant probability \( p \). When it does, consumption \( C_{t+1} \) falls to \( Be^{\mu_c}C_t \) with \( B \in (0, 1) \).

\(^1\)As discussed in Gabaix (2012), noise can be added to the end without changing stock price as long as the noise is uncorrelated with disaster.
Given the setup, the marginal utility $M_t = \beta^t C_t^{-\gamma}$ satisfies

$$\frac{M_{t+1}}{M_t} = e^{-\delta} \times \begin{cases} 1 & \text{No disaster in } t+1 \\ B^{-\gamma} & \text{If a disaster happens in } t+1 \end{cases} \tag{1}$$

where $\delta = -\log \beta + \gamma \mu_c$.

In this model, the agent has constant inter-temporal smoothing and precautionary saving motives, which implies a constant real interest rate

$$R_f = e^{\delta - \Psi} \tag{2}$$

where $\Psi = \ln (1 - p + pB^{-\gamma}) > 0$ measures the precautionary saving motive. With CRRA utility function, the elasticity of inter-temporal substitution $\psi = \frac{1}{\gamma}$ is low. Risk free rate needs to be unreasonably high for the agent to accept uneven consumption path that grows at $\mu_c \approx 2\%$. Disaster brings down the risk free rate thanks to the precautionary motive as short term bond provides high marginal utility in disaster state.

### A. Stock and Dividend Strip

In this model, stock is the claim to all future dividends $(D_{t+n})_{n>0}$, while $n$-year dividend strip is a claim to the dividend $D_{t+n}$\footnote{In reality, dividend strip pays out the sum of the declared ordinary gross dividends on index constituents that go ex-dividend between the third Friday of December of Year $t+n-1$ and the third Friday of December in Year $t+n$.}. I denote the price of stock as $P_t$ and the spot price of $n$-year dividend strip as $S^n_t$. Dividend is assumed to follow

$$\frac{D_{t+1}}{D_t} = e^{\mu_d} (1 + g_t) (1 + \sigma^D \epsilon^D_{t+1}) \times \begin{cases} 1 & \text{Disaster} \\ F_{t+1} & \text{No disaster} \end{cases} \tag{3}$$

I denote “resilience” as $H_t = p^E_{D_t} [(B^{-\gamma} F_{t+1} - 1) (1 + g_t)]$ and separate $H_t$ into $H_t = \hat{H} + \bar{H}_t$. A twist is added to the AR-coefficients to derive an exact formula for stock price.
Figure 1: **Recovery path of dividend following disaster.** The solid line is the average dividend growth path with no disaster. The dash line is with disaster at time 0. Parameters used here are $F = 0.35$, $g^D = 15\%$ and $\phi_g = 0.15$. Upon impact, dividend falls to 35\% of pre-disaster level and recovers fully in roughly 20 years.

Specifically, I assume that the process $g_t$ and $H_t$ follows

$$g_{t+1} = \begin{cases} 
\frac{1}{1+g_t}e^{-\phi_g g_t} & \text{No Disaster} \\
\frac{1}{1+g_t}e^{-\phi g_t} & \text{Disaster}
\end{cases}$$

$$H_{t+1} = \frac{1 + H_t}{1 + H_t + g_t}e^{-\phi_H H_t + \sigma^H \epsilon_{t+1}}$$

The jump of $g_t$ to $g^D$ reflects the recovery of dividend after disaster.

In Figure 1, I plot the typical recovery path following disaster, with the annual calibration of $F = 0.35$, $g^D = 15\%$ and $\phi_g = 0.15$ which I’ll elaborate in Section III.A. Disaster hits at time 0. In about 20 years, dividend will recover to the pre disaster level.
The spot price of dividend strip as well stock price can be solved explicitly.

**Proposition 1.** The (ex-dividend) price-dividend ratio of stock is

\[
P_t = \left( 1 \ 0 \ 0 \right) \Omega_D \left( e^{\delta - \mu_d - h} I_3 - \Omega_D \right)^{-1} \left( 1, g_t, \hat{H}_t \right)'
\]

The spot price of dividend strip is

\[
S_t^{(n)} = D_t \left( 1 \ 0 \ 0 \right) e^{-n(\delta - \mu_d - h)} \Omega_D^n \left( 1, g_t, \hat{H}_t \right)'
\]

where \( I_3 \) is the identity matrix of size 3 and

\[
\Omega_D = \begin{pmatrix}
1 & e^{-h} & e^{-h} \\
ge^{D} \left[1 - (1 - p) e^{-h} \right] & e^{-h} \left[ p e^{D} + (1 - p) e^{-\phi_s} \right] & e^{-h} g^{D} \\
0 & 0 & e^{-\phi_H}
\end{pmatrix}
\]

with \( h = \ln \left(1 + \bar{H} \right) \).

Although the topic of this paper is the holding period return, it is also useful to look at the dividend yield \(-\frac{1}{n} \ln \frac{S_t^{(n)}(\hat{H}_t)}{D_t}\) to have a better understanding of the result. In Figure 2, I plot the shape of dividend yield curve when \( \hat{H}_t = 0.3 pB^{-\gamma}, \hat{H}_t = 0 \) and \( \hat{H}_t = -0.3 pB^{-\gamma} \). In the first case, dividend falls to \( \bar{F} + 0.3 \) in disaster. Although the yield curve is downward sloping on average, the model predicts an upward sloping yield curve in some periods when dividend is resilient.

In reality, we don’t observe resilience \( \hat{H}_t \). However, we know from Equation (6) that the price to dividend ratio is higher when resilience \( \hat{H}_t > 0 \), which means “good times”. If disaster hits in good times, the damage is smaller. As resilience \( \hat{H}_t \) mean-reverts to 0, long maturity claims is riskier than short maturity claims. Figure 2 compares this implication with Figure 1 of van Binsbergen et al. (2013), from which we can observe a pro cyclical slope of dividend yield curve.

To gain further insight into Proposition 1, I define

\[
(PD_0, PD_g, PD_H) \equiv (1, 0, 0) \Omega_D \left( e^{\delta - \mu_d - h} I_3 - \Omega_D \right)^{-1}
\]
Figure 2: Model implied (Top) and empirical dividend yield curve (Bottom). Top left is dividend yield curve in “good times” when $\hat{H}_t = 0.3pB^{-\gamma}$, and the top right is when resilience $\hat{H}_t = -0.3pB^{-\gamma}$. Dividend yield is calculated as $-\frac{\ln S^n_t (H_t)}{n}$ using the calibration of Table I. Bottom is Figure 1 of van Binsbergen et al. (2013) that plots the time series of forward equity yields (dividend yield minus risk free rate) for dividend strips with maturities 1, 2, 5 and 7 years.
so that \( PD_0 \) is the mean price-dividend ratio, and \( PD_g \) and \( PD_H \) are the loading of price-dividend ratio on \( g_t \) and \( \hat{H}_t \). With some algebra, we have

\[
\frac{1}{PD_0} = e^{\delta - h - \mu_d} \left[ e^{\delta - \mu_d} - (1 - p) e^{-\phi_d} - pgD \right] - 1 \\
\frac{PD_g}{PD_0} = e^{\delta - h - \mu_d} \left( e^{\delta - \mu_d} - (1 - p) (e^{-\phi_d} - (1 - e^{-h})gD) \right) \\
\frac{PD_H}{PD_0} = \left( e^{\delta - \mu_d} - e^{h - \phi_d} \right) \left[e^{\delta - \mu_d} - (1 - p) (e^{-\phi_d} - (1 - e^{-h})gD) \right]
\]

When \( g^D = 0 \), the average dividend to price ratio simplifies to \( \frac{1}{PD_0} = e^{\delta - \mu_d - h} - 1 \) and expected stock return becomes \( \left( \frac{1}{PD_0} + 1 \right) e^{\mu_d} - 1 \approx \delta - h \), as in Gabaix (2012). With \( g^D > 0 \), the expression is more complicated so I keep the notation \( (PD_0, PD_g, PD_H) \) instead. The stock return conditional on no \( t + 1 \) disaster is

\[
\tilde{r}_{t+1}^{ND} \approx \mu_d + \frac{PD_g}{PD_0} \phi_d g_t - \frac{PD_H}{PD_0} \phi_H \hat{H}_t + \sigma_D^D \tilde{\epsilon}_{t+1}^D + \frac{PD_H}{PD_0} \sigma_H^G \tilde{\epsilon}_{t+1}^H
\]

in which the unexpected movement of resilience \( \frac{PD_H}{PD_0} \sigma_H^G \epsilon_{t+1}^H \) generates excessive volatility. A high \( \epsilon_{t+1}^H \) increases price-dividend ratio and lowers the future expected return, allowing for predictability in this model. Using Equation (8), I move on to calculate the price of variance forward.

### B. Variance Forward

The payoff of \( n \)-month variance forward is the sum of squared daily log returns of stock in month \( t + n \), minus a fixed payment that is called the variance forward price.

In disaster, a large negative return is realized, which makes variance forward a good hedge against disaster risk. As a result, variance forward at any maturity is more expensive than realized variance. A secondary effect of disaster is the jump of \( g_t \) to \( g^D \), which quickly reverts to 0 and boosts realized variance in all future months. We can intuitively think of variance forward as insurance against disasters: \( n \) month variance forward provides excellent insurance against the disaster in month \( t + n \), as well as some insurance against disasters
before that. As a result, there is a large difference between realized volatility and one-month forward price, followed by small increase from one-month to two-month forward price.

In Figure 3 I plot the simulated variance forward price $F_t^n$ against maturity $n = 1, 2, \ldots, 12$ months, evaluated at $\hat{H}_t = 0$. The details of simulation is also in Appendix C where I also provide approximate expression for variance forward price. To help interpret the results, I transform $F_t^n$ into $100 \times \sqrt{12F_t^n}$ which is the implied annualized volatility in percentage terms, with the average annualized realized volatility at $n = 0$.

Figure 3: **Realized Volatility and Simulated Variance Forward Price.** Average realized volatility is plotted at month 0 followed by simulated variance forward price $F_t^n$, reported as $100 \times \sqrt{12F_t^n}$ for each maturity $n = 1, 2, \ldots, 12$ month.
C. Nominal Bond

I derive formula for stock and variance in real terms. This section is on nominal bond, which has no default risk but is exposed to inflation risk. The formulation is the same as Gabaix (2012) where inflation jumps during disaster, the size of which is time varying.

The real value of $1 at time $t$, denoted as $Q_t$, follows

$$Q_{t+1} = 1 - I_t.$$ 

Assume inflation $I_t = \bar{I} + \hat{I}_t$ follows twisted AR process in normal times plus a jump by $J_t$ in disaster.

$$\hat{I}_{t+1} = \frac{1 - \bar{I}}{1 - \bar{I} - \hat{I}_t} \left( e^{-\phi \hat{I}_t} \hat{I}_t + \begin{cases} 0 \\
J_t & \text{Disaster}
\end{cases} \right) + \sigma^I \epsilon^I_{t+1},$$

with the jump size $J_t = \bar{J} + \hat{J}_t$ itself following

$$\hat{J}_{t+1} = \frac{1 - \bar{J}}{1 - \bar{J} - \hat{J}_t} e^{-\phi \hat{J}_t} \hat{J}_t + \sigma^J \epsilon^J_{t+1}.$$ 

**Proposition 2.** The price of zero coupon bond can be calculated as

$$Z^n_t = \left( 1 0 0 \right) e^{-n(\delta + \pi - \Psi)} \Omega^n Z \left( 1 0 0 \right)^t,$$

where $\Omega_Z = \begin{pmatrix} 1 & -e^\pi & 0 \\
\bar{J} \left( 1 - \frac{1-p}{e^\psi} \right) & e^{-\phi \bar{J}} \left( 1 - \frac{1-p}{e^\psi} \right) & 0 \\
0 & 0 & e^{-\phi J} \end{pmatrix}$, $\pi = -\ln \left( 1 - \bar{I} \right)$, and $\Psi = \ln \left( 1 - p + pB^{-\gamma} \right)$.

In addition to the real risk free rate of Equation (2), I calculate the average nominal risk free rate from Equation (9)

$$R_{fNominal} = \frac{1}{\left( 1 0 0 \right) e^{-\delta + \pi - \Psi} \Omega_Z \left( 1 0 0 \right)} = e^{\delta + \pi - \Psi} \quad (10)$$

Now have all necessary information to calculate returns. For more on this, I turn to Section III.
III. Term structure of returns

With the formula of prices in hand, I can now derive the term structure of holding period return. Generally, the strategy is to buy \( n \) maturity claim at time \( t \) and sell it as \( n - 1 \) maturity claim at time \( t - 1 \). To calculate the Sharpe ratio, I use the real risk free rate of Equation (2) for dividend and variance which are derived in real terms, while using the nominal risk free rate of Equation (10) for nominal bond.

In addition, I’ll be evaluating everything at \( \hat{H}_t = 0 \), \( \hat{I}_t = 0 \) and \( \hat{J}_t = 0 \), so what I’m deriving is actually “conditional” Sharpe ratios unless otherwise noted. Conditional on no disaster at time \( t + 1 \) and no prior disaster, the dividend follows

\[
\frac{D_{t+1}}{D_t} = e^{\mu_d} (1 + \sigma^D \epsilon^D_{t+1})
\]

and three state variables follow

\[
\begin{align*}
\hat{H}_{t+1} &= \frac{1 + \tilde{H}}{1 + \hat{H}_t} e^{-\phi_H} \hat{H}_t + \sigma_H \epsilon_H^{H_{t+1}} \\
\hat{I}_{t+1} &= \frac{1 - \tilde{I}}{1 - \hat{I}_t} e^{-\phi_I} \hat{I}_t + \sigma_I \epsilon_I^{I_{t+1}} \\
\hat{J}_{t+1} &= \frac{1 - \tilde{J}}{1 - \hat{J}_t} e^{-\phi_J} \hat{J}_t + \sigma_J \epsilon_J^{J_{t+1}}
\end{align*}
\]

A. Calibration

Table[1] reports my choice of parameter values. The agent has CRRA utility function with risk aversion \( \gamma = 4 \). In normal times, consumption growth rate is 2% per year, while in disaster consumption falls by \( 1 - B = 30\% \). I adopt the same probability of disaster \( p = 3.63\% \) per year and calibrate time preference \( \beta \) so that annual real risk free rate \( r_f \) equals 1.2%.

The numbers on consumption are based on the empirical work of Barro and Ursúa (2008); Barro and Jin (2011). On the other hand, the impact of disaster on dividend is not so well documented. In Appendix E, I solve a similar model with neat expressions for the dividend yield, which is \( \delta - \mu_d - h \) for 1 year and \( \delta - \mu_d - h - \zeta \) for infinitely long maturity, where \( e^\zeta \geq 1 \) is parameterized by

\[
(e^\zeta - 1) (e^\zeta - e^{-\delta_y}) = [1 - (1 - p) e^{-h}] e^{-h} g^D
\]
with $\delta = -\log \beta + \gamma \mu_c$, $h = \ln (1 + \bar{H})$ and $\delta_g = \phi_g + h - \ln (1 - p)$. I choose to set $\bar{F} = h/2 = 0.35$ which implies that the dividend yield for one year claim is $\delta - \mu_d - \ln (1 + p (B^{-\gamma \bar{F}} - 1)) \approx 8.4\%$. For the long maturity yield, both the jump size $g_D$ and its reversion $\phi_g$ matter. A combination of $g_D = 0.15$ and $\phi_g = 0.15$ leads to a slope of $\zeta \approx 3.4\%$ which means that the long maturity yield is about $5\%$. In Figure 1, dividend recover to pre disaster level after 20 years, which makes intuitive sense. At this point, the parameters already determine the average price to dividend ratio ($\approx 17$) and equity premium ($\approx 6.6\%$).

For the time varying features, the key variable is $\phi_H$ which governs the mean reversion of resilience $H$, as well as that of price to dividend ratio. As argued by Lettau and Van Nieuwerburgh (2007), the persistence is much lower when allowing for structural break(s) in the time series. I set $\phi_H = 0.3$ which implies a first order auto-correlation of $e^{-\phi_H} = 0.74$. In light of recent empirical evidence, I set the volatility of dividend growth rate to be $\sigma_D = 18\%$. I choose the volatility of resilience $\sigma_F = 9\%$ to help me match the volatility of stock return.

Finally there are parameters for nominal bond. I model nominal bond separately so the parameter choices do not affect the results in stock and variance. The results are robust to different choice of parameters, except that the average jump of inflation $J^D$ has to be positive. I set $J^D = 4\%$ so inflation jump in disaster is $4\%$ on average. We observe high Sharpe ratio for holding short maturity bond. I achieve it by choosing a relatively low volatility of inflation $\sigma_I = 1\%$. At the same time, the Sharpe ratio declines quickly as maturity increases. To match this pattern, I choose a high volatility for the jump $\sigma_J = 5\%$ so that long maturity bond has “excess volatility” that is uncorrelated with pricing kernel. My choice of mean reversion $\phi_I = 0.4$ is a bit high yet recent evidence does suggest that inflation is not persistent. Historically, the size of disaster induced inflation jumps is highly volatile so I capture it by choosing $\phi_J = 0.8$. The mean inflation $\bar{I}$ is not relevant for excess return. I set $\bar{I} = 3\%$ to be consistent with historical average.

\textbf{B. Dividend Strips}

Since the dividend process is modeled in terms of real value, I subtract the real risk free rate to get excess return. For $n = 1$, excess return equals to yield plus dividend growth rate
Table I: Values of the Parameters
Annual values are used for stock and bond whereas monthly values are used for variance swap. The third column compares the numbers with Gabaix (2012) original calibration.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Annual</th>
<th>Monthly</th>
<th>Gabaix (2012)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preference</td>
<td>γ = 4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Risk Aversion</td>
<td>B = 0.7</td>
<td>0.7</td>
<td>0.66</td>
</tr>
<tr>
<td>Cons Resilience</td>
<td>p = 0.0363</td>
<td>p/12</td>
<td>0.0363</td>
</tr>
<tr>
<td>Disaster Probability</td>
<td>µc = 0.02</td>
<td>µc/12</td>
<td>0.025</td>
</tr>
<tr>
<td>Real Rf</td>
<td>rf = 0.012</td>
<td>rf/12</td>
<td>0.012</td>
</tr>
<tr>
<td>Calibrate β</td>
<td>( β = \exp(γµc−rf)) (/(pB^{−γ}+1−p))</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Dividend \( D_t \) Process**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Annual</th>
<th>Monthly</th>
<th>Gabaix (2012)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resilience for ( D )</td>
<td>( F = B/2 )</td>
<td>( B/2 )</td>
<td>0.66</td>
</tr>
<tr>
<td>Mean Growth</td>
<td>( µ_d = 0.02 )</td>
<td>( µ_d/12 )</td>
<td>0.025</td>
</tr>
<tr>
<td>Volatility</td>
<td>( σ_d = 0.18 )</td>
<td>( σ_d/\sqrt{12} )</td>
<td>0.11</td>
</tr>
<tr>
<td>Recovery</td>
<td>( g_D = 0.15 )</td>
<td>( g_D/12 )</td>
<td>N.A.</td>
</tr>
<tr>
<td>Mean Reversion</td>
<td>( φ_d = 0.15 )</td>
<td>( φ_d/12 )</td>
<td>N.A.</td>
</tr>
</tbody>
</table>

**Resilience \( H_t \) process**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Annual</th>
<th>Monthly</th>
<th>Gabaix (2012)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>( H = p(B^{−γ}F − 1) )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Volatility of ( F )</td>
<td>( σ_F = 0.09 )</td>
<td>( σ_F/\sqrt{12} )</td>
<td>0.1</td>
</tr>
<tr>
<td>Volatility of ( H )</td>
<td>( σ_H = pB^{−γ}σ_F )</td>
<td>( σ_H/12\sqrt{12} )</td>
<td>( pB^{−γ}σ_F )</td>
</tr>
<tr>
<td>Mean Reversion</td>
<td>( φ_H = 0.3 )</td>
<td>( φ_H/12 )</td>
<td>0.13</td>
</tr>
</tbody>
</table>

**Inflation \( I_t \) and \( J_t \) processes**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Annual</th>
<th>Monthly</th>
<th>Gabaix (2012)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>( I = 0.03 )</td>
<td>( I/12 )</td>
<td>0.037</td>
</tr>
<tr>
<td>Volatility</td>
<td>( σ_I = 0.01 )</td>
<td>( σ_I/\sqrt{12} )</td>
<td>0.015</td>
</tr>
<tr>
<td>Mean Reversion</td>
<td>( φ_I = 0.4 )</td>
<td>( φ_I/12 )</td>
<td>0.18</td>
</tr>
<tr>
<td>Mean Jump</td>
<td>( J^D = 0.04 )</td>
<td>( J^D/12 )</td>
<td>0.021</td>
</tr>
<tr>
<td>Volatility</td>
<td>( σ_J = 0.05 )</td>
<td>( σ_J/\sqrt{12} )</td>
<td>0.015</td>
</tr>
<tr>
<td>Mean Reversion</td>
<td>( φ_J = 0.8 )</td>
<td>( φ_J/12 )</td>
<td>0.92</td>
</tr>
</tbody>
</table>
minus risk free rate.

A few notations here are useful to keep the expressions simple. The spot price of dividend is linear in the state variable, so I define

\[ s_{n,0} \equiv \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} e^{-n\delta_D} \Omega^0_D (1, 0, 0)' \]

and

\[ s_{n,H} \equiv \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} e^{-n\delta_D} \Omega^0_D (0, 0, 1)' \]

where \( s_{n,0} \) is the constant term and \( s_{n,H} \) is the coefficient before \( \hat{H}_t \). When evaluated at \( \hat{H}_t = 0 \) and conditional on no disaster before and after, the excess return can be written as

\[
\begin{align*}
\mathbf{r}_n^S &= \mu_d \frac{s_{n-1,0}}{s_{n,0}} R_f + \mu_d \frac{s_{n-1,0}}{s_{n,0}} \sigma_D \epsilon_{t+1}^D + \mu_d \frac{s_{n-1,H}}{s_{n,0}} \sigma_H \epsilon_{t+1}^H \tag{11}
\end{align*}
\]

At the same time, the excess return of stock is

\[
\mathbf{r}_{t+1}^{ND} \approx 1 + \mu_d + \frac{1}{P_D 0} R_f + \sigma_D \epsilon_{t+1}^D + \frac{P_D H}{P_D 0} \sigma_H \epsilon_{t+1}^H
\]

Just like stock, the return of dividend strip displays a pattern of excess volatility unjustified by volatility of underlying dividend. Similar to the stock, the excess volatility is driven by resilience \( H_t \).

It is straightforward to derive Sharpe ratio. We have two shocks at the same time: dividend and resilience. From [11] the Sharpe ratio can be calculated as

\[
SR_n^S = \frac{s_{n-1,0} - e^{-\mu_d s_{n,0} R_f}}{\sqrt{(s_{n-1,0} \sigma_D)^2 + (s_{n-1,H} \sigma_H)^2}}
\]

In Figure [4] I plot the annual excess return, volatility and Sharpe ratio of the strategy. It is clear that both return and Sharpe ratio are downward sloping in maturity. The shape of volatility is less clear. From Equation [11], the volatility has two parts. As a property of the model, the \( e^{\mu_d \frac{s_{n-1,0}}{s_{n,0}} \sigma_D} \) part is linear in expected gross return and decreasing in \( n \), whereas second part \( e^{\mu_d \frac{s_{n-1,H}}{s_{n,0}} \sigma_H} \) is increasing in \( n \) as \( D_{t+n} \) is affected by resilience during.
Table II: Term Structure of Dividend Strip Return

I calculate the excess return, volatility and Sharpe ratio of the strategy that buys a n-year dividend strip and holds it for one year before selling it as (n-1)-year strip. Model is evaluated at mean resilience $H_t = 0$ assuming no disaster. Excess return is calculated over the real risk free rate of 1.2 percent. To compare the numbers to data, I annualize the numbers from Table 1 of van Binsbergen and Koijen (2017) (US:SPX Nov 2002-Jul 2014), which uses the same strategy except the holding period is one month. Column 1-5 are for dividend strips with different maturities and the last column is for stock.

<table>
<thead>
<tr>
<th>Maturity in years</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Stock</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excess return</td>
<td>0.098</td>
<td>0.089</td>
<td>0.083</td>
<td>0.077</td>
<td>0.073</td>
<td>0.067</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.200</td>
<td>0.199</td>
<td>0.199</td>
<td>0.200</td>
<td>0.201</td>
<td>0.190</td>
</tr>
<tr>
<td>Sharpe ratio</td>
<td>0.490</td>
<td>0.450</td>
<td>0.416</td>
<td>0.388</td>
<td>0.365</td>
<td>0.351</td>
</tr>
<tr>
<td>van Binsbergen and Koijen (2017)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excess return</td>
<td>0.055</td>
<td>0.071</td>
<td>0.080</td>
<td>0.086</td>
<td>0.101</td>
<td>0.037</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.115</td>
<td>0.140</td>
<td>0.167</td>
<td>0.172</td>
<td>0.178</td>
<td>0.189</td>
</tr>
<tr>
<td>Sharpe ratio</td>
<td>0.479</td>
<td>0.507</td>
<td>0.486</td>
<td>0.503</td>
<td>0.566</td>
<td>0.197</td>
</tr>
</tbody>
</table>

all potential disasters from $t + 1$ to $t + n$. Quantitatively, the theoretical volatility across all maturity is no more than a few basis points from 20%.

In Table II, I compare the values to data, which is annualized Table 1 of van Binsbergen and Koijen (2017) for dividend futures on S&P500 index from Nov 2002 to Jul 2014. During this period, return on US stock is relatively low, suggesting the average stock return 0.037 in the last column may not be a good benchmark. Instead, the model generates a 6.7% equity premium which is more consistent with historical evidence. In line with empirical evidence, return and Sharpe ratio of dividend strip is much higher than those of stock, even though evidence is less clear across maturities within dividend strips. Different from van Binsbergen et al. (2012), the model predicts a virtually flat curve for volatility.

C. Variance Forward

I consider the strategy of buying n-month variance forward and sell it one month later as an (n-1)-month forward, with the convention that 0-month forward being the realized variance.
Figure 4: **Property of dividend strategy** that buys $n$ period claim and sells it next period as $n - 1$ claim. The periods are in years. From top to bottom are excess return over real risk free rate, volatility and Sharpe ratio of excess return.
The evidence presented in Dew-Becker et al. (2017) challenges many asset pricing models: Sharpe ratio of the strategy is as low as $-1.5$ for $n = 1$, and falls to virtually zero for $n > 2$, which suggests that “only unexpected, transitory realized variance was significantly priced”. Similar pattern are found in other strategies that are exposed to variance risk, e.g. shorting a straddle in Andries et al. (2015) and van Binsbergen and Koijen (2017), in both aggregate stock market and single-names, e.g. Kelly et al. (2016). The result on the holding period return is robust in early years, as well as in more recent years when VIX futures become available.

Generally speaking, for a model to work here, the correlation between SDF and realized variance must fall rapidly as maturity increases, which means the correlation cannot come from a persistent process, being either disaster probability, long run risk or consumption habit. As explained in Dew-Becker et al. (2017), the variable disaster model of Gabaix (2012) can match most of the features. Adding recovery to Gabaix (2012) creates a mild positive correlation between realized variance at $t+2$ and disaster at $t+1$. This is because the mean reversion of the recovery process adds to the realized variance. The positive correlation predicts a negative expected return for two month variance forward.

It is inconvenient to write the expression for the moments of return because the price is no longer linear in state variables. Instead, I plot the simulated distribution of Sharpe ratio in Figure 5.

Since I base my calibration on the facts in dividend strips, it is comforting to see that my model predicts economically significant one month variance risk premium while respecting the moments of stock return. The immediate impact of disaster on dividend is larger in my calibration than in Gabaix (2012), leading to a more negative realized return upon disaster.

As I have pointed out, my model implies negative holding period return even for two month variance forward as clear in Figure 5 albeit smaller in magnitude. Even though there is some overlapping between my confidence interval and that of Dew-Becker et al. (2017), I consider my predictions inconsistent with the claim that the Sharpe ratio of longer maturity claims is zero. At the same time, I conclude that the predicted Sharpe ratio of holding period return is significantly lower for long maturity variance forward than for one month variance forward.
Figure 5: Theoretical and empirical Sharpe ratios of variance forward of different maturity. Top is the result from 400 simulations of the model, each of which starts at $\hat{H}_t = 0$ and contains 204 months to match [Dew-Becker et al. (2017)] in the bottom. In both graphs, the solid line is the mean and dashed lines are 95% confidence intervals.
Table III: **Term Structure of Variance Forward Return**

I report the first to fourth moments of return across simulations, and compare them to the empirical evidence of Dew-Becker et al. (2017). Returns are monthly and expressed in percentage points. The last panel is annualized Sharpe ratio provided by Itamar Drechsler using post crisis sample of 6/2009 - 9/2015.

<table>
<thead>
<tr>
<th>Maturity in Months</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>9</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Simulation of my model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>-46.5</td>
<td>-1.3</td>
<td>-1.3</td>
<td>-1.1</td>
<td>-1.2</td>
<td>-1.1</td>
<td>-0.9</td>
<td>-1.1</td>
</tr>
<tr>
<td>Std</td>
<td>79.7</td>
<td>7.8</td>
<td>7.7</td>
<td>7.5</td>
<td>7.4</td>
<td>7.3</td>
<td>6.9</td>
<td>6.5</td>
</tr>
<tr>
<td>Skew</td>
<td>3.3</td>
<td>0.5</td>
<td>0.5</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>Exc. Kurt.</td>
<td>16.9</td>
<td>2.0</td>
<td>1.9</td>
<td>1.9</td>
<td>1.9</td>
<td>1.9</td>
<td>1.8</td>
<td>1.8</td>
</tr>
<tr>
<td>Annualized SR</td>
<td>-2.10</td>
<td>-0.57</td>
<td>-0.55</td>
<td>-0.51</td>
<td>-0.56</td>
<td>-0.48</td>
<td>-0.44</td>
<td>-0.59</td>
</tr>
</tbody>
</table>

| **Dew-Becker et al. (2017) Table 2** |       |       |       |       |       |       |       |       |
| Mean               | -25.7 | -5.8  | 0.7   | 0.6   | 0.1   | 0.5   | 0.9   | 1.8   |
| Std                | 67.9  | 47.7  | 33.9  | 27.4  | 22.5  | 19.6  | 16.2  | 17.4  |
| Skew               | 6.2   | 3.9   | 2.7   | 2.0   | 1.6   | 1.3   | 0.9   | 1.0   |
| Exc. Kurt.         | 56.5  | 23.4  | 14.1  | 7.6   | 5.2   | 3.3   | 1.7   | 1.4   |
| Annualized SR      | -1.82 | -0.43 | -0.36 | -0.24 |

For more details, I report the first to fourth moments of return across the simulations, and compare them to the empirical evidence in Table III.

The model broadly matches the pattern that excess return is small except for the very negative return of the one month claim. The one month return is skewed and has fat tails because the realized variance is square of a normal distribution. The model does not match every aspect of the data. For example, the standard deviation of return drops too fast, and the Sharpe ratio doesn’t drop fast enough. To further development of this model and match the two patterns, it is possible to specify a heteroskedastic process for the log dividend growth, whose volatility is a constant $\sigma^D$ for now.
D. Nominal Bond

Similar as above, I define \((z_{n,0} z_{n,I} z_{n,J}) \equiv \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} e^{-n(\delta+\pi-\Psi)} \Omega^n_Z\). The excess return of buying \(n\) period bond and sell next period as \(n-1\) period bond can be written as

\[
 r^n_Z = \frac{z_{n-1,0}}{z_{n,0}} - R^\text{Nominal}_f + \frac{z_{n-1,I}}{z_{n,0}} \sigma_I \epsilon_{t+1}^I + \frac{z_{n-1,J}}{z_{n,0}} \sigma_J \epsilon_{t+1}^J
\]

and the Sharpe ratio is

\[
 SR^n_Z = \frac{z_{n-1,0} - z_{n,0} R^\text{Nominal}_f}{\sqrt{(z_{n-1,I} \sigma_I)^2 + (z_{n-1,J} \sigma_J)^2}}
\]

which is valid for \(n \geq 2\). I use \(n = 1\) as the risk free rate.

As shown in Figure 6, the return and volatility of the strategy is increasing in maturity, whereas the Sharpe ratio is decreasing in maturity. The downward sloping Sharpe ratio comes from the time varying jump size \(\hat{J}_t\) that is uncorrelated with SDF.

Table [V] compares the numbers with data. It is unfortunate that the Sharpe ratio is undefined for \(n = 1\). Empirically, short maturity Sharpe ratio is high, as pointed out in van Binsbergen and Koijen (2017). This fact is consistent with what the disaster model would predict. For short maturity bond in normal times, the volatility can be close to zero while the expected return is positive, reflecting the probability of disaster.

E. Comparison with other leading asset pricing models

It is natural to think if alterations can be made to other leading asset pricing models to match the facts of holding period return. In this section, I analyze the disaster model of Wachter (2013), the habit model of Campbell and Cochrane (1999), the behavioral model of Barberis, Huang, and Santos (2001) and the long run risk model of Bansal and Yaron (2004) in terms of the “discount rate” and “cash flow” effects.

The varying disaster probability model of Wachter (2013) and Hasler and Marfe (2016) already shows great promise in dividends and bonds. As explained earlier, the challenge is
Figure 6: **Sharpe ratio for nominal bond.** The strategy is to buy $n$ period zero coupon bond and sell the next period as $n - 1$ period bond. From top to bottom are excess returns over real risk free rate, volatility and Sharpe ratios of excess returns, which are calculated over the nominal risk free rate of $n = 1$. 
Table IV: Term Structure of Nominal Bond Return

I calculate the excess return, volatility and Sharpe ratio of the strategy that buys a n-year zero coupon bond and holds it for one year before selling it as (n-1)-year bond. I evaluate the model at mean inflation $\hat{I}_t = 0$ and mean inflation jump $\hat{J}_t = 0$ assuming no previous disaster. Excess return is calculated over nominal risk free rate. To compare the numbers to data, I attach Table 4 of van Binsbergen and Koijen (2017), which is annualized monthly return over one month risk free rate.

<table>
<thead>
<tr>
<th>Maturity in years</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excess return</td>
<td>N.A.</td>
<td>0.59%</td>
<td>0.99%</td>
<td>1.27%</td>
<td>1.46%</td>
<td>1.83%</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>1.08%</td>
<td>1.97%</td>
<td>2.83%</td>
<td>3.56%</td>
<td>5.29%</td>
<td></td>
</tr>
<tr>
<td>Sharpe ratio</td>
<td>0.54</td>
<td>0.50</td>
<td>0.45</td>
<td>0.41</td>
<td>0.35</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Maturity in months</th>
<th>1-12</th>
<th>13-24</th>
<th>25-36</th>
<th>37-48</th>
<th>49-60</th>
<th>61-120</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excess return</td>
<td>0.58%</td>
<td>1.03%</td>
<td>1.36%</td>
<td>1.56%</td>
<td>1.56%</td>
<td>1.83%</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.80%</td>
<td>2.05%</td>
<td>3.13%</td>
<td>3.95%</td>
<td>4.67%</td>
<td>5.76%</td>
</tr>
<tr>
<td>Sharpe ratio</td>
<td>0.73</td>
<td>0.50</td>
<td>0.43</td>
<td>0.40</td>
<td>0.33</td>
<td>0.32</td>
</tr>
</tbody>
</table>

the variance forward pattern of Dew-Becker, Giglio, Le, and Rodriguez (2017), as the varying probability of disaster creates too strong positive correlation between the expected payoff of variance forward and stochastic discount factor. One possible solution is to re-calibrate the model featuring a less persistent probability process, which has two consequences. On the “cash flow” effect, the time payoff of long term variance forward becomes less positively correlated with marginal utility at $t + 1$. On the “discount rate” effect, low persistent leads to fast mean reversion, which creates a more negative correlation between pricing kernel at $t + 1$ and marginal utility at $t + 1$. Both channels suggest that the holding period return of long term variance forward should have a lower Sharpe ratio with less persistent probability of disaster.

The time variation of Campbell and Cochrane (1999) habit model comes from the time varying consumption habit, which is a function of past consumption. Positive shock to consumption increases surplus consumption and reduces effective risk aversion, which has more pronounced “discount rate” effects for long maturity claims. On the other hand, positive shock to consumption is correlated with positive shock $w_{t+1}$ to dividend growth $\Delta d_{t+1} = g + w_{t+1}$, which has the same permanent “cash flow” effect on both short term and long term claims. As a result, the habit model implies a positive and upward sloping curve.
for the holding period return of dividend strip. Though the issue can be fixed (e.g. by adding a mean-reverting process into dividend growth, and/or reducing persistence of habit as in Lynch and Randall (2011)), it remains an open question whether habit model can quantitatively match dividend yield and volatility pattern.

In the habit model, the effective risk aversion $\gamma S_t$ is high as the surplus consumption ratio $S_t$ is lower than one. Barberis et al. (2001) achieve similar effects by incorporating the concept of loss aversion, which suggests that investors feel more strongly about loss than about gain, relative to a reference level. Over time, the real value of dividend grows away from the reference level and becomes less risky, whereas the real value of $1$ shrinks towards or past the reference level. As a result, the dividend yield curve is downward sloping while the nominal bond yield curve is upward sloping from the view of a “prospective theory” investor with a constant reference level. However, the curve for holding period return may look different from yield curve. To address the excess volatility, it is crucial for the reference level to vary over time, which exposes Barberis et al. (2001) to the same “discount rate” effect as the habit model. The net effect of the two channels is not clear and I look forward to seeing a quantitative assessment of its predictions.

The long run risk model of Bansal and Yaron (2004) distinguishes short-run consumption from wealth through a small and persistent component in the consumption growth. A perhaps unwanted implication is that the same persistent long run components strongly affects the stochastic discount factor and the cash flow of long maturity claims, yet has close to zero effect on the cash flow of short maturity claims. As a result, holding period return and its Sharpe ratio generally increase with maturity for all asset classes in long run risk models, including Drechsler and Yaron (2011). Fundamental changes are needed, e.g. agent cannot observe the long run component but instead has to (incorrectly) infer from past consumption and dividend data as in Croce et al. (2014).

Building on the disaster framework of Gabaix (2012), my model is the only model that matches the observed patterns of holding period return in stock, bond and derivative market. And it doesn’t stop here. In Section LV, I showcase its versatility by taking it abroad to address the puzzling finding of Lustig, Stathopoulos, and Verdelhan (2017).
IV. Carry Trade

In this section, I examine the implication of my model in an environment with \( n \) countries. Each country has the same setting as the one country economy in Section II. To tie the countries together, I assume the disaster is now a global disaster, independent in each period with constant probability \( p \). The laws of motion for real consumption and inflation follows

\[
\frac{C_{i,t+1}}{C_{i,t}} = e^{\mu_{i,c}} \times \begin{cases} 
1 & \text{Disaster} \\
B_i & \text{Disaster}
\end{cases}
\]

(12)

\[
\hat{I}_{i,t+1} = \frac{1}{1 - \bar{I}_i - \hat{I}_{i,t}} \left( e^{-\phi I} \hat{I}_{i,t} + \begin{cases} 
0 & \text{Disaster} \\
J_i & \text{Disaster}
\end{cases} \right) + \sigma I \epsilon_{i,t+1}
\]

(13)

When the global disaster hits, consumption in country \( i \) falls to \( B_i \) of pre-disaster level. The resilience \( B_i \) differs across country but is fixed in the time dimension. The short rate \( R_i^{(1)} \) is given by

\[
R_i^{(1)} = e^{\delta_i} - \Psi_i = 1 - p + pB_i^\gamma.
\]

In the cross section, a high resilience reduces precautionary saving motive, leading to a high short rate.

The second effect of disaster is on average positive jump in inflation by the size \( \bar{J}_i \). The time variation of jump size is not necessary for the intuition of this section so I shut it down to keep the expression simply, i.e. \( \hat{J}_{i,t} = 0 \) for all \( i \) and \( t \). For a country with high resilience, the jump of inflation during disaster hurts less, implying a more “flat” yield curve.

I study a typical foreign country \( i \) against base country 0 (e.g. US) and define the nominal exchange rate \( S_{i,t} \) as foreign currency per base currency. An increase in \( S_{i,t} \) is an appreciation of base currency. In a complete market, the following equation holds,

\[
\frac{S_{i,t+1}}{S_{i,t}} = \frac{M_{i,t+1}^{Nominal}}{M_{i,t}^{Nominal}} \times \frac{M_{0,t+1}^{Nominal}}{M_{0,t}^{Nominal}} = e^{-\delta_i} \left( 1 - \hat{I}_i - \hat{I}_{0,t} \right) \times \begin{cases} 
1 & \text{Disaster} \\
(B_0/B_i)^{-\gamma} & \text{Disaster}
\end{cases}
\]

(14)

In disaster, country with low resilience \( B_i \) realizes high \( \left( \frac{S_{i,t+1}}{S_{i,t}} \right)^{-1} \) (it is multiplied by

27
\((B_0/B_i)\gamma\), which means an appreciation of its currency as it can trade for higher marginal utility in disaster. This appreciation only happens in the period with global disaster.

I also define the carry trade as the activity of borrowing from the base country and investing in the foreign country for one period. The carry trade can be conducted with short term bonds with maturity of exactly one period. Ang and Chen (2010) show that the return of this short maturity carry trade can be predicted with foreign short rate and (negative) foreign bond term spread. Alternatively, investor can borrow \(n\)-period long term bond from base country, and invest in foreign \(n\)-period long term bond. In the next period, investor sell the foreign holding as \(n-1\)-period bond, and pays back the loan in base country which is now a \(n-1\)-period bond. Lustig, Stathopoulos, and Verdelhan (2017) show that neither foreign short rate nor term premium can predict the return this long maturity carry trade strategy. The finding poses a challenge for the leading models in international finance.

A. Term Structure of Carry Trade

Now I put all the pieces together and evaluate the model conditional on no disaster actually happening. This is in spirit of the finding of Burnside, Eichenbaum, Kleshchelski, and Rebelo (2010) that shows even partially hedged carry trade commands abnormal return, which can only be justified by the out-of-sample event with high marginal utility.

It is clear from (14) that the exchange rate is exposed to the disaster risk, whereas the short rate is obviously not. As a result, return to the short maturity carry trade compensates exposure to the disaster risk. In the cross section, the exposure is determined by the resilience, which itself is correlated with both short interest rate and (negative) term premium. This is why both short interest rate and (negative) term premium can be used to predict carry trade return.

For the long maturity carry trade of Lustig et al. (2017), we can think of investing in long term bond as investing in a risky exchange trade, a risk-free short maturity bond and a risky slope. The slope trade realizes less negative return for the more resilient countries. Suppose we invest in a country with high short rate, or high resilience \(B_i\), then in disaster, the strategy realizes relatively bad return from the exchange trade part together with relatively good return from the slope trade. I formally summarize the indeterminacy of the net effect in Proposition 3.
Proposition 3. Conditional on no disaster, the expected return of borrowing in short term bond (maturity \( n = 1 \)) from base country 0 and investing in short term bond in foreign country \( i \), is

\[
\mathbb{E}_t \left( \frac{S_{i,t+1}}{S_{i,t}} \right)^{-1} \frac{R^{(1)}_i}{R^{(1)}_0} = \frac{e^{-\Psi_i}}{e^{-\Psi_0}}
\]

The expected return of borrowing in long term bond (maturity \( n \to \infty \)) from base country 0 and investing in long term bond in foreign country \( i \) for one period, is

\[
\mathbb{E}_t \left( \frac{S_{i,t+1}}{S_{i,t}} \right)^{-1} \frac{R^{(\infty)}_i}{R^{(\infty)}_0} = \frac{e^{-\Psi_i + l_i}}{e^{-\Psi_0 + l_0}} \left[ \frac{1 + \frac{1-e^{-\Psi_i + e^{-\phi_l}(1-e^{\pi_i})}}{e^{-\Psi_i - I_{i,t}}}}{1 + \frac{1-e^{-l_0 + e^{-\phi_l}(1-e^{\pi_0})}}{e^{-\Psi_0 - I_{0,t}}}} \right]
\]

where \( e^{\Psi_i} = 1-p+pB_i^{-\gamma} \) and \( e^{-\pi_i} = 1-I_i \) as before, and \( e^{-l_i} \in \left( \frac{e^{-\phi_l}+1}{2}, 1 \right) \) is parameterized as

\[
(1 - e^{-l_i}) \left( e^{-l_i} - e^{-\phi_l} \right) = e^{\pi_i} \left[ 1 - (1-p)e^{-\Psi_i} \right] J_i
\]

I introduce \( l_i \) to obtain a clean formula for bond price \( Z_{i,t} \). As I show in the Appendix, when evaluated at \( \hat{I}_{i,t} = 0 \), the short term interest rate of country \( i \) is

\[
-\log Z^{(1)}_{i,t} \left( \hat{I}_{i,t} = 0 \right) = \delta_i + \pi_i - \Psi_i
\]

and the long term interest rate is

\[
\lim_{n \to \infty} -\frac{1}{n} \log Z^{(n)}_{i,t} \left( \hat{I}_{i,t} = 0 \right) = \delta_i + \pi_i - \Psi_i + l_i
\]

We can see that \( l_i \) measures the slope of bond yield curve of country \( i \). A higher resilience \( B_i \) (closer to 1) corresponds to less precautionary saving \( \Psi_i \) and a smaller term spread \( l_i \).

Corollary 1. In the cross section, return to carry trade with short maturity bond can be positively predicted using foreign short term interest rate, and negatively predicted using foreign bond term spread.

Corollary 1 speaks to the finding of Ang and Chen (2010). As the model suggests, the violation of uncovered interest rate parity condition is due to heterogeneous exposure to global disaster, which can be coarsely measured by short rate or term spread. Less resilient
countries have lower interest rate due to stronger precautionary saving motive. At the same time, less resilience countries have steeper slope of nominal bond yield curve, because the jump in inflation induced by disaster hurts more.

When the carry trade is implemented using long term bond, the return can’t be guaranteed for two reasons, (a) the expression has inflation $\hat{I}_{t,t}$ inside which makes it very noisy, and more fundamentally (b) $-\Psi_i + l_i$ is not necessarily monotonic in resilience $B_i$.

**Corollary 2.** In the cross section, holding period return to carry trade with long term bond is not predictable with either foreign short term interest rate or foreign bond term spread.

To evaluate the model quantitatively, I make the following parameter assumption.

**Assumption.** All countries have the same $\delta_i$. All countries except base country have the same $\bar{J}_i$ and $\bar{I}_i$.

Indeed relatively large jump size $\bar{J}_i$ is needed so that the two effects are comparable. Barro and Ursúa (2008) document 67 cases in which the consumption of a OECD country falls by more than 10% over the years between 1870 to 2000. Not counting German hyperinflation of 1922-23, the fall in consumption is accompanied by an average inflation rate of 17 percent. With the US, the picture is mixed: during 1917-21 the inflation rate was 13.9%, whereas in great depression (1929-33) it was a 6.4% deflation. In Section III, I calibrate $\bar{J}_0 = 0.04$ to match US yield curve. Here I set $\bar{J}_i = 0.17$ for all the other countries.

I calibrate the distribution of resilience $B_i$ based on the distribution of disaster size documented by Barro and Jin (2011). Countries are sorted by the slope of their yield curves into three portfolios, in which investment currencies have the flattest yield curves and funding currencies have the steepest slopes. The simulation, which is detailed in the appendix, is consistent with the empirical procedure of Lustig et al. (2017), except that my holding period is one year, whereas Lustig et al. (2017) report annualized monthly results. Figure 7 plots the average portfolio returns using bond of different maturities.

It is clear that the model generates a downward sloping term structure for carry trade. The “traditional” carry trade that uses short maturity bond yields about 3% risk premium. This is consistent with evidence in the literature. However, on the long end, the return is no longer significantly different from zero. As a comparison, in the left panel of Figure 7 I copy

---

3In the analysis, the inflation of base country $\bar{I}_0$ is irrelevant. Specifically, the result holds with $I_i \neq I_0$, in which case the nominal exchange rate is non-stationary.
Figure 7: **Theoretical and empirical carry trade risk premia in U.S. Dollars.** The left is model implied term structure of carry trade risk premia, with maturity in years. Vertical axis is dollar excess returns that correspond to the holding period returns expressed in U.S. dollars of investment strategies that go long and short bonds of maturities in the horizontal axis. The strategies sort countries by the slope of their yield curves into three portfolios, with investment currencies having flat yield curves. The strategies are held for one period. The right is Figure 1 of [Lustig et al. (2017)]( ), with maturity in quarters. In both figures, the shaded areas correspond to one standard deviation above and below each point estimate.

Figure 1 of [Lustig et al. (2017)]( ) that uses zero-coupon bonds obtained from the estimation of term structure curves using government notes and bonds and interest rate swaps, whose long end is negative and insignificant.

---

V. **Concluding Remarks**

This paper reconciles the puzzling facts of holding period return in stock, bond, derivative and currency market by modeling the impact of disaster in a more realistic manner. The model is solved in closed form, allowing for further extension along the following directions: (a) instead of holding period return, examine the model’s implication for the yield to maturity across assets, which is another important topic in the third generation asset pricing models,
and (b) speak to the fourth generation of topics on the time variation of the distribution of risk, e.g. cyclical
turn of slope of dividend yield curve as studied by [Gormsen (2017)], which add even more constraints on asset pricing models, as well as (c) study the cross asset predictability, such as bond to stock market as in [Cochrane and Piazzesi (2005)] and variance to stock market as in [Andersen et al. (2017)]. Finally, since disaster models are based on rare events, it is useful to (d) document the historical pattern of dividend and realized variance during the disaster episodes, in addition to the evidence on consumption and inflation in [Barro and Ursúa (2008)].

References


Lynch, Anthony W, and Oliver Randall, 2011, Why surplus consumption in the habit model may be less persistent than you think, Working paper.


Appendix A. Linearity Generating Process

The results here are from Gabaix (2007) Section 3.1.

**Definition.** The process \( M_tD_t (1, X_t') \) with \( M_tD_t \in \mathbb{R} \setminus \{0\} \) and \( X_t \in \mathbb{R}^n \), is a linearity-generating process if it is \( L^1 \) and there are constants \( \alpha \in \mathbb{R}, \gamma, \delta \in \mathbb{R}^n, \Gamma \in \mathbb{R}^{n \times n} \), such that the following relations hold at all \( t \in \mathbb{N} \):

\[
\mathbb{E}_t \left[ \frac{M_{t+1}D_{t+1}}{M_tD_t} \right] = \alpha + \delta' X_t \\
\mathbb{E}_t \left[ \frac{M_{t+1}D_{t+1}}{M_tD_t} X_{t+1} \right] = \gamma + \Gamma X_t
\]

In practice, we “twist” the \( X_{t+1} \) process so that the above expression holds. Denote

\[
\Omega \equiv \begin{pmatrix} \alpha & \delta' \\ \gamma & \Gamma \end{pmatrix}
\]

and the basic pricing properties are summarized in the following theorem.

**Theorem.** The price-dividend ratio of a zero-coupon equity or bond of maturity \( T \), \( Z_t(T) = \mathbb{E}_t[M_{t+1}D_{t+1}/(M_tD_t)] \) is

\[
Z_t(T) = \begin{pmatrix} 1 & 0_n \end{pmatrix} \Omega^T \begin{pmatrix} 1 \\ X_t \end{pmatrix}
\]

Suppose that \( \Omega \)’s eigenvalues have modulus less than 1 (finiteness of the price). Then, the price-dividend ratio of the stock, \( \frac{P_t}{D_t} = \mathbb{E}_t[\sum_{s=t+1}^{\infty} M_sD_s]/(M_tD_t) \), is

\[
\frac{P_t}{D_t} = \frac{1}{1 - \alpha - \delta' (I_n - \Gamma)^{-1} \gamma} \left( \alpha + \delta' (I_n - \Gamma)^{-1} (X_t + \gamma) \right)
\]

\[
= \begin{pmatrix} 1 & 0_n \end{pmatrix} \Omega (I_{n+1} - \Omega)^{-1} \begin{pmatrix} 1 \\ X_t \end{pmatrix}
\]

In this main text, I use the first half for bond and dividend strip, and second half for stock price. Variance forward are derived based on stock price which is linear in state variables.
Appendix B. Proofs

Proof of Proposition 1. To calculate the price dividend ratio, we need three moments. Denoting $\delta_D = \delta - \mu_d$, we have

$$
\mathbb{E}_t \frac{M_{t+1} D_{t+1}}{M_tD_t} = e^{-\delta + \mu_d} (1 + g_t) \left( \frac{H_t}{1 + g_t} + 1 \right)
$$

$$
\mathbb{E}_t \frac{M_{t+1} D_{t+1} g_{t+1}}{M_tD_t} = e^{-\delta + \mu_d} (1 + g_t) \left[ \frac{H_t}{1 + g_t} g^D + pg^D + (1 - p) \frac{1}{1 + g_t} e^{-\phi g} g_t \right]
$$

$$
\mathbb{E}_t \frac{M_{t+1} D_{t+1} \hat{H}_{t+1}}{M_tD_t} = e^{-\delta + \mu_d} (1 + g_t) \left( \frac{H_t}{1 + g_t} + 1 \right) \mathbb{E}_t \left[ \frac{1 + \hat{H}}{1 + H_t + g_t} e^{-\phi \hat{H} \hat{H}_t + \sigma^H \hat{H}_{t+1}} \right]
$$

To summarize, we can write everything in terms of state variables

$$
\mathbb{E}_t \frac{M_{t+1} D_{t+1}}{M_tD_t} \left( 1, g_{t+1}, \hat{H}_{t+1} \right) = e^{\delta - \mu_d} \Omega_D \left( 1, g_t, \hat{H}_t \right)
$$

where $\Omega_D = \begin{pmatrix} 1 + \hat{H} & 1 \\ (\hat{H} + p) g^D & pg^D + (1 - p) e^{-\phi g} & 1 \\ 0 & 0 & (1 + \hat{H}) e^{-\phi H} \end{pmatrix}$. From here, one simply plugs in $h = \ln (1 + \hat{H})$ and reorganizes a bit. \hfill \Box

Proof of Proposition 2. Using the previous result, the pricing kernel $M_{t+1}$ satisfies

$$
\frac{M_{t+1}}{M_t} = e^{-\delta} \times \begin{cases} 1 \\ B^{-\gamma} \text{ Disaster} \end{cases}
$$

The first moment $\frac{M_{t+1} Q_{t+1}}{M_t Q_t}$

$$
\mathbb{E}_t \frac{M_{t+1} Q_{t+1}}{M_t Q_t} = e^{-\delta} \mathbb{E}_t \left[ e^{\Psi (1 - I_t)} \right]
$$

$$
= e^{-\delta + \Psi - \pi} - e^{-\delta + \Psi} \hat{I}_t
$$

37
The second moment \( \frac{M_{t+1}Q_{t+1} \hat{I}_{t+1}}{M_t Q_t} \)

\[
\mathbb{E}_t \frac{M_{t+1}Q_{t+1} \hat{I}_{t+1}}{M_t Q_t} = \mathbb{E}_t \left[ e^{-\delta} (1 - I_t) \left( \frac{1 - \bar{I}}{1 - I_t} e^{\phi_I \hat{I}_t + \sigma^I \epsilon_{t+1}^I} \right) + e^{-\delta} pB^{-\gamma} (1 - \bar{I}) \left( J + \hat{J}_t \right) \right] \\
= e^{-\delta} pB^{-\gamma} e^{-\pi} J \\
+ e^{-\delta} e^{\Psi - \phi_I} \hat{I}_t \\
+ e^{-\delta} pB^{-\gamma} e^{-\pi} \hat{J}_t
\]

The third moment \( \frac{M_{t+1}Q_{t+1} \hat{J}_{t+1}}{M_t Q_t} \)

\[
\mathbb{E}_t \frac{M_{t+1}Q_{t+1} \hat{J}_{t+1}}{M_t Q_t} = e^{-\delta} \mathbb{E}_t \left[ (1 - p + pB^{-\gamma}) (1 - I_t) \left( \frac{1 - \bar{I}}{1 - I_t} e^{\phi_j \hat{J}_t + \sigma^j \epsilon_{t+1}^j} \right) \right] \\
= e^{-\delta} e^{\Psi - \phi_J} \hat{J}_t
\]

To summarize,

\[
\mathbb{E}_t \frac{M_{t+1}Q_{t+1} \hat{I}_{t+1}}{M_t Q_t} (1, \hat{I}_{t+1}, \hat{J}_{t+1})' = e^{-\delta - \pi + \Psi} \Omega_Z (1, \hat{I}_t, \hat{J}_t')
\]

where \( \Omega_Z = \begin{pmatrix} 1 & -e^\pi & 0 \\ (1 - \frac{1 - p}{e^\phi I}) \bar{J} & e^{\phi_I} & 1 - \frac{1 - p}{e^\phi I} \\ 0 & 0 & e^{\phi_J} \end{pmatrix} \), \( \pi = -\ln (1 - \bar{I}) \), and \( e^\Psi = 1 - p + pB^{-\gamma} \)

Proof of Proposition: The price of nominal bond, denoted in local currency, equals

\[
Z_{i,t}^{(n)} (\hat{I}_{i,t}) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} e^{-n(\delta_i + \pi_i - \Psi_i)} \Omega_{Z,i}^{n} (1, \hat{I}_{i,t})' 
\]

where \( \delta_i = -\log \beta + \gamma \mu_{i,c} \) and

\[
\Omega_{Z,i} = \begin{pmatrix} 1 & -e^{\pi_i} \\ [1 - (1 - p) e^{-\Psi_i}] \bar{J}_i & e^{\phi_i} \end{pmatrix}
\]

with \( \pi_i = -\ln (1 - \bar{I}_i) \) and \( \Psi_i \equiv \ln (1 - p + pB_i^{-\gamma}) \).
We can write a more explicit formula by eigen-decomposing $\Omega_{Z,i}$ into

$$\Omega_{Z,i} = \begin{pmatrix} e^{\pi_i} & e^{\pi_i} \\ 1 - e^{-l_i} & e^{-l_i} - e^{-\phi I} \end{pmatrix} \begin{pmatrix} e^{-l_i} & 0 \\ 0 & e^{-\phi I} + 1 - e^{-l_i} \end{pmatrix} \begin{pmatrix} e^{\pi_i} & e^{\pi_i} \\ 1 - e^{-l_i} & e^{-l_i} - e^{-\phi I} \end{pmatrix}^{-1}$$

where $e^{-l_i} \in \left(\frac{e^{-\phi I} + 1}{2}, 1\right)$ is parameterized as

$$(1 - e^{-l_i}) (e^{-l_i} - e^{-\phi I}) = e^{\pi_i} [1 - (1 - p) e^{-\Psi_i}] \hat{J}_i$$

The price of nominal bond is (after some algebra)

$$Z_{i,t}^{(n)} \left( \hat{I}_i \right) = e^{-n(\delta_i + \pi_i - \Psi_i + l_i)} \times \left[ 1 - \left( 1 - \frac{(e^{-l_i} - e^{-\phi I}) - (1 - e^{-l_i})}{e^{-l_i}} \right)^n \right] \frac{e^{\pi_i} \hat{I}_{i,t} - (1 - e^{-l_i})}{(e^{-l_i} - e^{-\phi I}) - (1 - e^{-l_i})}$$

Notice that $0 < 1 - \frac{(e^{-l_i} - e^{-\phi I}) - (1 - e^{-l_i})}{e^{-l_i}} < 1$. So the expected holding period return of long term bond, when $n$ is large and $\left( 1 - \frac{(e^{-l_i} - e^{-\phi I}) - (1 - e^{-l_i})}{e^{-l_i}} \right)^n \rightarrow 0$, can be expressed as

$$\mathbb{E}_t R_i^{(\infty)} = \lim_{n \to \infty} \mathbb{E}_t \frac{Z_{i,t+1}^{(n-1)} (\hat{I}_{i,t+1})}{Z_{i,t}^{(n)} (\hat{I}_{i,t})}$$

$$= \frac{1}{e^{-(\delta_i + \pi_i - \Psi_i + l_i)}} \left[ 1 + \frac{(e^{-\pi_i} - e^{-\phi I} - \hat{I}_{i,t}) \hat{I}_{i,t}}{(e^{-l_i} - e^{-\phi I}) e^{-\pi_i} - \hat{I}_{i,t} (e^{-\pi_i} - \hat{I}_{i,t})} \right]$$

and the rest follows (after some algebra).

In addition to the two corollaries, we can express the return of traditional carry trade differently in terms of [Fama (1984)](Fama1984) regression with country pairs,

$$\frac{S_{i,t+1}}{S_{i,t}} - \frac{S_{j,t+1}}{S_{j,t}} = \alpha + \beta \left( \frac{R_i^{(1)}}{R_0^{(1)}} - \frac{R_j^{(1)}}{R_0^{(1)}} \right) + \epsilon$$

39
which according to the model, is approximately

\[
\left( \hat{I}_{i,t} - \hat{I}_{j,t} \right) = \alpha + \beta \left[ (\Psi_j - \Psi_i) + \left( \hat{I}_{i,t} - \hat{I}_{j,t} \right) \right] + \epsilon
\]

It yields a \( \hat{\beta} < 1 \) as long as there is cross sectional variation in \( B \) and probability of disaster \( p \) is not 0.

**Corollary 3.** *In the cross section, the coefficient in Fama (1984) regression \( \hat{\beta} = \frac{\sigma^2(\hat{I}_{i,t})}{\sigma^2(\Psi_i) + \sigma^2(\hat{I}_{i,t})} \) is less than 1.*

When there is enough variation of \( B_i \), the coefficient falls well below 1. Even though a clear expression of \( \sigma^2(\Psi_i) \) is not convenient to derive, I find that in the simulation \( \sigma^2(\Psi_i) \) and \( \sigma^2(\hat{I}_{i,t}) \) are roughly comparable, which implies an estimated \( \hat{\beta} \) about 1/2.

**Appendix C. Detail of Simulations**

*Section II.B and III.C*

The simulation takes two steps. The first step delivers simulated spot price for variance forward of each maturity \( n = 1, 2, \ldots, 12 \) for a grid of \( \hat{H}_t \). In each run, I start at \( \hat{H}_t \) and simulate the economy forward for \( n \) period(s), allowing for disaster. At the end of the period, I have the consumption and realized variance, the former of which is then translated into pricing kernel and the latter is the sum of squared log daily return. If disaster doesn’t happen in the last period, I assume the return of the last period is realized in one day. If disaster happens, I assume the return is realized in \( n_D \) days, say \( n_D = 2 \), following the trick used in [Dew-Becker et al. (2017)](https://example.com) to avoid a gigantic realized variance. I repeat the experiment for 100 million times, and calculate the average product of pricing kernel and payoff. This gives me the spot price \( V^n(\hat{H}_t) \) of variance forward, which is then converted to forward price \( F^n(\hat{H}_t) \) by multiplying \( (1 + r_f)^n \). I then repeat it for all grid points of \( \hat{H}_t \times n \).

Now that I have forward price, the second step is to simulate holding period return. In each simulation, the economy starts from \( \hat{H}_t = 0 \) and continues for 204 months with no disaster. In each of the 204 months, I buy variance forward at \( F^n(\hat{H}_t) \) and sell it next
month at $F^{n-1}(\hat{H}_{t+1})$, except for $n = 1$ in which case I receive the realized variance in the next month. The forward price is linearly extrapolated from the grid of the first step. Then I calculate the Sharpe ratio of realized holding period return for each maturity during the 204 months. Since I'm using forward price, it is not necessary to subtract risk free rate when calculating Sharpe ratio. I repeat the process for 400 times to plot the average Sharpe ratio and its confidence interval.

In this section, I also derive an “approximate” expression for variance forward price. To this end, some simplification needs to be made. Most importantly, I ignore the possibility of having more than one disaster before month $t + n$. Such events are very unlikely, although it results in a larger pricing error for longer maturity forward. Under this assumption, for month $t + n$ in the future, there are up to three possible types of state: (1) No disaster throughout $t + 1$ to $t + n$; (2) No disaster at $t + n$ with one prior prior disaster; (3) Disaster at $t + n$ with no more than one prior disaster. I denote the realized variance as $V^{ND,n}(\hat{H}_t)$, $V^{NDc,n}(\hat{H}_t)$ and $V^{D,n}(\hat{H}_t)$ for each scenario. In addition, I use net return instead of log gross return.

**Proposition 4.** The spot price of variance forward is

$$V^n_t = e^{-n\delta} \left\{ \begin{array}{c} (1 - np) V^{ND}_n(\hat{H}_t) + \sum_{\tau=1}^{n-1} pB^{-\gamma} V^{ND\tau}_n(\hat{H}_t) \\ + p \left[ 1 - (n - 1)p + (n - 1)pB^{-\gamma} \right] B^{-\gamma} V^{D}_n(\hat{H}_t) \end{array} \right\}$$
with

\[ V^{ND}_n(\hat{H}_t) \equiv \left( \mu_d + \frac{1}{PD_0} - \frac{PD_H}{PD_0} \phi_H (1 - \phi_H)^{n-1} \hat{H}_t \right)^2 + (\sigma^D)^2 + \left( \frac{PD_H}{PD_0} \sigma_H \right)^2 \]

\[ V^{ND, r}_n(\hat{H}_t) \equiv \left( \mu_d + \frac{1}{PD_0} - \frac{PD_g}{PD_0} \phi_g (1 - \phi_g)^{n-\tau-1} g^D \right) \left( 1 - \phi_g \right)^{n-1} \hat{H}_t \left( \frac{1 + (1 - \phi_H) \hat{H}_t}{H + p} \right)^2 + (\sigma^D)^2 + \left( \frac{PD_H}{PD_0} \sigma_H \right)^2 \]

\[ V^{D}_n(\hat{H}_t) \equiv \frac{1}{n_D} \left[ \left( \frac{1 + (1 - \phi_H) \hat{H}_t}{H + p} \right)^2 - 1 \right]^2 \]

and \( R^D \equiv F^{\text{inv}}_{\mu g D} \left| _{\mu g=0} \right. \).

**Proof of Proposition 4.** The movement of \( H_t \) is slow, which means \( \phi_H \approx 0 \) and \( 1 - \phi_H \approx 1 \). I throw out terms of \( o(\phi_H) \) importance. Absent disaster, the summed squared return roughly equals to

\[ V^{ND, n}_n(\hat{H}_t) \]

\[ \approx \mathbb{E}_t \left[ (1 + \mu_d) (1 + \sigma^D e_{i+n}^D) \left( 1 + \frac{1}{PD_0} - \frac{PD_H}{PD_0} \phi_H \hat{H}_{t+n-1} + \frac{PD_H}{PD_0} \sigma_H e_{i+1}^H \right) - 1 \right]^2 \]

\[ \approx \mathbb{E}_t \left[ \left( \mu_d + \frac{1}{PD_0} - \frac{PD_H}{PD_0} \phi_H \hat{H}_{t+n-1} \right)^2 + (\sigma^D)^2 + \left( \frac{PD_H}{PD_0} \sigma_H \right)^2 \right] \]

\[ \approx \left( \mu_d + \frac{1}{PD_0} - \frac{PD_H}{PD_0} \phi_H (1 - \phi_H)^{n-1} \hat{H}_t \right)^2 + (\sigma^D)^2 + \left( \frac{PD_H}{PD_0} \sigma_H \right)^2 \]

where I ignored the twist \( \frac{1 + \phi_H}{1 + \phi_H + \phi_g} \) and use \( e^{\mu_d} \approx 1 + \mu_d \) and \( e^{-\phi_H} \approx 1 - \phi_H \) in the first step. In the last step I neglect all \( o(\phi_H) \) terms. Intuitively, the three terms corresponds to the expected return, unexpected dividend growth and unexpected resilience change.

If there was a disaster at time \( t + \tau \), where \( \tau < n \), the realized variance is slightly different. The difference comes from the mean-reverting \( g_t \) process. The realized variance in this case
can be calculated similarly as

\[ V^{ND,n} \left( \hat{H}_t \right) \approx \left( \mu_d + \frac{1}{PD_0} - \frac{PD_0}{PD_0} \phi_g (1 - \phi_g)^{n-\tau-1} g^D - \frac{PD_H}{PD_0} \phi_H (1 - \phi_H)^{n-1} \hat{H}_t \right)^2 + (\sigma^D)^2 + \left( \frac{PD_H}{PD_0} \sigma^H \right)^2 \]

Finally, given the relative magnitude, I assume variance comes solely from the realized return when disaster happens in that month.

\[ R_{t+n}^D = \frac{F_{t+n}^P \hat{F}_{t+n}^{PD} \mid g_{t+n} = g^D}{F \hat{F}^{PD} \mid g_t = 0} = \frac{H_{t+n} + p \hat{F}_{t+n}^{PD} \mid g_t = g^D}{H + p \hat{F}^{PD} \mid g_t = 0} \]

Call \( R^D = \hat{F}_{t+n}^{PD} \mid g_{t+n} = g^D \), that is, the return when resilience is at its mean. The realized variance induced by disaster \( V^{D,n} \) can be calculated as

\[ V^{D,n} \left( \hat{H}_t \right) \approx \frac{1}{n_D} \left[ \frac{\hat{H}_{t+n} + \hat{H} + p R^D - 1}{H + p} \right]^2 \approx \frac{1}{n_D} \left[ 1 + \left( \frac{(1 - \phi_H)^n}{H + p} \right) R^D - 1 \right]^2 \]

\[ \Box \]

Figure 8 plots the forward curve in good time when \( \hat{H}_t = 0.2 p B^{-\gamma} \) and in bad time when \( \hat{H}_t = -0.2 p B^{-\gamma} \). The curve is much higher and downward sloping in bad times, when disaster could induce a much larger realized volatility if it happens sooner than later. While in good times because of the exact opposite reason the curve is upward sloping.

Of course, the model is not a complete description of the world. During the Great Recession, the forward curve reached as high as 60 and Figure 2 suggests a 40% equity yield. My model struggles to achieve these numbers with a low annual disaster probability of 3.63%.

43
Section IV.A

Barro and Jin (2011) show that the size distribution of disaster, $z_i = \frac{1}{B_i}$, can be well approximated by a double power law

$$f(z_i) = \begin{cases} 
0 & \text{if } z_i < z_0 \\
A_2z_i^{-(\beta+1)} & \text{if } z_0 \leq z_i < \delta \\
A_1z_i^{-(\alpha+1)} & \text{if } z_i \geq \delta 
\end{cases}$$

with the following estimations $z_0 = 1.105$, $\alpha = 4.16$, $\beta = 10.10$, $\delta = 1.38$, with $A_1$ and $A_2$ given by $\frac{1}{A_1} = \frac{\delta^{\beta-\alpha}}{\beta} (z_0^\beta - \delta^{-\beta}) + \frac{\delta^{-\alpha}}{\alpha}$ and $A_2 = A_1 \delta^{\beta-\alpha}$.

The first step is to calculate and invert the CDF of double-power-law as

$$F^{-1}(p_i) = \begin{cases} 
(z_0^{-\beta} - \frac{\beta}{A_2}p_i)^{-\frac{1}{\beta}} & \text{if } p_i \leq \frac{A_2}{\beta} (z_0^{-\beta} - \delta^{-\beta}) \\
\left(\frac{\alpha}{A_1} - \frac{\alpha}{A_1}p_i\right)^{-\frac{1}{\alpha}} & \text{if } p_i > \frac{A_2}{\beta} (z_0^{-\beta} - \delta^{-\beta}) 
\end{cases}$$

based on which I select an equal-interval sample of 36 from 1st Quintile to 4th Quintile ($p_1 = 0.2$, $p_2 = 0.2 + \frac{0.8-0.2}{36-1}$, ..., $p_{36} = 0.8$). Although there are only 35 OECD countries, I add one to save me the trouble when forming portfolio. I note that this is a rather conservative selection with limited cross section variation, which undermines my argument. The smallest $B_i$ is 0.7055664 and the largest one is 0.8818293.

Using $B_i$, I calculate $l_i$ from $(1 - e^{-l_i}) (e^{-l_i} - e^{-\phi_i}) = e^{\pi_i} [1 - (1 - p) e^{-\psi_i}] \bar{J}_i$, or

$$l_i = -\ln \frac{1 + e^{-\phi_i} + \sqrt{(1 - e^{-\phi_i})^2 - 4e^{\pi_i} [1 - (1 - p) e^{-\psi_i}] \bar{J}_i}}{2}$$

where $\psi_i \equiv \ln (1 - p + p B_i^{-\gamma})$, $\pi_i = -\ln (1 - \bar{I}_i)$, $\bar{J}_i = 0.17$, $\bar{I}_i = 0.03$, and $\phi_i = 0.2$.

In Figure 5, I plot the value of $e^{\psi_i}$ and $e^{\psi_i - l_i}$ against $B_i$. As we can see, there is a limited cross sectional variation in $e^{\psi_i - l_i}$ with varying monotonicity. Carry trade with long term bond sorted by short rate may receive negative return instead.
Figure 8: Model Implied (top) and Empirical (Bottom) Variance Forward Price. Top left is good time when $\hat{H}_t = 0.2pB^{-\gamma}$ and right is bad time when $\hat{H}_t = -0.2pB^{-\gamma}$. Bottom is the historical variance swap curve from Ait-Sahalia et al. (2015). Rate is plotted as $100 \times \sqrt{12F_t}$. 

45
Figure 9: **Value of** $e^{-\Psi_i}$ (**top**) and $e^{-\Psi_i+l_i}$ (**bottom**) **against** $B_i$, plotted using the same calibration of $p$, $\delta$, $\phi_I$ and $I_i = \hat{I}$ as before, but with $\hat{J}_i = 0.17$ to fit international data. I define $\Psi_i \equiv \ln \left(1 - p + pB_i^{-\gamma}\right)$ and parameterize $e^{-l_i} \in \left(e^{-\phi_I+1}, 1\right)$ as $(1 - e^{-l_i}) (e^{-l_i} - e^{-\phi_I}) = e^{\pi_i} \left[1 - (1 - p) e^{-\Psi_i}\right] \hat{J}_i$. A higher $B_i$ corresponds to a smaller $\Psi_i$ and a small $l_i$, which implies that everything else equal, a more resilient country has higher risk free rate (short rate=$\delta_i + \pi_i - \Psi_i$) and relatively “flat” yield curve (slope=$l_i$).

Starting from $\hat{I}_{i,t} = 0$, I simulate a path of inflation for each country $i$ following

$$\hat{I}_{i,t+1} = \frac{1}{1 - \hat{I}_i - \hat{I}_{i,t}} e^{-\phi_I} \hat{I}_{i,t} + \sigma_I \epsilon_{i,t+1}$$

where $\epsilon_{i,t+1}$ is i.i.d. standard normal and $\sigma_I = 0.009$ as before. The simulation is annual for 30 years. At each time $t$, countries are equally divided into three portfolios based on the difference between one year and 15 year bond yield which can be interpreted as the slope.
For each bond with maturity \( n \) of country \( i \), I calculate the realized return in dollar as

\[
\left( \frac{S_{i,t+1}}{S_{i,t}} \right)^{-1} \frac{Z^{(n-1)}_{i,t+1}}{Z^{(n)}_{i,t}} \left( \hat{I}_{i,t+1} \right)
\]

The inflation process of base country is not relevant here as it gets canceled out in the next step when I calculate the return of investment strategies that long countries with flat curve and short countries with steep curve. So feel free to set the inflation of base country to zero as I do. Finally, I plot the mean return of the strategy from Year 2 to 29 (in total 28 years) in Figure 7 with the shaded area representing one standard deviation, and compare with Figure 1 of Lustig et al. (2017).

**Appendix D. New Trading Strategies**

Inflation moves around in normal times, and jumps up in disaster. Normal movement of inflation increases the volatility of a strategy without getting compensated with high return. To increase the Sharpe ratio, we do our best to minimize the influence of normal movement of inflation. To this end, I introduce two locally self financing synthetic assets that are mathematically shown to be free of “normal inflation”. I then show how to construct globally self financing strategies based on them.

**Two Locally Self-financing Synthetic Assets**

**Synthetic Asset One: Refined “Slope”**

Typically a “slope” strategy is a long short strategy with two bonds, whose weights are the reciprocals of the maturities so that the strategy is duration neutral. Such a strategy is approximately immune from small changes in the level of nominal yield curve. However, it is not necessarily inflation neutral as inflation changes the shape of yield curve.

The first synthetic asset is the refined “slope” which involves three bonds with maturity \( n_1, n_2 \) and 2. The weights in the three bonds are respectively \(-\frac{n_2}{n_2-n_1}\), \(-\frac{n_1}{n_2-n_1}\) and \(-1\). As such, this synthetic asset is locally self financing. Proposition 5 shows that it is inflation neutral.
Proposition 5. Consider a refine “slope” that uses three bonds of maturity \( n_1, n_2 \) and 2 with respective weight \( \frac{n_2}{n_2-n_1} \), \( -\frac{n_1}{n_2-n_1} \) and \(-1\). Neglecting the jump size \( \hat{J}_{i,t} \), suppose trend inflation changes from \( \hat{I}_{i,t} \) to \( \hat{I}_{i,t+1} = e^{\frac{\phi_i}{1-e^{\pi_i}} \hat{I}_{i,t} + \sigma_i \epsilon_{i,t+1}} \), then the return of the refine “slope” equals

\[
\frac{n_2}{n_2-n_1} \frac{Z_i^{(n_1-1)}(\hat{I}_{i,t+1})}{Z_i^{(n_1)}(\hat{I}_{i,t})} - \frac{n_1}{n_2-n_1} \frac{Z_i^{(n_2-1)}(\hat{I}_{i,t+1})}{Z_i^{(n_2)}(\hat{I}_{i,t})} \approx -2 \left(1 - e^{-\hat{I}_{i,t}}\right) \left(e^{-\hat{I}_{i,t}} - e^{-\phi_i}\right) + \left[2 - \left(\frac{n_1n_2}{2} - 1\right) \left(1 - e^{-\phi_i}\right)\right] e^{\pi_i} \left[\left(1 - \frac{e^{\phi_i}}{1-e^{\pi_i} \hat{I}_{i,t} + \sigma_i \epsilon_{i,t+1}}\right) \hat{I}_{i,t} + \sigma_i \epsilon_{i,t+1}\right]
\]

(D1)

where \( e^{-\hat{I}_{i,t}} \in \left(\frac{e^{-\phi_i} + 1}{2}, 1\right) \) is parameterized as

\[
(1 - e^{-\hat{I}_{i,t}}) \left(e^{-\hat{I}_{i,t}} - e^{-\phi_i}\right) = e^{\pi_i} \left[1 - (1 - p) e^{-\Psi_i}\right] \hat{J}_i
\]

Now the task boils down to finding a combination of \( n_1 \) and \( n_2 \) such that the \( 2 - \left(\frac{n_1n_2}{2} - 1\right) \left(1 - e^{-\phi_i}\right) \) term is close to 0. In the annual calibration, I set \( \phi_i = 0.4 \) which implies that \( n_1 = 2 \) years and \( n_2 = 5 \) years might be suitable candidates.

The refined “slope” can be easily transformed into a global strategy in the way that I’ll explain later. The Sharpe ratio is 1.3, almost twice as high as the traditional carry trade. I plot its cumulative risk premium and compare that of the traditional currency carry trade strategy as in Koijen, Moskowitz, Pedersen, and Vrugt (2017). To increase interpretability, I scale both so that ex post volatility is 6 percent.

**Synthetic Asset Two: CP-style “Slope”**

Alternatively, a CP-style “slope” that uses three bonds of maturity \( T_1, T_1 + T_2 \) and \( T_1 + 2T_2 \) with respective weight \(-1, 2 \) and \(-1\) is largely immune from the local trend inflation. The exposures to the jump of inflation for the CP factor and the CP-style “slope” line up well in the cross section of countries. Thus we can rank the countries by the CP factor, and trade the CP-style “slope”.

48
Figure 10: **Cumulative risk premium of the strategy** that sorts country by $CP_{i,t}$ (12, 12) and trades the refine "slope" which consists of three bonds of maturity $n_1$, $n_2$ and 2 with respective weight $\frac{n_2}{n_2-n_1}$, $-\frac{n_1}{n_2-n_1}$ and \-1. Dashed line plots the cumulative risk premium of traditional carry trade strategy that sorts on short term interest rate and trades short maturity bond. For ease of comparison, both are scaled to the same ex post annualized volatility of 6 percent.

To show why it is the case, I start with Lemma 1 which calculates the general form of bond return under varying jump size.

**Lemma 1.** With time varying jump size of inflation, the nominal holding period return of zero-coupon bond of maturity $T$ is

$$
r_{S,t+1}(T) = \frac{Z_{St+1}(T-1)}{Z_{St}(T)} 
\approx \frac{1}{Z_{St}(T)} \left[ 1 + \left( \frac{1 - e^{-\psi_I - 2\kappa}}{\psi_I} \right) \left( \frac{e^{-2\kappa} - 1}{e^{-\psi_I T}} \right) \left( \hat{I}_t - \kappa \right) \right. \\
\left. + \left( e^{-\psi_J - \kappa} \Lambda_T + \left( 1 - e^{-\psi_J - \kappa} \right) K_T \right) \pi_t \right. \\
\left. + \left( -1 + e^{\psi_J} e^{-\psi_J T} \right) \epsilon_{t+1}^J \right]
$$

where $K_T \equiv \frac{1-e^{-\psi_J T}}{\psi_J - \psi_I} - \frac{1-e^{-\psi_I T}}{\psi_I - \psi_J}$, $\Lambda_T \equiv \frac{e^{-\psi_J T} - e^{-\psi_I T}}{\psi_I - \psi_J}$, $\psi_I \equiv \phi_I - 2\kappa$, and $\psi_J \equiv \phi_J - \kappa$.  

49
While $K_T$ is increasing in $T$, $\Lambda_T$ is hump-shaped. Hence

\[-e^{-\psi T_1} + 2e^{-\psi(T_1+T_2)} - e^{-\psi(T_1+2T_2)} \approx 0 \quad (D2)\]
\[-K_T + 2K_{T_1+T_2} - K_{T_1+2T_2} \approx 0 \quad (D3)\]
\[-\Lambda_T + 2\Lambda_{T_1+T_2} - \Lambda_{T_1+2T_2} \approx (T_2)^2 (\psi_I + \psi_J) > 0 \quad (D4)\]

**Proposition 6.** Consider a CP-style “slope” that uses three bonds of maturity $T_1$, $T_1 + T_2$ and $T_1 + 2T_2$ with respective weight $-1$, $2$ and $-1$. To the first order approximation, the CP-style “slope” is immune from local trend inflation $\hat{I}_t$

*Proof.* The return of the CP-style “slope” is

\[-r_{\$t+1}(T_1) + 2r_{\$t+1}(T_1 + T_2) - r_{\$t+1}(T_1 + 2T_2)\]

Since each $r_{\$t+1}$’s loading of $\hat{I}_t$ or of $\epsilon_{t+1}$ takes the form of $A + B \times e^{-\psi T}$, by Equation (D2), the CP style strategy is largely immune from the local inflation. \hfill \Box

The CP factor, being the same linear combination of forward rates, measures the “carry” of this CP-style “slope”. In the global implementation, the weight on each country’s “slope” is based on the ranking of CP factor across countries. So it remains to show that the loading on the jump in inflation $\hat{J}_t$ has the same monotonicity in the consumption resilience $B_i$. To see this, I write the expected return explicitly, using Equation (D3) and (D4),

\[
\mathbb{E}_t [-r_{\$t+1}(T_1) + 2r_{\$t+1}(T_1 + T_2) - r_{\$t+1}(T_1 + 2T_2)] = (T_2)^2 e^{(\delta - Hs + I_s - \psi_J)} (\psi_I + \psi_J) \pi_t
\]

whereas CP factor itself is (using result from Proposition 8 of Gabaix (2012))

\[
CP_{t}^{EF}(T_1, T_2) \equiv \frac{-f_t(T_1) + 2f_t(T_1 + T_2) - f_t(T_1 + 2T_2)}{(T_2)^2} \approx (\psi_I + \psi_J) \pi_t
\]

To the leading order, the two line up well in the cross section of countries. Hence we can rank countries by the CP factor, and trade the strategy.
However, the strategy falls short of my expectation. Setting $T_1 = T_2 = 24$ months, the strategy (solid line) marginally outperforms the traditional carry trade (dashed line), which is mostly driven by post 2009 period. I note that before 2009 I use zero coupon data available from the website of Jonathan Wright, used initially in Wright (2011). From June 2009 onward, I use zero coupon data from Bloomberg.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{chart.png}
\caption{Graph showing performance comparison between strategies.}
\end{figure}

**A Note on Global Implementation**

Each of the strategies is locally self financing. Now I explain how to assign “weight” $w_{i,t}$ to each country to transform it into a globally self financing strategy. In general, I choose the weight so that it reflects the funding cost of the long position of the strategy. For a positive weight $w_{i,t}$, the investor buys $P_{i,t}^* = w_{i,t}S_{i,t}$ amount of foreign long position by borrowing $w_{i,t}$ from base country at the interest rate $1 + r_t$, instead of financing the long position with the short position. To hedge the currency exposure, the investor also enters a forward contract to sell $P_{i,t}^* (1 + r_{i,t}^*)$ amount of foreign currency at forward rate $F_{i,t}$. At the same time, the investor keeps all proceeds $P_{i,t}^*$ from short position in the margin account which earns interest rate $1 + r_{i,t}^*$. In the next period, the investor settles the currency forward contract, converts remaining foreign currency into base currency at exchange rate $S_{i,t+1}$, and repays the loan.
in the base country. The cash flow in time $t + 1$ is

$$
\frac{-P_{i,t}^*}{S_{i,t}} (1 + r_t) \quad \text{(repayment of loan)}
$$

$$
+ P_{i,t}^* (1 + r_{i,t}^*) \left( \frac{1}{F_{i,t}} - \frac{1}{S_{i,t+1}} \right) \quad \text{(settlement of forward)}
$$

$$
+ \frac{1}{S_{i,t+1}} \left( P_{i,t+1}^{\text{Long}} - P_{i,t+1}^{\text{Short}} \right) \quad \text{(strategy payoff)}
$$

$$
+ \frac{1}{S_{i,t+1}} P_{i,t}^* (1 + r_{i,t}^*) \quad \text{(margin account)}
$$

and indeed can be simplified, assuming $F_{i,t} = S_{i,t} \frac{1 + r_{i,t}^*}{1 + r_t}$, into

$$
 w_{i,t} \times \frac{S_{i,t}}{S_{i,t+1}} \frac{P_{i,t+1}^{\text{Long}} - P_{i,t+1}^{\text{Short}}}{P_{i,t}^*}
$$

(D5)

In the global implementation, the weighting scheme assigns negative weight $w_{i,t}$ for half of the countries. To interpret negative $w_{i,t}$, it involves shorting $-w_{i,t} S_{i,t}$ worth of the long part of the foreign strategy at time $t$ and put all proceeds in the margin account which earns interest rate $1 + r_{i,t}^*$. To be consistent, the investor then borrows $-w_{i,t} S_{i,t}$ from foreign country, converts to $-w_{i,t}$ worth of base currency, while entering a forward contract to buy $-w_{i,t} S_{i,t} (1 + r_{i,t}^*)$ foreign currency in time $t + 1$ at forward exchange rate $F_{i,t}$. At the same time, the investor longs $-w_{i,t} S_{i,t}$ worth of the short part of the strategy using loans from the foreign country at interest rate $1 + r_{i,t}^*$. In total, the strategy provides the investor with $-w_{i,t}$ free cash at time $t$, which can be invested at interest rate $1 + r_t$, or equivalently, be used to finance strategies with positive weights. One can check that the cash flow to the investor at time $t + 1$ indeed satisfies the same Equation (D5).

With this type of implementation, the strategy is not locally self financing anymore. A country with a positive weight $w_{i,t}$ requires borrowing $w_{i,t}$ from the base country at time $t$, while a country with a negative weight $w_{i,t}$ provides $-w_{i,t}$ to the base country. However it is indeed globally self financing as I introduce a weighting scheme such that the weights sum up to 0 across countries. The weight is explicitly

$$
w_{i,t} = z_t \left( \text{rank}(CP_{i,t}(a, b)) - \frac{N_t + 1}{2} \right)
$$

(D6)
where $N_t$ is the number of countries and the scalar $z_t$ also ensures that the sum of the long and short positions equals 1 and $-1$ in each period $t$, and $CP_{i,t} (a, b)$ is the Cochrane and Piazzesi (2005) style factor

$$CP_{i,t} (a, b) = -f^{(a)}_{i,t} + 2f^{(a+b)}_{i,t} - f^{(a+2b)}_{i,t}$$  \hspace{1cm} (D7)

with the forward rate $f^{(n)}_{i,t} = \frac{Z^{(n-12)}(i,t)}{Z^{(n)}(i,t)} - 1$. This is a “inflation-free” measure of a country’s exposure to the jump of inflation in disaster (once again, see Proposition 8 of Gabaix (2012) for proof).

To conclude, the global return can be simply calculated as

$$\Sigma^n_i w_{i,t} \times \frac{S_{i,t}}{S_{i,t+1}} R_{i,t+1}$$

where $R_{i,t+1} \equiv \frac{P^{*\text{Long}}_{i,t+1} - P^{*\text{Short}}_{i,t+1}}{P^t_{i,t+1}}$ is the return of the local strategy.

**Data Source**

**Data source for exchange rate**

I use Barclays Bank International (BBI) data prior to 1997 and WMR/Reuters thereafter. Later I use Bloomberg data for cleaning. BBI and WMR are downloaded from Datastream. For the Bloomberg tickers, append them with string “Curncy”, e.g. the actual ticker is “AUD BGN Curncy” instead of just “AUD BGN”. The forward exchange rate in Bloomberg is quoted as “fx points”, which is forward rate minus spot rate in basis point.
<table>
<thead>
<tr>
<th>Country</th>
<th>Datastream: BBI</th>
<th>Datastream: WMR</th>
<th>Bloomberg</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spot</td>
<td>Forward</td>
<td>Spot</td>
</tr>
<tr>
<td>Australia</td>
<td>BBAUDSP</td>
<td>BBAUD1F</td>
<td>USAUD1F</td>
</tr>
<tr>
<td>Austria</td>
<td>USATS1F</td>
<td>AUSTSC$</td>
<td>ATS BGN</td>
</tr>
<tr>
<td>Belgium</td>
<td>USBEF1F</td>
<td>BELGLU$</td>
<td>BEF BGN</td>
</tr>
<tr>
<td>Canada</td>
<td>BBCADSP</td>
<td>BBCAD1F</td>
<td>USCAD1F</td>
</tr>
<tr>
<td>Denmark</td>
<td>BBDKKSP</td>
<td>BBDKK1F</td>
<td>USDKK1F</td>
</tr>
<tr>
<td>Euro</td>
<td>BBEURSP</td>
<td>BBEUR1F</td>
<td>USEUR1F</td>
</tr>
<tr>
<td>France</td>
<td>BBFRFSHP</td>
<td>BBFRF1F</td>
<td>USFRF1F</td>
</tr>
<tr>
<td>Germany</td>
<td>BBDEMSP</td>
<td>BBDEM1F</td>
<td>USDEM1F</td>
</tr>
<tr>
<td>Ireland</td>
<td>USIEP1F</td>
<td>IPUNTE$</td>
<td>IEP BGN</td>
</tr>
<tr>
<td>Italy</td>
<td>BBITLSP</td>
<td>BBITL1F</td>
<td>USITL1F</td>
</tr>
<tr>
<td>Japan</td>
<td>BBJPYSP</td>
<td>BBJPY1F</td>
<td>USJPY1F</td>
</tr>
<tr>
<td>Netherlands</td>
<td>BBNLGSP</td>
<td>BBNLG1F</td>
<td>USNLG1F</td>
</tr>
<tr>
<td>NewZealand</td>
<td>BBNZDSP</td>
<td>BBNZD1F</td>
<td>USNZD1F</td>
</tr>
<tr>
<td>Norway</td>
<td>BBNOKSP</td>
<td>BBNOK1F</td>
<td>USNOK1F</td>
</tr>
<tr>
<td>Portugal</td>
<td>USPTE1F</td>
<td>PORTE$</td>
<td>PTE BGN</td>
</tr>
<tr>
<td>Spain</td>
<td>USESP1F</td>
<td>SPANPE$</td>
<td>ESP BGN</td>
</tr>
<tr>
<td>Sweden</td>
<td>BBSEKSP</td>
<td>BBSEK1F</td>
<td>USSEK1F</td>
</tr>
<tr>
<td>Switzerland</td>
<td>BBCHFSP</td>
<td>BBCHF1F</td>
<td>USCHF1F</td>
</tr>
<tr>
<td>UK</td>
<td>BBGBPSP</td>
<td>BBGBP1F</td>
<td>USGBP1F</td>
</tr>
</tbody>
</table>

**Data source for IBOR**

I use the one month IBOR rate from Global Financial Data when available, and supplement missing numbers with Bloomberg data when

- "Country"="Denmark"&Date="2015-01-31"
- "Country"="Norway","Switzerland","UK","US"&Date="2014-02-28"
- "Country"="NewZealand"
<table>
<thead>
<tr>
<th>Country</th>
<th>GFD</th>
<th>Bloomberg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>IBAUS1D</td>
<td>VIBO1M Index</td>
</tr>
<tr>
<td>Austria</td>
<td>IBAUT1D</td>
<td>BIBOR1M Index</td>
</tr>
<tr>
<td>Belgium</td>
<td>IBBEL1D</td>
<td>PIPOFF1M Index</td>
</tr>
<tr>
<td>Canada</td>
<td>IBCAN1D</td>
<td>DM0001M Index</td>
</tr>
<tr>
<td>Denmark</td>
<td>IBDNK1D</td>
<td>DIBO01M Index</td>
</tr>
<tr>
<td>Euro</td>
<td>IBEUR1D</td>
<td>RIBORM1M Index</td>
</tr>
<tr>
<td>France</td>
<td>IBFRA1D</td>
<td>AIBO1M Index</td>
</tr>
<tr>
<td>Germany</td>
<td>IBDEU1D</td>
<td>LIS21M Index</td>
</tr>
<tr>
<td>Ireland</td>
<td>IBIRL1D</td>
<td>MIBOR01M Index</td>
</tr>
<tr>
<td>Italy</td>
<td>IBITA1D</td>
<td>EUR001M Index</td>
</tr>
<tr>
<td>Japan</td>
<td>IBJPN1D</td>
<td>AU0001M Index</td>
</tr>
<tr>
<td>Netherlands</td>
<td>IBNLD1D</td>
<td>CD0001M Index</td>
</tr>
<tr>
<td>NewZealand</td>
<td>IBNZL1D</td>
<td>CIBO01M Index</td>
</tr>
<tr>
<td>Norway</td>
<td>IBNOR1D</td>
<td>JY0001M Index</td>
</tr>
<tr>
<td>Portugal</td>
<td>IBPRT1D</td>
<td>NZ0001M Index</td>
</tr>
<tr>
<td>Spain</td>
<td>IBESP1D</td>
<td>NIBOR1M Index</td>
</tr>
<tr>
<td>Sweden</td>
<td>IBSWE1D</td>
<td>STIB1M Index</td>
</tr>
<tr>
<td>Switzerland</td>
<td>IBCHE1D</td>
<td>SF0001M Index</td>
</tr>
<tr>
<td>UK</td>
<td>IBGBR1D</td>
<td>BP0001M Index</td>
</tr>
<tr>
<td>US</td>
<td>IBUSA1D</td>
<td>US0001M Index</td>
</tr>
</tbody>
</table>

**Data source for zero coupon bond yields**

I use the data from Wright (2011) until May 2009. It can be downloaded from AEA website https://www.aeaweb.org/articles?id=10.1257/aer.101.4.1514. Afterwards, I use Bloomberg data. The tickers need to be appended by “Y Index”, e.g. “F12701Y Index” instead of “F12701”.
Data cleaning

I clean the data in the following order

- In the paper, I define exchange rate as foreign currency per $US, hence I invert the quotes of
  - All sources for UK
  - WMR/Reuters quotes for Australia and New Zealand
  - WMR/Reuters forward quotes for Euro
- Add pseudo US/US rate with both spot and forward rate = 1
- Combine one month IBOR rate and ZCB yield to form the entire yield curve with linear interpolation on yield
  - I think IBOR is per annum rate, which needs to be divided by 12 to get monthly rate
  - I assume both Wright (2011) and Bloomberg are annualized percentage points so I convert yield to monthly by $(1+\text{yield\%})^{(1/12)}-1$
- Calculate the annual forward rate $f_{i,t}^{(n)} = \frac{Z_{i,t}^{(n-12)}}{Z_{i,t}^{(n)}} - 1$ which I use later for trading strategy.
Delete post-1999 data of the following countries: Austria, Belgium, France, Germany, Ireland, Italy, Netherlands, Portugal, Spain

Delete pre-1999 data of Euro

Using spot and forward exchange rates and US IBOR rate, I calculate the difference between C.I.P. implied foreign IBOR rates and actual foreign IBOR rates. When the difference is greater than 2 annual percentage points, I replace the default source (BBI or WMR) of exchange rate with the other source if the implied difference of which is smaller than 2. Or else, I replace the default with numbers from Bloomberg if feasible data can be found. So far, there are still 200 obs from New Zealand and 100 obs from the rest countries with difference>2. The IBOR quote in New Zealand, which I get from Bloomberg, is a little volatile. Removing New Zealand from the sample slightly reduces the performance of my strategy – but no more than the effect of removing any other random country.

There are alternative data sources that one can use. I stop my cleaning at this point as my graph for cumulative return for carry trade strategy is visibly indifferent from the green line of Figure II in Koijen et al (2017).

Proofs

Proof of Proposition 5. To calculate the return of “slope”, the first step is to approximate the price of nominal bond

\[
Z^{(n)}_i \left( \hat{I}_{i,t} \right) = e^{- n (\delta_i + \pi_i - \Psi_i)} \times \left[ e^{- nl_i} - \left( e^{- nl_i} - \left( 1 + e^{- \phi_I} - e^{- l_i} \right)^n \right) \frac{e^{\pi_i} \hat{I}_{i,t} - (1 - e^{- l_i})}{(e^{- l_i} - e^{- \phi_I}) - (1 - e^{- l_i})} \right] \approx e^{- n (\delta_i + \pi_i - \Psi_i)} \times \left[ 1 - ne^{\pi_i} \hat{I}_{i,t} + \frac{n (n - 1)}{2} \left( (1 - e^{- \phi_I}) e^{\pi_i} \hat{I}_{i,t} - (1 - e^{- l_i}) (e^{- l_i} - e^{- \phi_I}) \right) \right]
\]
thus the return can be approximately by

\[
\frac{Z_i^{(n-1)}}{Z_i^{(n)}} \left( \hat{I}_{i,t+1} = \frac{-e^{-\phi_I}}{1-e^{\pi_i}\hat{I}_{i,t}} + \sigma^I \epsilon_{i,t+1} \right)
\]

\[
\approx e^{(\delta_i + \pi_i - \Psi_i)} \times \left[ 1 + ne^{\pi_i} \hat{I}_{i,t} - (n - 1) e^{\pi_i} \left( \frac{e^{-\phi_I}}{1 - e^{\pi_i} \hat{I}_{i,t}} + \sigma^I \epsilon_{i,t+1} \right) \right.
\]

\[
+ \frac{(n - 1)(n - 2)}{2} (1 - e^{-\phi_I}) e^{\pi_i} \left( \frac{e^{-\phi_I}}{1 - e^{\pi_i} \hat{I}_{i,t}} + \sigma^I \epsilon_{i,t+1} \right)
\]

\[
- \frac{n(n - 1)}{2} (1 - e^{-\phi_I}) e^{\pi_i} \hat{I}_{i,t} + (n - 1) (1 - e^{-\phi_I}) (e^{-l_i} - e^{-\phi_I})
\]

\[
\approx 1 + (\delta_i + \pi_i - \Psi_i) + e^{-\phi_I} e^{\pi_i} \hat{I}_{i,t} - (1 - e^{-\phi_I}) (e^{-l_i} - e^{-\phi_I})
\]

\[
+ \left[ 1 + \left( \frac{n_1 n_2}{2} \right) (1 - e^{-\phi_I}) \right] e^{\pi_i} \left( \frac{e^{-\phi_I}}{1 - e^{\pi_i} \hat{I}_{i,t}} - 1 \right) \hat{I}_{i,t} + \sigma^I \epsilon_{i,t+1} \right]
\]
At this time, the term $e^{-\phi_t} e^{\pi t} \hat{I}_{i,t}$ is very annoying. I notice that a special case of Equation (D8) is for $n = 2$ is

$$Z_i^{(1)}(\hat{I}_{i,t+1}) = \frac{e^{-\phi_t} \hat{I}_{i,t} + \sigma^I t_{t+1}}{Z_i^{(2)}(\hat{I}_{i,t})}$$

$$= 1 + (\delta_i + \pi_i - \Psi_i) + e^{-\phi_t} e^{\pi t} \hat{I}_{i,t} + (1 - e^{-l_i}) (e^{-l_i} - e^{-\phi_i})$$

$$- e^{\pi t} \left[ \left( \frac{e^{-\phi_t}}{1 - e^{\pi i} \hat{I}_{i,t}} - 1 \right) \hat{I}_{i,t} + \sigma^I t_{t+1} \right] \tag{D10}$$

Combining Equation (D9) and (D10), we arrive at

$$\frac{n_2}{n_2 - n_1} Z_i^{(n_1-1)}(\hat{I}_{i,t+1}) - \frac{n_1}{n_2 - n_1} Z_i^{(n_2-1)}(\hat{I}_{i,t+1}) - \frac{n_1}{n_2 - n_1} Z_i^{(1)}(\hat{I}_{i,t+1})$$

$$\approx -2 \left(1 - e^{-l_i}\right) (e^{-l_i} - e^{-\phi_i})$$

$$+ \left[ 2 - \left( \frac{n_1 n_2}{2} - 1 \right) (1 - e^{-\phi_i}) \right] e^{\pi t} \left[ \left( \frac{e^{-\phi_t}}{1 - e^{\pi i} \hat{I}_{i,t}} - 1 \right) \hat{I}_{i,t} + \sigma^I t_{t+1} \right]$$

Normal movement of inflation

Proof of Lemma 4 To calculate the return, I start by copying Theorem 2 from "Ten-puzzle" paper. The price of nominal zero-coupon bond of maturity $T$ is

$$Z_{St}(T) = e^{-(\delta - H_s + J_s + \kappa) T} \left( 1 - \frac{1 - e^{-\psi_i T}}{\psi_i} \left( \hat{I}_t - \kappa \right) - K_T \pi_t \right)$$
The (nominal) holding period return is

\[
 r_{s,t+1} (T) = \frac{Z_{s,t+1} (T - 1)}{Z_{s,t} (T)} 
\]

\[
= e^{(\delta - H_s + I_s + \kappa)} \frac{1 - \frac{1 - e^{-\psi_I (T - 1)}}{\psi_I} (\hat{I}_{t+1} - \kappa)}{1 - \frac{1 - e^{-\psi_I T}}{\psi_I} (\hat{I}_t - \kappa)} - K_{T-1} \pi_{t+1} 
\]

(Using \( \phi_I \equiv \psi_I + 2\kappa \))

\[
\approx e^{(\delta - H_s + I_s + \kappa)} \frac{1 - \frac{1 - e^{-\psi_I (T-1)}}{\psi_I} \left( e^{-\psi_I - 2\kappa} \hat{I}_t + \epsilon_{t+1} - \kappa \right)}{1 - \frac{1 - e^{-\psi_I T}}{\psi_I} (\hat{I}_t - \kappa) - K_T \pi_t} - K_{T-1} \pi_{t+1} 
\]

(Using \( \frac{1 + a}{1 + b} \approx 1 + a - b \) then colleting terms)

\[
\approx e^{(\delta - H_s + I_s + \kappa)} \left[ 1 + \frac{1 - e^{-\psi_I T} - e^{-\psi_I - 2\kappa} + e^{-\psi_I T - 2\kappa}}{\psi_I} (\hat{I}_t - \kappa) 
\right. 
\]

\[
+ \frac{e^{-\psi_I (T-1)} - 1}{\psi_I} \epsilon_{t+1} + K_T \pi_t - K_{T-1} \pi_{t+1} \right] 
\]

\[
= e^{(\delta - H_s + I_s + \kappa)} \left[ 1 + \frac{(1 - e^{-\psi_I - 2\kappa}) + (e^{-2\kappa} - 1) e^{-\psi_I T}}{\psi_I} (\hat{I}_t - \kappa) 
\right. 
\]

\[
+ \frac{-1 + e^{\psi_I} e^{-\psi_I T}}{\psi_I} \epsilon_{t+1} + K_T \pi_t - K_{T-1} \pi_{t+1} \right] 
\]

(D11)

From Equation [D11], the last term in the return \( r_{s,t+1} (T) \) is

\[
K_T \pi_t - K_{T-1} \pi_{t+1} \approx \left( K_T - K_{T-1} e^{-\phi_I} \right) \pi_t - K_{T-1} \epsilon_{t+1} 
\]
where

\[ K_{T-1} = \frac{1-e^{-\psi_I(T-1)}}{\psi_I} - \frac{1-e^{-\psi_J(T-1)}}{\psi_J} \]
\[ = \frac{\psi_J - \psi_I}{\psi_I} \frac{1-e^{-\psi_I T}}{\psi_I} - \frac{1-e^{-\psi_J T}}{\psi_J} \]
\[ \approx \frac{\psi_J - \psi_I}{\psi_I} \frac{1-e^{-\psi_I T}}{\psi_I} - \frac{1-e^{-\psi_J T}}{\psi_J} + K_T \]
\[ \approx \frac{\psi_I(1-e^{-\psi_I T})}{\psi_I} - \frac{\psi_J(1-e^{-\psi_J T})}{\psi_J} + K_T \]
\[ = K_T - \frac{e^{-\psi_I T} - e^{-\psi_J T}}{\psi_J - \psi_I} \]

Call \( \Lambda_T \equiv \frac{e^{-\psi_I T} - e^{-\psi_J T}}{\psi_J - \psi_I} \), then \( K_{T-1} = K_T - \Lambda_T \). \( \square \)

**Appendix E. An “overly-twisted” specification**

To help me calibrate the model, I introduce a close sibling to the economy presented in the main text. The evolution of the process here might agitate some readers as I add many unjustifiable twists, the purpose of which is to provide intuitive expressions for many key outputs, such as price-dividend ratio and dividend yield of any maturity. I postulate that the twists doesn’t qualitatively change the results. In fact, since I can calculate the price of assets in both economies, the next step (if asked to) is to compare the pricing difference which I believe is minor.

I start with some familiar aspects. The marginal utility \( M_t \) follows

\[ \frac{M_{t+1}}{M_t} = e^{-\delta} \times \begin{cases} 
1 & \text{Disaster} \\
B^{-\gamma} & \text{Disaster}
\end{cases} \]

where \( \delta = -\log \beta + \gamma \mu_c \).
Dividend is assumed to follow

\[ \frac{D_{t+1}}{D_t} = e^{\mu_D} (1 + g_t) \left( 1 + \sigma^{D_{t+1}} \right) \times \begin{cases} 1 \\ F_{t+1} \text{ Disaster} \end{cases} \]

Same as before, I denote the resilience as \( H_t = p\mathbb{E}^{D_t} \left[ (B^{-\gamma} F_{t+1} - 1) (1 + g_t) \right] \), and separate \( H_t \) into \( H_t = \bar{H} + \hat{H}_t \). Different from the main text, I now assume that the recovery \( g_t \) and the resilience \( H_t \) follow

\[
g_{t+1} = \begin{cases} \frac{1}{1+g_t} e^{-\phi_g} g_t \\ \frac{p+\bar{H}}{p+H_t+pg_t} g^D \text{ Disaster} \end{cases}
\]

\[
\hat{H}_{t+1} = \frac{1 + \bar{H}}{1 + H_t + g_t} e^{-\phi_H} \hat{H}_t + \sigma^H \hat{H}_{t+1}
\]

The weird recovery process is optimized after extensive reverse engineering. Basically, I want the LG moments to be as clean as possible, as we need three moments to calculate the price-to-dividend ratio, which is summarized in Proposition 7.

**Proposition 7.** In a economy where recovery process \( g_t \) jumps to \( \frac{p+\bar{H}}{p+H_t+pg_t} g^D \) in disaster and otherwise the same as the economy in the main text, the price-dividend ratio of stock satisfies

\[
\frac{P_t}{D_t} \big|_{g_t=0, \hat{H}_t=0} = \frac{e^{\delta_D} - (1 + \zeta) \left( e^{-\delta_g} - \zeta \right)}{[e^{\delta_D} - (1 + \zeta)] [e^{\delta_D} - (e^{-\delta_g} - \zeta)]}
\]

and the spot price of dividend satisfies

\[
\frac{S^n_t}{D_t} \big|_{g_t=0, \hat{H}_t=0} = e^{-n(\delta_D - h)} \left( \zeta + 1 - e^{-\delta_g} \right) (1 + \zeta)^n + \zeta \left( e^{-\delta_g} - \zeta \right)^n \\
(1 + \zeta) - (e^{-\delta_g} - \zeta)
\]

where \( h = \ln \left( 1 + \bar{H} \right) \), \( \delta_g = \phi_g + h - \ln (1 - p) \) and \( e^\zeta \geq 1 \) is a parameter given by

\[(e^\zeta - 1) (e^\zeta - e^{-\delta_g}) = \left[ 1 - (1 - p) e^{-h} \right] e^{-h} g^D
\]

Introduction of \( \zeta \) is an alternative way to capture the concept of recovery. When there is no recovery, \( g^D = 0 \) and \( \zeta = 0 \). Otherwise, a higher \( g^D \) corresponds to a larger \( \zeta \). When we evaluate the model conditional on \( g_t = 0 \) and \( \hat{H}_t = 0 \), introducing \( \zeta \) leads to simple
expressions for a few objects of interest.

1. Real risk free rate, as before, equals to

\[ R_f = e^{\delta - \Psi} \]

where \( \Psi \equiv \ln (1 - p + pB^{-\gamma}). \)

2. The expected return of stock is

\[
\mathbb{E}_t \left( \frac{P_{t+1}}{D_{t+1}} + 1 \right) \frac{D_{t+1}}{D_t} - 1 \\
= \left( \frac{e^{\delta - \mu_d - h} - e^\zeta}{e^{\delta - \mu_d - h} - e^\zeta (e^{-\delta_g} - e^\zeta + 1)} + 1 \right) e^{\mu_d} - 1 
\]

3. The yield of one-year dividend strip is

\[
-\ln \frac{S^1_t}{D_t} = -\ln \left( \frac{e^\zeta - e^{-\delta_g}}{e^{\delta - \mu_d - h} [2e^\zeta - 1 - e^{-\delta_g}]} \right) = \delta - \mu_d - h 
\]

whereas the yield of n-year dividend strip is, using \( e^\zeta > 1 + e^{-\delta_g} - e^\zeta, \)

\[
\lim_{n \to \infty} -\frac{1}{n} \log \frac{S^1_t}{D_t} = \delta - \mu_d - h - \zeta 
\]

4. If there is no recovery (\( g^D = 0 \) so \( \zeta = 0 \)), average pd ratio equals to \( \frac{1}{e^{\delta - \mu_d - h} - 1}, \) stock return can be simplified as \( e^{\delta - h} - 1, \) and dividend yield is \( \delta - \mu_d - h \) for any maturity.

Proof of Proposition 7. I rewrite the definition of resilience \( H_t \) as

\[
p \mathbb{E}_t B^{-\gamma} F_{t+1} = \frac{H_t}{1 + g_t} + p 
\]
The first moment $M_{t+1}D_{t+1}$ follows

$\mathbb{E}_t \frac{M_{t+1}D_{t+1}}{M_t D_t} = e^{-\delta + \mu_d} (1 + g_t) \left( 1 - p + p \mathbb{E}_t^D B^{-\gamma} F_{t+1} \right)$

$= e^{-\delta + \mu_d} (1 + g_t) \left( 1 - p + \frac{H_t}{1 + g_t} + p \right)$

$= e^{-\delta + \mu_d} (1 + g_t + H_t)$

The second moment $M_{t+1}D_{t+1} g_{t+1}$ is simplified thanks to the twists

$\mathbb{E}_t \frac{M_{t+1}D_{t+1} g_{t+1}}{M_t D_t} = e^{-\delta + \mu_d} (1 + g_t) \times \left( \left( 1 - p \right) e^{-\phi g} g_t + \left( \frac{H_t}{1 + g_t} + p \right) \left( \frac{p + \bar{H}}{p + H_t + pg_t g^D} \right) \right)$

$= e^{-\delta + \mu_d} \left[ (p + \bar{H}) g^D + (1 - p) e^{-\phi g} g_t \right]$

And finally the third moment

$\mathbb{E}_t \frac{M_{t+1}D_{t+1} \hat{H}_{t+1}}{M_t D_t} = \mathbb{E}_t \frac{M_{t+1}D_{t+1}}{M_t D_t} \mathbb{E}_t \hat{H}_{t+1}$

$= e^{-\delta + \mu_d} (1 + g_t + H_t) \left( 1 + \bar{H} \right) \frac{e^{-\phi h}}{1 + H_t + g_t} \hat{H}_t$

$= e^{-\delta + \mu_d} (1 + \bar{H}) e^{-\phi h} \hat{H}_t$

To summarize, we can write everything in terms of state variables

$\mathbb{E}_t \frac{M_{t+1}D_{t+1}}{M_t D_t} \left( 1, g_{t+1}, \hat{H}_{t+1} \right)' = e^{-\delta + \mu_d + h} \Omega_D \left( 1, g_t, \hat{H}_t \right)'$

where

$\Omega_D = \begin{pmatrix} 1 & e^{-h} & e^{-h} \\ \left[ 1 - (1 - p) e^{-h} \right] g^D & e^{-\delta_g} & 0 \\ 0 & 0 & e^{-\phi H} \end{pmatrix}$

with $h \equiv \ln (1 + \bar{H}) > 0$ and $\delta_g \equiv \phi g + h - \ln (1 - p) > 0$. 
With the help of $\zeta$, I decompose $\Omega_D$ to

$$\Omega_D = \begin{pmatrix}
1 & e^{-h} & e^{-h} \\
e^{h} (e^{\zeta} - 1) (e^{\zeta} - e^{-\delta_g}) & e^{-\delta_g} & 0 \\
e^{h} & e^{-\delta_g} & 0 \\
e^{-\phi_H} & e^{-\delta_g} & 0 \\
0 & 0 & e^{-\phi_H}
\end{pmatrix}$$

$$= \Gamma \begin{pmatrix}
e^{\zeta} - 1 & 0 & 0 \\
e^{-\delta_g} - e^{\zeta} + 1 & 0 \\
e^{-\phi_H} & e^{-\delta_g} + e^{\zeta} - 1
\end{pmatrix} \Gamma^{-1}$$

with

$$\Gamma = \begin{pmatrix}
1 & e^{-h} & e^{-h} (e^{-\phi_H} - e^{-\delta_g}) \\
(e^{\zeta} - 1) e^{h} & e^{-h} & - (e^{\zeta} - 1) (e^{-\delta_g} - e^{\zeta}) \\
0 & 0 & (e^{-\phi_H} - e^{\zeta}) (e^{-\phi_H} - e^{-\delta_g} + e^{\zeta} - 1)
\end{pmatrix}$$

The (ex-dividend) price-dividend ratio can be explicitly written as

$$\frac{P_t}{D_t} = \left( \begin{array}{ccc}
1 & 0 & 0
\end{array} \right) \Omega_D (e^{-\delta_D - \delta_H I_3} - \Omega_D)^{-1} \left( \begin{array}{c} 1, g_t, \hat{H}_t \end{array} \right)'$$

$$= \frac{e^{\delta - \mu_d - h} - e^{\zeta} (e^{-\delta_g} - e^{\zeta} + 1)}{e^{\delta - \mu_d - h} - e^{\zeta} (e^{-\delta_g} - e^{\zeta} + 1)}$$

$$+ \frac{e^{\delta - \mu_d - h} - e^{\zeta} (e^{-\delta_g} - e^{\zeta} + 1)}{e^{\delta - \mu_d - h} - e^{\zeta} (e^{-\delta_g} - e^{\zeta} + 1)} g_t$$

$$+ e^{\delta - \mu_d - 2h} \left( e^{-\delta_g} - e^{-\delta_D} e^{-\delta_H} \right) \hat{H}_t$$

We know that the spot price of dividend strip is

$$\frac{S^n}{D_t} = \left( \begin{array}{ccc}
1 & 0 & 0
\end{array} \right) e^{-n(\delta - \mu_d - h)} \Gamma \left( \begin{array}{ccc}
(e^{\zeta})^n & 0 & 0 \\
0 & (e^{-\delta_g} - e^{\zeta} + 1)^n & 0 \\
0 & 0 & e^{-n\phi_H}
\end{array} \right) \Gamma^{-1} \left( \begin{array}{c} 1, g_t, \hat{H}_t \end{array} \right)'$$

After some tedious algebra, the constant term equals

$$\frac{S^n}{D_t} \bigg|_{g_t=0, \hat{H}_t=0} = e^{-n(\delta - \mu_d - h)} \times \frac{(e^{\zeta} - e^{-\delta_g}) (e^{\zeta})^n + (e^{\zeta} - 1) (e^{-\delta_g} - e^{\zeta} + 1)^n}{e^{\zeta} - (e^{-\delta_g} - e^{\zeta} + 1)}$$
The loading in front of $g_t$ is

$$\frac{\partial S^n_t}{\partial g_t} = e^{-n(\delta-\mu_d-h)-h} \times \frac{(e^\zeta)^n - (e^{-\delta_g - e^\zeta + 1})^n}{e^\zeta - (e^{-\delta_g - e^\zeta + 1})}$$

And the loading of $\hat{H}_t$ is

$$\frac{\partial S^n_t}{\partial \hat{H}_t} = e^{-n(\delta-\mu_d-h)-h} \times \left\{ \frac{1 - e^\zeta}{[e^\zeta - (e^{-\delta_g - e^\zeta + 1})](e^{-\phi_H - (e^{-\delta_g - e^\zeta + 1})})} \right\}$$

- \left\{ \frac{(1 - e^\zeta)(e^{-\delta_g - e^\zeta + 1})^n}{(1 - (e^{-\delta_g - e^\zeta + 1}))(e^\zeta)^n} \right\}

+ \left\{ \frac{(e^{-\phi_H - e^\zeta})}{(e^{-\phi_H - e^{-\delta_g}})(e^{-\phi_H - (e^{-\delta_g - e^\zeta + 1})})} \right\}$$