Abstract

This thesis studies currency crashes and risk premia from the perspective of options, as well as the pricing of Contingent Convertible Bonds (CoCo).

In the first part, we set up an arbitrage-free time-changed Lévy model inspired by Carr and Wu (2004) which generates yield curves and exchange rate dynamics consistent with currency option prices in G10 countries. In this framework, we investigate the pricing of cross-pair G10 currency options, and construct estimates of conditional currency risk premia at each point in time in our sample from January 1999 to June 2012. We discover the option-implied currency risk premia are not distinguishable from their realized counterpart, suggesting the realized returns do not suffer from peso problems. We also decompose the total risk premium into their contributions from variance and higher order cumulants, and discover the higher cumulants’ contribution merely represents 15% of the total risk premium.

In the second part, we study a typical hybrid instrument (CoCo) designed to relieve banks from the pressure of raising new capital during crisis periods. CoCos convert into equities upon a contractual trigger event, and are exposed to different sources of risk: interest rate risk, equity risk and conversion risk. We develop a general framework for their pricing and hedging that can be specified in different ways. We focus on reduced-form and structural models driven by a finite-dimensional Markov process. But both allow to price CoCos and calculate dynamic hedging strategies with holdings in related instruments such as fixed income products, the issuing company’s stock and credit default swaps. We conclude the study of CoCo’s with a discussion on several issues in its design using the framework of Leland’s model (1996).
Acknowledgements

It is a pleasure to thank those who made this thesis possible. I would like to thank my Ph.D. supervisor, Prof. Patrick Cheridito. His great enthusiasm and extensive knowledge about financial mathematics is tremendously encouraging even since the beginning of my graduate studies. I sincerely appreciate his patience and effort in guiding me. This thesis would not have been possible without him.

It is difficult to overstate my gratitude to my co-adviser Prof. Jakub Jurek. I have learned a lot from his insightful thinking and profound understanding of finance. His passion about empirical evidence and "engineering" desire greatly influenced me and his collaborative teaching has made my way becoming an empiricist possible.

I wish to express my sincere thanks to Prof. Ronnie Sircar for reading my thesis and his constructive comments. I thank Prof. René Carmona for being in my thesis committee. I am also grateful to Prof. Jianqing Fan for encouraging and personal guidance.

My warm thanks are due to Matt Lorig for his extensive discussions on various topics in financial mathematics. I wish to extend my thanks to all those who have helped me with my thesis at Princeton University.

The financial support of the Gordon Y.S. Wu Fellowship in Engineering is gratefully acknowledged. I am lucky to meet these warm-hearted and intelligent people at the Department of Operations Research and Financial Engineering, who have helped me through the difficult times.

Lastly, and most importantly, I wish to thank my parents. They bore me, raised me, supported me, taught me, and loved me. To them I dedicate this thesis.
To my parents.
Contents

Abstract .......................................................... iii
Acknowledgements .................................................. iv
List of Tables ....................................................... x
List of Figures ...................................................... xi

1 Introduction ...................................................... 1

2 Time-changed Lévy Pricing Kernel, Exchange Rates and Term Structure .................................................. 4

2.1 Pricing Kernels, Currency Risk Premia, and FX Option Pricing ..................................................... 4

2.2 Static Model .................................................... 10

2.2.1 Pricing kernels .............................................. 10

2.2.2 Exchange rates and risk premia ........................ 12

2.2.3 FX option pricing .......................................... 19

2.2.4 Recovering currency risk premia ....................... 21

2.2.5 Shock parametrization .................................... 27

2.3 Data and Model Calibration .................................. 29

2.3.1 Implementation .......................................... 30

2.3.2 Model parameters ........................................ 33

2.3.3 Fitting errors ............................................. 36

3 Currency Risk Premia ........................................ 39
<table>
<thead>
<tr>
<th>Section Number</th>
<th>Section Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Option-Implied Currency Risk Premia</td>
<td>39</td>
</tr>
<tr>
<td>3.1.1</td>
<td>Forecasting currency returns</td>
<td>40</td>
</tr>
<tr>
<td>3.1.2</td>
<td>The $HML_{FX}$ and short dollar risk premia</td>
<td>42</td>
</tr>
<tr>
<td>3.1.3</td>
<td>The factor structure of currency returns</td>
<td>49</td>
</tr>
<tr>
<td>3.2</td>
<td>Discussion</td>
<td>53</td>
</tr>
<tr>
<td>3.2.1</td>
<td>The dynamics of global factor loadings</td>
<td>53</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Option-pricing implications of linking loadings to interest rates</td>
<td>55</td>
</tr>
<tr>
<td>3.3</td>
<td>Conclusion</td>
<td>58</td>
</tr>
<tr>
<td>4</td>
<td>Pricing and Hedging CoCos with Intensity-Based Models and First-Passage Time Models</td>
<td>60</td>
</tr>
<tr>
<td>4.1</td>
<td>General formulas for pricing, calibration and hedging</td>
<td>64</td>
</tr>
<tr>
<td>4.2</td>
<td>Intensity-based models</td>
<td>70</td>
</tr>
<tr>
<td>4.2.1</td>
<td>CoCo and CDS pricing with intensities</td>
<td>72</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Hedging in intensity-based models</td>
<td>76</td>
</tr>
<tr>
<td>4.3</td>
<td>First-passage time models</td>
<td>81</td>
</tr>
<tr>
<td>4.3.1</td>
<td>CoCo and CDS pricing with PDEs</td>
<td>83</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Hedging in first-passage time models</td>
<td>86</td>
</tr>
<tr>
<td>5</td>
<td>Case Study: Rabobank and Lloyds Banking Group CoCos</td>
<td>89</td>
</tr>
<tr>
<td>5.1</td>
<td>Rabobank and Lloyds Banking Group CoCos</td>
<td>89</td>
</tr>
<tr>
<td>5.2</td>
<td>Intensity-based model with a CIR process</td>
<td>90</td>
</tr>
<tr>
<td>5.3</td>
<td>First-passage time model with an exponential OU accounting ratio</td>
<td>94</td>
</tr>
<tr>
<td>5.4</td>
<td>Conclusion</td>
<td>97</td>
</tr>
<tr>
<td>6</td>
<td>CoCo Design and Financial Stability</td>
<td>98</td>
</tr>
<tr>
<td>6.1</td>
<td>Bank with CoCos and Regulator</td>
<td>99</td>
</tr>
<tr>
<td>6.1.1</td>
<td>Unlevered Asset Value</td>
<td>99</td>
</tr>
<tr>
<td>6.1.2</td>
<td>Liabilities</td>
<td>100</td>
</tr>
</tbody>
</table>
List of Tables

6.1 Table I Calibrated Model Parameters .............................. 135
6.2 Table II Forecasting Currency Excess Returns .................. 136
6.3 Table III Empirical Factor Mimicking Portfolios ............... 137
6.4 Table IV Model Risk Premium Decompositions .................. 138
6.5 Table V The Factor Structure of Currency Returns ............... 139
6.6 Table VI The Dynamics of Global Factor Loadings .............. 140
6.7 Table VII Linking Global Factor Loadings to Interest Rate

Differentials ............................................................... 141
6.8 Table VIII Risk-free Yields and CDS Spreads .................... 142
6.9 Table IX Calibrated Parameters .................................... 143

A.1 Table A.I Country-Specific Innovation Parameters .......... 145
A.2 Table A.III Forecasting Currency Excess Returns with

Option-Implied Risk Premia (Specification II) ..................... 147
A.3 Table A.IV Empirical $HML_{FX}$ Factor Mimicking Portfolio

Risk Premium Decomposition: Alternative specifications .... 148
List of Figures

6.1 Figure 1. The Time Series of Fitted Volatilities. .................. 125
6.2 Figure 2. The Cross-Section of Fitted Option-Implied Volatilities. ................................................................. 126
6.3 Figure 3. The Cross-Section of Model and Realized Currency Risk Premia. .......................................................... 127
6.4 Figure 4. Option-Implied Currency Risk Premia .................. 128
6.5 Figure 5. Decompositions of Option-Implied Currency Risk Premia. ........................................................ 129
6.6 Figure 6. \(\Phi\)-Survival Probabilities, CDS Spreads and CoCo Prices for Intensity-Based Model. ......................... 130
6.7 Figure 7. CoCo Value Decompositions in the Intensity-Based Model with CIR Jump Intensity. ......................... 131
6.8 Figure 8. \(\Phi\)-Survival Probabilities, CDS Spreads and CoCo Prices for First Passage Time Model. ......................... 132
6.9 Figure 9. CoCo Value Decompositions in the First Passage Time Model with Exponential OU Accounting Ratio. 133
Chapter 1

Introduction

The financial world has been violently shaken by the financial crisis in 2008. The crisis found its origin in defaults in the mortgage market, and had rapidly spread via securitized products in the credit markets and was followed by the failure of numerous large financial institutions (AIG, Lehman Brothers and Bear Stearns). Structural changes entailed by this crisis in various markets can still be seen today: the crisis officially set an end to the glory of Collaterized Debt Obligations (CDOs) and other structured products of similar flavor, and nominal interest rates in most of the developed economies stagger close to zero. Meanwhile, both commodity and currency markets have become more correlated with equity markets, and variance risk premia have become significantly higher in currency options since 2007.

Investors are more concerned about the so-called “Black Swan” event, paying more attention to tail risks they bear either consciously or unconsciously in their portfolio. The crucial characteristics of an investment opportunity are its expected return and its underlying risk. While variance, as a proxy for risk, can be tightly estimated from the data, the expected return - risk premium - suffers from more serious estimation problems. One of the highlighted difficulties in estimation is the peso problem (i.e. the observed sample being too specific to contain a large enough
number of bad realizations). This mission is almost impossible when it comes to the estimation of time-varying conditional risk premium. The first part of this thesis addresses this issue for currencies. We analyze the cross-section of currency options to extract estimates of conditional currency risk premia without reliance on realized returns in history. The methodology is directly comparable to Bakshi et al. (2008), but is in contrast to Fahri et al. (2013), and also related to the literature studying equity index options (Pan (2002), Santa Clara and Yan (2010)). We show that the historical realized return for the currency carry trade - a popular currency trading strategy which consists in longing high interest currencies by shorting low interest ones - agrees with our option-implied risk premia, and thus does not suffer from peso problem. We identify that the premium contributed by crash risks accounts for no more than 20% of the total risk premium despite the fat-tailed nature of currency returns.

On the other hand, regulators implement harsher capital requirements and envision new financial instruments to reduce systemic risk in financial system and attempt to avoid, or at least mitigate the Too-Big-To-Fail (TBTF) problem for systematically important financial institutions (SIFI). One of, if not the most, popular new instrument is the contingent convertible bond, adopted by many European banks and warmly welcomed by the investors seeking high coupon rates. CoCo falls into the category of hybrid products. It is exposed to interest rate risk, credit risk as well as equity risk of the issuer, posing new challenges for pricing and hedging. In the second part of this thesis, we present a general framework to solve this problem. While existing literature on CoCo pricing mainly focuses on the structural model approach, our framework can be flexibly adapted to both the structural approach and reduced-form approach. The latter enjoys huge popularity among practitioners thanks to the seminal works by Jarrow and Turnbull (1996), Duffie (1998) and Duffie and Singleton (1999).
The thesis is organized as follows: Chapter 2 presents our FX option pricing model and derives its implications on term structure, option prices and currency risk premia; Chapter 3 discusses the calibration procedure and empirical results; Chapter 4 introduces our CoCo pricing framework; Chapter 5 provides several possible model implementations within the general framework and Chapter 6 addresses concerns about the design of CoCo’s.
Chapter 2

Time-changed Lévy Pricing
Kernel, Exchange Rates and Term Structure

2.1 Pricing Kernels, Currency Risk Premia, and FX Option Pricing

We develop a parsimonious model of exchange rates based on a specification of pricing kernel dynamics driven by a combination of global and country-specific shocks, and derive its implications for currency risk premia and the prices of foreign exchange (FX) options. The underlying model parameters can be identified using cross-sectional FX option data and historical variance forecasts, allowing us to compute \textit{ex ante} estimates of option-implied currency risk premia. Our approach provides a novel perspective on the dynamics of exchange rates, complementing a vast literature using historical currency returns.

According to Carr and Wu (2004), a simple way to construct a dynamic model which allows to incorporate both stochastic volatility and skewness is to use time-
changed Lévy process. Bakshi, et al. (2008) employ this technique to build a two-factor model for pricing kernels and investigate international risk-sharing between United States, United Kingdom and Japan. Our model is a static one inspired by their work, and this section reviews their convenient dynamic framework. Our exposition of this dynamic framework is slightly different from their original paper as we allow for stochastic interest rates and endogenize the possible correlation between interest rates and exchange rates. This feature has been extensively documented in a large body of literature on uncovered interest rate parity (UIP).

The dynamic model begins with a specification of country-level log pricing kernel dynamics of the form:

\[ m^i_t = -\alpha^i t - \xi^i L^g_{\Pi^i_t} - L^i_{\Lambda^i_t}, \quad (2.1.1) \]

where \( L^g_{\Pi^i_t} \) and \( L^i_{\Lambda^i_t} \) are independent time-changed Lévy processes which model respectively a global risk factor commonly shared across all the countries and a country-specific risk factor. Different countries have different loadings \( \xi^i \) on the global factor.

The time change \( \Pi^i_t \) is common to all the countries, corresponding to the stochastic variance in the global component \( L^g \), and the instantaneous stochastic variance process \( Z_t \) is modeled by a square-root process (CIR process as in Cox, et al. (1985)),

\[ \Pi^i_t = \int_0^t Z_s ds \]

\[ dZ_t = \kappa_Z (\theta_Z - Z_t) dt + \sqrt{Z_t} dW^Z_t, \]

where \( W^Z_t \) is a Brownian Motion.

Similarly, the time change \( \Lambda^i_t \) is country-specific, modeled by

\[ \Lambda^i_t = \int_0^t Y^i_s ds \]

\[ dY^i_t = \kappa^i_Y (\theta^i_Y - Y^i_t) dt + \sqrt{Y^i_t} dW^Y^i_t, \]
where \((W_t^Z)_{t \geq 0}\) (resp. \((W_t^{Y,i})_{t \geq 0}\)) is independent from \((L_t^i)_{t \geq 0}\) and \((W_t^{Y,i})_{t \geq 0}\) (resp. \((L^g_t)_{t \geq 0}\) and \((W_t^Z)_{t \geq 0}\)), but can be correlated with the diffusive part of \((L_t^g)_{t \geq 0}\) (resp. \((L_t^i)_{t \geq 0}\)) and \((W_t^Z)_{t \geq 0}\) (resp. \((W_t^{Y,i})_{t \geq 0}\)). We suppose that the correlation between the diffusive part of \((L_t^g)_{t \geq 0}\) (resp. \((L_t^i)_{t \geq 0}\)) and \((W_t^Z)_{t \geq 0}\) (resp. \((W_t^{Y,i})_{t \geq 0}\)) is constant \(\rho_Z\) (resp. \(\rho_{Y}^i\)).

We denote the pricing kernel by \(M_t^i := \exp(m_t^i)\). As in Bakshi, et al. (2008), the exchange rate dynamics are determined by the ratio of pricing kernels, mathematically,

\[
 s_{t+\tau}^{ji} - s_t^{ji} = (m_{t+\tau}^i - m_t^i) - (m_{t+\tau}^i - m_t^i) \\
 = (\xi^i - \xi^j)(L_t^{g} - L_t^{g}) + (L_t^{i} - L_t^{i}) - (L_t^{j} - L_t^{j}),
\]

where \(s_t^{ji}\) denotes the log exchange rate between country \(J\) and \(I\) (the price of currency \(J\) in units of currency \(I\)).

The model derivation relies on the knowledge of the cumulant generating functions \(^1\) of \(L_t^g\) and \(L_t^i\). As they are closely linked to characteristic functions, FX options can also be priced using generalized Fourier transform method given the cumulant generating functions. We define the following notations for cumulant generating functions to characterize the innovations:

\[
 k_{L^g,t+\tau}^z[u] = \ln E_t^p [e^{u(L_{t+\tau}^g - L_t^g)}] = \ln E_t^p [e^{u(L_{t+\tau}^g - L_t^g)}| \mathcal{Z}_t = z] \\
 k_{L^i,t+\tau}^y[u] = \ln E_t^p [e^{u(L_{t+\tau}^i - L_t^i)}] = \ln E_t^p [e^{u(L_{t+\tau}^i - L_t^i)}| \mathcal{Y}_t = y].
\]

\(^1\)Recall that the cumulant generating function of a random variable, \(\epsilon_{t+1}\), is defined as follows:

\[
 k_{\epsilon} [u] = \ln E_t [\exp(u \cdot \epsilon_{t+1})] = \sum_{j=1}^{\infty} \frac{k_j \cdot u^j}{j!}
\]

where \(k_j\) are the cumulants of the random variable, which can be computed by taking the \(j\)-th derivative of \(k_{\epsilon} [u]\) and evaluating the resulting expression at zero. The cumulant generating function of the sum of two independent random variables is equal to the sum of their cumulant generating functions.
The yield from \( t \) to \( t + \tau \) can be expressed in terms of cumulant generating functions, since the specified pricing kernel needs to price risk-free claims:

\[
y^{i}_{t,t+\tau} = -\frac{1}{\tau} \ln \frac{E^{\mathbb{P}}[M_{i}^{t+\tau}]}{M_{i}^{t}}
\]

\[
= -\frac{1}{\tau} \ln E^{\mathbb{P}}[M_{i}^{t+\tau}|Z_{t}, Y^{i}]
\]

\[
= \alpha^{i} - \frac{1}{\tau} k_{L^{z},\tau}^{i}[\xi] - \frac{1}{\tau} k_{L^{y},\tau}^{i}[-1]. \quad (2.1.2)
\]

We will show, under model assumption, that \( k_{L^{z},\tau}^{i}[u] \) (resp. \( k_{L^{y},\tau}^{i}[u] \)) is affine in \( z \) (resp. \( y \)). Therefore, the yield curves generated by this model are driven by two independent CIR state variables. The term structure falls in the \( A_{2}(2) \) family under Dai and Singleton (2000)’s classification.

As for option pricing, we need to proceed under the forward measure as interest rates are stochastic. For an option on the exchange rate \( S_{t}^{ji} \) with maturity \( \tau \), the payoff will be denominated in currency \( I \) at maturity \( t + \tau \). To price this contingent claim, we can define country \( I \)’s \( \tau \)−forward measure \( \mathbb{F}^{i} \) to be:

\[
d\mathbb{F}^{i} = \frac{M_{i}^{t+\tau}}{M_{i}^{t}} e^{y^{i}_{t,t+\tau}.\tau}.
\]

Under this forward measure, the cumulant generating functions can be computed as a function of cumulant generating functions under the historical measure.

For the global factor,

\[
k_{L^{z},\tau}^{F^{i}}[u] = \ln E^{\mathbb{F}^{i}}[e^{uL_{tt}^{0}}|Z_{0} = z]
\]

\[
= \ln E^{\mathbb{F}^{i}}[e^{y_{t,\tau}^{i}+m_{t}^{i}+uL_{tt}^{0}}|Z_{0} = z]
\]

\[
= \ln E^{\mathbb{F}^{i}}[e^{-k_{L^{z},\tau}^{i}[\xi]-\xi L_{tt}^{0} + uL_{tt}^{0}} | Z_{0} = z]
\]

\[
= k_{L^{z},\tau}^{i}[u - \xi] - k_{L^{z},\tau}^{i}[-\xi].
\]
For country $I$’s country-specific factor, 

\[
\begin{align*}
    k_{L_i,\tau}^{y_i}[u] &= \ln E^{\mathbb{F}_t} \left[ e^{u L_{i}\tau} | Y_0^i = y \right] \\
    &= \ln E^{\mathbb{P}} \left[ e^{y_i \tilde{\tau} + m_i^i + u L_{i}\tau} | Y_0^i = y \right] \\
    &= \ln E^{\mathbb{P}} \left[ e^{-k_{L_i,\tau}^{y}[-1] - L_{i}\tau + u L_{i}\tau} | Y_0^i = y \right] \\
    &= k_{L_i,\tau}^{y}[-1] - k_{L_i,\tau}^{y}[-1].
\end{align*}
\]

And by independence, all the other country-specific factors are unaffected by the measure change. Therefore, we obtain the cumulant generating function for the log exchange rate $s_{t+\tau}^{ji}$ under the forward measure as:

\[
\begin{align*}
    k_{s^{ji},\tau}^{F}[u] &= k_{L^{ji},\tau}^{Z}[(\xi^{i} - \xi^{j})u - \xi^{i}] - k_{L^{ji},\tau}^{Y}[-\xi^{i}] + k_{L^{ji},\tau}^{Y}[-1] - k_{L^{ji},\tau}^{Y}[-1] + k_{L^{ji},\tau}^{Y}[-u].
\end{align*}
\] (2.1.3)

The knowledge of cumulant generating function of log exchange rate under forward measure is sufficient to derive option prices by using generalized Fourier transform methods (e.g. Carr and Madan (1999)). It only remains to compute the cumulant generating functions $k_{L^{i},\tau}^{x}[u]$ and $k_{L^{i},\tau}^{y}[u]$. The computation is made possible thanks to Theorem 1 in Carr and Wu (2004). The theorem shows that for any time-changed Lévy process $L_{T_{\tau}}$, if the cumulant generating function of the process $L_{t}$ is $k_{L}[u] = \frac{1}{2} \ln E[e^{u L_{t}}]$, then the cumulant generating function of the process $L_{T_{\tau}}$ is given by $k_{T_{\tau}}^{u}[k_{L}[u]]$ where $k_{T_{\tau}}^{u}$ denotes the cumulant generating function of the time-change process $(T_{t})_{t\geq0}$ under the measure $\mathbb{P}^{u}$ defined by:

\[
\frac{d\mathbb{P}^{u}}{d\mathbb{P}} = \exp(u' L_{T_{\tau}} - T_{\tau} k_{L}[u]).
\]
Applying this result to our case,

\[ k_{L^s,\tau}[u] = k_{T,\tau}^{z,P^u} [k_{L^s}[u]] \]
\[ k_{L^l,\tau}[u] = k_{T,\tau}^{y,P^u} [k_{L^l}[u]], \]

where \( k_{T,\tau}^{z,P^u} \) (resp. \( k_{T,\tau}^{y,P^u} \)) denotes the cumulant generating of \( \Pi_\tau \) (resp. \( \Lambda^i_\tau \)) under measure \( P^u \) (resp. \( P^u \)).

By Girsanov’s theorem, under \( P^u \), the \( (Z_t)_{t\geq 0} \) process remains a CIR process, with parameters \( \kappa_Z(u) = \kappa_Z - \rho_Z \omega_Z u \), \( \theta_Z(u) = \frac{\kappa \theta \omega}{\kappa_Z(u)} \), and \( \omega = \omega_Z \). Similarly, under \( P^u \), \( (Y^i_t)_{t\geq 0} \) is still a CIR process, with parameters \( \kappa_Y^i(u) = \kappa_Y^i - \rho_i^Y \omega_Y^i u \), \( \theta_Y^i(u) = \frac{\kappa_Y^i \theta}{\kappa_Y^i(u)} \), and \( \omega = \omega_Y^i \).

Hence \( k_{T,\tau}^{z,P^u} \) and \( k_{T,\tau}^{y,P^u} \) are cumulant generating functions of time-changes defined as integral of CIR state variable. For such time-changes, if the CIR state variable’s dynamic is characterized by parameters \( \kappa \) (mean-reversion coefficient), \( \theta \) (long term mean), and \( \omega \) (volatility), its cumulant generating function is known in closed-form as:

\[ k_{T,\tau}^x[u] = -b_r(u)x - c_r(u), \]

with

\[ b_r(u) := \frac{-2u(1 - e^{-\eta(u)\tau})}{2\eta(u) - (\eta(u) - \kappa)(1 - e^{-\eta(u)\tau})}, \]
\[ c_r(u) := \frac{\kappa \theta}{\omega^2} \left( 2 \ln \left( 1 - \frac{\eta(u) - \kappa}{2\eta(u)(1 - e^{-\eta(u)\tau})} \right) + (\eta(u) - \kappa)\tau \right) \]
\[ \eta(u) := \sqrt{\kappa^2 - 2\omega^2u}. \]
Therefore, the cumulant generating function of the time-changed Lévy innovations can be obtained in closed form when $(L^g_t)_{t \geq 0}$ and $(L^i_t)_{t \geq 0}$ are chosen to be Lévy processes with closed-form cumulant generating function.

This refinement of the model in Bakshi, et al. (2008) captures the joint dynamic of the interest rate and exchange rate, and allows us to price FX options. To construct estimates of currency risk premia for the G10 currencies, the model will be required to price options on as many as 45 currency pairs simultaneously. For parsimony, we simplify the time-changed model by shutting down the time change, turning the dynamic model into a static model which leaves the state variables’ dynamics unspecified.

2.2 Static Model

2.2.1 Pricing kernels

Our model begins with a specification of country-level log pricing kernel dynamics of the form:

\[
\begin{align*}
    m^i_{t+1} - m^i_t &= - \left( y^i_{t,t+1} \cdot \Delta + k_{L^g_i t+1} [-\xi^i_t] + k_{L^i_i t+1} [-1] \right) - \xi^i_t \cdot L^g_{t+1} - L^i_{t+1} \\
\end{align*}
\]  

(2.2.4)

where $y^i_{t,t+1}$ is the yield on a one-period, risk free bond, $\Delta$ is the time increment, and $L^g_{t+1}$ and $L^i_{t+1}$ are independent non-Gaussian shocks. The $L^g_{t+1}$ shock is global, and is common to all countries; the $L^i_{t+1}$ shocks are country-specific, and are cross-sectionally independent. Since our model is written in discrete-time, we remain agnostic whether the non-Gaussianity of the pricing kernel innovations is due to jumps or stochastic volatility within the observation interval. The terms $k_{L_{t+1} [u]}$ represent convexity corrections, which ensure that the pricing kernel correctly prices the one-period bond: $E_t [M^i_{t+1}] = M^i_t \cdot \exp (-y^i_{t,t+1} \cdot \Delta)$, where $M^i_{t+1} = \exp (m^i_{t+1})$. They are given by
the cumulant generating functions (CGF), \( k_{L_{t+1}}[u] \), of the shocks, \( L^g_{t+1} \) and \( L^i_{t+1} \), evaluated at \( u = -\xi^i_t \) and \( u = -1 \), respectively. Recall, the CGF of a random variable, \( L_{t+1} \), is defined as follows:

\[
k_{L_{t+1}}[u] = \ln E_P^t[\exp(u \cdot L_{t+1})]
\] (2.2.5)

For example, if \( L_{t+1} = \sigma \cdot \sqrt{\Delta} \cdot \varepsilon_{t+1} \), and \( \varepsilon_{t+1} \) is a standardized Gaussian innovation, we recover the “typical” Jensen term: \( k_{L_{t+1}}[u] = \frac{\sigma^2 \Delta}{2} \). For parsimony, CGFs of random variables under the objective (\( \mathbb{P} \)) measure are reported without superscripts.

Aside from the factor structure, the other key ingredients of our specification are: (a) asymmetric country-level loadings on global innovations, \( \xi^i_t \); and, (b) pricing kernel innovations with time-varying distributions. Asymmetric global factor loadings contribute to the determination of currency risk premia, and play a role in rationalizing violations of uncovered interest rate parity (Backus, et al. (2001)). They have recently been incorporated into exchange rate option pricing models (Bakshi, Carr, and Wu (2008)), observed in the factor structure of realized currency returns (Lustig, et al. (2011, 2014)), and microfounded in the context of models with imperfect risk-sharing (Verdelhan (2010) and Ready, et al. (2013)). Finally, the distributions of the pricing kernel shocks, \( L^g_{t+1} \) and \( L^i_{t+1} \), are assumed to be controlled by global \( (Z_t) \) and country-specific \( (Y^i_t) \) state variables, respectively. The time-varying state variables measure the periodic variance of the shocks, and also influence their higher-order moments. Specifically, our parametrization assumes:

\[
\begin{align*}
    k_{L^g_{t+1}}[u] &= k_{\tilde{L}^g_{t+1}}[u] \cdot Z_t \\
    k_{L^i_{t+1}}[u] &= k_{\tilde{L}^i_{t+1}}[u] \cdot Y^i_t
\end{align*}
\] (2.2.6a) (2.2.6b)

where \( \tilde{L}^g_{t+1} \) and \( \tilde{L}^i_{t+1} \) are non-Gaussian increments with unit variance. This specification is formally known as a time-change (Carr and Wu (2004)), and is distinct
from a multiplicative scaling, \( L^g_{t+1} = \sqrt{Z_t} \cdot \tilde{L}^g_{t+1} \) and \( L^i_{t+1} = \sqrt{Y^i_t} \cdot \tilde{L}^i_{t+1} \). Although the multiplicative scaling also produces shocks with periodic variances of \( Z_t \) and \( Y^i_t \), respectively, the state variables do not affect the higher moments of the shocks. Since our derivations are largely independent of the technical details of the specification, we defer their discussion to Appendix A. The dynamics of the global factor loadings and state variables are left unspecified, and are recovered period-by-period from cross-sectional asset price data. Taken together, these features create a channel for a risk-based explanation of UIP violations within an affine pricing kernel framework, while generalizing it to account for salient features of exchange rate data, such as non-Gaussian option-implied distribution and stochastic variation in volatility and higher moments.\(^2\)

2.2.2 Exchange rates and risk premia

No arbitrage requires that the price of an asset whose payoff is denominated in currency \( I \) be identical when valued by an investor in country \( I \), and an investor in country \( J \), once the payoff has been converted to units of currency \( J \). A symmetric restriction applies to assets denominated in currency \( J \), such that the exchange rate, \( S^{ji}_{t+1} \) – measured as the currency \( I \) price of currency \( J \) – and the asset payoffs must simultaneously satisfy:

\[
E^p_t \left[ \frac{M^i_{t+1}}{M^j_t} \cdot x^i_{t+1} \right] = E^p_t \left[ \frac{M^j_{t+1}}{M^j_t} \cdot \left( \frac{S^{ji}_{t+1}}{S^{ji}_t} \right)^{-1} \cdot x^i_{t+1} \right] \quad (2.2.7a)
\]

\[
E^p_t \left[ \frac{M^i_{t+1}}{M^j_t} \cdot \frac{S^{ji}_{t+1}}{S^{ji}_t} \cdot x^j_{t+1} \right] = E^p_t \left[ \frac{M^j_{t+1}}{M^j_t} \cdot x^j_{t+1} \right] \quad (2.2.7b)
\]

\(^2\)Our model can formally be viewed as a simplified version of the time-changed Lévy framework of Bakshi, et al. (2008), in which state variables are constant within each observation interval. In particular, it allows for non-Gaussian innovations in both global and idiosyncratic shocks, unlike the conditionally Gaussian setup in Lustig, et al. (2011, 201), or the jump-diffusive framework in Farhi, et al. (2014), in which idiosyncratic innovations are Gaussian. Section 1 of the supplementary online Technical Appendix derives the continuous-time analog of our model.
where \( x_{i}^{t+1} \) and \( x_{j}^{t+1} \) are payoffs of the traded assets denominated in currencies \( I \) and \( J \), respectively. We model the exchange rate as the ratio of the pricing kernels in the two economies, \( \frac{M_{j}^{t+1}}{M_{i}^{t+1}} \), which is sufficient (necessary) to satisfy no arbitrage when markets are incomplete (complete). As a result, log currency returns are given by:

\[
 s_{i}^{j} - s_{i}^{i} = (m_{j}^{t+1} - m_{j}^{t}) - \left( m_{i}^{t+1} - m_{i}^{t} \right) \\
= -\alpha_{i}^{j} + \alpha_{i}^{i} - (\xi_{i}^{j} - \xi_{i}^{i}) \cdot L_{i}^{g} - L_{i}^{j} + L_{i}^{i} 
\]

where: \( \alpha_{i}^{i} = y_{i}^{t+1} \cdot \Delta + k_{L_{i}^{t+1}} [ -\xi_{i}^{i} ] + k_{L_{i}^{t+1}} [-1] \). Thus, the model generates a factor structure in currency returns. In particular, if currency returns were all measured with respect to a single reference currency (e.g. the U.S. dollar), a principal components analysis would recover a reference currency factor (\( L_{i}^{j+1} \)) and the global factor (\( L_{i}^{i+1} \)), as the two major drivers of currency returns. These features are consistent with empirical evidence established by Lustig, et al. (2011). Finally, the conditional variance of the log currency return under the objective measure (\( \mathbb{P} \)) is given by:

\[
 Var_{t}^{p} [ s_{i}^{j} - s_{i}^{i} ] = (\xi_{i}^{j} - \xi_{i}^{i})^{2} \cdot Z_{t} + Y_{t}^{j} + Y_{t}^{i} 
\]

The central quantity of interest in our investigation is the expected excess return (risk premium) earned by an investor who borrows funds in currency \( I \) and invests

---

3When markets are complete the two restrictions must hold state-by-state (i.e. for each Arrow-Debreu security), such that the exchange rate is uniquely pinned down by the ratio of the pricing kernels (Fama (1984), Dumas (1992), Backus, et al. (2001), Brandt, et al. (2006), Bakshi, et al. (2008)). Graveline and Burnside (2013) highlight the sensitivity of inferences about real exchange rate determination and risk-sharing in models where the exchange rate is represented as a ratio of pricing kernels (“asset-market view”) to the underlying assumptions about frictions and preferences.
them in currency $J$ for one period:

$$
E_{t}^P \left[ x_{s,t,t+1}^{ji,i} \right] = E_{t}^P \left[ \exp \left( y_{t,t+1} \cdot \Delta \right) \cdot \frac{S_{t+1}^{ji}}{S_{t}^{ji}} - \exp \left( y_{t,t+1} \cdot \Delta \right) \right] \\
= \exp \left( \ln E_{t}^P \left[ \frac{S_{t+1}^{ji}}{S_{t}^{ji}} \right] + y_{t,t+1} \cdot \Delta \right) - \exp \left( y_{t,t+1} \cdot \Delta \right) \\
= \exp \left( y_{t,t+1} \cdot \Delta \right) \cdot \left( \exp \left( k_{t+1}^{ji} [1] - s_{t}^{ji} + (y_{t,t+1} - y_{t,t+1}) \cdot \Delta \right) - 1 \right)
$$

which depends on the $\mathbb{P}$-measure cumulant generating function of the exchange rate at time $t+1$, $k_{t+1}^{ji} [1]$, and the observable bond yields. The notation $x_{s,t,t+1}^{ji,i}$ emphasizes that this is an excess return for a long/short portfolio involving currencies $J$ (long) and $I$ (short), whose payoff is denominated in units of currency $I$. This notational flexibility subsequently allows us to consider portfolios in which the long and short currencies are distinct from the denomination of the final payoff (e.g. the USD payoff of a AUD/JPY carry trade). We choose to work with simple excess returns, rather than log excess returns, since they can be linearly aggregated using portfolio weights to produce model risk premia for factor mimicking portfolios of empirical interest. To facilitate intuition about the determinants of currency risk premia in our framework, it will be convenient to work with an approximate version of (2.2.10). Specifically, let:

$$
\lambda_{t}^{ji,i} = k_{t+1}^{ji} [1] - s_{t}^{ji} + (y_{t,t+1} - y_{t,t+1}) \cdot \Delta
$$

(2.2.11)

\footnote{Backus, et al. (2001) and Lustig, et al. (2014) measure currency risk premia using the mean log excess return, which is related to the CGFs of the pricing kernel innovations via:

$$
E_{t}^P \left[ s_{t+1}^{ji} - s_{t}^{ji} \right] + \left( y_{t,t+1} - y_{t,t+1} \right) = k_{L_{t+1}} \left[ -\xi_{t} \right] - k_{L_{t+1}} \left[ -\xi_{t} \right] + k_{L_{t+1}} [1] - k_{L_{t+1}} [1]
$$

Unlike simple excess returns, this measure cannot be aggregated to produce model-implied estimates of portfolio risk premia.}
and note that $E_t \left[ x_{t+1}^{ji,i} \right] \approx \lambda_t^{ji,i}$. The (approximate) risk premium can be expressed entirely in terms of the CGFs of the pricing kernel innovations:

\[
\lambda_t^{ji,i} = \left( k_L^g \left[ \xi_t^i - \xi_t^j \right] + k_{L^g_{t+1}} \left[ -\xi_t^i \right] - k_{L^g_{t+1}} \left[ -\xi_t^j \right] \right) + \left( k_{L^i_{t+1}}[1] + k_{L^i_{t+1}}[-1] \right)
\]

\[
= \left( k_{L^g_{t+1}} \left[ \xi_t^i - \xi_t^j \right] + k_{L^g_{t}} \left[ -\xi_t^i \right] - k_{L^g_{t+1}} \left[ -\xi_t^j \right] \right) \cdot Z_t
\]

\[
+ \left( k_{L^i_{t+1}}[1] + k_{L^i_{t+1}}[-1] \right) \cdot Y_t^i
\]

Equation (2.2.12) illustrates that the expected excess return on an individual currency pair is comprised of two components:

\[
\lambda_{HML,t}^{ji,i} = \left( k_{L^g_{t+1}} \left[ \xi_t^i - \xi_t^j \right] + k_{L^g_{t}} \left[ -\xi_t^i \right] - k_{L^g_{t+1}} \left[ -\xi_t^j \right] \right) \cdot Z_t
\]

\[
\lambda_{refFX,t}^{ji,i} = \left( k_{L^i_{t+1}}[1] + k_{L^i_{t+1}}[-1] \right) \cdot Y_t^i
\]

The first component represents compensation for exposure to the global risk factor. To the extent that global factor loadings are (inversely) linked to interest rate levels, the $HML_{FX}$ ("slope") factor identified by Lustig, et al. (2011), represents a mimicking portfolio for the global risk factor. We therefore label this component of the risk premium as $\lambda_{HML,t}^{ji,i}$. Crucially, note that our model does not impose a mechanical link between global factor loadings, $\xi_t$, and yields, $y_{t+1}^i$, instead identifying this relationship from the data. The second component – controlled by the country-
specific state variable, \( Y_t^i \) – represents the compensation demanded by an investor in country \( I \) for being short his own reference (local) currency.6

Finally, the premium demanded by an investor in country \( I \) for holding a currency pair \( J/K \), not involving his home currency, are readily obtained from, \( \lambda_{t}^{j,k,i} = \lambda_{t}^{ji,i} - \lambda_{t}^{ki,i} \), since a long position in pair \( J/K \) is equivalent to a portfolio which is long the pair \( J/I \) and short the pair \( K/I \). Consequently, the risk premium demanded by an investor in country \( I \) for exposure to the pair \( J/K \) is:

\[
\lambda_{t}^{j,k,i} = \left( k_{L_{t+1}^{j}} \left[ \xi_{t}^{j} - \xi_{t}^{i} \right] - k_{L_{t+1}^{k}} \left[ \xi_{t}^{j} - \xi_{t}^{k} \right] 
+ k_{L_{t+1}^{k}} \left[ -\xi_{t}^{k} \right] - k_{L_{t+1}^{j}} \left[ -\xi_{t}^{j} \right] \right) \cdot Z_{t} \tag{2.2.15}
\]

Example: Gaussian innovations

To provide some intuition, consider the case where both the global and country-specific innovations are Gaussian, such that the cumulant generating function of the increments is given by, \( k_{L_{t}}[u] = \frac{u^2}{2} \). In this case, the two risk premia are equal to:

\[
\lambda_{HML,t}^{ji,i} = \xi_{t}^{i} \cdot (\xi_{t}^{i} - \xi_{t}^{j}) \cdot Z_{t} \quad \lambda_{refFX,t}^{ji,i} = Y_{t}^{i} \tag{2.2.16}
\]

The first risk premium reflects the compensation demanded by an investor in country \( I \) for the exchange rate’s exposure to the global innovation, which is measured by its factor loading \( (\xi_{t}^{i} - \xi_{t}^{j}) \). The loading of the investor’s pricing kernel onto this shock is \(-\xi_{t}^{i}\), and the magnitude of the risk premium further depends on the quantity of global risk, as determined by the level of the global state-variable, \( Z_{t} \). The second component reflects compensation for being short exposure to his local (reference) currency shocks, \( L_{t+1}^{i} \). The loadings of the pricing kernel and exchange rate onto this

---

6In the continuous-time version of the model with parametric state-variable dynamics, derived in Section 1 of the Technical Appendix, there is a full term-structure of option-implied currency risk premia. Lustig, et al. (2013) study the corresponding term-structure of realized risk premia.
shock are $-1$ and 1, respectively, such that this risk premium is controlled entirely by the local state variable, $Y^i_t$. Finally, the risk premia demanded by an agent in country $I$ for a generic currency pair, $J/K$, not involving his home currency are:

$$\lambda^{j,k,i}_{HML,t} = \xi^i_t \cdot (\xi^k_t - \xi^j_t) \cdot Z_t \text{ and } \lambda^{j,k,i}_{refFX,t} = 0.$$ 

### Decompositions

In our preferred model implementation, the pricing kernel innovations are non-Gaussian. In this case, the parsimonious expressions for the currency risk premia, (2.2.13) and (2.2.14), given in terms of the cumulant generating functions of the underlying pricing kernel innovations, allow us to decompose the risk premium in terms of the distributional features of the corresponding log pricing kernel shocks ($L^g_{t+1}$ and $L^i_{t+1}$). These decompositions take advantage of the fact that the cumulant generating function of the pricing kernel innovations can be re-expressed in terms of their corresponding cumulants, which themselves are related to the moments of the innovations. Specifically, the cumulant generating function of a random variable, $\varepsilon$, has the following series expansion:

$$k_{\varepsilon}[u] = \sum_{n=1}^{\infty} \frac{\kappa_{\varepsilon,n} \cdot u^n}{n!} \text{ where } \kappa_{\varepsilon,n} = \left. \frac{\partial^n k_{\varepsilon}[u]}{\partial u^n} \right|_{u=0}$$  \hspace{1cm} (2.2.17)

where the series coefficients, $\kappa_{\varepsilon,n}$, are known as cumulants. The relation between cumulants and the mean, variance, skewness, and kurtosis of the random variable, $\varepsilon$, are given by: $\mathcal{M} = \kappa_{\varepsilon,1}$, $\mathcal{V} = \kappa_{\varepsilon,2}$, $\mathcal{S} = \kappa_{\varepsilon,3} \cdot (\kappa_{\varepsilon,2})^{-\frac{3}{2}}$, and $\mathcal{K} = \kappa_{\varepsilon,4} \cdot (\kappa_{\varepsilon,2})^{-2}$.

We decompose the risk premia across the even and odd cumulants, as well as, the moments of the log pricing kernel innovations, in order to isolate how various non-Gaussianities in the innovations contribute to the determination of currency risk premia. Following Backus, et al. (2011), the first decomposition emphasizes the role of distributional asymmetries, which are implicit in disaster risk models and are
controlled by the odd cumulants. To compute this decomposition, we take advantage
of the observation that the contribution from the odd cumulants can be computed
directly from the knowledge of the cumulant generating function as,
\[ k_{\text{odd}}^t[u] = \frac{1}{2} \cdot (k_{\epsilon_{t+1}}[u] - k_{\epsilon_{t+1}}[-u]) \]. The second decomposition highlights the overall role of higher-
order moments (skewness, kurtosis, etc.) in the log pricing kernel shocks, in the sprit
of a Gram-Charlier/Edgeworth series expansion of the innovation density.

Using the series expansion above, the contribution to the excess return of currency
pair \( J/I \) from exposure to the global factor (\( HML_{FX} \)) is given by:

\[
\lambda_{HML,t}^{ij} = \sum_{n=2}^{\infty} \frac{(\xi^i_{t} - \xi^j_{t})^n + (-\xi^i_{t})^n - (-\xi^j_{t})^n}{n!} \cdot \kappa_{L_{t+1}^g,n} \cdot Z_t
\] (2.2.18)

where \( \kappa_{L_{t+1}^g,n} \) denotes the \( n \)-th cumulant of the non-time-changed global increment,
\( L_{t+1}^g \). In particular, notice that the global component of the risk premium only de-
pends on cumulants of order two and higher. Taking advantage of the definitional
link between cumulants and moments, and the fact that the non-time-changed innova-
tions are standardized to have unit variance (Appendix A), we can re-write this
expression with terms up to order three as:

\[
\lambda_{HML,t}^{ij} \approx \xi^i_{t} \cdot (\xi^i_{t} - \xi^j_{t}) \cdot Z_t - \frac{1}{2} \cdot \xi^i_{t} \cdot \xi^j_{t} \cdot (\xi^i_{t} - \xi^j_{t}) \cdot S^g_{t} \cdot Z_t + O(\xi^4)
\] (2.2.19)

Comparing this expression to the one obtained for the Gaussian case, we immediately
recognize the first term as representing compensation for bearing the variance of the
global innovations. The compensation for a pair’s \( HML_{FX} \) risk is high whenever:
(a) the loading differential is large and positive \((\xi^i_{t} - \xi^j_{t} \gg 0)\); (b) the price of risk
is high (i.e. the volatility of global pricing kernel shocks, \( \sqrt{Z_t} \), is high); and, (c) the
skewness of the pricing kernel innovations is negative \((S^g_{t} < 0)\). The skewness of the
global shock, \( S^g_{t} \), is determined by parameters known one period in advance, and is
therefore subscripted with \( t \).
We obtain a similar risk premium decomposition for the short-reference component of a currency pair’s risk premium, \(\lambda^j_{refFX,t}\), by applying the same series expansion to the cumulant generating function of the country-specific shock, \(\tilde{L}^i_t\):

\[
\lambda^j_{refFX,t} = \sum_{k=1}^{\infty} \frac{1}{k!} \cdot \kappa_{\tilde{L}^i_{t+1},k} \cdot Y^i_t = \sum_{n=1}^{\infty} \frac{1}{(2n)!} \cdot \kappa_{\tilde{L}^i_{t+1},2n} \cdot Y^i_t (2.2.20)
\]

A stark feature of the reference currency premium is that it does not depend on the odd moments of the country-specific shock, such that asymmetries in the distribution of the country-specific shock play no role in its determination. In other words, one-sided events, such as crashes or sporadic flights to quality, cannot rationalize the short reference risk premium in the context of our model. Since our model does not ascribe a special role to any particular reference currency, we compute short reference risk premia for all currencies in our sample. Empirically, Lustig, et al. (2011, 2014) find evidence of large compensation for short dollar exposure.

### 2.2.3 FX option pricing

Given the choice of our model parametrization, option pricing can be efficiently accomplished using standard Fourier transform option pricing methods described in Carr and Madan (1999). The central input to this pricing methodology is the characteristic function of the log exchange rate, \(s^i_{t+1} = \log S^i_{t+1}\), computed under the pricing measure. Since the characteristic function of the log exchange rate is equivalent to the cumulant generating function evaluated at \(i \cdot u\), \(\phi(u) = \exp \left( k_{s^i_{t+1}} [i \cdot u] \right)\), the problem of option pricing is equivalent to deriving the cumulant generating function of the exchange rate under the pricing measure. We price assets under the risk-forward measure, \(\mathbb{F}^i\), whose numeraire is the one-period zero coupon bond in country \(I\). Finally, it is important to emphasize that the pricing measure depends on the home country of the investor, since investors in different countries have distinct pricing kernels.
The risk-forward measure for an investor from country \( I \), associated with a zero-coupon bond maturing at time \( t + 1 \), is defined as follows:

\[
\frac{d\mathbb{F}^i}{d\mathbb{P}} = \frac{M^i_{t+1}}{M^i_t} \cdot \exp \left( y^i_{t,t+1} \cdot \Delta \right) = \exp \left( m^i_{t+1} - m^i_t + y^i_{t,t+1} \cdot \Delta \right) \quad (2.2.21)
\]

where \( \mathbb{P} \) denotes the objective measure. The virtue of working under the risk-forward measure is that we can price claims denominated in units of currency \( I \) (e.g. bonds or exchange rate options on the currency pairs \( J/I \)) as the product of their expected payoff, \( x(s^i_{t+1}) \), under the \( \mathbb{F}^i \) measure multiplied by the value of the one-period zero-coupon bond in country \( I \). To see this, recall that any payoff satisfies the following pricing equation:

\[
V^i_t = \mathbb{E}_t^\mathbb{P} \left[ \frac{M^i_{t+1}}{M^i_t} \cdot x(s^i_{t+1}) \right].
\]

Dividing both sides by the value of the zero-coupon maturing at \( t + 1 \), and recognizing the measure change we obtain:

\[
V^i_t = B^i_{t,t+1} \cdot \mathbb{E}_t^\mathbb{P} \left[ \frac{M^i_{t+1}}{M^i_t} \cdot \exp \left( y^i_{t,t+1} \cdot \Delta \right) \cdot x(s^i_{t+1}) \right] = B^i_{t,t+1} \cdot \mathbb{E}_t^\mathbb{F}^i \left[ x(s^i_{t+1}) \right] \quad (2.2.22)
\]

where \( B^i_{t,t+1} = \exp \left( -y^i_{t,t+1} \cdot \Delta \right) \) is the value of the numeraire, zero-coupon bond. In order to apply this valuation approach to exchange rate options, we rely on Fourier pricing methods and characterize the distribution of the exchange rate at \( t + 1 \) using its cumulant generating function under \( \mathbb{F}^i \).

The cumulant generating function for the exchange rate at time \( t + 1 \) under the risk forward measure, \( \mathbb{F}^i \), is obtained by substituting (2.2.8), into the definition of the CGF:

\[
k^{\mathbb{F}^i}_{s^i_{t+1}} [u] = (s^i_{t+1} - \alpha^i_t + \alpha^i_t) \cdot u + k^{\mathbb{F}^i}_{L^i_{t+1}} [(\xi^i_t - \xi^i_t) \cdot u] + k^{\mathbb{F}^i}_{L^i_{t+1}} [u] + k^{\mathbb{F}^i}_{L^i_{t+1}} [-u] \quad (2.2.23)
\]

where the \( \alpha^i_t \) and \( \alpha^i_t \) are the log pricing kernel drifts. In turn, the risk-forward CGFs, \( k^{\mathbb{F}^i}_{L^i_{t+1}} [u] \), of the pricing kernel innovations can be linked their counterparts under
the objective measure, by proceeding directly from the definition of the cumulant generating function and the measure change (Appendix B), to obtain:

\[
k_F^{L_{i+1}} [u] = \ln E^F_t \left[ \exp \left( u \cdot L_{i+1}^g \right) \right] = \ln E^P_t \left[ \exp \left( y_{i,t+1}^i + \left( m_{i,t+1}^i - m_t^i \right) + u \cdot L_{i+1}^g \right) \right]
\]

\[
= \left( k_{L_{i+1}}^g \left[ u - \xi_t^i \right] - k_{L_{i+1}}^g \left[ -\xi_t^i \right] \right) \cdot Z_t
\]

(2.2.24a)

\[
k_L^{L_{i+1}} [u] = \left( k_{L_{i+1}} \left[ u - 1 \right] - k_{L_{i+1}} \left[ -1 \right] \right) \cdot Y_t
\]

(2.2.24b)

\[
k_F^{L_{j+1}} [u] = k_{L_{j+1}}^F \left[ u \right] \cdot Y_j^i
\]

(2.2.24c)

The change of measure results in: (1) a change in the distribution of the global factor dependent on country \(i\)'s loading on the global factor, \(\xi_t^i\); (2) a change in the distribution of the local (reference) shock, \(L_{i+1}^i\); and, (3) leaves the distribution of foreign, country-specific shocks unchanged. Putting these results together, the CGF of the exchange rate under the \(F_i\) pricing measure – which provides a complete characterization of the exchange rate distribution – is:

\[
k_{s_t}^{L_{i+1}} [u] = \left( s_t^i - \alpha_t^i + \alpha_{i,t}^i \right) \cdot u + \left( k_{L_{i+1}}^g \left[ \left( \xi_t^i - \xi_t^j \right) \cdot u - \xi_t^i \right] - k_{L_{i+1}}^g \left[ -\xi_t^i \right] \right) \cdot Z_t
\]

\[
+ \left( k_{L_{i+1}} \left[ u - 1 \right] - k_{L_{i+1}} \left[ -1 \right] \right) \cdot Y_t^i + k_{L_{j+1}}^F \left[ u \right] \cdot Y_j^i
\]

(2.2.25)

From here, we exploit the link between the cumulant generating function and the characteristic function, and price options using the numerical Fourier inversion techniques in Carr and Madan (1999). Bates (1996) provides an early application of these methods in the context of exchange rate option pricing.

\subsection*{2.2.4 Recovering currency risk premia}

In order to recover currency risk premia we calibrate the proposed model to a combination of cross-sectional exchange rate option data, and exchange rate variance forecasts computed using historical currency returns, similar to the approach
previously applied to equity indices (Pan (2002), Santa-Clara and Yan (2010), and Andersen, et al. (2013)). In general, there are multiple sets of model parameters that generate identical risk-forward distributions for exchange rates, but have distinct implications for quantities under the objective measure (risk premia, variances, etc.). Combining these two datasets allows us to: (a) identify the currency return factor structure under the risk-forward measure (i.e. disentangle the role of global and country-specific innovations) by matching option prices; and, (b) select the set of model parameters, which simultaneously generates a model forecast of exchange rate variance under the objective measure, \( \xi_{ref}^{fx} \), that is consistent with the empirical forecast based on historical data. Interestingly, we show that while recovering the \( HML_{FX} \) risk premium requires data on both FX option prices and variances under the objective measure, the short reference risk premium, \( \lambda_{ref}^{fx,t} \), within our model can be recovered exclusively using exchange rate option data.

**The role of cross-rate options**

The primary dataset in our calibration consists of exchange rate option prices for all of the \( \frac{N \cdot (N-1)}{2} \) possible exchange rates involving \( N \) currencies, commonly referred to as *cross-rates*. At each point in time, these data reveal the risk-forward (i.e. option-implied) distributions of exchange rates, \( s_{t+1}^{ji} \), and play a crucial role in identifying the underlying factor structure. Specifically, they help pin down the relative roles of the global, \( g_{t+1}^{ji} = (\xi^j_t - \xi^i_t) \cdot L^{g}_{t+1} \), and country-specific, \( L^i_{t+1} \) and \( L^j_{t+1} \), components in driving currency returns.

To illustrate the role of cross-rate options, we show that in their absence it is impossible to distinguish between our proposed factor specification, \( \xi_{ref}^{fx} \), and a model with no global factor, without imposing strong auxiliary distributional assumptions. Specifically, suppose the only available options are for exchange rates measured relative to a single currency \( I \) (e.g. the US dollar, \( I = USD \)), and compare the prices of
exchange rate options generated by our baseline specification, (2.2.8), and an alternative model where the log currency return is given by:

\[
s_{t+1}^j - s_t^j = -\alpha_t^j + \alpha_t^i - \hat{L}_{t+1}^j + L_{t+1}^i
\] (2.2.26)

A notable feature of this specification is that there is no shared global component, such that the global (HML_{FX}) risk premium is counterfactually absent. The shocks \( \hat{L}_{t+1}^j \) and \( L_{t+1}^i \) continue to be cross-sectionally independent, and the distribution of \( \hat{L}_{t+1}^j \) is assumed to be given by the following CGF:

\[
k_{L_{t+1}^j}^i[u] = k_{L_{t+1}^i}^i \left[ (\xi_t^i - \xi_t^j) \cdot u \right] + k_{L_{t+1}^i}^i \left[ -u \right] = \left( k_{L_{t+1}^i}^q \left[ (\xi_t^i - \xi_t^j) \cdot u - \xi_t^i \right] - k_{L_{t+1}^i}^q \left[ -\xi_t^i \right] \right) + k_{L_{t+1}^i}^q \left[ -u \right] \) (2.2.27)

Since the shock, \( \hat{L}_{t+1}^j \), is idiosyncratic with respect to investor \( I \)’s pricing kernel, we additionally have: \( k_{L_{t+1}^j}^i[u] = k_{L_{t+1}^j}^i[u] \). Substituting, (2.2.26), into the definition of the cumulant generating function, one immediately recovers (2.2.25). These starkly different models of currency returns result in the same exchange rate CGFs under the pricing measure, \( F^i \), and therefore, cannot be differentiated using FX option prices. Consequently, the factor structure of currency returns cannot be identified solely from options on exchange rates measured relative to a single reference currency, \( I \).

To identify the currency return factor structure, we exploit the availability of options on a full panel of cross-rates \((N-(N-1)/2)\) exchange rate pairs. Specifically, our model and the alternative specification above have distinct implications for the prices of cross-rate options (Appendix C.1), such that these data can be used to isolate the role of the global component in driving currency returns. In contrast to us, Farhi, et al. (2014) only use options on exchange rates relative to the U.S. dollar \((N-1)\) exchange rate pairs), but additionally assume that country-specific innovation are Gaussian. This distributional assumption effectively forces their model to
match any non-normalities present in exchange rate option data using a combination of global factor loadings and the distribution of the global pricing kernel innovation. Importantly, the results in Bakshi, et al. (2008) and our empirical calibration, indicate that country-specific components in exchange rate option data are significantly non-Gaussian. Similarly, the Jarque-Bera test rejects the Gaussianity of the (standardized) returns to strategies interpretable as factor mimicking portfolios for the country-specific innovations, which short each G10 currency against an equally-weighted basket of foreign currencies (Jurek (2014); Table A.I).

The global \((HML_{FX})\) risk premium

While cross-sectional exchange rate option data pin down the risk-forward distributions of the global and country-specific components of the log currency return for each pair, \(\{g_{it+1}^j, L_{it+1}^i, L_{it+1}^j\}\), this is generally insufficient to recover risk premia. For example, recovering the global \((HML_{FX})\) component of the currency risk premium, \(2.2.13\), requires that the distribution of the global factor innovation, \(L_{it+1}^0\), and the global factor loadings, \(\xi^i_t\), be separably identifiable. This is not possible using option data alone, and requires incorporating forecasts of the \(\mathbb{P}\)-measure exchange rate variance in the calibration.

To emphasize the link between risk premia and the quantities which are identifiable from option data, we re-write the expression for the risk premium on currency pair \(J/K\), \(2.2.15\), in terms the risk-forward CGFs of the global components of the log currency returns of pairs \(J/I\) and \(K/I\), \(g_{it+1}^j\) and \(g_{it+1}^k\) (Appendix C.2):

\[
\lambda_{HML,t}^{jk,i} = k_{g_{it+1}^j}^{\xi_i} \left[ \frac{2 \cdot \xi_t^i - \xi_t^j}{\xi_t^j - \xi_t^i} \right] - k_{g_{it+1}^k}^{\xi_i} \left[ \frac{2 \cdot \xi_t^i - \xi_t^k}{\xi_t^k - \xi_t^i} \right] + k_{g_{it+1}^j}^{\xi_i} [-1] - k_{g_{it+1}^k}^{\xi_i} [-1] \quad (2.2.28)
\]
In particular, if $\xi^i_t$ and $\xi^k_t$ are taken to represent the global factor loadings of asymptotically diversified portfolios of high- and low-interest rate currencies (i.e. such that idiosyncratic risks are negligible), the above expression can be interpreted as the risk premium on the global risk ($HML_{FX}$) factor mimicking portfolio. Option data pin down the risk-forward CGFs of $g^i_{jt+1}$ and $g^k_{jt+1}$, and therefore, the final two terms in the above expression. However, the first two terms additionally require knowledge of the global factor loadings, \{$\xi^i_t, \xi^j_t, \xi^k_t$\}, which are not identifiable separably from the distribution of the global pricing kernel innovation, $L^g_{t+1}$.

Appendix C.2 illustrates the identification problem by showing that the perturbed set of loadings, $\hat{\xi}^i_t = \xi^i_t - \delta^i + \delta^j$, can be combined with a perturbed set of model parameters, \{$\hat{Z}_t(\delta), \hat{\eta}_t(\delta), \hat{G}_t(\delta)$\}, to generate a shock: $\hat{g}^i_{jt+1} = (\hat{\xi}^i_t - \hat{\xi}^j_t) \cdot \hat{L}^g_{t+1}$, which is identical in distribution to the original shock $g^i_{jt+1}$ (i.e. has an identical cumulant generating function). Since the two sets of parameters generate identical risk-forward distributions irrespective of the choice of $\delta$, they cannot be distinguished using option data.

At the same time, the value of, $\delta$, affects risk premia, since replacing the global factor loadings with their perturbed counterparts affects the argument of the first two CGF terms in $[2.2.28]$. To resolve this identification problem, we exploit the fact that the choice of $\delta$ impacts the model-implied periodic variance of currency return under the $\mathbb{P}$-measure. Specifically, we proceed in a manner similar to Pan (2002), Santa-Clara and Yan (2010), and Andersen, et al. (2013) for the equity index, and augment cross-sectional exchange rate option data with variance forecasts obtained from backward-looking returns on the underlying assets. Taken together, these data allow us to empirically pin down the level of the $HML_{FX}$ risk premium.$^7$

$^7$Using only option data (i.e. leaving $\delta$ unidentified), it is possible to recover the relative magnitudes of the approximate $HML_{FX}$ risk premia for individual currencies. Specifically, if we only retain the leading order term in the expression for the global component of a currency pair’s risk premium, we have: \( \lambda^{i,j}_{HML,t} \approx \xi^i_t \cdot (\xi^i_t - \xi^j_t) \cdot Z_t \). Ratios of these approximate risk premia do not depend on $\delta$, and can be identified from exchange rate options alone. This facilitates the formation of a candidate factor replicating portfolio, though the level of its model-implied risk premium is unknown.
The short reference risk premium

In our model, the country-specific component of the currency risk premium for currency $I$ is determined by the sum of the $\mathbb{P}$-measure cumulant generating function of the country-specific shock, $k_{L_{t+1}^i}[u]$, evaluated at $u = 1$ and $u = -1$, (2.2.14). The parameters of this CGF can be obtained directly from options on exchange rates where currency $I$ is the investment (long) currency, $I/K$. Recall that the payoffs of these options are priced under measure $\mathbb{F}^k$, and that investors from country $K$ do not price exposure to the country-specific $I$. As a result, the distribution of $L_{t+1}^i$ under the pricing, $\mathbb{F}^k$, and objective, $\mathbb{P}$, measures is identical, such that $k_{L_{t+1}^i}[u] = k_{L_{t+1}^i}[u]$. Consequently, we can recover the short reference risk premium from option data alone.

Another way to see this, is to note that the short-reference currency risk premium can be written in terms of the risk-forward CGF of the country-specific shocks, $k_{L_{t+1}^i}[u]$, as follows (Appendix C.2):

$$\lambda_{\text{refFX},t}^{ji,i} = k_{L_{t+1}^i}^{F^i}[2] - 2 \cdot k_{L_{t+1}^i}^{F^i}[1] \quad (2.2.29)$$

The risk-forward cumulant generating function, $k_{L_{t+1}^i}^{F^i}[u]$, can be identified from exchange rate option data for currency pairs $J/I$, where currency $I$ is the funding (short) currency, since their payoffs are priced under measure $\mathbb{F}^i$. Therefore, option data on the full panel of exchange rates, reveal both the objective and risk-forward distributions of the country-specific innovations, either of which is sufficient to pin down the short reference risk premium for each currency.$^8$

---

$^8$Formally, identification requires not only the subset of options $J/I$ or $I/K$, but also options not involving $I$. Without such options the country-specific component, $L_{t+1}^i$, would not be distinguishable from the global component, $L_{t+1}^g$, as demonstrated in the section establishing the role of cross-rate options. However, once this condition is satisfied, the short reference risk premium is identifiable exclusively from exchange rate option data.
2.2.5 Shock parametrization

Finally, to bring the model to the data, we need to specify a parametric distribution for the innovations driving the pricing kernels and exchange rates. Empirical evidence from currency options indicates that exchange rate distribution are typically non-Gaussian, and exhibit stochastic variation in variance and skewness (Carr and Wu (2007)). To accommodate these features we parameterize the innovations using a discrete-time equivalent of the time-changed Lévy model of Bakshi, et al. (2008). The distribution of each shock, \( L_{t+1}^g \) and \( L_{t+1}^i \), is non-Gaussian, and time-series variation in their moments (variance, skewness, etc.) is controlled by the model state variables, \( Z_t \) and \( Y_t^i \), respectively.

The increments, \( L_{t+1}^g \) and \( L_{t+1}^i \), driving the pricing kernel dynamics and exchange rates are obtained by applying a time-change transformation to the random variables, \( \tilde{L}_{t+1}^g \) and \( \tilde{L}_{t+1}^i \), each of which is normalized to have unit variance (Appendix A). The time-change is controlled by the state variables, \( Z_t \) and \( Y_t^i \), respectively, which are assumed to be constant within the time interval \([t, t+1)\), but can change from period to period. The effect of the time-change is to scale the CGFs of these increments, (2.2.6a) and (2.2.6b), and is the consequence of Theorem 1 in Carr and Wu (2004) applied to our setting. The multiplicative scaling of the CGF induces variation in all moments of the distribution (variance, skewness, etc.), with the level of the state variable interpretable as the periodic variance of the time-changed increment.

Let \( \tilde{L}_{t+1}^\phi \) denote a generic, non-time-changed increment, e.g. either the global increment, \( \tilde{L}_{t+1}^g \), or one of the country-specific increments, \( \tilde{L}_{t+1}^i \). We model the increment \( \tilde{L}_{t+1}^\phi \) as a combination of a Gaussian innovation, \( W_{t+1}^\phi \), with variance \((1 - |\eta_t^\phi|)\), and a non-Gaussian innovation, \( X_{t+1}^\phi \), with variance \(|\eta_t^\phi|\), with \( \eta_t^\phi \in [-1, 1] \):

\[
\tilde{L}_{t+1}^\phi = W_{t+1}^\phi + \text{sign}(\eta_t^\phi) \cdot X_{t+1}^\phi \quad (2.2.30)
\]
Aside from controlling the variance shares, \( \eta^\phi_t \), controls the sign of the increment’s loading on the non-Gaussian component, thereby affecting the sign of its skewness. The non-Gaussian component has a CGMY distribution, introduced by Carr, et al. (2002), which is characterized by a quartet of potentially time-varying parameters, \( \{C, G, M, Y\}^\phi_t \):

\[
\mu^i_t[dx] = \begin{cases} 
C^\phi_t \cdot \exp\left(G^\phi_t \cdot x\right) \cdot |x|^{-Y^\phi_t - 1} \cdot dx & x \leq 0 \\
C^\phi_t \cdot \exp\left(-M^\phi_t \cdot x\right) \cdot x^{-Y^\phi_t - 1} \cdot dx & x > 0 
\end{cases}
\tag{2.2.31}
\]

To maintain consistency with the state-variable notation, we subscript the parameters governing the one-step ahead distribution of the shocks with time \( t \). The \( C^\phi_t \) parameter is a scaling factor, which is set such that the variance of the jump component is \( |\eta^\phi_t| \); \( G^\phi_t \) and \( M^\phi_t \) determine the exponential dampening of the distribution for negative and positive shocks, respectively. The \( Y^\phi_t \) parameter can be interpreted as measuring the degree of similarity between the jump process and a Brownian motion. The CGMY process nests compound Poisson jumps \((-1 \leq Y < 0)\), infinite-activity jumps with finite variation \((0 \leq Y < 1)\), as well as, infinite-activity jumps with infinite variation \((1 \leq Y < 2)\). Specifically, we assume that: (1) global shocks are negatively skewed \((\eta^g_t > 0 \text{ and } M^g_t = \infty)\); (2) country-specific shocks are two-sided, capturing both positive and negative idiosyncratic shocks; and, (3) individual currencies have different exposures, \( \{\eta^i_t\}_{i=1}^N \), to the non-Gaussian innovation, but the parameters governing its distribution, \( \{G_t, M_t, Y_t\} \), are shared across all countries. By allowing \( \eta^i_t \) to be country specific, the skewness of the country-specific innovations in each country can have arbitrary sign, irrespective of the skewness of \( X_{t+1} \).

Since the key expressions in the paper (risk premia, option prices, etc.) are given in terms of the CGFs of the pricing kernel innovations, \( L^g_{t+1} \) and \( L^i_{t+1} \), we report them here for completeness. Specifically, for the empirically relevant case, where
$Y \neq \{0, 1\}$, they are given by:

$$k_{L_{t+1}}[u] = \left(1 - \eta_t^g\right) \cdot \frac{u^2}{2} + \eta_t^g \cdot \frac{(G_t^g + u)^{Y_t^g} - (G_t^g)^{Y_t^g}}{Y_t^g \cdot (Y_t^g - 1) \cdot (G_t^g)^{Y_t^g - 2}} \cdot Z_t \quad (2.2.32a)$$

$$k_{L_{t+1}}[u] = \left(|\eta_t^i| \cdot \frac{(M_t \mp u)^{Y_t^i} - (M_t)^{Y_t^i} + (G_t \pm u)^{Y_t^i} - (G_t)^{Y_t^i}}{Y_t \cdot (Y_t - 1) \cdot ((M_t)^{Y_t^i - 2} + (G_t)^{Y_t^i - 2})} + + (1 - |\eta_t^i|) \cdot \frac{u^2}{2} \right) \cdot Y_t^i \quad (2.2.32b)$$

where $\mp^i = -\text{sign} (\eta_t^i)$ and $\pm^i = \text{sign} (\eta_t^i)$. The time-subscripts emphasize that our empirical calibration allows the state variables, $Z_t$ and $Y_t^i$, and the parameters governing the shock distribution – the share of variance attributable to the non-Gaussian component, $\eta$, and the CGMY distribution parameters, $\{C, G, M, Y\}$ – to change across periods. The corresponding CGFs under the risk-forward measure, $F_t^i$, are derived in Appendix B.

### 2.3 Data and Model Calibration

The key dataset used to calibrate the model includes price data on foreign exchange options for the full cross-section of 45 G10 exchange rate cross-pairs, spanning the period from January 1999 to June 2012 ($T = 3520$ days; $N = 10$ currencies). The dataset provides daily price quotes in the form of implied volatilities for European options at constant maturities and five strikes, and was obtained via J.P. Morgan DataQuery. FX option prices are quoted in terms of their Garman-Kohlhagen (1983) implied volatilities, which correspond to Black-Scholes (1973) implied volatilities adjusted for the fact that both currencies pay a continuous “dividend” given by their

---

9The G10 currency set is comprised of the Australian dollar (AUD), Canadian dollar (CAD), Swiss franc (CHF), Euro (EUR), U.K. pound (GBP), Japanese yen (JPY), Norwegian kronor (NOK), New Zealand dollar (NZD), Swedish krone (SEK), and the U.S. dollar (USD). There are a total of 45 possible cross-pairs.
respective interest rates. We focus attention on constant-maturity one-month exchange rate options. For each day and currency pair, we have quotes for five options at fixed levels of option delta (10δ puts, 25δ puts, 50δ options, 25δ calls, and 10δ calls), which correspond to strikes below and above the prevailing forward price. In standard FX option nomenclature an option with a delta of δ is typically referred to as a $|100 \cdot \delta|$ option; we adopt this convention throughout. The specifics of foreign exchange option conventions are further described in Wystup (2006), Carr and Wu (2007), and Jurek (2014). In general, a call (put) option on the pair J/I gives its owner the right to buy (sell) currency J at option expiration at an exchange rate corresponding to the strike price, which is expressed as the currency I price of one unit of currency J. The payoff of this option is denominated in units of currency I. The remaining data we use are one-month Eurocurrency (LIBOR) rates and daily exchange rates for the nine G10 currencies versus the U.S. dollar obtained from Reuters via Datastream. The exchange rate data are used to value the options, and to construct exchange rate variance forecasts.

2.3.1 Implementation

We separately fit the model on each day in the sample, such that the calibration produces a time series of global factor loadings, state variables, and distribution parameters for the pricing kernel innovations. After normalizing one of the loadings ($\xi^US_t = 1$), there are $N - 1$ country loadings and $N + 1$ state-variables (global, $Z_t$, and country-specific, $\{Y^i_t\}_{i=1}^{10}$) to be calibrated, in addition to the parameters of the global and country-specific shocks. Our parametrization assumes that global innovations are one-sided (three global parameters, $\{\eta^g_t, G^g_t, Y^g_t\}$), and that the parameters governing the country-specific non-Gaussian shocks are identical across all countries, with the exception of $\eta^i_t$ ($N + 3$ parameters, $\{\eta^i_t\}_{i=1}^{N}, G_t, \mathcal{M}_t, \mathcal{Y}_t\}$). The $3 \cdot N + 6$ model parameters are pinned down by matching up to $5 \cdot \frac{N \cdot (N-1)}{2}$ option quotes (five strikes
times the number of currency pairs) and up to \( \frac{N(N-1)}{2} \) \( \mathbb{P} \)-measure exchange rate variance forecasts (one per currency pair).

The goal of the calibration is to match both option-implied volatilities (pricing measure) and forecasts of the monthly realized exchange rate variance (objective measure). Formally, on each day \( t \) in our sample (\( T = 3520 \) days; 1999:1-2012:6), the calibration minimizes the sum of squared relative errors for volatilities under the two measures:

\[
\min_{Z_t, \{Y^i_t\}_{i=1}^N, \{\xi_t^i\}_{i=1}^{N-1}, \{\eta_t^i\}_{i=1}^N, \{g_t\}, \{y_t\}, \{G_t\}, \{M_t\}, \{Y_t\}, \{\hat{g}_t\}, \{\hat{G}_t\}, \{\hat{Y}_t\}} \left( \text{SSE-IV}_t + \text{SSE-RV}_t \right) \tag{2.3.33}
\]

where:

\[
\text{SSRE-IV}_t = \sum_{m=1}^{M} \sum_{k=1}^{5} \left( \frac{\hat{\sigma}_{m,k,t}^F - \hat{\sigma}_{m,k,t}^P}{\hat{\sigma}_{m,k,t}^F} \right)^2 \tag{2.3.34a}
\]

\[
\text{SSRE-RV}_t = \sum_{m=1}^{M} \left( \frac{\sigma_{m,t}^P - \hat{\sigma}_{m,t}^P}{\sigma_{m,t}^P} \right)^2 \cdot 1_{\hat{\sigma}_{m,t}^P > \sigma_{m,t}^P} + \left( \frac{\sigma_{m,t}^P - \hat{\sigma}_{m,t}^P}{\sigma_{m,t}^P} \right)^2 \cdot 1_{\hat{\sigma}_{m,t}^P < \sigma_{m,t}^P} \tag{2.3.34b}
\]

\( \hat{\sigma}_{m,k,t}^F \) (\( \hat{\sigma}_{m,k,t}^P \)) are the model and observed values of the option-implied volatility for currency pair \( m \) at strike \( k \) at time \( t \), respectively. The corresponding quantities, \( \hat{\sigma}_{m,t}^P \) (\( \sigma_{m,t}^P \)), denote the square roots of the forecasts of one-month exchange rate variance under the objective measure. The model forecast of the variance under the objective measure, \( (\hat{\sigma}_{m,t}^P)^2 \), is given by the expression, \( (2.2.9) \). Unlike under the pricing measure, we do not observe empirical forecasts of \( \mathbb{P} \)-measure variance, \( (\sigma_{m,t}^P)^2 \), and instead have to estimate this quantity based on historical data (e.g. realized variance, GARCH, etc.). For parsimony, we set the forecast equal to the variance of daily log currency returns computed using a 63-day, backward-looking window. Since this quantity is subject to measurement error, we set the fitting error to zero when-
ever the model implied quantity, (2.2.9), is within the 95% confidence interval for the historical variance estimate, whose lower and upper bounds are given by \( (\sigma_{m,t}^p)^2 \) and \( (\sigma_{m,t}^p)^2 \), respectively.\(^{10}\) In order to avoid a discontinuous penalty, we measure the deviation of the model implied quantity from the lower and upper bounds of the empirical estimator. If option data included information on bid-ask spreads, a similar truncation could have been applied to the implied volatility fitting error (SSRE-IV\(_t\)), whenever the model option price was within the bid-ask spread.

We repeat the calibration on each day in the sample using two sets of exchange rate options. The first set \((M = 24 \text{ pairs})\) includes options on currency pairs with the highest liquidity in FX markets, as well as, “typical” carry trade currencies. It is comprised of: (a) the X/USD currency pairs (9 pairs); and, (2) cross-rates formed on the basis of currencies which had the highest or lowest interest rates in the G10 set at some point in our sample (15 pairs).\(^{11}\) The second set uses the complete set of G10 exchange rate options \((M = 45 \text{ pairs})\). Depending on the option set, the calibration involves pricing 120 (225) individual options on each day. Both sets of test assets emphasize information on cross-rates, consistent with their central role in model identification (Section 1.4). In particular, X/USD pairs account for only 37.5% of the pairs in the first set, and 20% – in the second. Finally, notice that since the calibrations only use information available as of time \(t\) to estimate the time \(t\) model

\[^{10}\text{For example, under the null of constant volatility and Gaussian daily log returns, the lower and upper bounds of the 95\% confidence interval for the variance estimate, } (\sigma_{m,t}^p)^2, \text{ obtained using the sum of } \tau \text{ squared returns are:}
\]

\[
(\sigma_{m,t}^p)^2 = (\sigma_{m,t}^p)^2 \cdot \left( \frac{\chi_{\tau-1}^{-1}(0.025)}{\tau - 1} \right) \quad (\sigma_{m,t}^p)^2 = (\sigma_{m,t}^p)^2 \cdot \left( \frac{\chi_{\tau-1}^{-1}(0.975)}{\tau - 1} \right)
\]

where \( (\sigma_{m,t}^p)^2 \) is the variance point estimate, and \( \chi_{K}^{-1}(p) \) is the inverse of the CDF of the \( \chi^2 \) distribution with \( K \) degrees of freedom, evaluated at \( p \). This computation is conservative in that, when return innovations are non-Gaussian – as is empirically the case – the confidence interval becomes broader.

\[^{11}\text{X/USD pairs accounted for 68.5\% of all trading in spot exchange rate markets in 2013 (Table 3, BIS (2013)). The subset of unique currencies which had the highest one-month LIBOR rate includes: AUD, GBP, NOK, and NZD. The corresponding subset of unique currencies with the lowest interest rates includes CHF and JPY, yielding a total of 15 cross-pairs for use in estimation.}\]
parameters, they have no look-ahead bias and could have been constructed by an investor in real time.

2.3.2 Model parameters

We calibrate our model using options on two sets of exchange rates: the set of high/low interest rate cross pairs, combined with the nine X/USD pairs (24 pairs; HLX + X/USD), and all G10 cross pairs (45 pairs; ALL). We refer to these specifications as I and II, respectively. Although the inclusion of a greater number of cross-rates theoretically aids in the identification of the underlying factor structure, these options also tend to be less liquidly traded. Consequently, we focus our discussion on the output of Specification I to ensure our findings are not driven by comparatively less liquidly traded cross-rate options. The key empirical results regarding currency risk premia remain qualitatively unaffected under Specification II, and are reported in the data appendix.

Panel A of Table I reports the mean values of the global factor loadings, $\xi_t^i$, along with estimates of their time series volatilities. Throughout our implementation the U.S. global factor loading is normalized to one. Panel B reports the time series mean and volatility of the parameters governing the distribution of the global pricing kernel innovation. The latter include the global state variable, $Z_t$, the share of variance attributable to the non-Gaussian component of the innovation, $\eta_t^g$, and estimates of the parameters of the global CGMY component, $G_t^g$ (dampening coefficient) and $Y_t^g$ (power coefficient). To facilitate interpretation we also report the mean skewness and kurtosis of the monthly global innovation, $L_{t+1}^g$, induced by variation in the distribution parameters and the global state variable, $Z_t$. Panel C reports the quality of the model’s fit to the data in the form of the root mean squared fitting error (measured in volatility points) for option-implied volatilities and volatility forecasts under the objective measure.
Global factor loadings, $\xi_i^t$

The mean global loading parameters (Table I; Panel A) range from 0.77 (NZD) to 1.04 (JPY) under Specification I, and from 0.86 (NZD) to 1.12 (CHF) under Specification II. These values are broadly consistent with previously reported estimates from models allowing for asymmetries in exposures to global risks, though our estimates of loadings for “low” interest rates currencies are somewhat smaller. For example, using three years of data on a single currency triangle (JPY/USD, GBP/USD, GBP/JPY), Bakshi, et al. (2008), find loadings of 1.53 (JPY) and 1.01 (GBP). Our mean estimate of the loading for the British Pound is comparable and ranges from 0.99 to 1.03. These estimates can also be contrasted with the parameters used in the simulation framework in Lustig, at al. (2011), which was calibrated to match the historical risk and return properties of currency returns, risk-free rates, and inflation. Under their “restricted” model with constant loadings, the calibration requires loadings to range from $\sqrt{\frac{\delta}{\delta^*}} = 0.81$ (high interest rate currencies) to $\sqrt{\frac{\delta}{\delta^*}} = 1.16$ (low interest rate currencies), when measured as a fraction of the loading of the reference currency ($\delta^*$). Our estimates of global factor loadings based on option data fall into a comparable range.

We find a strong unconditional relationship between the mean global factor loading differentials relative to the US, and the mean one-month interest rate differentials relative to the US (Table I; Panel A). The slope of this relation in the baseline calibration is -4.6 (t-stat: -3.8). We also observe a negative conditional, cross-sectional relation between loading differentials at time $t$ and the contemporaneous interest rate differentials. These findings indicate that interest rate differentials are proxies for risk factor loadings, and provide independent empirical support for the (“unrestricted”)
specification postulated in Lustig, et al. (2011). Section 4 provides a detailed discussion of the relation between global factor loadings and interest rate differentials.

**Pricing kernel innovations, \( L_{t+1}^g \) and \( L_{t+1}^i \)**

Panel B of Table I reports the distributional properties of the global factor innovations, \( L_{t+1}^g \) – their variance (\( Z_t \)), the fraction of variance attributable to non-Gaussian shocks (\( \eta^g_t \)), and the parameters of the one-sided CGMY distribution (\( G^g_t, Y^g_t \)). The average annualized variance of the global pricing kernel innovation ranges from 0.27 to 0.33 across the two calibrations, with 30-40% of the total variance (\( \eta^g_t \)) contributed by the non-Gaussian component. Correspondingly, the price of risk for the global risk factor (i.e. the \( HML_{FX} \) Sharpe ratio) – which is approximately given by \( \sqrt{Z_t} \) at each point in time – averages 0.43 and 0.47 in Specifications I and II, respectively. The time-series standard deviations of the global price of risk are approximately 0.30 in both calibrations.

In each of the specifications, the skewness of the global innovation, \( L_{t+1}^g \), is time-varying and depends on the level of the global state variable, \( Z_t \), and the parameters of the CGMY distribution. The mean skewness (kurtosis) of the global factor innovation is approximately equal to -0.6 (5.5) under Specification I, and increases in magnitude as the set of test assets is expanded. These values are comparable to, though somewhat less extreme, than the realized moments of currency carry trade portfolios reported in Jurek (2014), which represent empirical factor mimicking portfolios for the global \( (HML_{FX}) \) factor in G10 currencies.

Appendix Table A.I reports an analogous summary for the distributions of the country-specific innovations, \( L_{t+1}^i \). The innovations have mean volatilities, \( \sqrt{Y^i_t} \), between 6-10% per annum, and are roughly five times less volatile than the global factor innovation. The country-specific innovations exhibit meaningful departures from Gaussianity for all G10 currencies, with average skewness ranging from -0.20
(SEK) to -1.13 (AUD); and kurtosis – from 3.96 (SEK) to 6.92 (CHF). This feature is consistent with earlier evidence from Bakshi, et al. (2008), and challenges the identifying restriction in Farhi, et al. (2014), which imposes that country-specific innovations are Gaussian. When evaluated at the averages of the parameter estimates, the model produces an annualized exchange rate volatility, \(2.29\), of 11% per annum for a currency pair with a global factor loading spread of zero, and – 15% for a currency pair with a loading spread of 0.20 (e.g. a typical carry trade pair). These values are in line with the realized volatilities of G10 cross rates in our sample, which average 11% per annum, and range from 6% (EUR/CHF) to 17% (NZD/JPY).

2.3.3 Fitting errors

Panel C of Table I reports the results of fitting the option-implied volatilities, and the one-month ahead \(\mathbb{P}\)-measure volatility forecasts. Specifically, the table reports the root mean squared fitting errors (in volatility points) based on the calibration fitting criterion. For option-implied volatilities, we report the fit by strike, and across all strikes jointly. Recall that, for each exchange rate pair included in the calibration set, we price options on each day (1999:1-2012:6; \(T = 3520\) days) at each of the five quoted strikes (10\(\delta\) put to 10\(\delta\) call). The model option pricing RMSE stands at roughly 0.8 volatility points for the combined HLX (high-low cross pairs) and X/USD option set; and at 1.1 volatility points – for the full panel of 45 G10 cross-rate options. These errors are generally within typical bid-ask spreads in FX option markets (Jurek (2014)), indicating that the model is doing a reasonable job of matching the option prices. The fit between the model and empirical volatility forecasts under the objective measure is comparable, with a fitting criterion RMSE of 0.8 volatility points. Appendix Table A.II explores the sensitivity of each calibrations’ fitting error to imposing: (a) Gaussian idiosyncratic innovations \(\eta^i_t = 0\); (b) Gaussian global innovations \(\eta^g_t = 0\); (c) Gaussian global and idiosyncratic innovations \(\eta^i_t = \eta^g_t = 0\); and, (d) fixing the
global pricing kernel innovations at their time-series means \( (\xi_i = \frac{1}{T} \cdot \sum \xi_i) \). We find that forcing global pricing kernel loadings to be constant results in the greatest deterioration in the fit of the model, with the RMSE increasing by a factor of 4-5x. By comparison, setting the distribution of the global (country-specific) innovation to be Gaussian increases the RMSE by 40-60% (15-35%) depending on the set of test assets. Moreover, we observe that the non-Gaussianity of the global factor innovation is particularly important for matching option prices during 2008.

Figure 1 illustrates the time-series match between the observed and fitted estimates of volatility under the pricing (top panel) and objective (bottom panel) measures. To compactly summarize the quality of the fit, we extract the first principal components from the panels of the observed and fitted at-the-money option-implied volatilities for the set of 24 exchange rate pairs from Specification I. The top panel of Figure 1 plots the first principal components from each panel, scaled to match the moments of the underlying panel data. Specifically, we set the mean (volatility) of each principal component equal to the time-series mean (volatility) of the cross-sectional average of the series in the corresponding panel. Consistent with the low FX option fitting RMSE reported in Panel C of Table I, we find that the calibrated model faithfully reproduces the aggregate variation in the level of option-implied volatilities observed in the data. The second panel repeats this exercise replacing implied volatilities with model and empirical forecasts of the monthly return volatility under the objective measure. Again, we find the time-series fit is very good, indicating the calibrated model is able to accurately capture the volatility dynamics both under the pricing and objective measures.\(^{13}\)

Figure 2 illustrates the quality of the fit under our baseline calibration across strikes, by plotting the mean observed option-implied volatilities (blue) and their

\(^{13}\)Since the first principal component accounts for over 80% of the aggregate variation in the panels of raw and fitted volatilities, both under the pricing and objective measures, the fit of the calibration remains very good at the level of individual currency pairs. Appendix Figure A.1 (A.2) illustrates this by plotting the fit for volatilities of individual X/USD pairs under the pricing (objective) measure.
fitted counterparts (dashed red) for cross pairs formed by combining two high interest rate currencies (AUD, NOK) with two low interest rate currencies (CHF, JPY), as well as, pairs involving each of these currencies against the U.S. dollar. We additionally plot a typical bid-ask spread (dashed blue lines) equal to 0.1 times the mean quoted implied volatility at each strike. Taken together with the evidence in Figure 1 and Table 1, this plot indicates our calibration is able to match both the time-series and cross-sectional dimensions of FX implied volatility data.
Chapter 3

Currency Risk Premia

3.1 Option-Implied Currency Risk Premia

The calibration yields a time series of model parameters and state variables, which we use to compute conditional currency risk premia for individual currency pairs, as well as, commonly-used empirical factor mimicking portfolios. These estimates represent \textit{ex ante} measures of currency excess returns over the subsequent month, and are free from peso problems. We demonstrate that the option-implied risk premia provide unbiased forecasts of future currency returns in the cross-section and time-series, and achieve superior explanatory power relative to \textit{ad hoc} specifications based on interest rate differentials. A structural decomposition of these risk premia across distributional features of the pricing kernel innovations indicates that asymmetries (e.g. disasters) play a limited role in the determination of currency risk premia in developed economies. Finally, we contrast the dynamics of global factor loadings recovered from exchange rate options with those obtained from \textit{ex post} realized currency returns.
3.1.1 Forecasting currency returns

The basic unit of observation in our empirical analysis is a currency excess return, capturing the net return to a zero-investment portfolio which borrows one unit of currency $K$, at interest rate $y_{t,t+1}^k$ and lends in currency $J$ at the rate $y_{t,t+1}^j$. The short-term yields are expressed in annualized terms, and each interval of time, $\Delta$, equals one month. Letting $S_{t}^{jk}$ denote the currency $K$ price of one unit of currency $J$, one unit of borrowed currency $K$ buys $\frac{1}{S_{t}^{jk}}$ units of currency $J$ at time $t$. At time $t+1$ the trade is unwound and the proceeds converted back to currency $K$, generating an excess return of:

$$xs_{t+1}^{jk,k} = \exp(y_{t,t+1}^j \cdot \Delta) \cdot \frac{S_{t+1}^{jk}}{S_{t}^{jk}} - \exp(y_{t,t+1}^k \cdot \Delta)$$  \hspace{1cm} (3.1.1)

When reporting historical realized returns, we take the perspective of a U.S. dollar investor, and report U.S. dollar denominated returns. Consequently, if $K$ is not the investor’s home currency, the above return needs to be converted to the home currency, $I$, by multiplying it by $\frac{S_{t+1}^{ki}}{S_{t}^{ki}}$, to obtain $xs_{t+1}^{jk,i}$. The forecast of this quantity within our model is given by (2.2.15). If $K$ is the investor’s home currency, the model estimate of the excess return is given by (2.2.12), reflecting the added compensation necessary for being short one’s home currency.

Figure 3 summarizes the unconditional relationship between the mean ex ante option-implied risk premium and the mean subsequent realized one-month currency excess return at the pair level. The cross-sectional fit is impressive, delivering an 88% adjusted $R^2$ (left panel) for Specification I. To put this result into perspective, we repeat the same exercise for the random walk model of exchange rate dynamics, in which the ex ante interest rate differential provides the estimate of the subsequent currency excess return, and find an unconditional cross-sectional adjusted $R^2$ of 60% (right panel). The results are only modestly weaker under Specification II, which
produces an 82% adjusted $R^2$ for the model, and essentially no change for the random
walk model.

To explore the ability of option-implied risk premia to forecast currency returns
in more detail, we consider two sets of predictive regressions in Table II: (a) repeated
cross-sectional regressions (Fama-MacBeth; Panel A); and (b) pooled panel regres-
sions (Panel B). In both instances the dependent variable is the simple one-month
ahead excess return, $r_{t+1}$, and the regressions are based on the calibration output
for Specification I. The data are sampled weekly. For the cross-sectional regressions,
we report the time-series means of the coefficient estimates and the adjusted $R^2$s,
with standard errors adjusted for time-series autocorrelation. Coefficient standard
errors in panel regressions are adjusted for cross-sectional correlation, as well as time
series auto- and cross-correlations using the methodology from Thompson (2011).

The cross-sectional regressions identify a positive and statistically significant re-
lation between the ex ante option-implied risk premium, measured at time $t$, and the
subsequent realized excess return, measured between $t$ and $t+1$ (regression (1)). The
hypothesis that the model correctly prices the cross-section of currency returns (i.e.
the intercept is zero and the slope coefficient is one) has a $p$-value of 0.49. Simulta-
neously, we can reject the alternative that the model has no explanatory power (i.e.
all the coefficients with the exception of the intercept are zero) with a $p$-value of 0.03.
The adjusted $R^2$ from the cross-sectional regression is 29%, and is slightly higher
than the 26% adjusted $R^2$ from an ad hoc predictive regression using the interest
rate differential (regression (3)). When the pair-level risk premium is disaggregated
into the model-implied $HML_{FX}$ and short reference components of the risk premium
(regression (2)), the regression adjusted $R^2$ rises to 44%. The $p$-value of the null
hypothesis of the model is 0.59, while that of the alternative of no predictability is
0.05. We find similarly strong results in regressions excluding currency excess returns
from 2008 (regressions (4)-(6)).
Panel B of Table II examines the predictive ability of the model-implied risk premia in the context of panel regressions. The results clearly indicate that the time series dimension poses a significant challenge both for the model, as well as, ad hoc regressions using interest rate differentials. Although the coefficients on the model-implied variables are positive, they are statistically indistinguishable from both zero and one. As such, even though the model is not rejected in the full sample, neither is the null of no predictability (regressions (1)-(3)). This echoes the features of equity return data, where the identification of a positive time series risk-return tradeoff has remained elusive. However, this conclusion is driven entirely by data from 2008, and once these data are excluded the null of no-predictability is rejected with p-values below 0.01. At the same time, the hypothesis that the model provides an unbiased description of expected currency returns is not rejected (p-values: 0.39 and 0.59 in regressions (4) and (5), respectively). Finally, the results of the cross-sectional and panel regressions are qualitatively unaffected when the model is calibrated to the full cross-section of 45 G10 cross-rates (Specification II), as reported in Appendix Table A.III.

3.1.2 The $HML_{FX}$ and short dollar risk premia

Having verified that the calibrated model provides a reasonable description of the time-series and cross-sectional properties of excess returns for individual currency pairs, we turn to its implications for the $HML_{FX}$ and short dollar risk factors introduced by Lustig, et al. (2011). First, we compare the historical returns to the factor mimicking portfolios with the corresponding time series of option-implied risk premia to examine whether the measured risk premia were unusually high or low, relative to the risks perceived by option markets. Second, we decompose the model risk premia

---

across the moments of the underlying pricing kernel innovations, similar to the equity risk premium decompositions in Backus, et al. (2011). Specifically, we decompose risk premia into contributions from even (symmetric) and odd (asymmetric) cumulants of the pricing kernel innovations, as well as, across the moments (variance, skewness, and higher moments) of the innovations. These decompositions allow us to evaluate the channel through which non-Gaussian innovations contribute to the determination of currency risk premia, and provide new evidence on the disaster risks view of currency returns (Farhi and Gabaix (2014), Farhi, et al. (2014)), complementing earlier reduced-form studies (Burnside, et al. (2011), Jurek (2014)).

Factor mimicking portfolios

To construct the $HML_{FX}$ factor replicating portfolio we use a dollar-neutral, spread-weighted portfolio of currency excess returns, which is long (short) the G10 currencies with the highest (lowest) one-month LIBOR interest rates. Following Lustig, et al. (2011), we consider two sorting schemes: a conditional portfolio sort based on the prevailing interest rates, and an unconditional portfolio sort based on backward-looking average of the one-month interest rates. Imposing dollar-neutrality ensures the portfolio loads exclusively on the global factor innovations, $L_{t+1}^g$, which allows us to cleanly isolate the pricing of the global ($HML_{FX}$) risk factor from the effects of short dollar exposure. Finally, the short dollar mimicking portfolio is short the U.S. dollar currency against an equally-weighted basket of foreign currencies independent of the level of interest rates.

\footnote{Frameworks with (time-varying) disaster risks have been used to match the equity risk premium in consumption models (Martin (2013)), explain aggregate stock market volatility (Wachter (2013)), and as a mechanism for generating violations of uncovered interest rate parity (Farhi and Gabaix (2014)). Gabaix (2012) examines the role of disaster risks in the context of macro-finance puzzles.}

\footnote{We construct the spread-weighted portfolio by assigning portfolio weights on the basis of the absolute distance of country $i$’s interest rate from the average of the interest rates in countries with ranks five and six, as in Jurek (2014). The spread-weighting procedure is similar in spirit to forming portfolios of currencies based on interest rate sorts, and computing a long-short return between the extremal portfolios, but is more pragmatic given the small cross-section. Dollar-neutrality is imposed by constraining the sum of the nine remaining country weights to equal zero.}
The realized time series of the returns to our replicating portfolios closely match those reported by Lustig, et al. (2011), over the time period in which our data overlap (Jan. 1999 - Mar. 2010). For example, our G10 $HML_{FX}$ factor replicating portfolio return series has a 95% correlation with their developed-market $HML_{FX}$, and explains 90% of its time series variation during the time period in which our data overlap. The mean return on our spread-weighted G10 currency portfolio during this time period is 5.49% per annum, which is also in line with values reported in Lustig, et al. (2011). Their values range from 4.86% to 7.26% per annum depending on whether the replicating portfolio is constructed only using developed market currencies, or their full currency dataset (15 developed, 20 emerging). Similarly, the short dollar factor implemented in G10 currencies has a 99% correlation with the developed-market $RX$ factor return series reported by Lustig, et al. (2011), and explains 99% of its time series variation when our data overlap. These facts lead us to conclude that forming the empirical factor replicating portfolios using the G10 currency set, rather than the broader set of developed market currencies used by Lustig, et al. (2011), is unlikely to affect subsequent inference.

The annualized volatilities of the factor mimicking portfolios are all approximately equal to 9%, and the monthly returns are negatively skewed (Table III). The realized skewness values range from -0.17 for the short dollar factor to -1.07 for the conditional $HML_{FX}$ portfolio. All historical point estimates lie within their bootstrapped confidence intervals obtained by drawing repeated samples of equal length from the historical record. Appendix Figure A.3 plots the time series of the cumulative realized returns of each of the factor replicating portfolios. The monthly returns exhibit stochastic volatility, and the Jarque-Bera test rejects Gaussianity both for the returns and standardized innovations (Z-scores), consistent with previously reported empirical evidence (e.g. Brunnermeier, et al. (2009), Chernov, et al. (2014), Jurek (2014)). These features substantiate the use of non-Gaussian distributions to model
the periodic global and country-specific pricing kernel innovations, $L_{t+1}^g$ and $L_{t+1}^i$, as well as, the use of the global, $Z_t$, and country-specific state variables, $Y_t^i$, to capture time variation in second moments.

Realized vs. option-implied risk premia

Table III reports the mean realized returns on the factor mimicking portfolios, and the corresponding model-implied risk premia based on the calibrated model parameters from Specification I. In the historical sample (1999:1-2012:6; 162 months), the mean returns for the conditional and unconditional $HML_{FX}$ factor mimicking portfolios were 4.96% (t-stat: 1.92) and 3.32% (t-stat: 1.29) per annum, respectively. The short dollar factor mimicking portfolio earned 3.12% per annum (t-stat: 1.29). The low significance of the estimates is driven by the short sample period, and once the sample is not constrained to include FX option data, the mean risk premia become statistically significant (1990:1-2012:6; Jurek (2014)).

For comparison, the calibration generates mean model-implied risk premia for the conditional and unconditional $HML_{FX}$ portfolios of 3.87% (t-stat: 7.35) and 3.80% (t-stat: 7.53) per annum, respectively (Table III). Since the empirical factor replicating portfolios are dollar-neutral, the entirety of the model-implied premium reflects compensation for exposure to the global pricing kernel shocks, $L_{t+1}^g$. The presence of a statistically significant model-implied risk premium for the $HML_{FX}$ replicating portfolios reflects the strong cross-sectional alignment between mean interest rate differentials and mean global factor loadings (Table 1). As a result, sorting currencies into long and short portfolios on the basis of historical average interest rates generates a spread in global factor loadings ($\xi^{long} = 0.86, \xi^{short} = 1.02$; conditional $HML_{FX}$), and a significant model risk premium. We find that the wedge between the historically realized and model-implied risk premia on the $HML_{FX}$ portfolios is statistically indistinguishable from zero, and is equal to 1.09% (t-stat: 0.38) for the conditional...
portfolio and -0.48% (t-stat: -0.17) for the unconditional portfolio. Notably, since the model estimates of currency risk premia are obtained on the basis of a calibration to option prices, they are free from peso problems. The lack of a statistically significant wedge between option-implied and realized excess returns, argues against the presence of peso concerns in historical estimates of risk premia in developed markets (Burnside, et al. (2011)).

The mean option-implied risk premium for the short dollar replicating portfolio is 1.97% (t-stat: 9.64) per annum, and is statistically indistinguishable from its realized counterpart (difference t-stat: 0.46). Importantly, neither of these quantities is interpretable as a pure measure of the premium for short dollar exposure, since the replicating portfolio is positively exposed to the $HML_{FX}$ factor ($\xi^{long} = 0.96$, $\xi^{short} = 1.00$). After accounting for the $HML_{FX}$ exposure, we find that only 0.54% per annum (t-stat: 22.89) of the portfolio risk premium represents compensation for short exposure to U.S. country-specific shocks (Table III). This is qualitatively consistent with the features of the historical returns of the factor replicating portfolios. Specifically, the short dollar factor replicating portfolio has a 0.43 beta (t-stat: 6.55) onto the conditional $HML_{FX}$ factor replicating portfolio, such that less than one third of the historical portfolio excess return is due to short dollar exposure.

The mean values reported in Table III mask a considerable amount of time series variation in the conditional risk premia of the factor replicating portfolios. To visualize this, Figure 4 plots the daily time series of the model-implied risk premium for the conditional $HML_{FX}$ factor replicating portfolio (top panel) and the component of the model-implied risk premium for the short dollar replicating portfolio, which is unrelated to $HML_{FX}$ exposure (bottom panel). The model risk premium for the conditional $HML_{FX}$ portfolio had a volatility of 7.59% per annum (Table IV; Panel A), and ranged from -0.73% (5th percentile) in the early part of the sample, to 13.83% (95th percentile) around the fall of 2008. The occasionally negative values
of the model risk premium are due to the fact that the cross-sectional interest rate sort doesn’t always line up with the global factor loadings obtained from the option pricing calibration. The bottom panel illustrates the time series variation in the short dollar risk premium, which ranged from 0.06% to 3.34% per annum, and had an annualized volatility of 0.31% (Table IV; Panel B).

Decomposing conditional currency risk premia

Figure 5 plots the time series decompositions of the model-implied $HML_{FX}$ and short dollar risk premia. The left panels display the risk premium contributions from the even (symmetric) and odd (asymmetric) cumulants of the underlying pricing kernel innovation; the right panels display the decompositions across the moments of the innovations (variance, skewness, and other higher-order moments). Table IV reports the corresponding summary statistics for the decompositions. We find that the even cumulants, on average account for 91.8% of the option-implied risk premium for the conditional $HML_{FX}$ replicating portfolio, with the vast majority of this (88.7%) due to the second cumulant (variance). The balance of the $HML_{FX}$ risk premium – or 0.32% per annum ($= 0.0387 \cdot 0.0818$), when measured in levels – is attributed to asymmetries in the global factor innovation. When viewed from the perspective of the moment decomposition, the share of the $HML_{FX}$ risk premium attributable to the variance of the global innovation averages nearly 90% (Table IV; Panel A), and ranges from 68.1% (5th-percentile) to 99.4% (95th-percentile) of the risk premium. By contrast, the average share of the total risk premium attributable to skewness is only 7%, rising to 16% at its 95th-percentile. For comparison, the contribution of

---

4Panel B of Table A.III reports summary statistics for a model-based factor replicating portfolio in which currencies are sorted on the basis of their calibrated global factor loadings. By construction, this portfolio has: (a) a higher model-implied risk premium; and, (b) a conditional risk premium which is strictly positive, at all points in time. The time series of returns to this portfolio has a 92% correlation with the interest-rate sorted portfolio, and a 90% correlation with the developed-market $HML_{FX}$ reported by Lustig, et al. (2011).

5Under Specification II the average contribution of asymmetries in the global pricing kernel innovation to the $HML_{FX}$ risk premium is 11.8%, or 45bps per annum (Table A.IV). Skewness and
skewness and higher-order moments to equity risk premia is on average 35% in models calibrated to historical consumption disasters (Barro (2006), Barro and Ursua (2008), Barro, et al. (2013)), but only 2% in a model calibrated to match the pricing of equity index options (Backus, et al. (2011)).

Our findings are robust to the choice of the option set in the calibration, as well as, forming the global factor replicating portfolio using the calibrated global factor loadings, rather than the interest rate differentials (Table A.IV). They are also consistent with the analysis of returns to crash-hedged currency carry trades (Jurek (2014)), which provide a non-parametric estimate of the contribution of tail risk premia to historical currency excess returns. Taken together, the data indicate that disaster risks do not appear to be a major driver of the HMLFX risk premium in developed economies. While disaster risks are inherently asymmetric, our calibrations indicate that: (a) the skewness of the global factor innovation is modest; and, (b) its contribution to the determination of risk premia is further attenuated by the small differentials in global factor loadings across G10 economies.

Other higher order moments of the innovation together account for 17.2% of the model risk premium.

Section 2 of the supplementary Technical Appendix explores the mechanism through which disasters contribute to the determination of equity risk premia. The expressions for equity risk premia are obtained from Martin (2013), and the parameters values describing the consumption growth process in macro models and implied by equity index options are from Backus, et al. (2011).

The HMLFX risk premium can also be decomposed into contributions from the Gaussian, \( W_{t+1}^g \), and non-Gaussian, \( X_{t+1}^g \), components of the global pricing kernel innovation, \( L_{t+1}^g \). Specifically, the share of the HMLFX risk premium for exchange rate J/I attributable to the non-Gaussian component is:

\[
\phi_{HML,t}^{i,j} = \frac{k_{X,t}^g \left[ \xi_i^t - \xi_j^t \right] + k_{X,t}^g \left[ -\xi_i^t \right] - k_{X,t}^g \left[ -\xi_j^t \right]}{k_{L,t}^g \left[ \xi_i^t - \xi_j^t \right] + k_{L,t}^g \left[ -\xi_i^t \right] - k_{L,t}^g \left[ -\xi_j^t \right]} \cdot \eta_t^g
\]

Importantly, since the non-Gaussian shock captures both the effects of stochastic volatility and jumps within a one-month interval, the share of the risk premium attributable to the non-Gaussian component overstates the role of jumps. We find that the risk premium attributable to the non-Gaussian component of the global innovation represents 43% (33%) of the total model (realized) HMLFX factor mimicking portfolio risk premium. However, the primary channel through which this component affects the HMLFX risk premium is by increasing the variance, rather than the skewness, of the global pricing kernel innovation.
The bottom panels of Figure 5 plot the decomposition of the short dollar risk premium time series. Recall that, within the model, this risk premium is determined by: (a) the average level of the country-specific state variable, $Y_{i,t}$; and, (b) the even cumulants/moments of the country-specific pricing kernel innovation, $(2.2.20)$. We find that the share of short dollar risk premium attributable to the variance of the innovations is consistently above 99%, indicating a very limited role for higher, even moments (e.g. kurtosis). We conduct similar decompositions for other currencies in our dataset, since our model does not assign a special role to the U.S. dollar, and investors in each country will demand compensation for being short their local currency (Table IV; Panel B). In the cross section, the mean level of the short reference currency risk premium varies between 0.38% per annum for the Australian dollar (AUD) to 1.07% for the Swedish krona (SEK). The time series variation of the risk premium within each country is generally modest, with volatilities an order of magnitude lower than for the $HML_{FX}$ risk premium. The decomposition of the short-reference risk premia across the features of the distribution governing the country-specific pricing kernel innovations, $L_{i,t+1}$, again highlights variance as the key determinant of risk premia. Across all the G10 countries, the share attributable to higher moments is low, and generally below 1% of the prevailing risk premium.

3.1.3 The factor structure of currency returns

Existing empirical insights regarding the cross section of currency returns have been established using panel data on historical returns from currency spot and futures markets. The calibration complements these results by revealing the factor structure of currency returns from the perspective of the FX option market. A distinct feature of this approach is that the factor structure is pinned down from ex ante estimates of the higher-order moments of currency returns (i.e. variance forecasts and option prices), rather than the path of realized returns. Pragmatically, the calibration provides a
parametric description of the conditional distribution of monthly currency returns – both under the objective and risk-neutral measures – for each day in our sample, which we utilize to simulate currency returns. Using this information, we are able to evaluate whether the realized sample was statistically anomalous (e.g., affected by peso problems), and contrast the option-implied factor structure with estimates derived from realized returns.

The empirical analysis in this section relies on simulated monthly panel data, which is generated as follows: (1) for each month in the simulation, we draw a set of parameters describing the pricing kernels from the time series of the calibrated model parameters and observed yields; (2) using these parameters we simulate realizations of the monthly global and country-specific innovations, $L_{i+1}$ and $\{L_i\}_{i=1}^{N=10}$, and compute the corresponding one-month currency returns. Following this recipe, we generate panels ($K = 10,000$ draws) of one-month currency returns matching the size of the historical sample (10 currencies, 162 months). Within each of these simulated panels, we construct the returns to the conditional $HML_{FX}$ and short dollar replicating portfolios. Table V reports the summary statistics for the factor mimicking portfolios (Panel A), and reports the results of factor regressions (Panel B).

For both risk factors, we find that the mean estimates of the risk premium, volatility, skewness and Sharpe ratio, obtained via the model simulation are all statistically indistinguishable from their historical point estimates ($t(data)$). The close match between the historical and simulated quantities, suggests that the realized sample is unlikely to have been affected by peso problems. For example, the historical mean returns to the conditional $HML_{FX}$ factor and the short dollar factor, correspond to

---

50

---

Rather than independently simulate the Gaussian and CGMY increments of pricing kernel shocks, we draw innovations from a normal inverse Gaussian distribution with matched mean, variance, skewness and kurtosis. To ensure consistency in the computation of the pricing kernel drifts, $\alpha^i_t$, we replace the model distribution CGF with the NIG CGF. We find that this adjustment speeds up the simulation and is inconsequential to the subsequent empirical results.
the 59th and 69th percentiles of their simulated distributions, respectively. Moreover, the skewness of the realized returns is actually more negative than the mean skewness observed in the model simulation, suggesting that the realized sample wasn’t unusually “lucky.”

To parsimoniously investigate the implications of our calibrated model for the factor structure of currency returns, we contrast the results of factor regressions performed using realized and simulated currency returns. In the spirit of the specifications in Lustig, et al. (2011) and Verdelhan (2013), we regress simple currency excess returns (measured relative to the U.S. dollar) onto the returns to the dollar-neutral, spread-weighted $HML_{FX}$ factor mimicking portfolio, and the returns to an equally-weighted short dollar factor mimicking portfolio ($RX$), each of which is constructed using the underlying (simulated or historical) currency pair excess returns. The panel regressions also include the $HML_{FX}$ factor interacted with: (1) the mean currency pair interest rate differential (capturing unconditional variation in loadings in the cross-section); and, (2) the pairwise demeaned one-month interest rate differential (capturing conditional variation in loadings in the time series):

$$x_{s,t+1}^j = \phi_0^j + \left( \phi_1^j + \phi_{11}^j \cdot \frac{y_{t,t+1}^j - y_{t+1}^S}{y_{t,t+1}^j - y_{t+1}^S} \right) \cdot HML_{FX,t+1}$$

The results of these regression in the historical sample, and in data simulated from the calibrated option model, are reported in Panel B of Table V.

The panel regression using historical data confirms evidence previously reported by Lustig, et al. (2011) and Verdelhan (2013). We find a strong cross-sectional relation between the loadings of currency excess returns onto the $HML_{FX}$ factor and mean interest rate differentials, consistent with the success of unconditional currency
carry trades ($\hat{\phi}_{11} = 23.93; \text{t-stat: 7.36}$). There is also a weaker, conditional link between the $HML_{FX}$ loading and the prevailing interest rate differential ($\hat{\phi}_{12} = 2.78; \text{t-stat: 1.97}$). Lastly, the regression confirms the presence of a short dollar factor ($\hat{\phi}_2 = 0.99; \text{t-stat: 13.29}$), and achieves an adjusted $R^2$ of 71.1%. We then repeat these regressions in panel data simulated from the calibrated model. We find that exchange rate loadings onto the $HML_{FX}$ factor are strongly positively tied to mean interest rate differentials in the cross-section ($\hat{\phi}_{11} = 24.28; \text{t-stat: 18.40}$), and exhibit positive, though weak, time series covariation with the prevailing interest rate differentials ($\hat{\phi}_{12} = 2.83; \text{t-stat: 1.39}$). Finally, since the excess returns of each currency pair are measured relative to the U.S. dollar, we recover a unit exposure to the short dollar factor. Strikingly, the coefficients governing the factor structure of currency returns recovered from FX option data are statistically indistinguishable from their counterparts in historical spot returns (t-statistics of differences in curly braces). The factor regressions generate a mean adjusted $R^2$ value of 58.4%, which is lower than – but statistically indistinguishable from – the goodness-of-fit observed in the historical sample.

The factor loading dynamics uncovered from the model simulation, support the results in Lustig, et al. (2011), and confirm that interest rate differentials act as proxies for exposure to a global risk factor ($HML_{FX}$). However, despite a strong cross-sectional relation between global factor loadings and interest rates, the link in the time series dimension appears to be weaker. The ensuing section presents additional implications of this finding for: (a) the model’s ability to generate a spread between the risk premia of the conditional and unconditional factor replicating portfolios; and, (b) the dynamics of option-implied exchange rate moments.
3.2 Discussion

Our methodology provides a novel vantage point – the FX option market – for assessing the role of time-varying factor loadings for explaining the cross-section of currency returns. Conceptually, the existence of a link between the prevailing global factor loadings and interest rate differentials plays a central role in generating a wedge between the risk premia on conditional and unconditional $HML_{FX}$ factor mimicking portfolios. For example, Lustig, et al. (2011) link global factor loadings, $\xi^i_t$, to the country-specific state variable, $Y^i_t$, such that the yield – which is affected by $Y^i_t$ – carries additional information about the conditional global factor loading, and therefore conditional currency risk premia.\footnote{Their preferred (“unrestricted”) specification can be interpreted as a special case of our model in which: (a) the non-Gaussian component of the pricing kernel innovations has been shut down ($\eta^g_t = \eta^i_t = 0$); and, (b) the global factor loadings have been parameterized as a function of the country-specific state variables via $\xi^i_t = \xi^i \cdot \sqrt{1 + \frac{\gamma^1_i}{\gamma^2} \cdot \frac{Y^i_t}{Z_t}}$.}

We examine these effects in the context of our calibration, and find that: (a) the time-series variation is loadings calibrated from option data is only weakly linked to variation in interest rates; and, (b) imposing a strong link between loadings and interest rates induces counterfactual implications for option price dynamics.

3.2.1 The dynamics of global factor loadings

Table VI characterizes the dependence of the calibrated global factor loading differentials measured relative to the U.S. ($\xi^i_t - \xi^{US}_t$) vis-a-vis differentials in interest rates and the state variables governing the volatility of the country-specific shocks. Specifically, we report the results of panel regressions of global factor loading differentials onto the contemporaneous differentials in interest rates ($y^i_{t,t+1} - y^{US}_{t,t+1}$) and country-specific state variables ($Y^i_t - Y^{US}_t$), as well as, versions of those variables which have been re-scaled by their contemporaneous cross-sectional dispersion. The data are sampled weekly. The table reports coefficient estimates along with standard
errors which have been adjusted for cross-sectional correlation, as well as time series auto- and cross-correlations using the methodology from Thompson (2011). Regressions (1)-(4) include currency-pair fixed effects; regressions (5)-(8) parameterize the pair fixed effects as a function of the time-series mean of the one-month interest rate differential. To highlight the fraction of explanatory power due to the time series variation in the explanatory variable, we report the raw adjusted $R^2$, along with a goodness-of-fit measure net of the explanatory power of the fixed effects (Adj. $R^2$ [NFE]).

The regressions indicate that the global factor loadings of foreign countries tend to decline when their interest rate differential–measured relative to the U.S. one-month interest rate–is above its time-series average (regressions (1)-(2)). The link is borderline significant, but becomes somewhat stronger when the loading differentials are related to the differential in the quantity of country-specific risk (regressions (3)-(4)). While these results are consistent with the qualitative features of the unrestricted specification of the affine model in Lustig, et al. (2011), the magnitude of this covariation appears to be too weak to generate a meaningful wedge between the model-implied risk premium for the conditional and unconditional trades (Table III). The panel regressions with currency-pair fixed effects uniformly deliver adjusted $R^2$ values of approximately 50%. However, the bulk of this explanatory power comes from the inclusion of fixed effects. Once the fixed effects are excluded, the adjusted $R^2$ drops to under 5% for the regressions using interest rates, and under 8% for regressions using the country-specific state variables.

To explore the determinants of the loading differentials in the cross-section in greater detail, we repeat the panel regressions parameterizing the currency pair fixed effects as a function of the time-series mean of each pair’s interest rate differential (regressions (5)-(8)). These regressions point to a strong negative, unconditional association between the mean global factor loading and interest differential in the
cross-section, with the coefficient estimate on the mean interest rate differential ranging from -3.23 (t-stat: -2.37) to -4.83 (t-stat: -4.23). In economic terms, this implies that for every one-percentage point in the mean interest rate differential, a pair’s required $HML_{FX}$ risk premium rises by roughly 108 basis points per annum.\(^{10}\) The overall regression $R^2$'s decline to slightly above 30%, and – as before – are weakly related to the time-series variation in the explanatory variable, such that the $R^2$ net of the parametric currency fixed effect ranges from 2.4% to 7.3%. The results of the regressions indicate that, from the perspective of the FX option market, global factor loadings appear to be a fixed characteristic of currency pairs, rather than a function of the prevailing differential in interest rates or the quantity of country-specific risk. While mean loading differentials line up closely with the mean historical interest rate differentials in the cross-section, the time series covariation between loadings and interest rate differentials or the quantity of country-specific risk plays a minor role in the dynamics of global factor loadings.

### 3.2.2 Option-pricing implications of linking loadings to interest rates

Risk-based explanations of deviations between the returns to conditional and unconditional factor mimicking portfolios suggest linking global factor loadings to interest rate differentials. We show that in the context of models featuring asymmetries in global factor loadings (Backus, et al. (2001), Bakshi, et al. (2008), Lustig, et al. (2011, 2014)), such a link would induce covariation between the cumulants of the option-implied distribution and interest rate differentials. This provides a novel

\(^{10}\)This illustrative computation assumes that: (a) a one percentage point increase in the interest rate differential decreases a currency’s loading by -0.04 (-4·0.01), based on a crude average of the coefficient estimates from the panel regression; (b) the global state variable, $Z_t$, is at its long-run mean (0.27; Table I); and, (c) the distribution of the global shock is Gaussian, such that the $HML_{FX}$ risk premium is given by $\xi_t^{US} \cdot (\xi_t^{US} - \xi_t) \cdot Z_t$. 

55
approach to assessing the presence of such a mechanism using exchange rate option data.

To quantitatively illustrate the impact of linking global factor loadings on risk premia and option prices, we parameterize the global factor loadings as follows:

\[
\tilde{\xi}_t - \xi_t^{US} = \phi_0 + \phi_1 \cdot \left(y_{t,t+1}^i - y_{t,t+1}^{US}\right) + \phi_2 \cdot \left(y_{t,t+1}^i - y_{t,t+1}^{US} - y_{t+1,t}^i - y_{t+1,t}^{US}\right)
\]

(3.2.3)

Specifically, we set \(\phi_0\) and \(\phi_1\) equal to their point estimates from regression specification (5) in Table VI, and mechanically vary the magnitude of \(\phi_2\); \(\xi_t^{US}\) is normalized to one for all \(t\). We leave the dynamics of the global and country-specific state variables, as well as, the parameters governing the distribution of the shocks unchanged from Specification I. We then reprice the exchange rate options and compute: (1) the option pricing errors; (2) the sensitivity of option-implied moments (variance and skewness) to the interest rate differentials; and, (3) the model-implied risk premia for conditional and unconditional factor mimicking portfolios. We contrast these values with those observed in the raw data, and under our preferred calibration in which global factor loadings vary non-parametrically (Specification I). Table VII reports the results of the analysis.

Imposing a link between global factor loadings and interest rate differentials generally leads to an increase in the option pricing RMSE, relative to that obtained under the non-parametric loading dynamics of Specification I (Model). Moreover, as \(\phi_2\) increases, panel regressions indicate a significant increase in the covariation of the option-implied moments (variance and skewness) and interest rate differentials. For example, in the data the covariation of option-implied variance with the interest rate differentials is insignificant, such that the panel \(R^2\) is due entirely to the inclusion of currency pair fixed effects (Panel A). This feature is reproduced in our calibration
(Model). However, as we increase $\phi_2$, the slope coefficient on the interest rate differential becomes significant, and the $R^2$ net of the fixed effects rises. These effects are even more pronounced for the option-implied skewness (Panel B). The preferred calibration already induces a stronger link between the model option-implied skewness and interest rate differentials than is observed in the data. Forcing an even tighter link between loadings and interest rate differentials by increasing $\phi_2$, leads to a large increase in the slope coefficient, and the panel adjusted $R^2$. In particular, the $R^2$ net of the fixed effects rises dramatically to over 60%, whereas the corresponding value in the data is only 0.3%. An interesting corollary of this experiment, is that testing for covariation between option-implied moments and other observables may provide an effective method for identifying the drivers of global factor loadings.

Finally, Panel C of Table VII reports the implications of the various loading specifications for estimates of option-implied risk premia for the $HML_{FX}$ and short dollar factor replicating portfolios implemented in G10 currencies. Indeed, we find that forcing a tighter link between the global factor loadings and interest rate differentials results in an increase in the wedge between the option-implied risk premia on the conditional and unconditional factor mimicking portfolios. For example, as $\phi_2$ ranges from -1 to -10, the difference between the mean model risk premia for the conditional and unconditional $HML_{FX}$ portfolios rises from 0.36% to 2.41% per annum. However, the magnitude of $\phi_2$ necessary to bring the model wedge in line with the point estimate observed in the historical sample, would result in a counterfactually high degree of covariation between the option-implied moments of exchange rates and interest rate differentials.
3.3 Conclusion

We derive *ex ante* estimates of currency risk premia using cross-sectional information on option-implied exchange rate distributions and currency return variance forecasts. Unlike estimates obtained from panel data on realized currency returns, option-implied risk premia are free from peso problems, and can be constructed in real time. To obtain these values we assemble a non-Gaussian factor model of exchange rate dynamics, and calibrate it using options on G10 exchange rates. The availability of cross-rate options (i.e. options on all possible currency pairs) enables the identification of the complete factor structure of currency returns, yielding conditional risk premium estimates for the full cross-section of currency pairs. The results of our empirical analysis point to a risk-based explanation of the returns to commonly used empirical factor mimicking portfolios (e.g. $HML_{FX}$ and short dollar; Lustig, et al. (2011, 2014)), and facilitate a structural decomposition of currency risk premia, yielding new evidence on the role of disaster risks in currency markets.

We find that option-implied risk premia provide unbiased forecasts of subsequent currency excess returns at the one-month horizon, and the null hypothesis that the model accurately describes the data is never rejected in cross-sectional and panel contexts. The model achieves a mean adjusted $R^2$ of up to 44% in cross-sectional (Fama-MacBeth) regressions, exceeding the 26% adjusted $R^2$ obtained from *ad hoc* regressions based on interest rate differentials. When the data are viewed unconditionally, the mean model risk premium explains nearly 88% of the cross-sectional variation in mean G10 currency excess returns.

Aggregating the pair-level option-implied risk premia, we produce the corresponding time series of conditional risk premia for the empirical $HML_{FX}$ and short dollar factor replicating portfolios (Lustig, et al. (2011, 2014)) in developed market currencies. We find that the model risk premia for these portfolios are statistically indistinguishable from their realized counterparts indicating that historical estimates
of currency returns in developed economies are unlikely to suffer from peso problems. Across the two sets of test assets used to calibrate the model, the share of the $HML_{FX}$ risk premium attributable to the higher-order moments (skewness, kurtosis, etc.) of global risks is below 20%, or 70bps per annum in levels. In particular, asymmetries in the distribution of the global pricing kernel innovation – captured by its odd cumulants – on average account for roughly 10% of the risk premium. These results are consistent with the observation that crash-hedged currency carry trades continue to deliver positive excess returns (Jurek (2014)), and suggest that extreme outcomes play a secondary role in determining currency risk premia in developed markets. The modest risk premium contribution from the higher-order moments of the aggregate shock parallels the results for equity risk premia, when models are calibrated to match equity index option prices (Backus, et al. (2011)), but contrasts with inferences based on models calibrated to consumption data (Barro and Ursua (2008), Farhi and Gabaix (2014)).

Finally, the factor structure calibrated from option data reveals a strong unconditional cross-sectional relation between the $HML_{FX}$ factor loadings of individual currency pairs and their mean interest rate differentials. This finding is consistent with a risk-based explanation of the positive excess returns earned by carry trades, which sort currencies using their mean historical interest rates (Lustig, et al. (2011), Hassan and Mano (2014)). However, evidence of a conditional, time-series link between global factor loadings and the prevailing interest rates is generally weaker. As a result, the wedge between the model-implied risk premia for the conditional and unconditional $HML_{FX}$ factor mimicking portfolios is considerably smaller than its historical estimate. Put differently, in the presence of such a link, the model would predict considerably stronger covariation between the moments of option-implied exchange rate distributions and interest rate differentials than observed in the data.
Chapter 4

Pricing and Hedging CoCos with Intensity-Based Models and First-Passage Time Models

A typical contingent convertible bond (CoCo) is a corporate bond that converts into equity of the issuing firm if a prespecified trigger event occurs. However, there exist different variants. Some convert into a cash payment (write-down), and others just become worthless (write-off). Motivations for issuing CoCos vary. But since the financial crisis of 2007–2009 they have been offered by a number of financial institutions to protect their capital buffers in difficult times. For instance, they have been issued by Lloyds Banking Group, Rabobank, Credit Suisse, Bank of Cyprus, Australia and New Zealand Banking Group, UBS, Zürcher Kantonalbank and Macquarie. The purpose of this paper is to develop a theoretical framework for the pricing, calibration and hedging of CoCos from the point of view of a market participant. A CoCo is specified by the following three characteristics:
- Maturity, principal and coupons. Like a standard corporate bond, a CoCo promises to make coupon payments and redeem the principal at maturity. The coupon rate can be fixed or floating.

- A trigger event causing the CoCo to convert. Different trigger mechanisms are possible. Accounting triggers are based on accounting measures of capital adequacy. Market triggers are set off by market events such as declines in stock prices or indexes. Regulatory triggers allow regulators to impose conversion on firms in financial distress. Decision triggers leave it to the firm’s management to decide when to convert the CoCo.

- A conversion mechanism describing what happens when the trigger event occurs. A typical CoCo converts into a prespecified number of equity shares. Write-down CoCos lose part of their principal, and write-off CoCos become worthless. Normally, a CoCo also pays accrued interest at conversion.

Most existing CoCos have an accounting trigger. Some have an additional regulatory trigger. Others are callable by the issuer, and some include an option for the holder to convert the CoCo early. As case studies we focus on issuances by Lloyds Banking Group in December of 2009 and Rabobank in March of 2010. Lloyds Banking Group issued different CoCos at the end of the year 2009. We consider one that was issued on the 1st of December with a maturity of 10 years. It makes fixed coupon payments and converts into a prespecified number of common equity shares if the Core Tier 1 Ratio of Lloyds Banking Group falls below 5%. The Rabo CoCo was issued on March 19, 2010 with a maturity of 10 years and fixed coupon payments. Rabobank is a cooperative society without publicly traded stock for the CoCo to turn into. Instead, it converts into an immediate cash payment of 25% of the principal amount if Rabobank’s equity capital ratio falls below 7%. For our theoretical analysis we consider the following prototype:
• The maturity is $T$, the principal amount is $F$, and there is a stream of fixed coupon payments $c_i$ at times $0 < t_1 < \cdots < t_n = T$.

• The trigger event occurs at a random time $\tau$.

• If triggered, the CoCo converts into $G$ shares of equity.

This covers the Lloyds CoCo. The Rabo CoCo is simpler. Instead of $G$ shares of equity it converts into a payment of $G$ units of currency. All our formulas can easily be adjusted to this case. A CoCo is a typical hybrid product in that it is exposed to different types of risk:

• Interest rate risk. Before conversion, a CoCo is a fixed income product and therefore sensitive to movements of the risk-free yield curve. This exposure can be hedged with government bonds or interest rate swaps.

• Conversion risk. If the Coco has a market trigger based on one or more liquidly traded securities, they can be used to hedge conversion risk. Otherwise, we propose to hedge it with CDS’s. Conversion risk is related to default risk, and CDS’s are issued with long maturities. Many of them are liquidly traded.

• Equity risk. A CoCo that potentially converts into equity is exposed to equity risk. It can be hedged with equity shares, futures or options. Write-down and write-off CoCos have no direct exposure to equity risk if they are not triggered by the stock price.

A reasonable CoCo model should include all relevant sources of risk. Moreover, for calibration and hedging purposes it should lend itself to the efficient valuation of related instruments. For a CoCo that is triggered by the market value of the issuing firm’s equity, it is enough to model the firm’s stock price and risk-free interest

\footnote{For simplicity we only consider fixed coupon payments. Floating coupon rates can be covered with minor modifications.}
rates. Other trigger mechanisms require additional model components. If conversion is not triggered by liquidly traded instruments, we propose to hedge it with CDS’s. To price them, one must describe the firm’s default time. Two kind of models have been studied extensively in the credit risk literature: first-passage time models, which go back to Merton (1974) and describe bankruptcy as an event in which the value of a firm’s total assets falls below that of its liabilities, and intensity-based models, originated by Jarrow and Turnbull (1995), in which credit events happen according to a given intensity.

We investigate both approaches to modeling the trigger event and describe default in the same way. Intensity-based models are more wide-spread in the current credit risk literature since they usually offer a better fit to the term structure of credit spreads. They can be used to describe any trigger mechanism not directly based on liquidly traded securities. In our specification they model conversion and default as jumps of a time-changed Poisson process. Therefore, both come as a surprise, and prices of the firm’s stock, the CoCo and CDS’s are expected to jump. In the first-passage time approach the trigger process can be an accounting ratio or a market price. We assume it to be continuous and observable. Then the prices of the issuing firm’s stock and the CoCo will not jump at conversion since even if the firm’s equity is diluted at conversion, investors continuously take this into account before it happens. A CoCo can be hedged by dynamically investing in enough non-redundant securities. In practice it is important that they be liquidly traded and offer exposure to the same types of risk as the CoCo. As case studies we price the CoCos issued by Lloyds Banking Group on Dec 1, 2009 and Rabobank on March 19, 2010 by calibrating an intensity-based and first-passage time model to market quotes of equity shares, risk-free yields and CDS spreads.

Most of the existing quantitative CoCo studies take a first-passage time approach. Raviv (2004) as well as Hilscher and Raviv (2012) use the barrier approach of Black
and Cox (1976) to price CoCos. De Spiegeleer and Schoutens (2012) and its generalization Corcuera et al. (2012) focus on modeling the issuing firm’s equity value and approximate accounting triggers by an event where the stock price falls below a given level. Several papers use a structural model for studying CoCo designs; see e.g. Pennachi (2010), Albul et al. (2010), McDonald (2010), Glasserman and Nouri (2010), Koziol and Lawrenz (2011), Bolton and Samama (2012), Buergi (2012), Berg and Kaserer (2012), Brigo et al. (2013), Metzler and Reesor (2013), Pennachi et al. (2013). For a critical assessment of some of the existing CoCo pricing models we refer to Wilkens and Bethke (2012).

The contribution of this paper consists in a general CoCo framework that can be specified in different ways. Its structure is as follows. In Section 4.1 we provide formulas for the pricing and calibration of CoCos in the case where conversion happens at a general stopping time. In Section 4.2 we study intensity-based models and in Section 4.3 first-passage time models. In both cases we assume the underlying uncertainty to be generated by a finite-dimensional Markov process. Then a CoCo can be hedged with a dynamic trading strategy if there exist enough liquidly traded securities with exposure to the same sources of risk. In Section 5.1 we use an intensity-based and a first-passage time model to value CoCos issued by Lloyds Banking Group and Rabobank.

4.1 General formulas for pricing, calibration and hedging

We consider a financial institution with outstanding CoCos and assume that its equity shares pay dividends at a constant rate $q \geq 0$. The market price of an equity

$\text{2}$The model could be extended to include stochastic dividends. But this would only have a minor influence on the price of a CoCo.
share and the instantaneous risk-free interest rate are modeled with stochastic processes \((S_t)_{t \geq 0}\) and \((r_t)_{t \geq 0}\). The filtration generated by all observable events is denoted by \((\mathcal{F}_t)_{t \geq 0}\), and discounted prices of future cash-flows are assumed to be martingales under a risk neutral probability measure \(Q\). In particular, \(\tilde{S}_t := e^{-\int_0^t r_s ds} e^{q t} S_t\) is a \(Q\)-martingale, and the time-\(t\) price of a risk-free zero-coupon bond with maturity \(s\) is given by \(P(t, s) := \mathbb{E}_t^Q \left[ e^{-\int_t^s r_v dv} \right]\), where \(\mathbb{E}_t^Q\) denotes the conditional expectation with respect to \(\mathcal{F}_t\). We denote the conversion time by \(\tau\) and the default time by \(\theta\). We assume both of them to be stopping times with respect to \((\mathcal{F}_t)\) and \(\tau \leq \theta\), that is, the firm cannot go into bankruptcy before conversion has been triggered. In Section 4.2, we model \(\tau\) and \(\theta\) as jump times of a time-changed Poisson process and in Section 4.3 as first-passage times of an underlying state process.

A standard CoCo can be viewed as the sum of a defaultable bond and an option delivering a fixed number of equity shares if the trigger event occurs. If the CoCo has not converted by time \(t < T\) and can be hedged with liquid instruments, its unique arbitrage-free price is

\[
C_t = \sum_{t_i > t} c_i \mathbb{E}_t^Q \left[ e^{-\int_{t_i}^t r_s ds} 1_{\{\tau > t_i\}} \right] + \sum_{t_i > t} c_i \mathbb{E}_t^Q \left[ e^{-\int_{t_i}^{\tau} r_s ds} \frac{\tau - t_{i-1}}{t_i - t_{i-1}} 1_{\{t_{i-1} < \tau \leq t_i\}} \right] + F \mathbb{E}_t^Q \left[ e^{-\int_t^\tau r_s ds} 1_{\{\tau > T\}} \right] + G \mathbb{E}_t^Q \left[ e^{-\int_{\tau}^T r_s ds} S_\tau 1_{\{\tau \leq T\}} \right].
\]

(4.1.1)

The first two terms represent the value of future coupon payments together with accrued interest paid upon conversion. The third term is the value of the principal and the last term the value of a possible conversion into equity. If the CoCo converts into a cash payment instead of equity shares, formula (4.1.1) is easily adjusted by replacing the last term with

\[
G \mathbb{E}_t^Q \left[ e^{-\int_{\tau}^T r_s ds} 1_{\{\tau \leq T, \tau < \theta\}} \right].
\]

(4.1.2)
Here, we assume the cash payment of the CoCo to be subordinate to other debt. Then, if there is a conversion before maturity, the payment only occurs if the firm does not go bankrupt at the same time.

The following result gives the price in more convenient form ($Q_t, Q^i_t, Q^*_t$ denote conditional probabilities with respect to $\mathcal{F}_t$).

**Theorem 4.1.1.** If by time $t < T$ the CoCo has not converted yet, its price $C_t$ can be written as

$$
\sum_{t_i > t} c_i P(t, t_i) Q^i_t[\tau > t_i] + \sum_{t_i > t} \frac{c_i}{t_i - t_{i-1}} E_t^Q \left[ e^{-\int_{t_{i-1}}^{\tau} r_s ds} (\tau - t_{i-1}) 1_{\{t_{i-1} < \tau \leq t_i\}} \right] + FP(t, T) Q^n_t[\tau > T] + GS_t E_t^{Q^*} \left[ e^{-q(\tau - t)} 1_{\{\tau \leq T\}} \right],
$$

(4.1.3)

where the measures $Q^i$ and $Q^*$ are defined by

$$
\frac{dQ^i}{dQ} = \frac{e^{-\int_0^{t_i} r_s ds}}{E_t^Q \left[ e^{-\int_0^{t_i} r_s ds} \right]} \quad \text{and} \quad \frac{dQ^*}{dQ} = \frac{\tilde{S}_{\tau \wedge T}}{S_0}.
$$

(4.1.4)

If the CoCo converts into a cash payment instead of equity, the last term of formula (4.1.3) has to be replaced with

$$
GE_t^Q \left[ e^{-\int_t^{\tau} r_s ds} 1_{\{\tau \leq T, \tau < \theta\}} \right].
$$

(4.1.5)

If $\tau$ is independent of $(r_s)_{t \leq s \leq T}$ with respect to $Q_t$, the first three terms of formula (4.1.3) simplify to

$$
\sum_{t_i > t} c_i P(t, t_i) Q_t[\tau > t_i], \quad \sum_{t_i > t} \frac{c_i}{t_i - t_{i-1}} E_t^Q \left[ P(t, \tau)(\tau - t_{i-1}) 1_{\{t_{i-1} < \tau \leq t_i\}} \right],
$$

(4.1.6)

$$
FP(t, T) Q^n_t[\tau > T],
$$
and the expression \([4.1.5]\) to

\[
G \mathbb{E}^Q_t [P(t, \tau) 1_{\{\tau \leq T, \tau < \theta\}}].
\]

**Proof.** It is clear that the first and third term of \([4.1.1]\) can be written as

\[
\sum_{t_i > t} c_i \mathbb{E}^Q_t \left[ e^{-\int_{t}^{t_i} r_s ds} 1_{\{\tau > t_i\}} \right] = \sum_{t_i > t} c_i P(t, t_i) Q^\tau_{t_i}[\tau > t] \tag{4.1.7}
\]

and

\[
F \mathbb{E}^Q_t \left[ e^{-\int_{T}^{T} r_s ds} 1_{\{\tau > T\}} \right] = F P(t, T) Q^\tau_{t}[\tau > T]. \tag{4.1.8}
\]

To transform the last term, one uses that \((\tilde{S}_t)\) is a \(Q\)-martingale. Therefore conditioned on \(t < \tau\), one has

\[
G \mathbb{E}^Q_t \left[ e^{-\int_{\tau}^{T} r_s ds} S_{\tau} 1_{\{\tau \leq T\}} \right] = GS_t \mathbb{E}^Q_t \left[ \frac{\tilde{S}_\tau}{S_t} e^{-q(\tau-t)} 1_{\{\tau \leq T\}} \right] = GS_t \mathbb{E}^Q_{t} \left[ e^{-q(\tau-t)} 1_{\{\tau \leq T\}} \right].
\]

If \(\tau\) is independent of \((r_s)_{t \leq s \leq T}\) with respect to \(Q_t\), the measures \(Q^\tau_{t}\) in \([4.1.7] - [4.1.8]\) can be replaced with \(Q_t\), and by first conditioning on \(\tau\), one obtains

\[
\sum_{t_i > t} \frac{c_i}{t_i - t_{i-1}} \mathbb{E}^Q_t \left[ e^{-\int_{t}^{t_i} r_s ds} (\tau - t_{i-1}) 1_{\{t_{i-1} < \tau \leq t_i\}} \right]
= \sum_{t_i > t} \frac{c_i}{t_i - t_{i-1}} \mathbb{E}^Q_t \left[ P(t, \tau)(\tau - t_{i-1}) 1_{\{t_{i-1} < \tau \leq t_i\}} \right],
\]

\[
G \mathbb{E}^Q_t \left[ e^{-\int_{T}^{T} r_s ds} 1_{\{\tau \leq T, \tau < \theta\}} \right] = G \mathbb{E}^Q_{t} \left[ P(t, \tau) 1_{\{\tau \leq T, \tau < \theta\}} \right].
\]

\[
\square
\]

**Remark 4.1.2.** It is possible that if a CoCo converts, there is a structural break in the dynamics of the stock price \(S_t\). The capital structure of the firm changes, and the management might change the strategy. However, for the valuation of CoCos it is enough to specify the stock price \(S_t\) up to time \(\tau \wedge T\) and make sure that \(\tilde{S}_{t \wedge \tau \wedge T}\)
is a $\mathbb{Q}$-martingale. If one extends the process $\tilde{S}_t$ beyond $\tau \land T$ in such a way that it stays a $\mathbb{Q}$-martingale until time $T$, one can define $\mathbb{Q}^*$ by $d\mathbb{Q}^*/d\mathbb{Q} = \tilde{S}_T/S_0$. Then formula (4.1.3) still holds since the term to which the conditional expectation $\mathbb{E}_t^{\mathbb{Q}^*}$ is applied is $\mathcal{F}_{\tau \land T}$-measurable. If the CoCo converts into a cash payment, the stock price $S_t$ does not have to be modeled, except in the case where the conversion trigger depends on it or it is needed to model CDS’s.

A meaningful model should also be able to price liquidly traded instruments that are related to CoCos such as the issuing firms’s stock, fixed income products and CDS’s. The stock price $S_t$ is a building block of our model, and prices of risk-free zero-coupon bonds are given by $P(t, s) = \mathbb{E}_t^\mathbb{Q} \left[ e^{-\int_t^s r_v dv} \right]$. Moreover, we assume that at the pricing date $t < T$, there exist $K$ liquidly traded CDS contracts with maturities $T_1 < \cdots < T_K$ lying on an equally spaced grid $t < s_1 < s_2 < \cdots$ such that $T_K \geq T$. If the CoCo has not converted until time $t$, default has not occurred either. Therefore, the time-$t$ value of a protection buyer position in a CDS with coupon times $s_i$ and maturity $T_k$ is $PL_t - CL_t$, where

$$PL_t = (1 - R) \mathbb{E}_t^\mathbb{Q} \left[ e^{-\int_t^\theta r_v ds} 1_{\{\theta \leq T_k\}} \right]$$

(4.1.9)

is the value of the protection leg and

$$CL_t = \delta \sum_{t < s_i \leq T_k} \Delta s \mathbb{E}_t^\mathbb{Q} \left[ e^{-\int_t^{s_i} r_v ds} 1_{\{\theta > s_i\}} \right] + \mathbb{E}_t^\mathbb{Q} \left[ e^{-\int_t^\theta r_v ds} (\theta - s_{i-1}) 1_{\{s_{i-1} < \theta \leq s_i\}} \right]$$

(4.1.10)

the value of the coupon leg. $R$ is the recovery rate. In reality it is random. But for simplicity, practitioners usually assume it to be constant\footnote{More generally, one can assume the recovery rate to be a random variable independent of other components of the model. This adds the flexibility that its expectation can be different under the physical and pricing measure. Instead of a fraction of the principal, one could also model the recovery amount as a fraction of the discounted principal or market value; we refer to Duffie and Singleton (1999) for more details.}(often around 40%). It
could be estimated from past defaults or with the method of Pan and Singleton (2008) from time series of CDS spreads and then use it for CoCo pricing. \( \Delta s \) denotes the time between coupon payments in years. It normally is 1/4 or 1/2. \( \delta \) is the spread and specified in the contract. Formula (4.1.10) gives the time-\( t \) value of future coupons payments including accrued interest. Before CDS’s were standardized, the spread was usually set so that no initial cash-flow was necessary. Now this is in general no longer possible. But the price of a CDS with maturity \( T_k \) is still quoted in terms of the spread that would make its current price equal to zero:

\[
\delta(t, T_k) = \frac{(1 - R) \mathbb{E}_t^Q \left[ e^{-\int_t^\theta r_s ds} 1_{\{\theta \leq T_k\}} \right]}{\sum_{t < s_i \leq T_k} \Delta s \mathbb{E}_t^Q \left[ e^{-\int_t^{s_i} r_s ds} 1_{\{s_i > s_i\}} \right] + \mathbb{E}_t^Q \left[ e^{-\int_t^\theta r_s ds} (\theta - s_{i-1}) 1_{\{s_{i-1} < \theta \leq s_i\}} \right]}. 
\]

(4.1.11)

In the models of Sections 4.2 and 4.3 below, \( \tau \) and \( \theta \) are independent of the short rate \( (r_t) \). Then it follows as in the proof of Theorem 4.1.1 that \( \delta(t, T_k) \) can be written as

\[
\delta(t, T_k) = \frac{(1 - R) \mathbb{E}_t^Q \left[ P(t, \theta) 1_{\{\theta \leq T\}} \right]}{\sum_{t < s_i \leq T_k} \Delta s P(t, s_i) \mathbb{Q}_t[\theta > s_i] + \mathbb{E}_t^Q \left[ P(t, \theta)(\theta - s_{i-1}) 1_{\{s_{i-1} < \theta \leq s_i\}} \right]}. 
\]

(4.1.12)

So to price all relevant instruments in the models of Sections 4.2 and 4.3, it will be enough to compute the following quantities:

\[
\begin{align*}
\mathbb{Q}_t[\tau > t_i], & \quad \mathbb{E}_t^Q \left[ P(t, \tau) 1_{\{\tau \leq t_i\}} \right], & \quad \mathbb{E}_t^Q \left[ P(t, \tau) 1_{\{\tau \leq T, \tau < \theta\}} \right], & \quad \mathbb{E}_t^Q \left[ P(t, \tau) 1_{\{\tau \leq t_i\}} \right] \\
\mathbb{Q}_t[\theta > s_i], & \quad \mathbb{E}_t^Q \left[ P(t, \theta) 1_{\{\theta \leq s_i\}} \right], & \quad \mathbb{E}_t^Q \left[ P(t, \theta) 1_{\{\theta \leq s_i\}} \right], & \quad \mathbb{E}_t^Q \left[ e^{-q_T} 1_{\{T \leq T\}} \right].
\end{align*}
\]

(4.1.13)

Since we are only interested in the price of a CoCo as long as it has not converted, we just have to calculate the expressions of (4.1.13) in the case \( t < \tau \). Moreover, for pricing and calibration we take the whole curve \( P(t, .) \) as given. For finitely many tenors the prices of zero-coupon bonds can be deduced from market data of government bonds or interest rate swaps. From there the curve can be completed by
interpolation. So for pricing and calibration, one does not need an explicit model for
the short rate ($r_t$). To hedge interest rate risk one can either specify the dynamics
of ($r_t$) or just immunize against the most common movements of the risk-free yield
curve. The hedging aspect will be discussed in more detail in the frameworks of the
next two sections.

### 4.2 Intensity-based models

Let ($X_t$) be a $d$-dimensional diffusion process that can be realized as the unique
strong solution of an SDE of the form

$$dX_t = a(X_t)dt + b(X_t)dW_t,$$  \hspace{1cm} (4.2.1)

where ($W_t$) is a $d$-dimensional Brownian motion and $a : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $b : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ are
deterministic functions. We assume that the instantaneous risk-free interest rate is
of the form

$$r_t = r(X_t) \quad \text{for a continuous function} \quad r : \mathbb{R}^d \rightarrow \mathbb{R}.$$  \hspace{1cm} (4.2.2)

In addition, we introduce an intensity

$$\lambda_t = \lambda(X_t) \quad \text{for a continuous function} \quad \lambda : \mathbb{R}^d \rightarrow \mathbb{R}$$  \hspace{1cm} (4.2.3)

such that $\Lambda_t := \int_0^t \lambda_s ds$ is almost surely finite. In this section we model the conversion
time $\tau$ as the first jump time of ($N_{\Lambda_t}$) and the default time as $\theta = 1_{\{\xi=0\}} \tau + 1_{\{\xi=1\}} \nu$,
where

- ($N_t$) is a standard Poisson process independent of ($W_t$).
• $\xi$ is a Bernoulli random variable independent of $(W_t, N_t)$ with distribution $Q[\xi = 0] = \alpha, \quad Q[\xi = 1] = 1 - \alpha$.

• $\nu$ is the second jump time of the process $(N_{\Lambda_t})$, where $\beta > 0$ is a constant,

$$
\lambda_t^\beta := 1_{\{t < \tau\}} \lambda(X_t) + 1_{\{t \geq \tau\}} \beta \lambda(X_t) \quad \text{and} \quad \Lambda_t^\beta := \int_0^t \lambda_s^\beta ds.
$$

We assume $(\lambda_t)$ and $(r_t)$ to be independent. Then $\tau$ and $\theta$ are independent of $(r_t)$. Moreover, we suppose that for $t < \tau$, the stock price $S_t$ equals $X_t^1$, and the first component of the SDE (4.2.1) reads as

$$
\frac{dX_t^1}{X_t^1} = (r_t - q + [\alpha(1 + \gamma) - \gamma] \lambda_t) dt + \sigma(X_t)^T dW_t
$$

for a volatility function $\sigma : \mathbb{R}^d \to \mathbb{R}^d$ regular enough such that

$$
\exp \left( \int_0^t \sigma(X_s)^T dW_s - \frac{1}{2} \int_0^t \sigma(X_s)^T \sigma(X_s) ds \right)
$$

is a $\mathbb{Q}$-martingale. Then, by conditioning on the Brownian motion $W$, one obtains that

$$
\tilde{S}_{\tau \wedge t} = \begin{cases}
S_0 \exp \left( \int_0^t \sigma(X_s)^T dW_s - \frac{1}{2} \int_0^t \sigma(X_s)^T \sigma(X_s) ds \right) & t < \tau \\
+ [\alpha(1 + \gamma) - \gamma] \Lambda_t, & t = \tau \\
0, & t \geq \tau \text{ and } \xi = 0 \\
(1 + \gamma) \tilde{S}_{\tau -}, & t \geq \tau \text{ and } \xi = 1
\end{cases}
$$

is a martingale under $\mathbb{Q}$. This model has the following features:

• The conversion time $\tau$ corresponds to the first jump time of the time-changed Poisson process $N_{\Lambda_t}$. 

71
• With probability $\alpha$ the default time $\theta$ coincides with $\tau$, and with probability $1 - \alpha$ it is equal to the second jump time of $N_{\Lambda^\beta}$.

• At the conversion time $\tau$, the jump intensity jumps from $\lambda_{\tau^-}^\beta$ to $\lambda_{\tau^+}^\beta = \beta \lambda_{\tau^-}^\beta$.

• Before conversion, the stock price $(S_t)$ is continuous, and at time $\tau$ it either jumps to 0 (default) or $(1 + \gamma)S_{\tau^-}$ (conversion without default).

• $(r_t)$ and $(\lambda_t)$ are independent, but both can be correlated with $(S_t)$.

In reality, it is difficult to predict whether and by how much the stock price will jump upon conversion. It will depend on the specifics of the CoCo contract and the situation the company will be in. Furthermore, the state of the financial sector and the broader economy might play a role. If, for instance, conversion happens during a systemic crisis, the stock price might react differently than in normal times. For simplicity we assume $\gamma$ to be a constant. Then one can consider the conclusions of the model for different values of $\gamma$, or one can estimate $\gamma$ from the composition of the balance sheet.

4.2.1 CoCo and CDS pricing with intensities

In the following proposition we provide formulas for the quantities (4.1.13). For $(t, x) \in [0, T] \times \mathbb{R}^d$, we denote by $Q_{t,x}$ the probability $\mathbb{P}$ conditioned on $X_t = x$ and $t < \tau$. For the corresponding conditional expectation we write $\mathbb{E}^Q_{t,x}$. 

72
Proposition 4.2.1. Assume $\tau$ has not occurred until time $t < T$, then

\[
\begin{align*}
\mathbb{Q}_{t,x}[\tau > t_i] & = \mathbb{E}_{t,x}^Q[e^{-(\Lambda_t - \Lambda_i)}] \\
\mathbb{E}_{t,x}^Q[P(t, \tau)1_{\{\tau \leq t_i\}}] & = \int_t^{t_i} P(t, s) \mathbb{E}_{t,x}^Q[\lambda_s e^{-(\Lambda_s - \Lambda_t)}] \, ds \\
\mathbb{E}_{t,x}^Q[P(t, \tau)1_{\{\tau \leq T, \tau < \theta\}}] & = (1 - \alpha) \int_t^T P(t, s) \mathbb{E}_{t,x}^Q[\lambda_s e^{-(\Lambda_s - \Lambda_t)}] \, ds \\
\mathbb{E}_{t,x}^Q[P(t, \tau)\tau 1_{\{\tau \leq t_i\}}] & = \int_t^{t_i} P(t, s)s \mathbb{E}_{t,x}^Q[\lambda_s e^{-(\Lambda_s - \Lambda_t)}] \, ds \\
\mathbb{Q}_{t,x}[\theta > s_i] & = \mathbb{E}_{t,x}^Q[\psi_t(s_i)] \\
\mathbb{E}_{t,x}^Q[P(t, \theta)1_{\{\theta \leq s_i\}}] & = \int_t^{s_i} P(t, s) \mathbb{E}_{t,x}^Q[\varphi_t(s)] \, ds \\
\mathbb{E}_{t,x}^Q[P(t, \theta)\theta 1_{\{\theta \leq s_i\}}] & = \int_t^{s_i} P(t, s)s \mathbb{E}_{t,x}^Q[\varphi_t(s)] \, ds \\
\mathbb{E}_{t,x}^Q[\sigma_q 1_{\{T \leq T\}}] & = (1 - \alpha)(1 + \gamma) \int_t^T e^{-qs} \mathbb{E}_{t,x}^Q[\lambda_s e^{-(1-\alpha)(1+\gamma)(\Lambda_s - \Lambda_t)}] \, ds.
\end{align*}
\]

where

\[
\varphi_t(s) = \begin{cases} 
\alpha\lambda_s e^{-(\Lambda_s - \Lambda_t)} + (1 - \alpha)\frac{\beta}{\beta - 1}\lambda_s (e^{-(\Lambda_s - \Lambda_t)} - e^{-\beta(\Lambda_s - \Lambda_t)}) & \text{if } \beta \neq 1 \\
\alpha\lambda_s e^{-(\Lambda_s - \Lambda_t)} + (1 - \alpha)\lambda_s(\Lambda_s - \Lambda_t)e^{-(\Lambda_s - \Lambda_t)} & \text{if } \beta = 1
\end{cases}
\]

and

\[
\psi_t(s) = \begin{cases} 
\alpha e^{-(\Lambda_s - \Lambda_t)} + \frac{(1 - \alpha)}{\beta - 1}(\beta e^{-(\Lambda_s - \Lambda_t)} - e^{-\beta(\Lambda_s - \Lambda_t)}) & \text{if } \beta \neq 1 \\
\alpha e^{-(\Lambda_s - \Lambda_t)} + (1 - \alpha)(1 + \Lambda_s - \Lambda_t)e^{-(\Lambda_s - \Lambda_t)} & \text{if } \beta = 1.
\end{cases}
\]

Proof. Under $\mathbb{Q}$, conditioned on $t < \tau$ and $\lambda$, the density of $\tau$ is given by $\lambda_s e^{-(\Lambda_s - \Lambda_t)}$ and the one of $\theta$ by $\varphi_t$. Accordingly, one has

\[
\mathbb{Q}_{t,x}[\tau > t_i \mid \lambda] = e^{-(\Lambda_t - \Lambda_i)} \quad \text{and} \quad \mathbb{Q}_{t,x}[\theta > s_i \mid \lambda] = \psi_t(s_i).
\]

From here, all equalities except the last one follow by first conditioning on $\lambda$. 

73
To derive the last equality, we note that

\[
\mathbb{E}_t^Q \left[ e^{-q\tau} 1_{\{\tau \leq T\}} \right] = \mathbb{E}_t^Q \left[ \frac{\tilde{S}_{\tau \wedge T}}{S_t} e^{-q\tau} 1_{\{\tau \leq T\}} \right] = \mathbb{E}_t^Q \left[ \frac{Z_{\tau \wedge T}}{Z_t^1} \frac{Z_{\tau \wedge T}}{Z_t^2} e^{-q\tau} 1_{\{\tau \leq T\}} \right],
\]

where the dynamics of \((Z_s^1)\) and \((Z_s^2)\) are given by

\[
d\frac{Z_s^1}{Z_t^1} = \sigma(X_s)^T dW_s \quad \text{and} \quad d\frac{Z_s^2}{Z_t^2} = -[\gamma - \alpha(1 + \gamma)]\lambda_s ds + [(1 + \gamma)\xi - 1]dN_{\Lambda_s}.
\]

In the next step we condition on the Brownian motion \(W\). Then \(N_{\Lambda_s} - N_{\Lambda_t}, s \geq t, \) is a point process with deterministic compensator \(\Lambda_s - \Lambda_t, s \geq t, \) and one obtains from Girsanov’s theorem for general martingales (see e.g. Theorem III.3.11 in Jacod and Shiryaev, 2003) that

\[
N_{\Lambda_{\tau \wedge s} - N_{\Lambda_t} - (1 - \alpha)(1 + \gamma)(\Lambda_{\tau \wedge s} - \Lambda_t), \ t \leq s \leq T,
\]

is a martingale under the measure \(Z_{\tau \wedge T}^2 / Z_t^2 \cdot Q_{t,x}\). It follows from the martingale characterization of Poisson processes (see e.g. Theorem II.4.5 of Jacod and Shiryaev, 2003) that conditioned on \(W, \tau\) has density

\[
(1 - \alpha)(1 + \gamma)\lambda_s e^{(1-\alpha)(1+\gamma)(\Lambda_s - \Lambda_t)}
\]

on the interval \([t, T]\) under \(Z_{\tau \wedge T}^2 / Z_t^2 \cdot Q_{t,x}\). This gives

\[
\mathbb{E}_t^Q \left[ \frac{Z_{\tau \wedge T}^1}{Z_t^1} \frac{Z_{\tau \wedge T}^2}{Z_t^2} e^{-q\tau} 1_{\{\tau \leq T\}} \right] W = \frac{Z_{\tau \wedge T}^1}{Z_t^1} (1 - \alpha)(1 + \gamma) \int_t^T e^{-qs} \lambda_s e^{(1-\alpha)(1+\gamma)(\Lambda_s - \Lambda_t)} ds,
\]

and therefore,

\[
\mathbb{E}_t^Q \left[ e^{-q\tau} 1_{\{\tau \leq T\}} \right] = (1 - \alpha)(1 + \gamma) \int_t^T e^{-qs} \mathbb{E}_t^Q \left[ \lambda_s e^{(1-\alpha)(1+\gamma)(\Lambda_s - \Lambda_t)} \right] ds.
\]
The next result is useful for calculating the expectations on the right side of the equations in Proposition 4.2.1.

**Proposition 4.2.2.** Denote

\[
\phi_{t,x}(s, z) := \mathbb{E}^Q_{t,x}[e^{-z(\Lambda_s - \Lambda_t)}], \quad \phi^*_{t,x}(s, z) := \mathbb{E}^{Q^*}_{t,x}[e^{-z(\Lambda_s - \Lambda_t)}], \quad s \geq t, \quad z > 0,
\]

and assume that \(\mathbb{E}^Q_{t,x}[\sup_{t \leq v \leq s} \lambda_v] < \infty\) as well as \(\mathbb{E}^{Q^*}_{t,x}[\sup_{t \leq v \leq s} \lambda_v] < \infty\) for all \(s \geq t\). Then

\[
- \frac{\partial \phi_{t,x}}{\partial s}(s, z) = \mathbb{E}^Q_{t,x}[z\lambda_s e^{-z(\Lambda_s - \Lambda_t)}]
\]

\[
- \frac{\partial \phi^*_{t,x}}{\partial s}(s, z) = \mathbb{E}^{Q^*}_{t,x}[z\lambda_s e^{-z(\Lambda_s - \Lambda_t)}]
\]

\[
- \frac{\partial \phi_{t,x}}{\partial z}(s, z) = \mathbb{E}^Q_{t,x}[(\Lambda_s - \Lambda_t)e^{-z(\Lambda_s - \Lambda_t)}]
\]

\[
\frac{\partial^2 \phi_{t,x}}{\partial s \partial z}(s, 1) - \frac{\partial \phi_{t,x}}{\partial s}(s, 1) = \mathbb{E}^Q_{t,x}[\lambda_s (\Lambda_s - \Lambda_t)e^{-(\Lambda_s - \Lambda_t)}].
\]

**Proof.** One has

\[
\phi_{t,x}(s, z) = \mathbb{E}^Q_{t,x}[e^{-z(\Lambda_s - \Lambda_t)}] = 1 - \int_t^s \mathbb{E}^Q_{t,x}[z\lambda_v e^{-z(\Lambda_v - \Lambda_t)}] \, dv,
\]

and since \(\mathbb{E}^Q_{t,x}[\sup_{t \leq v \leq s} \lambda_v] < \infty\), one deduces from Lebesgue’s dominated convergence theorem that \(\mathbb{E}^Q_{t,x}[z\lambda_s e^{-z(\Lambda_s - \Lambda_t)}]\) is continuous in \(s\). So \(\phi_{t,x}(s, z)\) is continuously differentiable in \(s\) with derivative \(-\mathbb{E}^Q_{t,x}[z\lambda_s e^{-z(\Lambda_s - \Lambda_t)}]\). The same argument yields that \(\phi^*_{t,x}(s, z)\) is continuously differentiable in \(s\) with derivative \(-\mathbb{E}^{Q^*}_{t,x}[z\lambda_s e^{-z(\Lambda_s - \Lambda_t)}]\).

For \(z' > z \geq 0\), one has

\[
\phi_{t,x}(s, z') - \phi_{t,x}(s, z) = -\int_z^{z'} \mathbb{E}^Q_{t,x}[(\Lambda_s - \Lambda_t)e^{-v(\Lambda_s - \Lambda_t)}] \, dv,
\]

75
and \( \mathbb{E}^Q_{t,x} \left[ (\Lambda_s - \Lambda_t) e^{-v(\Lambda_s - \Lambda_t)} \right] \) is continuous in \( v \). Therefore, \( \phi_{t,x}(s,z) \) is continuously differentiable in \( z \) with derivative \(-\mathbb{E}^Q_{t,x} \left[ (\Lambda_s - \Lambda_t) e^{-z(\Lambda_s - \Lambda_t)} \right] \). Similarly, it follows that \(-\mathbb{E}^Q_{t,x} \left[ z \lambda_s e^{-z(\Lambda_s - \Lambda_t)} \right] \) is continuously differentiable in \( z \) with derivative \( \mathbb{E}^Q_{t,x} \left[ z \lambda_s (\Lambda_s - \Lambda_t) e^{-z(\Lambda_s - \Lambda_t)} - \lambda_s e^{-z(\Lambda_s - \Lambda_t)} \right] \). Hence,

\[
\frac{\partial^2 \phi_{t,x}(s,1)}{\partial s \partial z}(s,1) - \frac{\partial \phi_{t,x}(s,1)}{\partial s} = \mathbb{E}^Q_{t,x} \left[ \lambda_s (\Lambda_s - \Lambda_t) e^{-(\Lambda_s - \Lambda_t)} \right],
\]

and the proof is complete. \( \square \)

### 4.2.2 Hedging in intensity-based models

We now address the problem of hedging a CoCo in an intensity-based model. The goal is to replicate the CoCo by investing in liquidly traded contracts. This hedges a short CoCo position. A long position is hedged with the opposite strategy. Let us assume that there exists a money market account with instantaneous return rate \( r_t \) and an additional \( J \) liquid securities. A unit of currency invested in the money market grows like \( \pi^0_t = \exp(\int_0^t r_s ds) \), and since \( X \) is a Markov process, prices of the CoCo and all other securities at time \( t < \tau \land T \) are given by \( \pi(t, X_t) \) and \( \pi^j(t, X_t) \) for functions \( \pi, \pi^j : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R} \). We assume these functions to be regular enough to admit all the derivatives we need. The following two quantities are measures of sensitivity with respect to an immediate default and conversion without default:

- The jump-to-default \( \Delta^d \pi^j(t, X_t) \) is the price jump of the \( j \)-th security if default were to occur at time \( t \).

- The jump-to-conversion \( \Delta^c \pi^j(t, X_t) \) is the price jump of the \( j \)-th security if the CoCo were to convert without default at time \( t \).

For the CoCo one has \( \Delta^d \pi(t, X_t) = -\pi(t, X_t) \) and

\[
\Delta^c \pi(t, X_t) = G(1 + \gamma)X_t^1 - \pi(t, X_t) \quad \text{or} \quad \Delta^c \pi(t, X_t) = G - \pi(t, X_t).
\]
depending on whether it converts into equity or cash.

Perfect hedge

To completely hedge a CoCo in the model of this section one needs a money market account and \(d+2\) non-redundant liquid instruments since the process \((X_t)\) has \(d\) components and there is additional uncertainty coming from the Poisson process \((N_t)\) and the Bernoulli random variable \(\xi\). By Itô’s lemma, a CoCo can perfectly be hedged with a dynamic strategy \(\vartheta^j_t, j = 0, \ldots, d+2\), satisfying for all \(t < \tau \wedge T\),

\[
\frac{\partial \pi}{\partial x^i}(t, X_t) = \sum_{j=1}^{d+2} \vartheta^j_t \frac{\partial \pi^j}{\partial x^i}(t, X_t), \quad i = 1, \ldots, d \tag{4.2.4}
\]

\[
-\pi(t, X_t) = \sum_{j=1}^{d+2} \vartheta^j_t \Delta^d \pi^j(t, X_t) \tag{4.2.5}
\]

\[
\Delta^e \pi(t, X_t) = \sum_{j=1}^{d+2} \vartheta^j_t \Delta^e \pi^j(t, X_t) \tag{4.2.6}
\]

\[
\pi(t, X_t) = \vartheta^0_t \pi^0 + \sum_{j=1}^{d+2} \vartheta^j_t \pi^j(t, X_t). \tag{4.2.7}
\]

This is a system of \(d+3\) linear equations in \(\vartheta^0_t, \ldots, \vartheta^{d+2}_t\), which, if it is regular, has a unique solution. 

\(\text{(4.2.4)–(4.2.6)}\) match the sensitivities of the replicating portfolio with those of the CoCo, and \(\text{(4.2.7)}\) determines how much cash has to be held in the money market account to make the strategy self-financing.

In theory, any \(d+2\) non-redundant contracts can be used as hedging instruments. However, in practice one would naturally try to use instruments with exposure to the same risk factors as the CoCo. We propose to hedge a CoCo that is not directly triggered by the issuing firm’s stock price with equity, CDS’s and interest rate swaps. Let us assume the following:

a) \(\pi^1\) is the stock \(S\).
b) $\pi^2, \ldots, \pi^{m+1}$ are $m \geq 2$ CDS’s. For $m = 2$, $\lambda_t$ is deterministic whereas for $m \geq 3$, it only depends on $X_t^2, \ldots, X_t^{m-1}$, and the coefficients of the SDEs for $X_t^2, \ldots, X_t^{m-1}$ do not depend on $X_t^1$.

c) $\pi^{m+2}, \ldots, \pi^{d+2}$ are $d-m+1$ interest rate swaps, $r_t$ only depends on $X_t^m, \ldots, X_t^d$, the coefficients of the SDEs for $X_t^m, \ldots, X_t^d$ do not depend on $X_t^1$, and $(X_t^2, \ldots, X_t^{m-1})$ is independent of $(X_t^m, \ldots, X_t^d)$.

Since interest rate swaps are not affected by a conversion, the system (4.2.4)–(4.2.7) then becomes

\[
\frac{\partial \pi}{\partial x^1}(t, X_t) = \vartheta^1_t \\
\frac{\partial \pi}{\partial x^i}(t, X_t) = \sum_{j=2}^{m+1} \vartheta^j_t \frac{\partial \pi^j}{\partial x^i}(t, X_t) \quad \text{for } i = 2, \ldots, m - 1 \\
\frac{\partial \pi}{\partial x^i}(t, X_t) = \sum_{j=2}^{d+2} \vartheta^j_t \frac{\partial \pi^j}{\partial x^i}(t, X_t) \quad \text{for } i = m, \ldots, d \\
-\pi(t, X_t) = -\vartheta^1_t X_t^1 + \sum_{j=2}^{m+1} \vartheta^j_t \Delta^d \pi^j(t, X_t) \\
\Delta^c \pi(t, X_t) = \vartheta^1_t \gamma X_t^1 + \sum_{j=2}^{m+1} \vartheta^j_t \Delta^c \pi^j(t, X_t) \\
\pi(t, X_t) = \vartheta^0_t \pi^0_t + \sum_{j=1}^{d+2} \vartheta^j_t \pi^j(t, X_t).
\]

Under the assumptions a)–c) one has

\[
\frac{\partial \pi}{\partial x^1}(t, X_t) = G \mathbb{E}^Q_t \left[ e^{-q(\tau-t)} 1_{\{\tau \leq T\}} \right] \quad \text{or} \quad \frac{\partial \pi}{\partial x^1}(t, X_t) = 0,
\]

depending on whether the CoCo converts into equity or cash. Since stock shares are only used to hedge equity risk, the hedging position in the stock $\vartheta^1_t$ never exceeds the number of equity shares $G$ the CoCo converts into if triggered. If conversion
is triggered by the stock price, one invests in equity shares to hedge equity and conversion risk. So investments in the stock might exceed G; see e.g. De Spiegeleer and Schoutens (2012).

**Hedging without specifying the interest rate**

Instead of specifying the full model and trying to implement a complete hedge, one can invest in the stock and different CDS’s to hedge the equity and conversion risk and then immunize against movements of the risk-free yield curve. This has the advantage that the short rate \( r_t \) does not have to be modeled explicitly. For simplicity we here only immunize against parallel shifts of the risk-free yield curve. But one can also take other movements, like twists and changes in curvature, into account. As before, we consider a CoCo that is not triggered by the issuing company’s stock price and assume a)–b) from above. But now we do not make any assumptions on the interest rate \( r_t \), except that it is a Markov process independent of \( \tau \) and \( \theta \).

Then for \( t < \tau \wedge T \), one holds \( \vartheta_1^t = G \mathbb{E}^Q_t \left[ e^{-q(\tau-t)} 1_{\{\tau \le T\}} \right] \) (or \( \vartheta_1^t = 0 \) if the CoCo converts into cash) stock shares and invests in \( m \) CDS’s such that

\[
\frac{\partial \pi}{\partial x^i}(t,X_t) = \sum_{j=2}^{m+1} \vartheta_j^t \frac{\partial \pi^j}{\partial x^i}(t,X_t) \quad \text{for} \quad i = 2, \ldots, m-1,
\]

\[
-\pi(t,X_t) = -\vartheta_1^t \pi^1(t,X_t) + \sum_{j=2}^{m+1} \vartheta_j^t \Delta^c \pi^j(t,X_t),
\]

and

\[
\Delta^c \pi(t,X_t) = \sum_{j=1}^{m+1} \vartheta_j^t \Delta^c \pi^j(t,X_t).
\]

After that one invests in fixed income products to immunize against parallel shifts of the yield curve. The sensitivity of the CoCo to a parallel shift of the yield curve
is the negative of the absolute Fisher–Weil duration (AFW)\(^4\) which for a standard CoCo, is

\[
Y_t = \sum_{t_i > t} c_i (t - t_i) P(t, t_i) Q_{t,x}[\tau > t_i] + F(T - t) P(t, T) Q_{t,x}[\tau > T] \\
+ \sum_{t_i > t} \frac{c_i}{t_i - t_{i-1}} E^{Q}_{t,x} [(\tau - t) P(t, \tau)(\tau - t_{i-1}) 1_{\{t_{i-1} < \tau \leq t_i\}}]
\]

(4.2.8)

and for a CoCo converting into cash,

\[
Y_t = \sum_{t_i > t} c_i (t - t_i) P(t, t_i) Q_{t,x}[\tau > t_i] + F(T - t) P(t, T) Q_{t,x}[\tau > T] \\
+ \sum_{t_i > t} \frac{c_i}{t_i - t_{i-1}} E^{Q}_{t,x} [(\tau - t) P(t, \tau)(\tau - t_{i-1}) 1_{\{t_{i-1} < \tau \leq t_i\}}] \\
+ (1 - \alpha) G E^{Q}_{t,x} [P(t, \tau) 1_{\{\tau \leq T\}}].
\]

(4.2.9)

\(^4\)Absolute Fisher–Weil duration measures the sensitivity to a parallel shift of the zero coupon rate curve. With our notation,

\[
P(t, s) = e^{-Z(t,s)(s-t)}
\]

where \(Z(t,s)\) denotes the zero coupon rate for \(t\) to \(s\) observed at time \(t\). For example, assume a stream of risk-free cash flows \(c_i\)'s on time \(t_i\)'s, the total value is

\[
\pi(t) = \sum_{i=1}^{N} c_i P(t, t_i) = \sum_{i=1}^{N} c_i e^{-Z(t,t_i)(t_i-t)}
\]

If the zero coupon rate curve if shift up by \(y\), then the price of this stream of cash flows becomes

\[
\pi(t, y) = \sum_{i=1}^{N} c_i e^{-(Z(t,t_i)+y)(t_i-t)}.
\]

Hence, the AFW is

\[
-\frac{\partial \pi}{\partial y} \bigg|_{y=0} = \sum_{i=1}^{N} c_i (t_i - t) e^{-Z(t,t_i)(t_i-t)}.
\]
Investments in the stock are not sensitive to movements of the risk-free yield curve, and the AFW of a protection buyer position in a CDS with maturity \( T_k \) is

\[
Y^k_t = (1 - R)\mathbb{E}_{t,x}^Q \left[ ((\theta - t)P(t,\theta)1_{\{\theta \leq T\}}) \right] \\
- \sum_{t < s_i \leq T_k} \Delta s (s_i - t)P(t,s_i)\mathbb{Q}_{t,x}[^{\theta > s_i}] + \mathbb{E}_{t,x}^Q \left[ ((\theta - t)P(t,\theta)(\theta - s_i)1_{\{s_i-1 < \theta \leq s_i\}}) \right].
\]

(4.2.10)

So parallel shifts of the yield curve can be neutralized by investing in a fixed income portfolio with AFW \( Y_t - \sum_{j=2}^{m+1} \delta^j Y^j_t \).

The durations (4.2.8)–(4.2.10) are readily computed if one knows the quantities (4.1.13) together with

\[
\mathbb{E}_{t,x}^Q \left[ P(t,\tau)\tau^21_{\{\tau \leq t_i\}} \right] \quad \text{and} \quad \mathbb{E}_{t,x}^Q \left[ P(t,\theta)\theta^21_{\{\theta \leq s_i\}} \right].
\]

It can be shown as in Proposition 4.2.1 that

\[
\mathbb{E}_{t,x}^Q \left[ P(t,\tau)\tau^21_{\{\tau \leq t_i\}} \right] = \int_t^{t_i} P(t,s)s^2\mathbb{E}_{t,x}^Q \left[ \lambda_s e^{-(\Lambda_s - \Lambda_t)} \right] ds,
\]

and

\[
\mathbb{E}_{t,x}^Q \left[ P(t,\theta)\theta^21_{\{\theta \leq s_i\}} \right] = \int_t^{s_i} P(t,s)s^2\mathbb{E}_{t,x}^Q \left[ \varphi_t(s) \right] ds.
\]

Together with Proposition 4.2.2 this allows to calculate the AFWs of the CoCo and the CDS’s.

### 4.3 First-passage time models

In this section we assume that all noise is generated by a \( d \)-dimensional diffusion process \((X_t)\) of the form (4.2.1) above. As in Section 4.2 we model the instantaneous risk-free interest rate as \( r_t = r(X_t) \) for a continuous function \( r : \mathbb{R}^d \to \mathbb{R} \). But now the conversion time \( \tau \) and default time \( \theta \) are modeled as first-passage times:
\[ \tau = \inf \{ t \geq 0 : H_t \leq h^* \}, \quad \theta = \inf \{ t \geq 0 : H_t \leq h_* \}, \quad (4.3.1) \]

where \( h^* > h_* \) are constants and \( H_t \) is a process of the form \( H_t = h(X_t) \) for a continuous function \( h : \mathbb{R}^d \to \mathbb{R} \). If \( H_t \) is the stock price \( S_t \), this becomes a model for a CoCo with a stock price trigger. But since most existing CoCos have an accounting trigger, we here concentrate on the case where \( H_t \) is a capital ratio and discuss stock price triggers in the appendix. To keep the model tractable, we assume that \((H_t)\) is continuous, observable and independent of \((r_t)\). Then \( \tau \) and \( \theta \) are predictable and independent of \((r_t)\). This means that conversion does not come as a surprise, and therefore, even though the number of equity shares increases at time \( \tau \), it should not induce a jump in the stock price since investors continuously take the possibility of conversion into account before the trigger event happens.

We assume that for \( t \leq \tau \), the stock price \( S_t \) equals \( X_t^1 \), which solves the SDE

\[ dX_t^1 = X_t^1(r_t - q)dt + X_t^1 \sigma(X_t)^T dW_t \]

for a volatility function \( \sigma : \mathbb{R}^d \to \mathbb{R}^d \) such that

\[ \tilde{S}_t = \exp \left( \int_0^t \sigma(X_s)^T dW_s - \frac{1}{2} \int_0^t \sigma(X_s)^T \sigma(X_s) ds \right) \]

is a martingale under \( Q \). Theoretically, one could add jumps to the process \((H_t)\) or assume it to be only partially observable. In this case, there would be an element of surprise if the trigger event occurs, and one would expect the stock price to jump. But it would make the model much more complex. As it is, the model has the following properties:
Conversion occurs at the first time $\tau$ when the process $(H_t)$ hits the level $h^*$ and default at the first time $\theta$ when $(H_t)$ hits $h_*$. Since $(H_t)$ is continuous, one always has $\tau < \theta$.

- The stock price $(S_t)$ is continuous before and at conversion.
- $(r_t)$ and $(H_t)$ are independent, but both can be correlated with $(S_t)$.

### 4.3.1 CoCo and CDS pricing with PDEs

The distributions of first-passage times are known in closed form only in special cases. But the quantities (4.1.13) can always be obtained by solving simple parabolic PDEs with Dirichlet boundary conditions.

Let $D := \{ x \in \mathbb{R}^d : h(x) > h^* \}$ and denote for all $(t, x) \in [0, T] \times D$ by $Q_{t,x}$ the probability $Q$ conditioned on $X_t = x$ and $t < \tau$. For the corresponding conditional expectation we write $E_{Q_{t,x}}$. The following four propositions are Feynman–Kac type results giving PDEs for $Q_{t,x}[\tau > t_i]$, $E_{Q_{t,x}}[P(t, \tau)1_{\{\tau \leq t_i\}}]$, $E_{Q_{t,x}}[P(t, \tau)\tau 1_{\{\tau \leq t_i\}}]$ and $E_{Q_{t,x}}[e^{-q\tau}1_{\{\tau \leq T\}}]$. $E_{Q_{t,x}}[P(t, \tau)1_{\{\tau \leq T, \tau < \theta\}}]$ equals $E_{Q_{t,x}}[P(t, \tau)1_{\{\tau \leq T\}}]$ in our first-passage time model since $\tau$ cannot coincide with $\theta$. The PDEs for $Q_{t,x}[\theta > s_i]$, $E_{Q_{t,x}}[P(t, \theta)1_{\{\theta \leq s_i\}}]$ and $E_{Q_{t,x}}[P(t, \theta)\theta 1_{\{\theta \leq s_i\}}]$ are the same as for $Q_{t,x}[\tau > t_i]$, $E_{Q_{t,x}}[P(t, \tau)1_{\{\tau \leq t_i\}}]$ and $E_{Q_{t,x}}[P(t, \tau)\tau 1_{\{\tau \leq t_i\}}]$ except that the conversion level $h^*$ has to be replaced with the default level $h_*$.  

**Proposition 4.3.1.** Fix $i = 1, \ldots, n$, and assume there exists a bounded function $u : [0, t_i] \times \bar{D} \to \mathbb{R}$ satisfying the PDE

$$
u_t(t, x) + Au(t, x) = 0, \quad u(t, x) = 0 \text{ for } x \in \partial D, \quad u(t_i, x) = 1_D(x), \quad (4.3.2)$$
where

\[ A u(t, x) := \sum_j a_j(x) \frac{\partial u}{\partial x_j}(t, x) + \frac{1}{2} \sum_{j,k} \alpha_{jk}(x) \frac{\partial^2 u}{\partial x_j \partial x_k}(t, x), \quad \alpha(x) := b(x)b(x)^T. \]

Then

\[ u(t, x) = \mathbb{Q}_{t,x}[\tau > t_i] \quad \text{for all} \quad (t, x) \in [0, t_i] \times D. \]

**Proof.** If \( u \) is a bounded solution of the PDE (4.3.2), then the process \( M_t = u(t \wedge \tau, X_{t \wedge \tau}) \) is a bounded martingale. So on the set \( \{ X_t = x, \tau > t \} \), one has

\[ u(t, x) = M_t = \mathbb{E}_{t,x}^Q[M_t] = \mathbb{Q}_{t,x}[\tau > t_i]. \]

\[ \square \]

**Proposition 4.3.2.** Fix \( i = 1, \ldots, n, \ t < t_i \) and \( x \in D \). If \( u : [t, t_i] \times D \to \mathbb{R} \) is a bounded solution of the PDE

\[ u_t(s, y) + Au(s, y) = 0, \quad u(s, y) = P(t, s) \ \text{for} \ y \in \partial D, \quad u(t, y) = P(t, t_i)1_{\partial D}(y), \quad (4.3.3) \]

then

\[ u(t, x) = \mathbb{E}_{t,x}^Q[P(t, \tau)1_{\{\tau \leq T\}}]. \]

**Proof.** If \( u \) is bounded solution of the PDE (4.3.3), then \( M_s = u(s \wedge \tau, X_{s \wedge \tau}) \) is a bounded martingale. So on the set \( \{ X_t = x, \tau > t \} \) one has

\[ u(t, x) = M_t = \mathbb{E}_{t,x}^Q[M_T] = \mathbb{E}_{t,x}^Q[P(t, \tau)1_{\{\tau \leq T\}}]. \]

\[ \square \]
Proposition 4.3.3. Fix $i = 1, \ldots, n$, $t < t_i$ and $x \in D$. If $u : [t, t_i] \times \overline{D} \to \mathbb{R}$ is a bounded solution of the PDE

$$u_t(s, y) + Au(s, y) = 0, \quad u(s, y) = P(t, s) s \text{ for } y \in \partial D, \quad u(t_i, y) = P(t, t_i) t_i \mathbb{1}_{\partial D}(y),$$

then

$$u(t, x) = \mathbb{E}^Q_{t, x} \left[ P(t, \tau) \mathbb{1}_{\{\tau \leq T\}} \right].$$

Proof. The proof of this proposition is the same as the one of Proposition 4.3.2.

Proposition 4.3.4. Assume $u : [0, T] \times \overline{D} \to \mathbb{R}$ is a bounded solution of the PDE

$$u_t(t, x) + A^* u(t, x) = 0, \quad u(t, x) = e^{-qt} \text{ for } x \in \partial D, \quad u(T, x) = e^{-qT} \mathbb{1}_{\partial D}(x),$$

where

$$A^* u := \sum_j a_j^*(x) \frac{\partial u}{\partial x_j}(t, x) + \frac{1}{2} \sum_{j,k} \alpha_{jk}(x) \frac{\partial^2 u}{\partial x_j \partial x_k}(t, x), \quad a^*(x) := a(x) + b(x) \sigma(x).$$

Then

$$u(t, x) = \mathbb{E}^{Q^*}_{t, x} \left[ e^{-qt} \mathbb{1}_{\{\tau \leq T\}} \right] \text{ for all } (t, x) \in [0, T] \times D.$$

Proof. By Girsanov’s theorem, $W^*_t = W_t - \int_0^{t \wedge \tau} \sigma(X_s) ds$ is a Brownian motion under $Q^*$, and one can write $dX_t = a^*(X_t) dt + b(X_t) dW^*_t$, $t \leq \tau$. So if $u$ is a bounded solution of the PDE (4.3.5), the process $M_t = u(t \wedge \tau, X_{t \wedge \tau})$ is a bounded $Q^*$-martingale. On the set $\{X_t = x, \tau > t\}$, this gives

$$u(t, x) = M_t = \mathbb{E}^{Q^*}_{t, x}[M_T] = \mathbb{E}^{Q^*}_{t, x}[e^{-qt} \mathbb{1}_{\{\tau \leq T\}}].$$
4.3.2 Hedging in first-passage time models

To study the hedging of a CoCo in a first-passage time model, we assume as above that there exists a money market account with instantaneous return \( r_t \). Then a unit of currency in the money market grows like 
\[
\pi_t^0 = \exp(\int_0^t r_s ds),
\]
and since \( X_t \) is a Markov process, the prices of the CoCo and the hedging instruments at time \( t < \tau \) are given by \( \pi(t, X_t) \) and \( \pi^j(t, X_t) \) for deterministic functions \( \pi, \pi^j : [0, T] \times \mathbb{R}^d \to \mathbb{R} \).
We assume them to admit all the derivatives we need.

Perfect hedge

For a perfect hedge, one needs a money market account and \( d \) liquid securities. From Itô’s lemma one obtains that the CoCo can be fully hedged with a dynamic trading strategy \( \vartheta^j_t, j = 0, \ldots, d \), satisfying
\[
\frac{\partial \pi}{\partial x^i}(t, X_t) = \sum_{j=1}^d \vartheta^j_t \frac{\partial \pi^j}{\partial x^i}(t, X_t), \quad i = 1, \ldots, d, \quad \text{and} \quad \pi(t, X_t) = \pi^0(t, X_t) + \sum_{j=1}^d \vartheta^j_t \pi^j(t, X_t)
\]
at all times \( t < \tau \land T \).

In the case where conversion is not caused by the stock price, the CoCo can be hedged with stock shares, CDS’s and interest rate swaps. To compute a realistic hedging strategy we assume the following:

a) \( \pi^1 \) is the stock \( S \)

b) \( \pi^2, \ldots, \pi^{m+1} \) are \( m \) CDS’s, \( H_t \) only depends on \( X_t^2, \ldots, X_t^{m+1} \), and the coefficients of the SDEs for \( X_t^2, \ldots, X_t^{m+1} \) do not depend on \( X_t^1 \)

c) \( \pi^{m+2}, \ldots, \pi^d \) are \( d-m-1 \) interest rate swaps, \( r_t \) only depends on \( X_t^{m+2}, \ldots, X_t^d \), the coefficients of the SDEs for \( X_t^{m+2}, \ldots, X_t^d \) do not depend on \( X_t^1 \), and (\( X_t^2, \ldots, X_t^{m+1} \)) is independent of (\( X_t^{m+2}, \ldots, X_t^d \)).
Then for $t < \tau \wedge T$ the system of equations (4.3.6) can be written as

$$
\frac{\partial \pi}{\partial x^1}(t, X_t) = \vartheta_1^1(t, X_t)
$$
$$
\frac{\partial \pi}{\partial x^i}(t, X_t) = \sum_{j=2}^{m+1} \vartheta^i_j \frac{\partial \pi^j}{\partial x^i}(t, X_t) \quad \text{for } i = 2, \ldots, m+1
$$
$$
\frac{\partial \pi}{\partial x^i}(t, X_t) = \sum_{j=2}^{d} \vartheta^i_j \frac{\partial \pi^j}{\partial x^i}(t, X_t) \quad \text{for } i = m+2, \ldots, d
$$
$$
\pi(t, X_t) = \vartheta^0_t \pi^0_t + \sum_{j=1}^{d} \vartheta^j_t \pi^j(t, X_t),
$$

where

$$
\frac{\partial \pi}{\partial x^1}(t, X_t) = G\mathbb{E}_t^{Q^r}[e^{-q(\tau-t)}1_{(\tau \leq T)}] \quad \text{or} \quad \frac{\partial \pi}{\partial x^1}(t, X_t) = 0,
$$

for a CoCo converting into equity or cash, respectively. For fixed $t < \tau \wedge T$, this is a system of $d + 1$ linear equations in $\vartheta^0_t, \ldots, \vartheta^d_t$. So if it is regular, it has a unique solution.

**Hedging without specifying the interest rate**

As in an intensity-based model, one can hedge the CoCo without specifying the short rate ($r_t$). Let us assume a) and b) from above. In addition, we suppose that the interest rate ($r_t$) is a Markov process independent of ($H_t$). Again, we only immunize against parallel shifts of the risk-free yield curve. Then one holds for every $t < \tau \wedge T$,

$$
\vartheta^1_t = G\mathbb{E}_t^{Q^r}[e^{-q(\tau-t)}1_{(\tau \leq T)}] \quad \text{(or } \vartheta^1_t = 0 \text{ if the CoCo converts into cash)}
$$

stock shares and invests in the $m$ CDS’s such that

$$
\frac{\partial \pi}{\partial x^i}(t, X_t) = \sum_{j=2}^{m+1} \vartheta^i_j \frac{\partial \pi^j}{\partial x^i}(t, X_t) \quad \text{for } i = 2, \ldots, m+1.
$$

Now parallel shifts of the yield curve can be neutralized by investing in a fixed income portfolio with AFW $Y_t - \sum_{j=2}^{m+1} \vartheta^j_t Y^j_t$, where $Y_t$ and $Y^j_t$ are given by (4.2.8)-(4.2.10).
The latter are easily calculated from the quantities in (4.1.13) together with

$$\mathbb{E}_{t,x}^{Q} \left[ P(t, \tau) \tau^2 1_{\{\tau \leq t_1\}} \right] \quad \text{and} \quad \mathbb{E}_{t,x}^{Q} \left[ P(t, \theta) \theta^2 1_{\{\theta \leq s_1\}} \right],$$

which can be obtained by solving the same PDE as in Proposition 4.3.3 with adjusted boundary conditions.
Chapter 5

Case Study: Rabobank and Lloyds Banking Group CoCos

5.1 Rabobank and Lloyds Banking Group CoCos

In this section we apply an intensity-based and a first-passage time model to price CoCos issued by Lloyds Banking Group and Rabobank. In both models, we use a small number of stochastic factors. This already seems to yield realistic results. But the models can easily be extended by adding more factors. The short rate does not have to be specified since it enters the pricing formulas only through the zero-coupon bond prices $P(t, s)$, which for finitely many $s$-values can be deduced from government bonds or interest rate swaps. From there the curve can be interpolated.

Like Corcuera et al. (2012) we choose Oct 14, 2011 as the pricing date and use the same data on risk-free yields and CDS spreads for calibration. The specifics of the two CoCos we consider are as follows:

- The Lloyds Banking Group’s Enhanced Capital Notes (ISIN XS0459089255) were issued on Dec 1, 2009 with a maturity of 10 years. They pay semi-annual coupons at an annual rate of 15%. If Lloyds Banking Group’s Core Tier 1 Ratio
falls below 5%, each ECN converts into $F/K$ ordinary equity shares, where $F$ is the principal amount of the note and $K$ equals 59 pence.

In 2011, Lloyds Banking Group reported a Core Tier 1 Ratio of 9.6%, while Basel III regulation requires it to be at least 4.5%. On the pricing date the ECNs had a time to maturity of 8.19 years and traded at 109.9% of their principal amount. The market price of Lloyds Banking Group’s stock was 33.25 pence.

- The Rabobank’s Senior Contingent Notes (ISIN XS0496281618) were issued on March 19, 2010 with a maturity of 10 years. They pay yearly coupons at a rate of 6.875%. They are triggered if Rabobank’s Equity Capital Ratio (equity capital/risk-weighted assets) falls below 7%. If triggered, they are written down by 75%, and each note converts into an immediate cash payment of 25% of the principal amount $F$.

Rabobank reported an Equity Capital Ratio of 14.7% in 2011, and Basel III requires a minimum of 4.5%. On the pricing date, the time to maturity of the SCNs was 8.44 years, and they traded at 88.84% of their principal amount.

The Table VIII shows risk-free yields and CDS spreads prevailing on Oct 14, 2011. Lloyds Banking Group’s stock and the ECNs are denominated in GBP. The risk-free yields were extracted from GBP interest rate swap data. The SCNs are denominated in EUR. The risk-free yields correspond to EUR interest rate swaps.

## 5.2 Intensity-based model with a CIR process

We consider an intensity-based model with a mean-reverting jump intensity. To ensure that it does not become negative, we model it as a Cox–Ingersoll–Ross process. It is then possible to deduce closed form expressions for the functions $\phi_{t,x}$ and $\phi_{t,x}^*$ introduced in Proposition 4.2.2. More precisely, we assume that the stock price follows
the dynamics

\[
\frac{dS_t}{S_t} = (r_t - q + [\alpha(1 + \gamma) - \gamma]\lambda_t) dt + \sigma_1 dW_t^1 + \sigma_2 \sqrt{\lambda_t} dW_t^2 + [(1 + \gamma)\xi - 1]dN_{\Lambda_t},
\]

and the jump intensity is a CIR process:

\[
d\lambda_t = \kappa(\mu - \lambda_t)dt + \eta \sqrt{\lambda_t} dW^2.
\] (5.2.1)

Then \(Q^*\) is given by

\[
\frac{dQ^*}{dQ} = \begin{cases} \exp \left( \sigma_1 W^1_T - \sigma_1^2 T/2 + \sigma_2 \int_0^T \sqrt{\lambda_s} dW_s^2 + [\alpha(1 + \gamma) - \gamma - 2\eta^2/2]\Lambda_T \right) \\ 0 \end{cases} \quad \text{if } \tau > T,
\]

\[
(1 + \gamma) \exp \left( \sigma_1 W^1_T - \sigma_1^2 T/2 + \sigma_2 \int_0^\tau \sqrt{\lambda_s} dW_s^2 + [\alpha(1 + \gamma) - \gamma - 2\eta^2/2]\Lambda_\tau \right) \quad \text{if } \tau \leq T, \xi = 1.
\]

It follows from Girsanov’s theorem that the dynamics of \((\lambda_t)\) can be written as

\[
d\lambda_t = [\kappa\mu + (\eta\sigma_2 - \kappa)\lambda_t]dt + \eta \sqrt{\lambda_t} dW^*_t,
\]

where \(W^*_t = W^2_t - \sigma_2 \int_0^t \sqrt{\lambda_s} ds\) is a Brownian Motion under \(Q^*\). To calculate the expressions of Proposition 4.2.1 with the methods of Proposition 4.2.2 we have to compute the Laplace transforms

\[
\phi_{t,x}(s, z) := \mathbb{E}^Q_{t,\lambda} \left[ e^{-z(\Lambda_s - \Lambda_t)} \right] \quad \text{and} \quad \phi^*_{t,x}(s, z) := \mathbb{E}^{Q^*}_{t,\lambda} \left[ e^{-z(\Lambda_s - \Lambda_t)} \right].
\]

It is well-known from Cox et al. (1985) that for the specification (5.2.1), one has

\[
\phi_{t,x}(s, z) = e^{A(s-t, z) + \lambda_t B(s-t, z)},
\]
where

\[ A(s, z) = \frac{\kappa \mu}{\eta^2} \left( 2 \ln \left( 1 - \frac{c(z)}{2c(z)} \left( 1 - e^{-c(z)s} \right) \right) + (c(z) - \kappa)s \right), \]

\[ B(s, z) = \frac{-2z(1 - e^{-c(z)s})}{2c(z) - (c(z) - \kappa)(1 - e^{-c(z)s})}, \]

\[ c(z) = \sqrt{\kappa^2 + 2\eta^2z}. \]

It follows that

\[
\frac{\partial \phi_{t,x}}{\partial s}(s, z) = \phi_{t,x}(s, z) \left( \frac{\partial A}{\partial s}(s - t, z) + \lambda_t \frac{\partial B}{\partial s}(s - t, z) \right) \\
\frac{\partial \phi_{t,x}}{\partial z}(s, z) = \phi_{t,x}(s, z) \left( \frac{\partial A}{\partial z}(s - t, z) + \lambda_t \frac{\partial B}{\partial z}(s - t, z) \right) \\
\frac{\partial^2 \phi_{t,x}}{\partial s \partial z}(s, z) = \phi_{t,x}(s, x) \left( \frac{\partial^2 A}{\partial s \partial z}(s - t, z) + \lambda_t \frac{\partial^2 B}{\partial s \partial z}(s - t, z) \right) + \frac{\partial \phi_{t,x}}{\partial s}(s, z) \left( \frac{\partial A}{\partial z}(s - t, z) + \lambda_t \frac{\partial B}{\partial z}(s - t, z) \right). 
\]

Analogous expressions can be derived for \( \phi_{t,x}^*(s, z) \) and \( \frac{\partial \phi_{t,x}^*}{\partial s}(s, z) \). For a given zero-coupon curve \( P(t, \cdot) \), the integrals of Proposition 4.2.1 can be computed numerically to obtain approximations to all the quantities of (4.1.13). For a given coefficient of jump in jump intensity \( \beta \), we chose the parameters \( \kappa, \mu, \eta \) and the initial value \( \lambda_t \) of the jump intensity \( (\lambda_a) \) so as to produce CDS spreads consistent with market quotes. However, the error function constructed only with CDS spreads is almost flat in some regions, causing problems for gradient based minimization algorithms. Therefore, we also inferred the Q-survival probabilities from the CDS spread data using the bootstrapping method of O’Kane and Turnbull (2003) and added them to the objective function. That is, we chose the parameters which minimized the criterion

\[
\sum_{i=2,3,4,5,7,10} \left( \frac{\delta_{t}^{\text{model}} - \delta_{t}^{\text{market}}}{\delta_{t}^{\text{market}}} \right)^2 + \left( \frac{q_{t}^{\text{model}} - q_{t}^{\text{OKT}}}{q_{t}^{\text{OKT}}} \right)^2, \quad (5.2.2)
\]

92
for the CDS spreads $\delta_i^{\text{model}}$, $\delta_i^{\text{market}}$ implied by the model and market data and the Q-survival probabilities $q_i^{\text{model}}$, $q_i^{\text{O'K}}$ produced by our model and the one of O’Kane and Turnbull (2003).

Panel A of Table IX gives the resulting parameters corresponding to the CoCo-implied jump in jump intensity $\beta$ (0.81 for Lloyds in the case $\gamma = -20\%$ and 1.36 for Rabobank). RMSE denotes the root mean squared error of the calibration.

The two panels at the top of Figure 6 show CoCo model prices as a function of the coefficient of jump in jump intensity $\beta$ chosen in the CDS calibration. For gamma we picked values of $0\%$, $+20\%$, $-20\%$, $-50\%$ and $-99\%$. The two panels in the middle of Figure 6 show the Q-survival probabilities produced by our intensity-based model and the ones extracted with the method of O’Kane and Turnbull (2003). They turned out to be virtually the same. The last two panels of Figure 1 show CDS spreads generated by our model compared to the market quotes.

Figure 7 shows the value decompositions of the ECNs and the SCNs into the parts stemming from the coupon payments, the redemption of the principal and a possible conversion. We let $\beta$ range from 0.5 to 2.0 and set the jump fraction $\gamma$ of the stock price of Lloyds Banking Group equal to $-20\%$. It can be seen that the prices of both CoCos are increasing in $\beta$, and the main contribution to their total values is coming from coupon payments and a possible redemption of the principal.

The conversion value of the ECNs depends on the product $\eta \sigma_2$, which cannot be inferred from CDS spreads. To obtain a negative correlation between the increments of the stock price and the conversion intensity, we chose $\sigma_2$ equal to $-30\%$. However, since in our example the calibrated value of $\eta$ is close to zero, the choice of $\sigma_2$ has practically no influence on the model price of the ECNs.
5.3 First-passage time model with an exponential OU accounting ratio

Now we model the accounting ratio as a diffusion process. The simplest model would be a (exponential) Brownian motion. But we obtained a better fit to market quotes of CDS spreads with an exponential Ornstein–Uhlenbeck process. So we here assume that the stock price follows a geometric Brownian motion

\[
dS_t = (r_t - q)S_t dt + \sigma S_t(\sqrt{1 - \rho^2}dW_1^t + \rho dW_2^t)
\]

and the logarithm of the accounting ratio an Ornstein–Uhlenbeck process

\[
dH_t = \kappa(\mu - H_t)dt + \eta dW_t^2.
\]

Then the measure \(Q^*\) is given by

\[
\frac{dQ^*}{dQ} = \frac{\tilde{S}_T}{S_0} = \exp\left(\sigma \left(\sqrt{1 - \rho^2}W_1^T + \rho W_2^T\right) - \frac{1}{2} \sigma^2 T\right),
\]

and it follows from Girsanov’s theorem that the dynamics of \((H_t)_{t \geq 0}\) can be written as

\[
dH_t = \kappa(\mu^* - H_t)dt + \eta dW_t^*,
\]

where \(\mu^* = \mu + \eta \sigma \rho / \kappa\) and \(W_t^* = W_t^2 - \sigma \rho t\) is a Brownian Motion under \(Q^*\). The problem of pricing and calibrating CoCos then reduces to solving linear PDEs with Dirichlet boundary conditions as described in Section 4.3.1. We solved the PDEs numerically with the Crank–Nicholson method\(^1\). In the case of Lloyds Banking Group, \(H_t\) is meant to model the logarithm of the Core Tier 1 Capital Ratio and

\(^1\) Göing-Jaeschke and Yor (2002) and Alili et al. (2005) derived formulas for hitting time distributions of Ornstein–Uhlenbeck processes. But the expressions given in these papers are so complicated that it was easier for us to solve the PDEs numerically.

94
in Rabobank’s case the logarithm of the equity capital ratio. \( \exp(h^*) \) is the contractual trigger level (5% for Lloyds and 7% for Rabobank) and \( \exp(h_s) \) the minimum capital ratio level required by Basel III (4.5% in both cases).

For a given value of the recovery rate \( R \) we calibrated the model by choosing the parameters \( \kappa, \mu, \eta \) and the starting point \( H_t \) of the process \( (H_s) \) so that the CDS spreads produced by the model were close to the market quotes. As in Subsection 5.2 above, we added the Q-survival probabilities extracted from CDS spreads with the method of O’Kane and Turnbull (2003) to the objective function and minimized the squared error function (5.2.2).

It is well-known that first-passage time models do not fit short term CDS spreads well; see for instance, Lando (2004). Therefore, we dropped the one year CDS spread from the calibration. The following table lists the optimal parameters corresponding to the CoCo-implied recovery rates, that is, the recovery rates \( R \) which made the model prices of the CoCos equal to their market prices (76.6% for Lloyds and 42% for Rabobank). Panel B of Table IX displays the calibrated parameters and RMSE for this model.

The two top panels of Figure 8 show CoCo prices produced by the model for different values of the recovery rate \( R \) chosen in the CDS calibration. The jaggedness of the curves is caused by numerical instabilities of the calibration. The two panels in the middle compare the Q-survival probabilities of Lloyds Banking Group and Rabobank implied by our model to the ones constructed with the O’Kane and Turnbull (2003) approach. It can be seen that they are almost indistinguishable. The last two panels of Figure 8 plot the CDS spreads generated by our model against the market quotes. The fit is not as good as for the intensity-based model of Subsection 5.2. Especially for Lloyds Banking Group, the 7- and 10-year CDS spreads of the model are too low. This is in line with, for instance, Eom et al. (2004) or Huang and Zhou (2012), which
find that standard first-passage time models do not fit CDS spread term structures well. In the case of Rabobank the fit is better, which is reflected in a smaller RSME.

In the case of Lloyds Banking Group, the CoCo price depends on $\sigma \rho$, which cannot be deduced from CDS data. We set it equal to 30% to have a positive correlation between the increments of the stock price and the accounting ratio. The calibration yielded a long term mean for the accounting ratio of 8.28% and an implied accounting ratio at the pricing date of 9.47%, slightly lower than the reported 9.6%. The CoCo-implied recovery rate $R$ is 76.0%, much higher than the usually assumed 40%. So according to the model, on Oct 14, 2011 the ECNs traded at a low price relative to the stock of Lloyds, interest rate swaps and CDS’s, or market participants were expecting a higher recovery rate than usual. In any case, it turned out that the price of the ECNs increased to around 130% of the principal amount during the first half of 2012.

For Rabobank, the estimated long term mean of the accounting ratio came out as 14.82%, and the implied capital ratio at the pricing date as 14.54%, a bit below the reported 14.7%. The CoCo-implied recovery rate for Rabobank was 53.0%. So under standard assumptions, the model would have priced the SCNs very close to their market value on Oct 14, 2011.

Figure 9 shows the value decompositions of the ECNs and the SCNs into the parts corresponding to the coupon payments, the principal redemption and a possible conversion. The recovery rate chosen in the calibration ranges from 10% to 90%. It can be seen that for both CoCos future coupon payments and a possible redemption of the principal account for most of the total value. For increasing values of the recovery rate, the model-implied $Q$-survival probabilities decrease. So a conversion becomes more likely, and correspondingly, the value of future coupons and the principal goes down while the conversion value increases.
5.4 Conclusion

We develop a framework for the pricing and hedging of CoCos. It introduces a general model that can price CoCos together with related products such as fixed income instruments, equity shares and CDS’s. We focused on two different specifications: intensity-based and first-passage time models. In the former, conversion comes as a surprise, and as a consequence, the prices of CoCos, the issuing firm’s stock and CDS’s jump at conversion. In the latter, the conversion time is predictable, and all prices are continuous at conversion. Both approaches can be taken to calculate CoCo prices and dynamic hedging strategies. As case studies, we calibrated an intensity-based and a first-passage time model to market quotes of equity, interest rate swaps and CDS’s to price CoCos issued by Lloyds Banking Group on Dec 1, 2009 and Rabobank on March 19, 2010. Consistently with the credit risk literature, we found that it was easier to reproduce the term structure of CDS spreads with an intensity-based model, and the calibration of a first-passage time model posed some numerical challenges.
Chapter 6

CoCo Design and Financial Stability

In this chapter, we tackle the problem of CoCo design. We analyze in-depth the key features making CoCos attractive for controlling systemic risk. Our approach is related to Raviv (2004), Abul et al. (2010), Koziol and Lawrenz (2010) and Chen et al. (2013); the main novelty is that we primarily focus on the asset substitution effect by solving the volatility control problem faced by the manger. To be more precise, the asset volatility before and after the conversion can be different as CoCos change the manager’s risk taking incentive. We derive explicit conditions in the framework of Leland’s model so that the presence of CoCos will induce the manager to be risk-averse while he is risk-taking if the bank is financed only by equity and straight debt. We recommend the CoCos satisfying these conditions referred to as well-designed CoCos. Our findings are in line with Berg and Kaserer (2012) and our well-designed CoCos are similar in spirit to their ’Convert-to-Surrender’ CoCo. Unfortunately, as pointed out by Berg and Kaserer (2012), the existing issuances do not qualify for well-designed CoCos.

We further illustrate how well-designed CoCos can
1. increase the total value of the bank even if CoCo coupons are not tax-deductible,

2. decrease both the risk-neutral and real probability of default,

3. and relieve the regulator from the concern of setting too high a capital ratio.

6.1 Bank with CoCos and Regulator

6.1.1 Unlevered Asset Value

The basic setting in this chapter is a generalization of Leland’s model. To adapt the model to study the capital structure of a systematically important financial institution (SIFI), we introduce two parties to the model: manager of the bank and the regulator to form a principal agent problem. The bank’s unlevered pre-tax asset process \( V_t \) is given as a diffusive process whose volatility is controlled by the manager. We assume the tax rate to be constant \( \theta \), and the presence of tax provides incentive to issue debts in order to benefit from their tax shield effect. We define \( V_t \) on probability space \((\Omega, (\mathcal{F}_t)_{t \geq 0}, Q)\), it satisfies

\[
\frac{dV_t}{V_t} = \mu dt + \sigma_t dW_t,
\]

with \((W_t)_{t \geq 0}\) a standard Brownian Motion defined on the filtration \((\mathcal{F}_t)_{t \geq 0}\) under \(Q\). At each moment \(t\), the manager can choose the instantaneous volatility \(\sigma_t\) within the interval \([\sigma_l, \sigma_h]\).

The bank pays out cash at the rate of \(\delta V_t\) to satisfy the coupon payments and remunerate the equity-holders (by dividend). As \(V_t\) models the pre-tax asset value, the after-tax cash flows generated by the bank is \((1 - \theta)\delta V_t\). Depending on different choices for the bankruptcy, we may or may not allow the dividend rate to be negative. A negative dividend rate corresponds to the case where equity holders are willing to
provide new capital to continue running the bank. We assume all the agents are risk-neutral under $\mathbb{Q}$, thus $\mu = r - \delta$.

As our model is time-homogeneous and the unlevered asset value process is Markovian, all the value functions will only depend on a single state variable $v$ corresponding to the asset level and other model parameters.

6.1.2 Liabilities

For simplicity, we assume all the liabilities to be with infinite maturity. In this case, equity holders need to pay a constant coupon rate to other claimholders. As shown in Chen et al. (2013), this situation can also be viewed as a stationary state where the bank renews a constant proportion of its debts with maturities following an exponential distribution.

A. Deposits Before bankruptcy, the equity holders pay a constant coupon rate $d$ to the depositors. The payment is tax-deductible. The deposit is guaranteed by the regulator and the depositors have to pay a premium for this government guarantee. We assume the premium rate to be constant $\varphi$, then the value function for the depositors is simply:

$$D(v) = \mathbb{E}_{V_0=v}^\mathbb{Q} \left[ \int_0^\infty e^{-rt} (d - \varphi) dt \right] = \frac{d - \varphi}{r} \quad (6.1.1)$$

In case of bankruptcy, the regulator will have to pay for the difference between the deposit value $\frac{d - \varphi}{r}$ and the residual asset value after deducting the bankruptcy cost if it is positive. The choice of $\varphi$ only affects the risk-sharing between the regulator and the depositors, and its value does not have any impact on the total value of the bank. We assume the constant $\varphi$ to be exogenous as a result of negotiation between depositors and the regulator.
B. Debts The bank pays out a constant coupon rate at $b$ for its debt-holders\textsuperscript{1}. The coupon payments are tax-deductible. Unlike the deposits, the debts are not guaranteed by the regulator. However, the regulator is sometimes forced to rescue the bank if its bankruptcy has a destructive effect to the entire financial system. We model this phenomenon by introducing bail-out solution such that with probability $p$, the regulator would pay for the difference between the debt face value $b$ and the residual asset value after deducing the bankruptcy cost along with deposit payments. $p$ can be interpreted as the the risk-neutral probability that the bankruptcy of a bank causes failure of the financial system. Therefore, $p$ can directly serve as a measure of how important a bank is in the financial system.

C. CoCos In addition to the traditional debt financing instruments, the bank also issues CoCos. The CoCos will be converted into equity upon a contractual trigger event, and upon conversion, the CoCo holders will receive a total of $n$ times the current outstanding share of the bank. The ratio $n$ will be referred to as conversion ratio. For example, if the bank has 1 million outstanding shares before the conversion, it will have $1+n$ million after the conversion, where 1 million (resp. $n$ million) shares are held by the old equity holders (resp. CoCo holders). Therefore, after the conversion, due to dilution effect, the old shareholders’ value (resp. CoCo holders’ value) will become $\frac{1}{1+n} E_{t_+}$ (resp. $\frac{n}{1+n} E_{t_+}$) where $E_{t_+}$ is the bank’s total equity value immediately after the conversion takes place. Another concept related to conversion ratio is conversion price, which is simply the face value of CoCo divided by the total number of shares received upon conversion.

\textsuperscript{1}Our model can be easily extended to incorporate different seniorities of debt issuance, but we believe adding subordinated debts will simply change the risk-sharing among different debt-holders, and our main results remain valid.
We consider here a trigger event based on the unlevered asset value \( V_t^2 \) - the conversion takes place as soon as \( V_t \) drops the first time below \( V_c \). We use \( \tau_c \) to denote this conversion time.

Similar to the debts, the equity holders have to pay a constant coupon rate at \( c \) to the CoCo holders before the conversion. CoCos can be either tax-deductible or not: in Switzerland where the CoCo is made mandatory, the taxation law has accordingly changed so that CoCo coupon payments are tax-deductible, while Sundaresan and Wang (2011) points out CoCo coupons are not eligible to tax deduction according to the US law in vigor. Hence, the tax-deductibility is left as an open choice for the regulator.

### 6.1.3 Bankruptcy

We consider a situation where the default is exogenous - similar to Koziol and Lawrenz (2011), we specify a default boundary \( V_b \). The bank would experience financial distress as soon as its asset falls below this level. We intentionally avoid an endogenous default chosen by the manager of the bank: in a scenario of credit crunch, bank’s bankruptcy decision belongs neither to the manager nor to the shareholders, but to the debt holders and the depositors. However, as pointed out in Leland’s original paper, the case of endogenous default corresponds to a particular choice of \( V_b \), providing an interesting benchmark \( V_{endo} \). The choice of \( V_b \) is constrained to be larger than \( V_{endo} \), since the equity holders can always voluntarily choose to file for bankruptcy.

To relate the default boundary to cash flows, we can define \( \xi := \frac{V_b}{b+d} \). Our default boundary can be interpreted as a form of bond convenant: when the cash-flow generated by the asset is lower than a proportion \( \xi \) of the coupon payments \( b + d \), the

---

2A trigger based on the total asset value can be assimilated as an accounting ratio trigger as shown in Pennachi (2010). Also Chen et al. (2013) adapted the same setting for the trigger event.
bank has to file for bankruptcy. We assume the default boundary can be chosen by
the regulator since it is closely related to the capital requirements. When \( \xi \) is large,
the situation will be referred to as \textit{bank with tight financial constraints}, as the default
occurs before the bank cannot afford coupon payments; in contrast, if \( \xi \) is small, the
situation will be referred to as \textit{bank with loose financial constraints}, as the bank can
continue to operate by allowing equity holders to receive a negative dividend rate for
a certain amount of time before bankruptcy.

We will use \( \tau_b \) to denote the first time the default boundary is hit. After the
bankruptcy, the bank liquidates all the assets subject to a bankruptcy cost \( \alpha \), so
the debt holders along with depositors receive \((1 - \alpha)V_b\). Depositors have higher
priority than the debt holders, and in the event of bankruptcy, we assume the default
boundary \( V_b \) is so low that equity holders cannot receive any residual value.

### 6.1.4 Manager

As our primary focus is CoCo and systemic risk, we do not model the manager’s
moral hazard problem. We assume the manager can control the riskiness of his
investments, and thus choose his desired volatility to maximize the equity holders’
value. In other words, the manager’s interest is perfectly in line with the equity
holders. Hence, the manager’s value is:

\[
E(v) = \begin{cases} 
0 & \text{for } v \leq V_b \\
\mathbb{E}_V^Q \left[ (1 - \theta) \int_0^{\tau_b} e^{-rt} (\delta V_t - (b + d)) dt \right] & \text{for } V_b < v \leq V_c \\
\mathbb{E}_V^Q \left[ (1 - \theta) \int_0^{\tau_c} e^{-rt} (\delta V_t - (b + c + d)) dt + e^{-r \tau_c} E(V_c) \right] & \text{for } v > V_c
\end{cases}
\]

(6.1.2)

if CoCo coupons are tax-deductible and for the last case \( v > V_c \),

\[
E(v) = \mathbb{E}_V^Q \left[ \int_0^{\tau_c} e^{-rt} ((1 - \theta)(\delta V_t - (b + d)) - c) dt + e^{-r \tau_c} E(V_c) \right]
\]

(6.1.3)
if CoCo coupons are not tax-deductible. With our benchmark CoCo, the manager’s only control is the volatility process. We will defer the discussion about the alternative design where the manager decides when to convert the CoCos.

6.1.5 Regulator

The role of regulator is crucial to the discussion of systemic risk. In our model, the regulator receives a premium $\varphi$ from the depositors and has to fully guarantee their deposits; upon default, he also has to guarantee the debts with probability $p$. Finally, he receive tax payments throughout the lifetime of the bank. His value can be computed as:

$$R(v) = \mathbb{E}_V^Q \left[ \int_0^{\tau_b} \varphi e^{-rt} dt - e^{-r\tau_b} \left( \frac{d - \varphi}{r} - (1 - \alpha) V_b \right) + \right]$$

$$- p \cdot \mathbb{E}_V^Q \left[ e^{-r\tau_b} \left( \frac{b}{r} - ((1 - \alpha) V_b - \frac{d - \varphi}{r})_+ \right) + \right]$$

$$+ \theta \cdot \mathbb{E}_V^Q \left[ \int_0^{\tau_c} e^{-rt} (\delta V_t - (b + d)) dt \right]$$

$$- \theta \cdot \mathbb{E}_V^Q \left[ \int_0^{\tau_c} e^{-rt} cd t \right]$$

(6.1.4)

where the last term would be zero if the CoCos are not tax-deductible. The deposit guarantee is modeled as a put option written by the regulator on the asset level upon default with a strike equal to $\frac{d - \varphi}{r}$, the default-free discounted value of the coupon stream. The bail-out is modeled as a put option written by the regulator which will be exercised with probability $p$ upon default (the exercise is independent from $V_t$). In other words, the regulator will have to guarantee the total face value of debts.
and deposits \((b+d-\varphi^r)\) upon default with probability \(p\), and just guarantee the deposit \((d-\varphi^r)\) with probability \(1-p\). The regulator can choose:

1. Whether the CoCo coupons are tax deductible or not.
2. The default boundary \(V_b\) or \(\xi\) equivalently.

Although the regulator cannot directly raise the default boundary, it is possible to make stringent capital requirement to modulate the default boundary. However, capital requirement is not perfectly equivalent to default boundary: theoretically, a bank holding inadequate capital should still be able to raise new capital or sell off assets to meet the capital requirement. Nonetheless, a bank faces capital constraints only because of losses, and new capital is proven to be extremely costly under such circumstances. The investors would be concerned about the debt overhang problem and reluctant to provide capital. It is straightforward to illustrate the intuition behind debt overhang in Merton’s framework: when the bank needs to raise new capital, the asset value is relatively low. An investor providing one additional dollar of capital will become an equity holder. The equity value is a call option on the asset value, and its delta is smaller than 1, implying equity value only increases by a fraction delta of the additional dollar. Hence, the investor only obtains delta dollar of equity by providing one dollar, and the rest of it results in an increase in debt holders' value. Due to the debt overhang effect, we believe the regulator can effectively change the default boundary by setting capital requirement.

6.2 Manager’s Problem

The manager is assumed to have the same interest perfectly aligned with the equity holders, so he chooses the volatility process to maximize the equity value of the bank. In this section, we will derive the manager’s optimal strategy and analyze his risk incentive when the bank is financed with and without CoCos.
6.2.1 After conversion

After conversion, the bank is only financed by debts, deposits and equities. Equivalently, the bank behaves almost exactly as one without CoCos. However, due to the dilution caused by conversion, the equity holders own a fraction \( \frac{1}{1+n} \) of the bank as the former CoCo holders own the remainder. The equity value is equal to the expected future discounted cash flows equity holders will receive. Hence, the manager’s maximization problem is:

\[
E^{ac}(v) = \frac{1}{1 + n} \sup_{(\sigma_t)_{t \geq 0} \in [\sigma_l, \sigma_h]} \mathbb{E}_V^{Q}_{V_0=v} \left[ (1 - \theta) \int_0^{\tau_b} e^{-rt} (\delta V_t + \int_0^{\tau_b} (\mu - \frac{1}{2} \sigma_t^2) dt + \int_0^{\tau_b} \sigma_t dw_t) - (b + d) dt \right].
\]

We can write the Hamilton-Jacobi-Bellman (HJB) equation as:

\[
\mu v E^{ac'} + \sup_{\sigma_t \in [\sigma_l, \sigma_h]} \frac{1}{2} \sigma_t^2 v^2 E^{ac''} - r E^{ac} \leq \frac{1 - \theta}{1 + n} (b + d - \delta v) \quad (6.2.5)
\]

\[
E^{ac}(V_b) = 0
\]

The boundary condition corresponds to the bankruptcy event. The manager’s decision clearly depends on whether the function \( v \mapsto E^{ac}(v) \) is convex or concave: if convex, then manager will always choose the high volatility and vice versa. The equation can be solved by a standard guess and verification argument: assuming the manager would always choose a constant volatility \( \sigma \), using the standard results on Brownian Motion’s hitting time we recapitulate in Appendix D.1, the solution can be obtained as:

\[
E^{ac}(v; \sigma) = \frac{1}{1 + n} \mathbb{E}_V^{Q}_{V_0=v, \sigma_t=\sigma} \left[ (1 - \theta) \int_0^{\tau_b} e^{-rt} (\delta V_t - (b + d)) dt \right]
\]

\[
= \begin{cases} 
\frac{1 - \theta}{1 + n} \left( v - \left( \frac{v}{V_b} \right)^{-\gamma(\sigma)} V_b \right) - \frac{b+d}{r} \left( 1 - \left( \frac{v}{V_b} \right)^{-\gamma(\sigma)} \right) & \text{for } v > V_b \\
0 & \text{for } v \leq V_b 
\end{cases}
\]

106
with \( m(\sigma) = \mu - \frac{1}{2}\sigma^2 \) and \( \gamma(\sigma) = \frac{m(\sigma) + \sqrt{m(\sigma)^2 + 2\sigma^2}}{\sigma^2} \).

Focusing only on the region \( v \neq V_b \) and by introducing the function \( \mathcal{V}(x, y) := x - \frac{y}{r} \) and \( \mathcal{D}_\gamma(x, y) := \left( \frac{x}{y} \right)^{-\gamma} \), \( E(v; \sigma) \) can be rewritten as:

\[
E^{ac}(v; \sigma) = \frac{1 - \theta}{1 + n} \left[ \mathcal{V}(v, b + d) - \mathcal{V}(V_b, b + d) \mathcal{D}_\gamma(v, V_b) \right] \tag{6.2.6}
\]

To verify the solution, we only need to check the convexity of the value function \( v \mapsto E(v; \sigma) \). As matter of fact, it is either strictly concave when \( \mathcal{V}(V_b, b + d) > 0 \) or strictly convex when \( \mathcal{V}(V_b, b + d) < 0 \). Therefore, the manager’s optimal strategy is always to choose a constant volatility, depending on the bank’s financial constraint:

- if \( \xi > \frac{\delta}{r} \), the financial constraint is tight, the manager will choose low volatility \( \sigma_l \) after the conversion; and the corresponding value function is \( E^{ac}(v) = E^{ac}(v; \sigma_l) \)
- if \( \xi < \frac{\delta}{r} \), the financial constraint is loose, the manager will choose high volatility \( \sigma_h \) after the conversion; and the corresponding value function is \( E^{ac}(v) = E^{ac}(v; \sigma_h) \)

The manager’s risk preference after the conversion is thus exclusively determined by the default boundary: facing tight financial constraints, the threat of losing residual value at bankruptcy induces the manager to be risk-averse, while facing loose financial constraints he would be risk-seeking. Finally, as in Leland (1994), the endogenous default boundary \( V_{endo} \) can be obtain by imposing the differentiability of \( E^{ac}(V) \) at \( V_b \) (or equivalently, the \( V_b \) maximizing the equity value), yielding:

\[
V_{endo} = \frac{\gamma(\sigma^{ac})}{1 + \gamma(\sigma^{ac})} \frac{b + d}{r}. \tag{6.2.7}
\]

This endogenous default boundary will lead to a loose financial constraint situation, where the manager is better off by choosing a higher volatility.
Before conversion: CoCos with tax-deductible coupons

Before the conversion, the equity value function equals to the expected discounted cash flows until conversion, plus equity holders’ discounted residual value when conversion takes place. In this section, we assume the CoCo coupons are tax-deductible. The case where CoCos are not tax-deductible is treated similarly in the next section. The manager try to maximize the equity value by choosing the volatility:

\[ E^{bc}(v) = \text{esssup}_{v \geq 0 \in [\sigma_l, \sigma_h]} \mathbb{E}^Q_{V_0=v} \left[ \int_0^{\tau_c} e^{-rs}(1 - \theta)(\delta V_s - (b + c + d)) ds + e^{-r\tau_c} E^{ac}(V_c) \right] \] (6.2.8)

Similarly, the HJB equation is:

\[ \mu v E^{bc'} + \sup_{\sigma_t \in [\sigma_l, \sigma_h]} \frac{1}{2} \sigma_t^2 v^2 E^{bc''} - r E^{bc} \leq (1 - \theta)(b + c + d - \delta v) \] (6.2.9)

\[ E^{bc}(V_c) = E^{ac}(V_c) \]

To solve the control problem, we proceed by guess and verification as for the after conversion value function. We first look for solutions where the manager chooses constant volatility \( \sigma_t = \sigma \), in this case the equity value can be obtained using the results in Appendix [D.1]

\[ E^{bc}(v; \sigma) = \mathbb{E}^Q_{V_0=v, \sigma_t=\sigma} \left[ (1 - \theta) \int_0^{\tau_c} e^{-rs}(\delta V_s - (b + c + d)) ds + e^{-r\tau_c} E^{ac}(V_c) \right] \]

\[ = (1 - \theta) \mathcal{V}(v, b + c + d) \]

\[ - [(1 - \theta) \mathcal{V}(V_c, b + c + d) - E^{ac}(V_c)] \mathcal{D}_\gamma(\sigma)(v, V_c) \] (6.2.10)

The value function is strictly concave if and only if \( E^{ac}(V_c) < (1 - \theta) \mathcal{V}(V_c, b + c + d) \). Hence, the manager’s value function before the conversion is:

\[ E^{bc}(v) = E^{bc}(v; \sigma^{bc}) \]
with $\sigma^{bc} = \sigma_l$ if $E^{ac}(V_c) < (1 - \theta)\mathcal{V}(V_c, b + c + d)$ and $\sigma^{bc} = \sigma_h$ otherwise. We can summarize the manager’s problem as follow:

**Proposition 6.2.1.** When CoCo coupon payments are tax-deductible, the equity value of a bank financed with CoCos, debts, deposits and equities can be written as:

$$E(v) = \begin{cases} 
(1 - \theta)\mathcal{V}(v, b + c + d) & \text{for } v > V_c \\
- [(1 - \theta)\mathcal{V}(V_c, b + c + d) - E(V_c)] \mathcal{D}_{\gamma(\sigma^{bc})}(v, V_c) & \text{for } V_c \geq v > V_b \\
\frac{1}{1+n} [\mathcal{V}(v, b + d) - \mathcal{V}(V_b, b + d)\mathcal{D}_{\gamma(\sigma^{ac})}(v, V_b)] & \text{for } V_c \geq v > V_b \\
0 & \text{for } v \leq V_b
\end{cases}$$

(6.2.11)

where

- $\sigma^{bc} = \sigma_l$ (resp. $\sigma_h$) when $\mathcal{V}(V_b, b + d) > \mathcal{D}_{\gamma(\sigma^{ac})}(V_b, V_c) \left( \frac{(1+n)c}{r} - n\mathcal{V}(V_c, b + d) \right)$ (resp. $\mathcal{V}(V_b, b + d) < \mathcal{D}_{\gamma(\sigma^{ac})}(V_b, V_c) \left( \frac{(1+n)c}{r} - n\mathcal{V}(V_c, b + d) \right)$);

- $\sigma^{ac} = \sigma_l$ (resp. $\sigma_h$) when $\mathcal{V}(V_b, b + d) > 0$ (resp. $\mathcal{V}(V_b, b + d) < 0$).

If the conversion ratio $n$ is chosen such that the conversion happens at par (i.e. $\frac{c}{r} = nE(V_c)$), then $\sigma^{bc} = \sigma^{ac}$.

**Before conversion: CoCos with non tax-deductible coupons**

If the CoCo coupons are not tax-deductible, the equity holders will have to pay $c$ instead of $(1 - \theta)c$. Hence, the results in the previous section hold true if we just replace $c$ by $c/(1 - \theta)$. We can summarize the manager’s volatility control problem as follow:
Proposition 6.2.2. When CoCo coupon payments are not tax-deductible, the equity value of a bank financed with CoCos, debts, deposits and equities can be written as:

\[
E(v) = \begin{cases} 
(1 - \theta)\mathcal{V}(v, b + d) - \frac{\xi}{r} & \text{for } v > V_c \\
- \left[ (1 - \theta)\mathcal{V}(V_c, b + d) - \frac{\xi}{r} - E(V_c) \right] D_{\gamma(\sigma_{bc})}(v, V_c) & \text{for } V_c \geq v > V_b \\
\frac{1 - \theta}{1 + n} \left[ \mathcal{V}(v, b + d) - \mathcal{V}(V_b, b + d) D_{\gamma(\sigma_{ac})}(v, V_b) \right] & \text{for } v \leq V_b 
\end{cases}
\]

(6.2.12)

where

- \( \sigma_{bc} = \sigma_l \) (resp. \( \sigma_h \)) when \( \mathcal{V}(V_b, b + d) > D_{\gamma(\sigma_{ac})}(V_b, V_c) \left( \frac{(1+n)c}{(1-\theta)r} - n\mathcal{V}(V_c, b + d) \right) \)
  (resp. \( \mathcal{V}(V_b, b + d) < D_{\gamma(\sigma_{ac})}(V_b, V_c) \left( \frac{(1+n)c}{(1-\theta)r} - n\mathcal{V}(V_c, b + d) \right) \));

- \( \sigma_{ac} = \sigma_l \) (resp. \( \sigma_h \)) when \( \mathcal{V}(V_b, b + d) > 0 \) (resp. \( \mathcal{V}(V_b, b + d) < 0 \)).

If the conversion ratio \( n \) is chosen such that the conversion happens at par (i.e. \( \frac{\xi}{r} = nE(V_c) \)), then \( \sigma_{bc} = \sigma_{ac} \).

### 6.2.2 Risk preference and default probability

An important question is whether the choice of a lower volatility necessarily implies a lower default probability. Under the risk-neutral measure, the answer is obvious: the lower the volatility is, the lower the hitting probability is as the expected rate of return does not depend on the choice of volatility. Unfortunately, the risk-neutral default probability does not mean much from the perspective of risk management. The historical default probability is a more meaningful metric. Under classical assumption of CAPM, the price of risk for the return on the bank’s asset depends on its covariance with the market portfolio, which is proportional to the volatility chosen by the manager. If we assume the price of risk to be a constant \( \lambda \) over time, the instantaneous risk premium needs to be \( \lambda \sigma_l \). Since the manager’s optimal choice for
volatility is always a constant (though it can be different before and after the conversion), the expected growth rate of the unlevered asset $V_t$ under historical measure is $r - \delta + \lambda \sigma$.

The default probability is the probability such that the bank’s asset level will eventually below $V_b$. Conditioning on the event the asset level falls below $V_c$, we obtain

$$P_{V_0=v}[\tau_b < \infty] = P_{V_0=v}[\tau_c < \infty] \cdot P_{V_0=v}[\tau_b < \infty] = P_{V_0=v}[\tau_c < \infty] \cdot P_{V_0=v}[\tau_b < \infty],$$

where $P$ denotes the historical probability measure and the second equality is result of the Markov property of $V_t$. The probability $P_{V_0=v}[\tau_c < \infty]$ equals to $\left(\frac{V_c}{v}\right)^{2(r - \delta + \lambda \sigma + 0.5(\sigma_{bc})^2)}$, and $P_{V_0=v}[\tau_b < \infty]$ equals to $\left(\frac{V_b}{V_c}\right)^{2(r - \delta + \lambda \sigma + 0.5(\sigma_{ac})^2)}$. As $\lambda > 0$ and $V_b \leq V_c \leq v$, both of these probabilities are decreasing in volatility, asserting a lower volatility will incur a lower probability of default. For the purpose of enhancing the stability of financial system, the CoCos should be designed to curb the risk-seeking behavior of the managers.

### 6.2.3 Manager’s risk appetite

The manager’s risk appetite can be reflected by his choice of volatility which a direct consequence from the solution to the control problem. Aversion towards risk can unsurprisingly translate into a lower volatility choice and in turn a lower probability of default. For a bank without CoCos, the manager’s attitude towards volatility is completely determined by the bank’s financial constraint: if the default boundary is high enough, under tight financial constraint the manager becomes risk-averse under the threat of losing residual value upon default. The manager’s choice is completely based on the residual value upon default $\mathcal{V}(V_b, b + d) = V_b - \frac{b+d}{r}$: the first term corresponds to the residual asset value and the second term correspond
to the residual liability value if the bank continues to pay coupons to depositors and
debt-holders till eternity. The critical point is $V_b = \frac{b + d}{r}$, denoted as $V_b^*$. To be precise,

- If $V_b > V_b^*$, the bank is under tight financial constraint, the manager prefers
  low volatility;

- if $V_b < V_b^*$, the bank is under loose financial constraint, the manager prefers
  high volatility;

- if $V_b = V_{endo} = \frac{\gamma(\sigma^ac)}{1+\gamma(\sigma^ac)} \frac{b + d}{r}$, the bank is under loose financial constrain and the
  manager prefers high volatility.

To ensure tight financial constraint, the only option is to impose a high capital require-
ment from the standpoint of a regulator. If the manager can choose to strategically
default to his discretion, then he would also optimally choose a high volatility.

For a bank financed with CoCo, the situation is more complex. We first deal with
the case where CoCo coupons are tax-deductible. At first glance, the condition

$$\mathcal{V}(V_b, b + d) > D_{\gamma(\sigma^ac)}(V_b, V_c) \left( \frac{(1 + n)c}{r} - n\mathcal{V}(V_c, b + d) \right)$$  \hspace{1cm} (6.2.13)

for manager to choose a low volatility may appear to be obscure. In addition to the
residual value $\mathcal{V}(V_b, b + d)$, manager’s attitude towards risk is jointly determined by
the size of CoCos (reflected by $c$), the conversion boundary $V_c$ as well as the conversion
ratio $n$.

Obviously, the sign of $\frac{(1 + n)c}{r} - n\mathcal{V}(V_c, b + d)$ is critical in determining whether
the manager is more likely to be risk-averse or risk-seeking compared to the case
without CoCos. If this term is negative, the condition for the manager to choose
lower volatility is relaxed compared to the case without CoCos, and the CoCo will be
considered to be well-designed; in contrast, if it is positive, then the manager is even
more likely to be risk-seeking all else being equal, the CoCo is ill-designed.
The CoCo is well-designed if $\frac{\xi}{r} \leq \frac{n}{1+n}(V_c - \frac{b+d}{r})$. The intuition behind this condition is that the CoCo coupon value cannot be too large relative to the residual value upon conversion. The r.h.s. reflects the fraction of residual value belonging to the CoCo holders upon conversion. This residual value can be thought of as the intrinsic value of the equity (CoCo holders become equity holders after conversion), which is a call option on the unlevered asset according to Merton (1976). If the CoCo coupons are expensive relative to the residual value, the manager would rather giving the CoCo holders equity than paying the coupons. This would then encourage the manager to be risk-seeking. On the other hand, a costly conversion transferring a large proportion of value to the CoCo holders would deter the manager from bearing excessive risks.

A well-designed CoCo needs to have the following features:

- High enough remaining asset value when conversion is triggered ($V_c - \frac{b+d}{r} > 0$): in other words, the residual value upon conversion needs to be positive, otherwise the r.h.s. can never exceed the l.h.s. which is positive. Furthermore, if $V_c > \frac{b+d+c}{r}$, the r.h.s. is an increasing function of the dilution ratio, this simply reflects bigger dilution threat would dissuade the manager from risk-seeking behavior.

- High dilution ratio ($\frac{n}{1+n}$): the r.h.s. is an increasing function of the dilution ratio, this simply reflects bigger dilution threat would dissuade the manager from risk-seeking behavior.

- Low coupon rate for CoCos ($c$): This implication suggests replacing a large proportion of debts by CoCos might not be a good design. If $c$ is large, the only solution to make sure the CoCos are well-designed is to increase $V_c$ and $n$. If $V_c$ is high, the conversion will always occur in a premature way ahead of time - new capital will be injected into the bank while it is unnecessary; if $n$ is high,
the bank will fully acquired by CoCo holders and the original equity holders will lose control of the bank. None of these designs appear to be reasonable.

If CoCo coupons are not tax-deductible, all the conditions remain unchanged except \( c \) needs to be replaced by \( \frac{c}{1-g} \). As discussed above, a well-designed CoCo needs to have small \( c \), but making the coupons non tax-deductible is equivalent to increase the value of \( c \). Consequently, all else being equal, a non tax-deductible CoCo is more likely to be ill-designed.

However, this criterion for well-designed CoCo can be misleading in that well-designed CoCos can still fail to induce the manager to take lower risks prior to conversion. To avoid such confusion, in the rest of this chapter, well-designed CoCo simply refers to CoCos such that the manager is threatened to be risk-averse prior to the conversion event.

The mathematics may first appear to be counter-intuitive: for a straight-debt financing bank, the shareholders’ position can be viewed to a call option on the total asset with debt nominal being its strike according to Merton’s model. Therefore the shareholders are better off when volatility is higher. A generalization to a bank financed by CoCo, equity and debt would be the following: as CoCo reduces the probability of default, shareholders desire an even higher volatility.

Nonetheless, such analysis neglects one of, if not the most important features of CoCo - the path dependency. Rather than being similar to the strike, the CoCo trigger level acts as the barrier for a down-and-out barrier call option, since the CoCos convert whenever the trigger level is attained. In this case, the equity holders’ position would be similar to holding a down-and-out barrier call option, and it is well known that in this case if the barrier level is sufficiently higher than the strike, equity holders are net short volatility. This leads to one of our key results: well-designed CoCos, instead of encouraging the manager to be more risk-taking, can effectively reduce his risk-taking incentives.
6.2.4 Value of debts, deposits, CoCos, and total value of the bank

Similar as the value of equity, we summarize the value of debts, deposits, as well as the total value of the bank given by the model as follow:

**Theorem 6.2.3.** For a bank financed by deposits, debts, CoCos and equity, assume its unlevered pre-tax asset level is $v > V_c$, if the CoCo coupons are tax-deductible, the equity value is given in (6.2.1) and the value of debts, deposits, CoCos and total value of the bank are:

- **Deposits:**
  $$D(v) = \frac{d - \varphi}{r}$$  \hspace{1cm} (6.2.14)

- **Debts:**
  $$B(v) = p \cdot D_{\gamma (\sigma_{ac})}(v, V_c) \cdot D_{\gamma (\sigma_{bc})}(V_c, V_b) \cdot \left( \frac{b}{r} - \left( \frac{(1 - \alpha) V_b - \frac{d - \varphi}{r} }{r} \right)_+ \right) +$$
  $$+ (1 - D_{\gamma (\sigma_{ac})}(v, V_c) \cdot D_{\gamma (\sigma_{bc})}(V_c, V_b)) \frac{b}{r}$$  \hspace{1cm} (6.2.15)

- **CoCos:**
  $$C(v) = (1 - D_{\gamma (\sigma_{bc})}(v, V_c)) \frac{c}{r} + nE(V_c)$$  \hspace{1cm} (6.2.16)

- **Total Value:**
  $$T(v) = \left\{ \begin{array}{l}
  (1 - \theta) \cdot v - \alpha (1 - \theta) D_{\gamma (\sigma_{ac})}(v, V_c) \cdot D_{\gamma (\sigma_{bc})}(V_c, V_b) \cdot V_b \\
  \text{After-tax asset} \\
  \text{Bankruptcy cost} \\
  + (1 - D_{\gamma (\sigma_{ac})}(v, V_c) \cdot D_{\gamma (\sigma_{bc})}(V_c, V_b)) \frac{\theta b + d}{r} \\
  \text{Tax shield from deposit and debt coupons} \\
  + (1 - D_{\gamma (\sigma_{bc})}(v, V_c)) \frac{\theta c}{r} \\
  \text{Tax shield from CoCo coupons} \\
  + p \cdot D_{\gamma (\sigma_{ac})}(v, V_c) \cdot D_{\gamma (\sigma_{bc})}(V_c, V_b) \cdot \left( \frac{b}{r} - \left( \frac{(1 - \alpha) V_b - \frac{d - \varphi}{r} }{r} \right)_+ \right) \\
  \text{Value of bail-out} \\
  + D_{\gamma (\sigma_{ac})}(v, V_c) \cdot D_{\gamma (\sigma_{bc})}(V_c, V_b) \cdot \left( \frac{d - \varphi}{r} - (1 - \alpha) V_b \right)_+ \\
  \text{Guarantee for deposits} \\
  - (1 - D_{\gamma (\sigma_{ac})}(v, V_c) \cdot D_{\gamma (\sigma_{bc})}(V_c, V_b)) \frac{\varphi}{r} \\
  \text{Premium from deposits}
  \end{array} \right.$$  \hspace{1cm} (6.2.17)
If the CoCo coupons are not tax-deductible, the equity value is given in 6.2.2 and the value of debts, deposits, CoCos and total value of the bank are

\[ \text{Deposits: } D(v) = \frac{d - \varphi}{r} \tag{6.2.18} \]

\[ \text{Debts: } B(v) = p \cdot D_{\gamma(\sigma^{ac})}(v, V_c) \cdot D_{\gamma(\sigma^{bc})}(V_c, V_b) \cdot \left( \frac{b}{r} - ((1 - \alpha)V_b - \frac{d - \varphi}{r})_+ \right) + \]
\[ + (1 - D_{\gamma(\sigma^{ac})}(v, V_c) \cdot D_{\gamma(\sigma^{bc})}(V_c, V_b)) \frac{b}{r} \tag{6.2.19} \]

\[ \text{CoCos: } C(v) = \left( 1 - D_{\gamma(\sigma^{bc})}(v, V_c) \right) \frac{c - nE(V_c)}{r} \tag{6.2.20} \]

\[ \text{Total Value: } T(v) = \left( 1 - \theta \right) \cdot v - \alpha (1 - \theta) D_{\gamma(\sigma^{ac})}(v, V_c) \cdot D_{\gamma(\sigma^{bc})}(V_c, V_b) \cdot V_b \]
\[ + (1 - D_{\gamma(\sigma^{ac})}(v, V_c) \cdot D_{\gamma(\sigma^{bc})}(V_c, V_b)) \frac{\theta(b + d)}{r} \]
\[ + p \cdot D_{\gamma(\sigma^{ac})}(v, V_c) \cdot D_{\gamma(\sigma^{bc})}(V_c, V_b) \cdot \left( \frac{b}{r} - ((1 - \alpha)V_b - \frac{d - \varphi}{r})_+ \right) + \]
\[ + D_{\gamma(\sigma^{ac})}(v, V_c) \cdot D_{\gamma(\sigma^{bc})}(V_c, V_b) \cdot \left( \frac{d - \varphi}{r} - (1 - \alpha)V_b \right)_+ \]
\[ - \left( 1 - D_{\gamma(\sigma^{ac})}(v, V_c) \cdot D_{\gamma(\sigma^{bc})}(V_c, V_b) \right) \frac{\varphi}{r} \tag{6.2.21} \]

It has been shown in Koziol and Lawrenz (2011) that CoCo can increase the total value of the bank. However, others argue that the increase in total value is no more than the tax deductible from CoCo coupons, and we should not forget that banks’ saving from tax deduction is taxpayers’ cost. In our case, the deposits are guaranteed by regulator, and the bail-out, if necessary, will also be financed by the regulator. Thus the three last terms of the bank’s total value also fall into the same category - a gain from the bank’s perspective results in a loss from the tax-payers perspective. Hence, the components of interest in the total value are 1) bankruptcy cost 2) tax shield from deposit and debt coupons. If the manager is allowed to choose volatility,
a higher value of volatility \( \sigma \) leads to a lower value of \( \gamma(\sigma) \), in turn increases the terms \( D_\gamma(\sigma)(v, V_c) \) and \( D_\gamma(\sigma)(V_c, V_b) \). Therefore, both the bankruptcy cost and tax shield from deposit and debt coupons are decreasing function of \( \sigma \), and the total value would increase when the manager becomes risk-averse. Intuitively, when the manager becomes more risk-averse, his choice would lower the probability of default. An immediate consequence is that the bank is less likely to suffer from the bankruptcy cost and will in expectation benefit from tax shield for a longer period.

*Well-designed* CoCos, if effectively force the manager to choose low volatility via dilution threat, increase the bank’s total value as well.

### 6.3 Regulator’s Problem

The regulator plays an important role in stabilizing the financial system - he naturally faces a principal-agent problem while setting the capital requirements. We consider the regulator’s value function as in 6.1.4 and compare his optimal choice for default boundary \( V_b \) in different cases: 1) the bank is financed only by deposits, debts and equities 2) the bank is financed by deposits, debts, CoCos, and equities.

#### 6.3.1 Without CoCos

The regulator’s value function is

\[
R(v) = \mathbb{E}_{V_0=v}^Q \left[ \int_0^{\tau_b} \varphi e^{-rt} dt - e^{-r\tau_b} \left( \frac{d - \varphi}{r} - (1 - \alpha)V_b \right) \right] + \text{Premium from deposits} \\
- p \cdot \mathbb{E}_{V_0=v}^Q \left[ e^{-r\tau_b} \left( \frac{b}{r} - \left( (1 - \alpha)V_b - \frac{d - \varphi}{r} \right)_+ \right) \right] + \text{Guarantee for deposits} \\
+ \theta \cdot \mathbb{E}_{V_0=v}^Q \left[ \int_0^{\tau_b} e^{-rt} (\delta V_t - (b + d)) dt \right] + \text{Cost for bailout} \\
+ \mathbb{E}_{V_0=v}^Q \left[ \int_0^{\tau_b} e^{-rt} (\delta V_t - (b + d)) dt \right] + \text{Tax income} \tag{6.3.22}
\]
For simplicity, we assume the deposit premium \( \varphi \) is chosen to be a fair price - the value of the deposit insurance is equal to the expected discounted value of the premium. With this simplifying assumption, the first two terms in the value function cancel out. Also, if the premium is non-zero, the residual value upon default is not enough to cover deposit payments. If the regulator decides to bail out the bank, he will have to pay \( \frac{b}{r} \) as the bonds issued by the bank becomes worthless. Hence, the regulator’s value can be written as:

\[
R(v) = \theta \cdot \mathbb{E}_{V_0=v}^{Q} \left[ \int_{0}^{T_b} e^{-rt} (\delta V_t - (b + d)) \, dt \right] - p \cdot \mathbb{E}_{V_0=v}^{Q} \left[ e^{-rT_b} \left( \frac{b}{r} - ((1 - \alpha)V_b - \frac{d - \varphi}{r})_+ \right) \right]
\]

We let \( R^{\text{noCoCo}}(v; V_b) \) denote the regulator’s value function in this case, and his objective is to choose \( V_b \) to maximize \( V_b \mapsto R^{\text{noCoCo}}(v; V_b) \). The choice of \( V_b \) will affect the value of \( \varphi \): the higher the default residual value is, the lower the premium required for deposit guarantee. We distinguish the following two cases:

- If \( V_b < \frac{d}{(1-\alpha)r} \), then \( \varphi > 0 \),

\[
R^{\text{noCoCo}}(v; V_b) = \theta \cdot (\mathcal{V}(v, b + d) - \mathcal{V}(V_b, b + d)D_{\gamma(\sigma)}(v, V_b)) - p \cdot \frac{b}{r} D_{\gamma(\sigma)}(v, V_b),
\]

- and if \( V_b \geq \frac{d}{(1-\alpha)r} \), then \( \varphi = 0 \),

\[
R^{\text{noCoCo}}(v; V_b) = \theta \cdot (\mathcal{V}(v, b + d) - \mathcal{V}(V_b, b + d)D_{\gamma(\sigma)}(v, V_b)) - p \cdot \left( \frac{b + d}{r} - (1 - \alpha)V_b \right) D_{\gamma(\sigma)}(v, V_b).
\]

Also, \( \sigma \) depends on \( V_b \) as shown in Section 6.2.3: if \( V_{\text{endo}} \leq V_b \leq V_b^\ast \) (resp. \( V_b > V_b^\ast \)), \( \sigma = \sigma_h \) (resp. \( \sigma = \sigma_h \)). We can derive the regulator’s optimal choice for \( V_b \) and obtain the following theorem:
Theorem 6.3.1. If \( \theta > p \cdot (1 - \alpha) \), the regulator’s value is a strictly decreasing function of \( V_b \) on the interval \([V_{\text{endo}}, \frac{b+d}{r}]\) and \([\frac{b+d}{r}, \infty)\). Therefore, the regulator’s optimal choice would either be \( V_b^* \) or \( V_{\text{endo}} \). In the first case, the manager would choose \( \sigma_h \) and in the second case the manager would choose \( \sigma_h \).

The condition \( \theta > p \cdot (1 - \alpha) \) guarantees the tax rate to be high enough so that the regulator has incentive to keep the bank running. This condition is consistent with empirical facts: the corporate tax rate in the United States vary from 15\% to 35\%. As for the bankruptcy cost, For example, in order to fit the data, Cremers et al. (2007) need bankruptcy costs to be approximately equal to 50\%. Furthermore, given Lehman Brothers’ extremely low recovery rate of 7\% (compared to an average of 40\%) in 2008, it is reasonable to assume that banks also have higher bankruptcy cost due to considerable price impact while selling off financial assets.

The theorem shows without CoCos, the regulator is always better off when the default boundary is lower. A lower default boundary would entail lower probability of default, thus increase the expected tax income while decreasing the expected cost of bail-out. Unfortunately, a low default boundary would give the manager wrong risk incentive, so if the manager is able to take up extremely risky opportunities (i.e. \( \sigma_h \gg \sigma_h \)), the regulator has to choose the critical \( V_b \) value to dissuade the manager from risk-taking activities, but choosing a high default boundary will increase the default probability as well as the expected cost of bail-out option. The regulator naturally faces the dilemma of giving the right risk-incentive while lowering the default probability.

\footnote{In fact, the regulator should choose \( V_b \) to be slightly higher than the critical value \( V_b^* \). If \( V_b = V_b^* \), the manager would be indifferent about the volatility level.}
6.3.2 With CoCos

We now assume the bank has issued well-designed CoCos. All else being equal, the regulator’s value can be written as:

\[
R(v) = \theta \cdot E_{V_0=v}^{Q} \left[ \int_0^{\tau_b} e^{-rt} (\delta V_t - (b + d)) \, dt \right] - p \cdot E_{V_0=v}^{Q} \left[ e^{-r\tau_b} \left( \frac{b}{r} - ((1 - \alpha)V_b - \frac{d - \varphi}{r})^+ \right) \right] - \theta \cdot E_{V_0=v}^{Q} \left[ \int_0^{\tau_c} e^{-rt} \, dt \right].
\]

The last term corresponds to losses related to tax deductibility of CoCo coupons, it is zero if CoCo coupons are not tax-deductible. We let \( R^{\text{CoCo}}(v; V_b) \) denote the regulator’s value function in this case, and his objective is to choose \( V_b \) to maximize \( V_b \mapsto R^{\text{CoCo}}(v; V_b) \). The losses incurred by tax-deductibility are not affected by the choice of \( V_b \), for brevity, in this section we only study the case where CoCo coupons are not tax-deductible. Nonetheless, the effect of tax deductible to the regulator’s value is not clear in general. Allowing CoCo coupons to be tax-deductible certainly reduces the regulator’s tax income, however we have shown it also decreases the threat of dilution upon conversion since it implicitly increases the cost of CoCo coupon payments, and in turn induces the manager to bear more risks before the conversion.

As CoCos are supposed to be well-designed, the manager chooses low volatility before the conversion (\( \sigma^{bc} = \sigma_h \)). The choice after conversion \( \sigma \) depends on whether \( V_b \) is larger or smaller than \( V_b^* \). Similar as in the previous case without CoCos, according to different values of \( \varphi \), we distinguish two cases and use the Laplace transform of \( \tau_b \) derived in the Proof of theorem 6.2.3 to compute the value function:
• If $V_b < \frac{d}{(1-\alpha)r}$, then $\varphi > 0$,

$$R^{CoCo}(v; V_b) = \theta \cdot (V(v, b + d) - V(V_b, V_b)D_{\gamma(\sigma_h)}(V_c, V_b)) - p \cdot \frac{b}{r} D_{\gamma(\sigma_h)}(v, V_c)D_{\gamma(\sigma)}(V_c, V_b),$$

• and if $V_b \geq \frac{d}{(1-\alpha)r}$, then $\varphi = 0$,

$$R^{noCoCo}(v; V_b) = \theta \cdot (V(v, b + d) - D_{\gamma(\sigma_h)}(v, V_c)D_{\gamma(\sigma)}(V_c, V_b)) - p \cdot \left(\frac{b + d}{r} - (1 - \alpha)V_b\right) D_{\gamma(\sigma_h)}(v, V_c)D_{\gamma(\sigma)}(V_c, V_b).$$

We have already shown well-designed CoCos induce manager to be more risk-averse, we should expect CoCo to resolve the regulator’s dilemma of tradeoff between risk-incentive and default probability while setting a capital requirement. The following proposition shows the situation can at least be alleviated with CoCo:

**Proposition 6.3.2.** When $V_b$ is chosen to be the endogenous default boundary $V_{endo}$, for a bank with well-designed CoCos, the regulator’s value ($R^{CoCo}(v; V_{endo})$) is strictly higher than that of a bank without CoCos ($R^{noCoCo}(v; V_{endo})$); when $V_b$ is chosen to be the critical value $V_b^*$, issuing CoCos or not does not affect the regulator’s value.

The proposition shows the regulator is more likely to choose $V_{endo}$ if CoCos are issued. If the regulator does not impose tight capital requirement, then the classical moral hazard arises as the manager’s risk incentive is distorted. Under the presence of well-designed CoCos, we show that it is possible to restore the manager’s risk incentive. Intuitively, if the conversion trigger level $V_c$ is close to $V_b$ (or $V_{endo}$), as the manager chooses low volatility before conversion, he would be risk-averse during the majority of time.
6.4 Conclusion

In our opinion, well-designed CoCo needs to be converted before the bankruptcy and prevent the manager from bearing high risks prior to the conversion. Ideally, it also helps the regulator gain more flexibilities in setting the rules for capital requirements. By solving the principal-agent problem between the regulator and the manager, we have established the following recommendation for CoCo design:

1. High conversion ratio $n$ or low conversion price,
2. Low and tax-deductible coupon payments $c$,
3. Conversion trigger level $V_c$ high relative to face value of liabilities $\frac{b+d+c}{r}$ but as low as possible.

These conditions guarantee CoCos to be well-designed, offering incentive of bearing lower risks to manager prior to conversion and increases the total value of the bank. The cost of creating such incentive is to the private sector - the equity holders, as lower volatility decreases their option value. The added value is transferred to both CoCo holders via a low conversion price and debt holders and depositors via a lower default probability. Unfortunately, the existing issuances of CoCos are far from being well-designed according to these criteria. For instance, some series of the Lloyds CoCos issued in 2009 have an annual coupon rate as high as 15%, while the prevailing LIBOR rate was merely 0.50% in GBP. Furthermore, the conversion price is likely to be at a large premium - the conversion price is determined by taking an average of stock prices one month prior to the issuance, as the conversion supposedly occurs when the bank suffers from losses, the stock price at conversion would be much lower than before issuance.

The instrument is structured in the opposite way as people are fascinated by the appealing feature of contingent capital. Indeed, issuers will receive capitals only
when they are in need, and investors can gain equity exposure via this fixed-income instrument. Issuers thus obtain cheap capital, while the cheapness is achieved through tax shield effect. However, as pointed out by Sundaresan and Wang (2011), this would be at taxpayers’ cost, not for free. We argue that well-designed CoCos can only create added value by rendering the banks safer, while increasing the regulator’s, debt holders’ and depositors’ value to the detriment of equity holders’ value.
Figures
This figure plots the time series of exchange rate volatility forecasts under the pricing (top panel) and objective (bottom panel) measures, based on the output of the calibrated model (Specification 1). The calibration spans the period from January 1999 to June 2012 ($T = 3520$ days), and covers a cross-section of 24 G10 currency pairs including: (a) all X/USD currency pairs (9 pairs); and, (2) cross-pairs formed on the basis of currencies which had the highest or lowest interest rates in the G10 set at any point in our sample (15 pairs). Under the pricing measure, the empirical volatility forecast ($\text{Data}$) is represented by the first principal component extracted from the panel of observed at-the-money option-implied volatilities. $\text{Model}$ is the first principal component extracted from the panel of fitted at-the-money volatilities from the calibrated model. The mean (volatility) of each principal component is set equal to the time-series mean (volatility) of the equal-weighted average of the series in the respective underlying panel. The bottom plot repeats the analysis for monthly volatility forecasts under the objective measure. To construct an empirical forecast of volatility for each currency pair at each point in time, we use its daily return standard deviation, computed using a backward-looking, 63-day window. $\text{Data}$ plots the first principal component extracted from the panel of these estimates for the 24 currency pairs. $\text{Model}$ is the first principal component of the one-month exchange rate volatility forecasts, under the objective measure, obtained from the calibrated model. The mean (volatility) of each principal component is set equal to the time-series mean (volatility) of the equal-weighted average of the series in the respective underlying panel.
Figure 2. The Cross-Section of Fitted Option-Implied Volatilities.

This figure illustrates the fit of the model to FX option implied volatilities across strikes for a subset of the currency cross-rate pairs used in the calibration. At each point in time (1999:1-2012:6) the calibration matches the prices of a cross-section of 120 FX option prices for 24 G10 currency pairs (Specification I). Each panel plots the mean actual option-implied volatility (blue) and its fitted counterpart (dashed red) for pairs formed by combining two “high” interest rate currencies (AUD, NOK) with two “low” interest rate currencies (CHF, JPY), as well as, pairs involving each of these currencies against the U.S. dollar. All values are annualized and reported in units of percent per annum.

To illustrate the economic quality of the fit, the actual mean implied volatility is plotted with a typical bid-ask spread (dashed blue lines), equal to 0.1 times the implied volatility at that strike; the mean at-the-money volatility is reported in the title of each subplot. For each currency pair we plot option-implied volatilities at the five quoted strikes (10δ put, 25δ put, at-the-money, 25δ call, 10δ call). The plot reports time-series means computed using daily data (T = 3520 days).
Figure 3. The Cross-Section of Model and Realized Currency Risk Premia.

The left panel plots the cross-sectional relation between the mean one-month-ahead realized currency pair excess return, and the mean option-implied currency pair risk premium obtained from the calibrated model (Specification I). The right panel plots the cross-sectional relation between the mean realized currency pair excess return and the mean one-month interest rate differential. Each subplot is equipped with a 45-degree line with an intercept of zero. In the left panel, this line corresponds to the hypothesis that the calibrated, option-implied risk premia are an unbiased predictor of currency excess returns. In the right panel, this line corresponds to a random walk model of exchange rate dynamics. The results are plotted for the 24 G10 currency pairs used in the model calibration. This set includes: (a) all X/USD currency pairs (9 pairs); and, (2) cross-pairs formed on the basis of currencies which had the highest or lowest interest rates in the G10 set at any point in our sample (15 pairs). X/USD pairs are additionally denoted with red dots. The plots report data spanning the period 1999:1-2012:6; all quantities are reported in annualized terms.
Figure 4. Option-Implied Currency Risk Premia

This figure illustrates the model-implied risk premia for the conditional $HML_{FX}$ (top panel) and short dollar (bottom panel) factor mimicking portfolios. The $HML_{FX}$ replicating portfolio is long (short) the currencies with the highest (lowest) prevailing one-month LIBOR interest rates. Positions are spread-weighted, and the portfolio is constrained to be dollar-neutral. The short dollar risk premium is computed as the model risk premium for a portfolio which is short the U.S. dollar against an equal-weighted basket of G10 currencies, net of the premium attributable to its $HML_{FX}$ exposure. The figures plot the 20-day moving average of the raw model-implied risk premia; all values are reported in annualized terms. The plots span the period 1999:1-2012:6 ($T = 3520$ days), and are based on the output of Specification I.
Figure 5. Decompositions of Option-Implied Currency Risk Premia.

This figure decomposes the model-implied risk premia for the conditional $HML_{FX}$ and short dollar factor replicating portfolios, across the distributional features of the corresponding pricing kernel innovations. The $HML_{FX}$ risk premium is computed as the model risk premium for a portfolio which is long (short) the currencies with the highest (lowest) prevailing one-month LIBOR interest rates. Positions are spread-weighted, and the portfolio is constrained to be dollar-neutral. The short dollar risk premium is computed as the model risk premium for a portfolio which is short the U.S. dollar against an equal-weighted basket of G10 currencies, net of the premium attributable to its $HML_{FX}$ exposure. The top (bottom) panels report decompositions for the $HML_{FX}$ (short dollar) risk premia. The left panels decompose the risk premia across the symmetric (even) and asymmetric (odd) cumulants of the global pricing kernel innovation, $L_{t+1}^q$ and $L_{t+1}^{US}$, respectively. The right panels decompose the risk premia into contributions from the variance, skewness, and higher moments (Other) of the innovations. The panels plot the 20-day moving average of the share of the risk premium due to each component, and report their means in the title. The plots span the period 1999:1-2012:6 ($T = 3520$ days), and are based on the output of Specification I.
Figure 6. \(Q\)-Survival Probabilities, CDS Spreads and CoCo Prices for Intensity-Based Model.

This figure illustrates the calibration to survival probabilities and CDS spreads as well as the choice of implied recovery rate for the intensity-based model with CIR jump intensity.

The top two panels show CoCo prices produced by the model for different values of the coefficient of jump in jump intensity (\(\beta\)). The price of Lloyds Banking Group ECNs depends on the jump size of equity price \(\gamma\) upon conversion, we plot the prices corresponding to \(\gamma = -99\%, -50\%, -20\%, 0\%\) and \(20\%\).

The panels in the middle compare the \(Q\)-survival probabilities of Lloyds Banking Group and Rabobank implied by our model to the ones constructed with the O’Kane and Turnbull (2003) approach.

The bottom panels plot the CDS spreads generated by our model against the market quotes.
Figure 7. CoCo Value Decompositions in the Intensity-Based Model with CIR Jump Intensity.

This figure decomposes the total value of CoCo into three parts: 1) the coupons 2) the principal 3) the equity payment (or the cash payment) upon conversion as a function of the recovery rate. The top panel shows the decomposition for Lloyds ECNs and the bottom panel shows the decomposition for Rabobank SCNs. The pricing model is chosen to be the intensity-based model with CIR jump intensity.
Figure 8. Q-Survival Probabilities, CDS Spreads and CoCo Prices for First Passage Time Model.

This figure illustrates the calibration to survival probabilities and CDS spreads as well as the choice of implied recovery rate in the first passage time model with exponential OU accounting ratio.

The top two panels show CoCo prices produced by the model for different values of the recovery rate $R$ chosen in the CDS calibration.

The two panels in the middle compare the Q-survival probabilities of Lloyds Banking Group and Rabobank implied by our model to the ones constructed with the O’Kane and Turnbull (2003) approach.

The bottom panels plot the CDS spreads generated by our model against the market quotes.
Figure 9. CoCo Value Decompositions in the First Passage Time Model with Exponential OU Accounting Ratio.

This figure decomposes the total value of CoCo into three parts: 1) the coupons 2) the principal 3) the equity payment (or the cash payment) upon conversion as a function of the recovery rate. The top panel shows the decomposition for Lloyds ECNs and the bottom panel shows the decomposition for Rabobank SCNs. The pricing model is chosen to be the first passage time model with exponential OU accounting ratio.
Tables
This table reports summary statistics for the calibrated model parameters. Results are reported for calibrations using two sets of test assets: (a) options on X/USD exchange rates and cross-pairs involving currencies with high/low interest rates (HLX + X/USD pairs; 24 pairs); and, (b) options on the full set of G10 cross rates (All pairs; 45 pairs). Calibrations use one-month FX option-implied volatilities at five option strikes (10\(\delta\) put, 25\(\delta\) put, at-the-money, 25\(\delta\) call, 104 call), as well as, forecasts of monthly exchange rate volatility under the objective measure. The data span the period from January 1999 to June 2012 (T = 3520 days). The model is fitted separately on each day in the sample by minimizing the sum of squared relative fitting errors for the option-implied volatilities and volatility forecasts. Panel A reports the time series means and volatilities of the country-specific loadings, \(\xi_t\). We also report the mean one-month LIBOR interest rate differential for each country relative to the U.S. \((y_{i,t+1} - y_{US,t+1})\). Panel B summarizes the characteristics of the global factor innovation, \(L_{t+1}\). We report the mean level of the global state variable, \(Z_t\), the share of variance of the global innovation due to the non-Gaussian component, \(\eta^g_t\), and estimates of the parameters of the global CGMY component, \(G^g_t\) (dampening coefficient) and \(Y^g_t\) (power coefficient). Finally, we compute the skewness (Skewness\(^g_t\)) and kurtosis (Kurtosis\(^g_t\)) of the conditional distribution of the global factor innovation. Panel C reports the fitting errors for matching option implied volatilities and forecasts of volatility under the objective measure, measured in volatility points. We compute root mean squared errors (RMSE) for each pair in the target option set, and report their mean pooled: (a) across all strikes and currency pairs; (b) across all pairs with a given option strike. \(\Phi\) Vol. Forecast reports the RMSE of the model-implied objective measure volatility forecast relative to its empirical estimate averaged across all pairs in the panel.

### Panel A: Global Factor Loadings, \(\xi_t\)

<table>
<thead>
<tr>
<th>Specification</th>
<th>Moment</th>
<th>AUD</th>
<th>CAD</th>
<th>CHF</th>
<th>EUR</th>
<th>GBP</th>
<th>JPY</th>
<th>NOK</th>
<th>NZD</th>
<th>SEK</th>
<th>USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>I (HLX and X/USD pairs)</td>
<td>Mean</td>
<td>0.80</td>
<td>0.99</td>
<td>1.03</td>
<td>1.00</td>
<td>0.99</td>
<td>1.04</td>
<td>1.00</td>
<td>0.77</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Volatility</td>
<td>0.08</td>
<td>0.07</td>
<td>0.13</td>
<td>0.10</td>
<td>0.08</td>
<td>0.07</td>
<td>0.14</td>
<td>0.09</td>
<td>0.10</td>
<td>-</td>
</tr>
<tr>
<td>II (All pairs)</td>
<td>Mean</td>
<td>0.88</td>
<td>0.95</td>
<td>1.12</td>
<td>1.09</td>
<td>1.03</td>
<td>1.06</td>
<td>1.06</td>
<td>0.86</td>
<td>1.06</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Volatility</td>
<td>0.11</td>
<td>0.06</td>
<td>0.17</td>
<td>0.18</td>
<td>0.11</td>
<td>0.07</td>
<td>0.18</td>
<td>0.11</td>
<td>0.19</td>
<td>-</td>
</tr>
</tbody>
</table>

\(y_{i,t+1} - y_{US,t+1}\) [%]

<table>
<thead>
<tr>
<th>Specification</th>
<th>Moment</th>
<th>AUD</th>
<th>CAD</th>
<th>CHF</th>
<th>EUR</th>
<th>GBP</th>
<th>JPY</th>
<th>NOK</th>
<th>NZD</th>
<th>SEK</th>
<th>USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>I (HLX and X/USD pairs)</td>
<td>Mean</td>
<td>2.46</td>
<td>0.20</td>
<td>-1.58</td>
<td>-0.15</td>
<td>1.06</td>
<td>-2.54</td>
<td>1.53</td>
<td>2.78</td>
<td>0.07</td>
<td>0.00</td>
</tr>
<tr>
<td>II (All pairs)</td>
<td>Mean</td>
<td>2.46</td>
<td>0.20</td>
<td>-1.58</td>
<td>-0.15</td>
<td>1.06</td>
<td>-2.54</td>
<td>1.53</td>
<td>2.78</td>
<td>0.07</td>
<td>0.00</td>
</tr>
</tbody>
</table>

### Panel B: Global Factor Summary Statistics

<table>
<thead>
<tr>
<th>Specification</th>
<th>Moment</th>
<th>(Z_t)</th>
<th>(\eta^g_t)</th>
<th>(G^g_t)</th>
<th>(Y^g_t)</th>
<th>Skewness(^g_t)</th>
<th>Kurtosis(^g_t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I Mean</td>
<td>0.27</td>
<td>0.38</td>
<td>9.47</td>
<td>-0.75</td>
<td>-0.64</td>
<td>5.45</td>
<td></td>
</tr>
<tr>
<td>Volatility</td>
<td>0.57</td>
<td>0.22</td>
<td>9.25</td>
<td>0.18</td>
<td>0.80</td>
<td>3.86</td>
<td></td>
</tr>
<tr>
<td>II Mean</td>
<td>0.33</td>
<td>0.33</td>
<td>5.74</td>
<td>-0.83</td>
<td>-0.67</td>
<td>6.59</td>
<td></td>
</tr>
<tr>
<td>Volatility</td>
<td>0.70</td>
<td>0.22</td>
<td>5.93</td>
<td>0.16</td>
<td>0.58</td>
<td>4.68</td>
<td></td>
</tr>
</tbody>
</table>

### Panel C: Fitting Errors (RMSE; volatility points)

<table>
<thead>
<tr>
<th>Specification</th>
<th>All</th>
<th>10(\delta)</th>
<th>25(\delta)</th>
<th>50(\delta)</th>
<th>104(\delta)</th>
<th>(\Phi) Vol. Forecast</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0.83</td>
<td>0.96</td>
<td>0.84</td>
<td>0.79</td>
<td>0.78</td>
<td>0.82</td>
</tr>
<tr>
<td>II</td>
<td>1.08</td>
<td>1.15</td>
<td>1.10</td>
<td>1.04</td>
<td>1.05</td>
<td>1.06</td>
</tr>
</tbody>
</table>
This table reports the results of regressions forecasting currency excess returns using option-implied currency risk premia and interest rate differentials. The dependent variable is the 21-day U.S. dollar simple excess return on currency pair $J/I$, and the explanatory variables are observed/computed as of the last day preceding the return measurement interval. The data are sampled weekly. We report two sets of results. The first is based on repeated cross-sectional regressions (Fama-MacBeth; Panel A); the second, is based on a pooled panel regression (Panel B). For cross-sectional regressions, the reported coefficients are time series averages ($N = 700$); Newey-West standard errors of the coefficient estimates are reported in parentheses. Adj. $R^2$ reports the average cross-sectional adjusted $R^2$. The second set of regressions is based on pooled panel regressions with currency-pair fixed effects ($N = 700 \text{ weeks} \cdot 24 \text{ pairs} = 16,800$). For panel regressions, standard errors are adjusted for cross-sectional correlation and time series auto- and cross-correlations using the methodology from Thompson (2011) with eight lags. The dependent variables are: the U.S. investor’s model-implied risk premium for exposure to the $J/I$ exchange rate ($\lambda_{ji,US}^t$), the global and country-specific components of the model-implied risk premium ($\lambda_{HML,t}^{ji,US}$ and $\lambda_{refFX,t}^{ji,US}$), and the differential in the one-month LIBOR rates ($y_{j,t,t+1} - y_{i,t,t+1}$). Model-implied variables are computed using the calibrated parameter values from Specification I. Finally, we report the $p$-value for two hypothesis tests. The first hypothesis ($H_0$) asserts that the model is correctly specified, such that the regression intercept is zero and the slope coefficients on the model-implied risk premia are one. If no model-implied variables are included in the regression, we do not report the result of the test. The second hypothesis ($H_1$) is that the included variables have no explanatory power (i.e. all the coefficients with the exception of the intercept are zero). The regressions are based on data from January 1999 to June 2012, unless noted otherwise.

### Panel A: Cross-Sectional Regressions

<table>
<thead>
<tr>
<th></th>
<th>Full sample</th>
<th>Excluding 2008</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Intercept (x100)</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>$\lambda_{ji,US}^t$</td>
<td>1.94</td>
<td>2.28</td>
</tr>
<tr>
<td></td>
<td>(0.87)</td>
<td>(0.92)</td>
</tr>
<tr>
<td>$\lambda_{HML,t}^{ji,US}$</td>
<td>2.30</td>
<td>2.68</td>
</tr>
<tr>
<td></td>
<td>(0.98)</td>
<td>(1.04)</td>
</tr>
<tr>
<td>$\lambda_{refFX,t}^{ji,US}$</td>
<td>1.70</td>
<td>1.65</td>
</tr>
<tr>
<td></td>
<td>(5.40)</td>
<td>(5.69)</td>
</tr>
<tr>
<td>$y_{j,t,t+1} - y_{i,t,t+1}$</td>
<td>0.14</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.29</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>N</td>
<td>700</td>
<td>700</td>
</tr>
<tr>
<td>$H_0$ p-value</td>
<td>0.49</td>
<td>0.59</td>
</tr>
<tr>
<td>$H_1$ p-value</td>
<td>0.03</td>
<td>0.05</td>
</tr>
</tbody>
</table>

### Panel B: Panel Regressions

<table>
<thead>
<tr>
<th></th>
<th>Full sample</th>
<th>Excluding 2008</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Intercept (x100)</td>
<td>0.05</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.10)</td>
</tr>
<tr>
<td>$\lambda_{ji,US}^t$</td>
<td>0.13</td>
<td>1.28</td>
</tr>
<tr>
<td></td>
<td>(0.45)</td>
<td>(0.37)</td>
</tr>
<tr>
<td>$\lambda_{HML,t}^{ji,US}$</td>
<td>0.10</td>
<td>1.24</td>
</tr>
<tr>
<td></td>
<td>(0.43)</td>
<td>(0.34)</td>
</tr>
<tr>
<td>$\lambda_{refFX,t}^{ji,US}$</td>
<td>2.35</td>
<td>3.95</td>
</tr>
<tr>
<td></td>
<td>(3.84)</td>
<td>(4.04)</td>
</tr>
<tr>
<td>$y_{j,t,t+1} - y_{i,t,t+1}$</td>
<td>0.09</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>16,800</td>
<td>16,800</td>
</tr>
<tr>
<td>$H_0$ p-value</td>
<td>0.10</td>
<td>0.07</td>
</tr>
<tr>
<td>$H_1$ p-value</td>
<td>0.87</td>
<td>0.92</td>
</tr>
</tbody>
</table>
### Table III

Empirical Factor Mimicking Portfolios

This table compares the realized returns to empirical factor mimicking portfolios with the corresponding option-implied risk premia. The factor mimicking portfolios are rebalanced monthly, and the data span the period from January 1999 to June 2012 ($N = 162$ months). The Conditional $HML_{FX}$ factor mimicking portfolio is a dollar-neutral portfolio formed by sorting currencies into long and short portfolios on the basis of their prevailing one-month LIBOR rates as of each month-end. Currencies are spread-weighted within the long and short portfolios using the absolute deviation of their respective interest rates from the mean of the interest rates in countries with ranks five and six. The portfolio is constrained to be dollar-neutral. The Unconditional $HML_{FX}$ portfolio is formed analogously, but with currencies sorted on the basis of their mean historical interest rates, computed using an expanding window starting in January 1990. The Short USD factor mimicking portfolio is long an equal-weighted basket of G10 currencies against the U.S. dollar. We report the moments of the realized factor mimicking portfolios returns and the $p$-value of the Jarque-Bera test of Gaussianity of the returns and Z-scores. Model risk premia are computed on the basis of Specification I, and are decomposed into contributions from exposure to the global ($HML$) and U.S. country-specific (Short reference) pricing kernel innovations. Portfolio $\xi^{long}$ and $\xi^{short}$ report the time-series mean of the global factor loading for the long and short legs of the factor mimicking portfolio, respectively. $t$-statistics of the mean realized and model risk premia, as well as, their differences are reported in square brackets.

<table>
<thead>
<tr>
<th></th>
<th>$HML_{FX}$</th>
<th>Conditional</th>
<th>Unconditional</th>
<th>Short USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Realized</td>
<td>Mean</td>
<td>4.96</td>
<td>3.32</td>
<td>3.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[1.92]</td>
<td>[1.32]</td>
<td>[1.29]</td>
</tr>
<tr>
<td></td>
<td>Volatility</td>
<td>9.51</td>
<td>9.26</td>
<td>8.90</td>
</tr>
<tr>
<td></td>
<td>Skewness</td>
<td>-1.07</td>
<td>-0.89</td>
<td>-0.17</td>
</tr>
<tr>
<td></td>
<td>Kurtosis</td>
<td>7.63</td>
<td>6.90</td>
<td>3.72</td>
</tr>
<tr>
<td></td>
<td>JB (returns)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>JB (Z-scores)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Model</td>
<td>Total risk premium</td>
<td>3.87</td>
<td>3.80</td>
<td>1.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[7.35]</td>
<td>[7.53]</td>
<td>[9.64]</td>
</tr>
<tr>
<td></td>
<td>Global (HML)</td>
<td>3.87</td>
<td>3.80</td>
<td>1.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[7.35]</td>
<td>[7.53]</td>
<td>[7.51]</td>
</tr>
<tr>
<td></td>
<td>Short reference (USD)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[-]</td>
<td>[-]</td>
<td>[22.89]</td>
</tr>
<tr>
<td></td>
<td>Portfolio $\xi^{long}$</td>
<td>0.86</td>
<td>0.90</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>Portfolio $\xi^{short}$</td>
<td>1.02</td>
<td>1.03</td>
<td>1.00</td>
</tr>
<tr>
<td>Difference</td>
<td>Mean</td>
<td>1.09</td>
<td>-0.48</td>
<td>1.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.38]</td>
<td>[-0.17]</td>
<td>[0.46]</td>
</tr>
</tbody>
</table>

137
Table IV
Model Risk Premium Decompositions

This table uses the calibrated model to compute and decompose the option-implied currency risk premia from the baseline calibration (Specification I). The model risk premia are computed daily and span from January 1999 to June 2012 (\(T = 3520\) days). Panel A decomposes the model premium for the \(HML_{FX}\) factor mimicking portfolio, which is a spread-weighted, dollar-neutral portfolio. As of each month-end, the portfolio sorts G10 currencies – excluding the U.S. dollar – into long and short legs using their yields, \(y_{i,t+1}\), and weighs each currency on the basis of the absolute deviation of its yield from the average yields of currencies with ranks five and six. We report the global factor loadings of the long (\(\xi^{long}_{i}\)) and short (\(\xi^{short}_{i}\)) portfolios, as well as, the mean portfolio risk premium (\(\lambda_{HML} \% per annum\)). We then decompose the portfolio risk premium into contributions from the even (symmetric) and odd (asymmetric) cumulants of the global factor innovation, \(L_{t+1}^{g}\). Finally, we decompose risk premia across the moments of the pricing kernel innovations (variance, skewness, etc.). For each quantity, we report its time-series mean, volatility, and the 5\(^{th}\) and 95\(^{th}\) percentiles of its distribution. Panel B reports the time series mean, volatility, and 5\(^{th}\) and 95\(^{th}\) percentiles of the model-implied short reference risk premium for each country (% \(per annum\)). The risk premium is then decomposed into: (a) contributions from even and odd cumulants of the country-specific pricing kernel innovation, \(L_{t+1}^{i}\); and, (b) contributions from variance and higher moments of the kernel innovations.

### Panel A: Global Factor Risk Premium [%]

<table>
<thead>
<tr>
<th></th>
<th>Loadings</th>
<th>Risk premium</th>
<th>By Cumulant</th>
<th>By Moment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\xi^{long}_{i})</td>
<td>(\xi^{short}_{i})</td>
<td>(\lambda_{HML} %)</td>
<td>Even</td>
</tr>
<tr>
<td>Mean</td>
<td>0.86</td>
<td>1.02</td>
<td>3.87</td>
<td>91.82</td>
</tr>
<tr>
<td>Volatility</td>
<td>0.10</td>
<td>0.07</td>
<td>7.59</td>
<td>9.11</td>
</tr>
<tr>
<td>5%</td>
<td>0.73</td>
<td>0.91</td>
<td>-0.73</td>
<td>79.23</td>
</tr>
<tr>
<td>95%</td>
<td>1.07</td>
<td>1.14</td>
<td>13.83</td>
<td>99.46</td>
</tr>
</tbody>
</table>

### Panel B: Short Reference Risk Premia [%]

<table>
<thead>
<tr>
<th>Country</th>
<th>(\lambda_{r_{FX}})</th>
<th>AUD</th>
<th>CAD</th>
<th>CHF</th>
<th>EUR</th>
<th>GBP</th>
<th>JPY</th>
<th>NOK</th>
<th>NZD</th>
<th>SEK</th>
<th>USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.38</td>
<td>0.47</td>
<td>0.45</td>
<td>0.64</td>
<td>0.49</td>
<td>0.79</td>
<td>0.56</td>
<td>0.49</td>
<td>1.07</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>Volatility</td>
<td>0.32</td>
<td>0.74</td>
<td>0.45</td>
<td>0.52</td>
<td>0.50</td>
<td>0.51</td>
<td>0.61</td>
<td>0.26</td>
<td>0.91</td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td>5%</td>
<td>0.13</td>
<td>0.01</td>
<td>0.09</td>
<td>0.11</td>
<td>0.13</td>
<td>0.28</td>
<td>0.11</td>
<td>0.17</td>
<td>0.34</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>95%</td>
<td>0.87</td>
<td>1.62</td>
<td>1.48</td>
<td>1.54</td>
<td>1.13</td>
<td>1.77</td>
<td>1.44</td>
<td>1.00</td>
<td>3.21</td>
<td>1.06</td>
<td></td>
</tr>
</tbody>
</table>

Even (Share [%]):

<table>
<thead>
<tr>
<th>Country</th>
<th>Mean</th>
<th>AUD</th>
<th>CAD</th>
<th>CHF</th>
<th>EUR</th>
<th>GBP</th>
<th>JPY</th>
<th>NOK</th>
<th>NZD</th>
<th>SEK</th>
<th>USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Volatility</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>5%</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>95%</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Odd (Share [%]):

<table>
<thead>
<tr>
<th>Country</th>
<th>Mean</th>
<th>AUD</th>
<th>CAD</th>
<th>CHF</th>
<th>EUR</th>
<th>GBP</th>
<th>JPY</th>
<th>NOK</th>
<th>NZD</th>
<th>SEK</th>
<th>USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volatility</td>
<td>0.78</td>
<td>0.42</td>
<td>0.70</td>
<td>0.46</td>
<td>0.43</td>
<td>1.15</td>
<td>0.35</td>
<td>0.87</td>
<td>0.37</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>5%</td>
<td>98.97</td>
<td>99.31</td>
<td>99.11</td>
<td>99.44</td>
<td>99.44</td>
<td>98.93</td>
<td>99.53</td>
<td>99.38</td>
<td>99.64</td>
<td>99.37</td>
<td></td>
</tr>
<tr>
<td>95%</td>
<td>99.99</td>
<td>100.00</td>
<td>99.99</td>
<td>100.00</td>
<td>100.00</td>
<td>99.99</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

Variance (Share [%]):

<table>
<thead>
<tr>
<th>Country</th>
<th>Mean</th>
<th>AUD</th>
<th>CAD</th>
<th>CHF</th>
<th>EUR</th>
<th>GBP</th>
<th>JPY</th>
<th>NOK</th>
<th>NZD</th>
<th>SEK</th>
<th>USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volatility</td>
<td>0.78</td>
<td>0.42</td>
<td>0.70</td>
<td>0.46</td>
<td>0.43</td>
<td>1.15</td>
<td>0.35</td>
<td>0.87</td>
<td>0.37</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>5%</td>
<td>98.97</td>
<td>99.31</td>
<td>99.11</td>
<td>99.44</td>
<td>99.44</td>
<td>98.93</td>
<td>99.53</td>
<td>99.38</td>
<td>99.64</td>
<td>99.37</td>
<td></td>
</tr>
<tr>
<td>95%</td>
<td>99.99</td>
<td>100.00</td>
<td>99.99</td>
<td>100.00</td>
<td>100.00</td>
<td>99.99</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

Other (Share [%]):

<table>
<thead>
<tr>
<th>Country</th>
<th>Mean</th>
<th>AUD</th>
<th>CAD</th>
<th>CHF</th>
<th>EUR</th>
<th>GBP</th>
<th>JPY</th>
<th>NOK</th>
<th>NZD</th>
<th>SEK</th>
<th>USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.24</td>
<td>0.16</td>
<td>0.23</td>
<td>0.14</td>
<td>0.13</td>
<td>0.31</td>
<td>0.12</td>
<td>0.20</td>
<td>0.10</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>Volatility</td>
<td>0.78</td>
<td>0.42</td>
<td>0.70</td>
<td>0.46</td>
<td>0.43</td>
<td>1.15</td>
<td>0.35</td>
<td>0.87</td>
<td>0.37</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>5%</td>
<td>0.01</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>95%</td>
<td>1.03</td>
<td>0.69</td>
<td>0.89</td>
<td>0.56</td>
<td>0.56</td>
<td>1.07</td>
<td>0.47</td>
<td>0.63</td>
<td>0.36</td>
<td>0.63</td>
<td></td>
</tr>
</tbody>
</table>

138
Table V
The Factor Structure of Currency Returns

This table contrasts the factor structure of currency returns in the historical sample with data simulated from the calibrated model (Specification I). To generate currency returns on the basis of the calibrated model, we sample the daily time series of calibrated parameters and observed one-month interest rates, and simulate the realizations of the global, \( L_{t+1} \), and country-specific, \( \{ L_{i,t+1} \}_{i=1}^{N=10} \), pricing kernel shocks. Using the simulated shocks, we construct 10,000 panels of currency returns whose dimension matches the empirical sample (\( T = 162 \) months, \( N = 10 \) currencies). Panel A reports summary statistics for the moments and Sharpe ratio of the conditional \( HML_{FX} \) and short dollar factor mimicking portfolios constructed within the historical and simulated datasets. For the historical sample, we report the point estimate of each statistic and the results of a \( t \)-test for a difference from zero under the assumption of i.i.d. returns (\( t(0, iid) \), in brackets). For the simulated samples, we report the average of each statistic across the simulated paths (\( \text{Mean} \)) along with its 95% confidence interval, as well as, the results of \( t \)-tests for a difference from zero (\( t(0) \)), and the historical point estimate (\( t(\text{data}) \), in curly braces). \( t \)-statistics for simulated samples are computed on the basis of the standard deviation of the simulated distribution of each summary statistic. Panel B reports results of panel regressions of monthly currency excess returns (measured relative to the U.S. dollar) onto the returns of the factor mimicking portfolios, as well as, interaction terms based on the average pairwise one-month interest rate differential \( (\tilde{y}_{t+1} - \tilde{y}_U) \), and the pairwise demeaned contemporaneous one-month interest rate differential \( (\tilde{y}_{t+1} - \tilde{y}_U) \). For the historical regressions, coefficient \( t \)-statistics (in brackets) are computed on the basis of standard errors, which have been adjusted for cross-sectional correlation and time series auto- and cross-correlations using the methodology from Thompson (2011) with three lags. For regressions using simulated data, \( t \)-statistics are computed on the basis of the standard deviation of the simulated distribution of the coefficient point estimate. We report the results of tests for a difference from zero (in brackets), and – a difference from the historical point estimate (in curly braces).

### Panel A: Factor mimicking portfolio returns

<table>
<thead>
<tr>
<th></th>
<th>( HML_{FX} )</th>
<th>( \text{Short USD} )</th>
<th>( HML_{FX} )</th>
<th>( \text{Short USD} )</th>
<th>( HML_{FX} )</th>
<th>( \text{Short USD} )</th>
<th>( HML_{FX} )</th>
<th>( \text{Short USD} )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Historical sample</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Point estimate</td>
<td>5.0%</td>
<td>3.3%</td>
<td>9.5%</td>
<td>9.3%</td>
<td>-1.07</td>
<td>-0.89</td>
<td>0.52</td>
<td>0.36</td>
</tr>
<tr>
<td>( t(0, iid) )</td>
<td>[1.92]</td>
<td>[1.29]</td>
<td>[18.00]</td>
<td>[18.00]</td>
<td>[-5.61]</td>
<td>[-4.67]</td>
<td>[6.23]</td>
<td>[4.42]</td>
</tr>
<tr>
<td><strong>Simulated (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>4.3%</td>
<td>2.1%</td>
<td>11.0%</td>
<td>8.4%</td>
<td>0.11</td>
<td>-0.41</td>
<td>0.40</td>
<td>0.26</td>
</tr>
<tr>
<td>( t(0) )</td>
<td>[1.40]</td>
<td>[0.93]</td>
<td>[4.38]</td>
<td>[6.25]</td>
<td>[0.07]</td>
<td>[-0.31]</td>
<td>[1.44]</td>
<td>[0.95]</td>
</tr>
<tr>
<td>( t(\text{data}) )</td>
<td>{-0.23}</td>
<td>{-0.51}</td>
<td>{0.58}</td>
<td>{-0.61}</td>
<td>{0.79}</td>
<td>{0.36}</td>
<td>{-0.45}</td>
<td>{-0.34}</td>
</tr>
<tr>
<td>2.5-pct</td>
<td>-1.6%</td>
<td>-2.4%</td>
<td>8.4%</td>
<td>6.7%</td>
<td>-2.80</td>
<td>-3.35</td>
<td>-0.14</td>
<td>-0.27</td>
</tr>
<tr>
<td>97.5-pct</td>
<td>10.3%</td>
<td>6.6%</td>
<td>14.9%</td>
<td>11.6%</td>
<td>3.20</td>
<td>2.04</td>
<td>0.94</td>
<td>0.82</td>
</tr>
</tbody>
</table>

### Panel B: Factor regressions

<table>
<thead>
<tr>
<th></th>
<th>( HML_{FX} ) return interacted with</th>
<th>( \tilde{y}_{t+1} - \tilde{y}_U )</th>
<th>( \tilde{y}_{t+1} - \tilde{y}_U )</th>
<th>( \text{Short USD} )</th>
<th>( \text{Adj.R}^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Historical sample</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Point estimate</td>
<td>1</td>
<td>23.93</td>
<td>2.78</td>
<td>0.99</td>
<td>71.1%</td>
</tr>
<tr>
<td>( t(0) )</td>
<td>-0.11</td>
<td>[7.36]</td>
<td>[1.97]</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Simulated (model)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( t(0) )</td>
<td>-0.11</td>
<td>[18.40]</td>
<td>[1.39]</td>
<td>[124.48]</td>
<td>58.4%</td>
</tr>
<tr>
<td>( t(\text{data}) )</td>
<td>{-0.28}</td>
<td>{0.27}</td>
<td>{0.02}</td>
<td>{0.34}</td>
<td>{-1.91}</td>
</tr>
</tbody>
</table>
Table VI
The Dynamics of Global Factor Loadings

This table reports the results from panel regressions of the calibrated global factor loading differentials ($\xi_i^t - \xi_i^{US}$) onto differentials in one-month LIBOR rates ($y_{i,t+1} - y_{i,t+1}^{US}$) and country-specific state variables $Y_i^t - Y_i^{US}$, as well as, specifications in which these variables have been scaled by their cross-sectional dispersion at each point in time. The calibrated factor loadings, interest rate differentials, and country-specific state variables are from Specification I. The data span the period from January 1999 to June 2012, and are sampled weekly ($N = 704$ weeks · $9 \times$ USD pairs = 6,336). We report results from a panel regression with country fixed effects (1-4), and a panel regression in which the country-fixed effects have been parameterized as a function of the time-series mean of the one-month LIBOR differential of each country relative to the U.S. dollar (5-8). Coefficient $t$-statistics are reported in brackets below, and are computed on the basis of standard errors, which have been adjusted for cross-sectional correlation and time series auto- and cross-correlations using the methodology from Thompson (2011) with eight lags. We report the regression adjusted $R^2$, and the adjusted $R^2$ net of the explanatory power of the fixed effects ([NFE]).

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>FE</td>
<td>FE</td>
<td>FE</td>
<td>FE</td>
<td>FE</td>
<td>FE</td>
<td>FE</td>
<td>FE</td>
</tr>
<tr>
<td>$y_{i,t+1} - y_{i,t+1}^{US}$</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.03</td>
<td>-1.30</td>
<td>-1.22</td>
<td>-1.13</td>
<td>-1.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-3.46</td>
<td>-3.23</td>
<td>-4.73</td>
<td>-4.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-2.63</td>
<td>-2.37</td>
<td>-3.97</td>
<td>-4.23</td>
</tr>
<tr>
<td>$Y_i^t - Y_i^{US}$</td>
<td>-1.10</td>
<td>-1.10</td>
<td>-1.87</td>
<td>-1.84</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_y(\eta_{t+1} - \eta_{t+1}^{US})$</td>
<td>0.02</td>
<td>0.02</td>
<td></td>
<td></td>
<td>-2.07</td>
<td>-2.25</td>
<td>-3.38</td>
<td>-1.97</td>
</tr>
<tr>
<td>$\sigma_Y(\eta_{t+1} - \eta_{t+1}^{US})$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-2.53</td>
<td>-2.53</td>
<td>-2.68</td>
<td>-2.71</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>49.9%</td>
<td>50.6%</td>
<td>50.5%</td>
<td>52.3%</td>
<td>32.9%</td>
<td>33.7%</td>
<td>32.7%</td>
<td>34.8%</td>
</tr>
<tr>
<td>Adj. $R^2$ [NFE]</td>
<td>3.4%</td>
<td>4.3%</td>
<td>4.4%</td>
<td>8.0%</td>
<td>5.0%</td>
<td>7.3%</td>
<td>2.4%</td>
<td>5.4%</td>
</tr>
<tr>
<td>$N$</td>
<td>6336</td>
<td>6336</td>
<td>6336</td>
<td>6336</td>
<td>6336</td>
<td>6336</td>
<td>6336</td>
<td>6336</td>
</tr>
</tbody>
</table>
This table characterizes the implications of imposing a parametric link between the global factor loadings, $\xi_{it}^i$, and interest rate differentials, on: (a) option-implied variance and skewness (Panels A and B); and, (b) model-implied risk premia for the $HML_{FX}$ and short dollar factor replicating portfolios (Panel C). The global factor loadings are parameterized as follows:

$$\hat{\xi}_{it}^i - \xi_{it}^{US} = \phi_0 + \phi_1 \cdot (y_{i,t+1} - y_{i,t+1}^{US}) + \phi_2 \cdot (y_{i,t+1} - y_{i,t+1} - y_{i,t+1}^{US})$$

where the coefficients $\phi_0$ and $\phi_1$ are set equal to their point estimates from regression specification (5) in Table VI, and $\phi_2$ is varied mechanically. Using the time series of the parameterized loadings, and the time series of remaining model parameters from Specification I, we recompute option prices as a function of the sensitivity of the loadings to the prevailing interest rate differential, $\phi_2$. Panel A (B) reports the results of panel regressions of the one-month option-implied variance (skewness) onto the prevailing one-month interest rate differential, $y_{i,t+1} - y_{i,t+1}^{US}$, in the raw data (Data), the preferred model calibration (Model), and after imposing the parametric link between the loadings and interest rate differentials. The data are sampled weekly, and include the 24 currency pairs used in Specification I ($N = 704$ weeks $\times$ 24 pairs = 16,896). The panel regressions include currency pair fixed effects. Coefficient $t$-statistics are reported in brackets below, and are computed on the basis of standard errors which have been adjusted for cross-sectional correlation and time series auto- and cross-correlations using the methodology from Thompson (2011) with eight lags. We report the report the regression adjusted $R^2$, and the adjusted $R^2$ net of the explanatory power of the fixed effects ($[NFE]$). Pricing error reports the RMSE option pricing error measured in units of volatility points. Panel C reports the realized excess returns of the factor mimicking portfolios (Data), estimates of the annualized model-implied risk premium under Specification I (Model), and after parametrically linking global factor loadings to interest rate differentials ($\phi_2$).
Table VIII
Risk-free Yields and CDS Spreads

This table summarizes risk-free yields and CDS spreads prevailing on Oct 14, 2011 we use to price CoCos issued by Lloyds Banking Group (ECN) and Rabobank (SCN). Panel A shows the risk-free yields of GBP, as Lloyds Banking Group’s stock and the ECNs are denominated in GBP. The risk-free yields were extracted from GBP interest rate swap data. Panel B shows the risk-free yields of EUR, as the SCNs are denominated in EUR. The risk-free yields correspond to EUR interest rate swaps.

| Panel A: GBP Risk-free Yields and Lloyds Banking Group CDS Spreads |
|-----------------|---|---|---|---|---|---|---|
| Tenor in years  | 1 | 2 | 3 | 4 | 5 | 7 | 10 |
| Risk-free yield in % | 1.73 | 1.38 | 1.53 | 1.73 | 1.94 | 2.35 | 2.81 |
| CDS spread in bps | 253.2 | 286.4 | 299.3 | 315.8 | 325.9 | 330.2 | 337.5 |

| Panel B: EUR Risk-free Yields and Rabobank CDS Spreads |
|-----------------|---|---|---|---|---|---|---|
| Tenor in years  | 1 | 2 | 3 | 4 | 5 | 7 | 10 |
| Risk-free yield in % | 2.12 | 1.62 | 1.75 | 1.93 | 2.13 | 2.47 | 2.77 |
| CDS spread in bps | 50.7 | 73.7 | 97.6 | 109.2 | 116.3 | 122.4 | 127.1 |
This table summarizes the calibrated parameters for an intensity-based model and a first-passage time model. Panel A summarizes the calibration results for the intensity-based model with CIR jump intensity. We report the value of $\beta$ (coefficient of jump in jump intensity upon conversion), $\kappa$ (mean reversion coefficient), $\mu$ (long-term mean of jump intensity), $\eta$ (volatility) and $\lambda$ (market-implied jump intensity), and $R$ (CoCo-implied recovery rate). Panel B summarizes the calibration results for the first-passage time model with exponential OU accounting ratio. We report the value of $\kappa$ (mean reversion coefficient), $\exp(m)$ (long-term mean of capital ratio), $\eta$ (volatility), $\exp(h)$ (market-implied capital ratio), and $R$ (CoCo-implied recovery rate). The root mean squared errors (RMSE) are reported in relative terms. The root mean squared errors (RMSE) are reported in relative terms.

| Panel A: Parameters of Intensity-Based Model with CIR Jump Intensity |
|-----------------|-----------------|-----------------|
| Parameters      | Lloyds Banking Group | Rabobank        |
| $\beta$        | 0.81             | 1.36            |
| $\kappa$       | 0.950            | 0.264           |
| $\mu$          | 0.130            | 0.040           |
| $\eta$         | 0.00035          | 0.00011         |
| $\lambda_t$    | 0.4715           | 0.1248          |
| RMSE            | 2.14%            | 2.82%           |

| Panel B: Parameters of First-Passage Time Model with Exponential OU Accounting Ratio |
|-----------------|-----------------|-----------------|
| Parameters      | Lloyds Banking Group | Rabobank        |
| $\kappa$       | 1.1804           | 0.2431          |
| $\exp(\mu)$    | 8.28%            | 14.82%          |
| $\eta$         | 0.5129           | 0.4467          |
| $\exp(H_t)$    | 9.47%            | 14.45%          |
| $R$             | 76.0%            | 53.00%          |
| RMSE            | 6.65%            | 4.00%           |
Appendix A

Appendix Tables
Table A.I
Country-Specific Innovation Parameters

This table reports detailed information on the state variable dynamics, \( Y^i_t \), and calibrated CGMY parameters under Specification I for each of the G10 currencies. The first column reports the time-series means (first row) and volatilities (second row) of the country-specific state variable (scaled by a factor of 100). The subsequent columns report the fraction of the country-specific innovation variance due to the non-Gaussian innovation, \( \eta^i_t \), and the calibrated CGMY parameters; for each country, the time-series mean (volatility) is reported in the first (second) row. The jump structures in the countries are assumed to share the same parameters, \( \{ \zeta, \lambda, \gamma \} \), which we report on the line corresponding to AUD, and omit elsewhere. Finally, we report the time series mean and volatility of the conditional skewness (Skewness\(^i\)) and kurtosis (Kurtosis\(^i\)) of the country-specific factor innovation, \( L^i_{t+1} \), induced by variation in the corresponding state variable.

<table>
<thead>
<tr>
<th>State-variable</th>
<th>( Y^i_t \times 100 )</th>
<th>( \eta^i_t )</th>
<th>( \zeta_t )</th>
<th>( \lambda_t )</th>
<th>( \gamma_t )</th>
<th>Skewness(^i)</th>
<th>Kurtosis(^i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUD</td>
<td>0.38</td>
<td>0.52</td>
<td>33.29</td>
<td>79.10</td>
<td>-0.58</td>
<td>-1.13</td>
<td>6.46</td>
</tr>
<tr>
<td></td>
<td>0.31</td>
<td>0.20</td>
<td>17.82</td>
<td>18.38</td>
<td>0.17</td>
<td>1.20</td>
<td>4.07</td>
</tr>
<tr>
<td>CAD</td>
<td>0.47</td>
<td>0.31</td>
<td></td>
<td></td>
<td>-1.03</td>
<td>6.77</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.73</td>
<td>0.36</td>
<td></td>
<td></td>
<td>1.62</td>
<td>4.27</td>
<td></td>
</tr>
<tr>
<td>CHF</td>
<td>0.44</td>
<td>0.55</td>
<td></td>
<td></td>
<td>-1.09</td>
<td>6.92</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.44</td>
<td>0.21</td>
<td></td>
<td></td>
<td>0.91</td>
<td>4.40</td>
<td></td>
</tr>
<tr>
<td>EUR</td>
<td>0.64</td>
<td>0.29</td>
<td></td>
<td></td>
<td>-0.58</td>
<td>5.29</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.52</td>
<td>0.25</td>
<td></td>
<td></td>
<td>0.76</td>
<td>3.37</td>
<td></td>
</tr>
<tr>
<td>GBP</td>
<td>0.49</td>
<td>0.26</td>
<td></td>
<td></td>
<td>-0.56</td>
<td>5.47</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.49</td>
<td>0.18</td>
<td></td>
<td></td>
<td>0.65</td>
<td>3.50</td>
<td></td>
</tr>
<tr>
<td>JPY</td>
<td>0.79</td>
<td>0.52</td>
<td></td>
<td></td>
<td>-0.86</td>
<td>5.33</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>0.27</td>
<td></td>
<td></td>
<td>1.12</td>
<td>3.40</td>
<td></td>
</tr>
<tr>
<td>NOK</td>
<td>0.55</td>
<td>0.27</td>
<td></td>
<td></td>
<td>-0.53</td>
<td>5.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.60</td>
<td>0.21</td>
<td></td>
<td></td>
<td>0.60</td>
<td>2.87</td>
<td></td>
</tr>
<tr>
<td>NZD</td>
<td>0.49</td>
<td>0.39</td>
<td></td>
<td></td>
<td>-0.69</td>
<td>5.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.26</td>
<td>0.20</td>
<td></td>
<td></td>
<td>0.92</td>
<td>3.06</td>
<td></td>
</tr>
<tr>
<td>SEK</td>
<td>1.06</td>
<td>0.13</td>
<td></td>
<td></td>
<td>-0.20</td>
<td>3.96</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>0.28</td>
<td></td>
<td></td>
<td>0.37</td>
<td>1.62</td>
<td></td>
</tr>
<tr>
<td>USD</td>
<td>0.54</td>
<td>0.27</td>
<td></td>
<td></td>
<td>-0.51</td>
<td>4.84</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.30</td>
<td>0.21</td>
<td></td>
<td></td>
<td>0.61</td>
<td>2.86</td>
<td></td>
</tr>
</tbody>
</table>
Table A.II
Comparative Statics of Option Pricing Errors

This table explores the sensitivity of the option fit to various changes in the model parametrization. Results are reported for calibrations using two sets of tests assets: (a) options on X/USD exchange rates and cross-pairs involving currencies with high/low interest rates (HLX + X/USD pairs; 24 pairs); and, (b) options on the full set of G10 cross rates (all pairs; 45 pairs). All calibrations are performed using one-month FX option quotes from January 1999 to June 2012 (T = 3520 days) at five individual option strikes (10δ put, 25δ put, at-the-money, 25δ call, 10δ call). Panel A reports the average level of FX option-implied volatilities across strikes for the currency pairs used in each specification over the full sample, and in 2008, separately. Panel B reports the fit of the baseline calibration, along with the effect of imposing: (a) Gaussian idiosyncratic innovations (η_i = 0); (b) Gaussian global innovations (η_g = 0); (c) Gaussian global and idiosyncratic innovations (η_i = η_g = 0); and, (d) fixing the global pricing kernel innovations at their time-series means (ξ_i = ξ_g = 1/T · Σ_t ξ_i). For each computation, all other model parameters and the time series dynamics of the state variables, Z_t and Y_{i,t}, are left unchanged at their values from the baseline calibration. Panel C reports the corresponding option fitting errors in 2008.

### Panel A: Average Implied Volatilities

<table>
<thead>
<tr>
<th>Specification</th>
<th>All</th>
<th>10δp</th>
<th>25δp</th>
<th>50δ</th>
<th>25δc</th>
<th>10δc</th>
</tr>
</thead>
<tbody>
<tr>
<td>I (HLX + X/USD)</td>
<td>12.06</td>
<td>12.84</td>
<td>12.00</td>
<td>11.54</td>
<td>11.67</td>
<td>12.24</td>
</tr>
<tr>
<td>II (ALL)</td>
<td>11.51</td>
<td>12.15</td>
<td>11.40</td>
<td>11.02</td>
<td>11.20</td>
<td>11.80</td>
</tr>
<tr>
<td>I (HLX + X/USD), 2008 only</td>
<td>16.91</td>
<td>18.34</td>
<td>16.96</td>
<td>16.15</td>
<td>16.15</td>
<td>16.93</td>
</tr>
<tr>
<td>II (ALL), 2008 only</td>
<td>15.78</td>
<td>16.86</td>
<td>15.70</td>
<td>15.08</td>
<td>15.22</td>
<td>16.04</td>
</tr>
</tbody>
</table>

### Panel B: Fitting Errors (RMSE; volatility points)

<table>
<thead>
<tr>
<th>Specification</th>
<th>All</th>
<th>10δp</th>
<th>25δp</th>
<th>50δ</th>
<th>25δc</th>
<th>10δc</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0.83</td>
<td>0.96</td>
<td>0.84</td>
<td>0.79</td>
<td>0.77</td>
<td>0.78</td>
</tr>
<tr>
<td>I (Gaussian L^i)</td>
<td>1.12</td>
<td>1.35</td>
<td>1.03</td>
<td>1.05</td>
<td>1.05</td>
<td>1.13</td>
</tr>
<tr>
<td>I (Gaussian L^g)</td>
<td>1.33</td>
<td>1.73</td>
<td>1.38</td>
<td>1.09</td>
<td>1.14</td>
<td>1.31</td>
</tr>
<tr>
<td>I (all Gaussian)</td>
<td>1.55</td>
<td>2.14</td>
<td>1.48</td>
<td>1.22</td>
<td>1.31</td>
<td>1.62</td>
</tr>
<tr>
<td>I (fixed loadings)</td>
<td>3.92</td>
<td>3.86</td>
<td>4.02</td>
<td>4.01</td>
<td>3.92</td>
<td>3.79</td>
</tr>
<tr>
<td>II</td>
<td>1.08</td>
<td>1.15</td>
<td>1.10</td>
<td>1.04</td>
<td>1.05</td>
<td>1.06</td>
</tr>
<tr>
<td>II (Gaussian L^i)</td>
<td>1.25</td>
<td>1.53</td>
<td>1.15</td>
<td>1.07</td>
<td>1.14</td>
<td>1.34</td>
</tr>
<tr>
<td>II (Gaussian L^g)</td>
<td>1.48</td>
<td>1.75</td>
<td>1.47</td>
<td>1.26</td>
<td>1.35</td>
<td>1.58</td>
</tr>
<tr>
<td>II (all Gaussian)</td>
<td>1.60</td>
<td>2.09</td>
<td>1.48</td>
<td>1.24</td>
<td>1.38</td>
<td>1.79</td>
</tr>
<tr>
<td>II (fixed loadings)</td>
<td>4.12</td>
<td>3.90</td>
<td>4.13</td>
<td>4.23</td>
<td>4.21</td>
<td>4.12</td>
</tr>
</tbody>
</table>

### Panel C: Fitting Errors in 2008 (RMSE; volatility points)

<table>
<thead>
<tr>
<th>Specification</th>
<th>All</th>
<th>10δp</th>
<th>25δp</th>
<th>50δ</th>
<th>25δc</th>
<th>10δc</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>1.07</td>
<td>1.32</td>
<td>1.02</td>
<td>0.91</td>
<td>0.95</td>
<td>1.13</td>
</tr>
<tr>
<td>I (Gaussian L^i)</td>
<td>1.60</td>
<td>1.87</td>
<td>1.44</td>
<td>1.48</td>
<td>1.53</td>
<td>1.66</td>
</tr>
<tr>
<td>I (Gaussian L^g)</td>
<td>2.63</td>
<td>3.46</td>
<td>2.79</td>
<td>2.22</td>
<td>2.22</td>
<td>2.45</td>
</tr>
<tr>
<td>I (all Gaussian)</td>
<td>2.99</td>
<td>4.14</td>
<td>2.95</td>
<td>2.34</td>
<td>2.49</td>
<td>3.01</td>
</tr>
<tr>
<td>I (fixed loadings)</td>
<td>7.71</td>
<td>7.77</td>
<td>8.01</td>
<td>7.90</td>
<td>7.60</td>
<td>7.28</td>
</tr>
<tr>
<td>II</td>
<td>1.46</td>
<td>1.57</td>
<td>1.45</td>
<td>1.37</td>
<td>1.39</td>
<td>1.51</td>
</tr>
<tr>
<td>II (Gaussian L^i)</td>
<td>1.73</td>
<td>2.09</td>
<td>1.60</td>
<td>1.50</td>
<td>1.62</td>
<td>1.86</td>
</tr>
<tr>
<td>II (Gaussian L^g)</td>
<td>2.56</td>
<td>3.09</td>
<td>2.56</td>
<td>2.16</td>
<td>2.32</td>
<td>2.68</td>
</tr>
<tr>
<td>II (all Gaussian)</td>
<td>2.78</td>
<td>3.64</td>
<td>2.64</td>
<td>2.14</td>
<td>2.41</td>
<td>3.06</td>
</tr>
<tr>
<td>II (fixed loadings)</td>
<td>6.92</td>
<td>6.61</td>
<td>6.91</td>
<td>7.13</td>
<td>7.00</td>
<td>6.88</td>
</tr>
</tbody>
</table>
Table A.III
Forecasting Currency Excess Returns with Option-Implied Risk Premia
(Specification II)

The table reports the results of regressions forecasting realized currency excess returns using option-implied currency risk premia and interest rate differentials. The dependent variable is the 21-day U.S. dollar simple excess return on currency pair J/I, and the explanatory variables are observed on the last day preceding the return measurement interval. The data are sampled weekly. We report two sets of results. The first is based on repeated cross-sectional regressions (Fama-MacBeth); the second, is based on a pooled panel regression. For cross-sectional regressions, the reported coefficients are time series averages (N = 700); Newey-West standard errors of the coefficient estimates are reported in parentheses. Adj. $R^2$ reports the average cross-sectional adjusted $R^2$. The second set of regressions is based on pooled panel regressions with currency-pair fixed effects ($N = 45 \cdot 700 = 16,800$). For panel regressions, standard errors are adjusted for cross-sectional correlation and time series auto- and cross-correlations using the methodology from Thompson (2011) with eight lags. The dependent variables are: the U.S. investor’s model-implied risk premium for exposure to the J/I exchange rate ($\lambda^{ji,US}_{HML,t}$ and $\lambda^{ji,US}_{refFX,t}$), and the differential in the one-month yields ($y^{j}_{t+1} - y^{i}_{t+1}$). Finally, we report the $p$-value for two hypothesis tests. Model-implied variables are computed using the calibrated parameter values from Specification II. The first hypothesis asserts that the model is correctly specified, such that the regression intercept is zero and the slope coefficients on the model-implied risk premia are one ($H_0$). If no model-implied variables are included in the regression, we do not report the result of the test. The second hypothesis ($H_1$) is that the included variables have no explanatory power (i.e. all the coefficients with the exception of the intercept are zero). The regressions are based on data from January 1999 to June 2012, unless noted otherwise.

<table>
<thead>
<tr>
<th>Panel A: Cross-Sectional Regressions</th>
<th>Full sample</th>
<th>Excluding 2008</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Intercept (x100)</td>
<td>0.00</td>
<td>-0.04</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>$\lambda^{ji,US}_t$</td>
<td>1.62</td>
<td>1.90</td>
</tr>
<tr>
<td></td>
<td>(0.64)</td>
<td>(0.67)</td>
</tr>
<tr>
<td>$\lambda^{ji,US}_{HML,t}$</td>
<td>1.43</td>
<td>1.72</td>
</tr>
<tr>
<td></td>
<td>(0.68)</td>
<td>(0.71)</td>
</tr>
<tr>
<td>$\lambda^{ji,US}_{refFX,t}$</td>
<td>0.08</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>(10.91)</td>
<td>(11.73)</td>
</tr>
<tr>
<td>$y^{j}<em>{t+1} - y^{i}</em>{t+1}$</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td></td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.23</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>$N$</td>
<td>700</td>
<td>700</td>
</tr>
<tr>
<td>$H_0$ p-value</td>
<td>0.62</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_1$ p-value</td>
<td>0.01</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Panel Regressions</th>
<th>Full sample</th>
<th>Excluding 2008</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Intercept (x100)</td>
<td>-0.03</td>
<td>-0.06</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>$\lambda^{ji,US}_t$</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.44)</td>
<td></td>
</tr>
<tr>
<td>$\lambda^{ji,US}_{HML,t}$</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.43)</td>
<td></td>
</tr>
<tr>
<td>$\lambda^{ji,US}_{refFX,t}$</td>
<td>3.64</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.02)</td>
<td></td>
</tr>
<tr>
<td>$y^{j}<em>{t+1} - y^{i}</em>{t+1}$</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td></td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>$N$</td>
<td>31,500</td>
<td>31,500</td>
</tr>
<tr>
<td>$H_0$ p-value</td>
<td>0.17</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td>$H_1$ p-value</td>
<td>0.74</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>
**Table A.IV**

Empirical $HML_{FX}$ Factor Mimicking Portfolio Risk Premium Decomposition: Alternative specifications

This table uses the calibrated pricing kernel model to compute and decompose the $HML_{FX}$ factor risk premium from the perspective of a U.S. dollar investor. In Panel A, we compute the model-implied $HML_{FX}$ factor risk premium by constructing a hypothetical, dollar-neutral factor mimicking portfolio on the basis of the one-month interest rates, $y_{i,t+1}$. At each point in time, we sort the G10 currencies – excluding the U.S. dollar – into long and short portfolios on the basis of their yields, and weight the currencies within each portfolio on the basis of the absolute deviation of their loading from the average yields of currencies with ranks five and six. In Panel B, the factor mimicking portfolio is formed on the basis of the calibrated time series of global factor loadings, $\xi_i^t$, rather than interest rates. The model risk premia are computed daily and span from January 1999 to June 2012 ($T = 3520$ days).

We report the mean global factor loadings of the long ($\xi_{long}$) and short ($\xi_{short}$) portfolios, which contain the high- and low-interest rate currencies, respectively, as well as the mean portfolio risk premium ($\lambda_{HML}$; % per annum). The portfolio risk premium is then decomposed across the even (symmetric) and odd (asymmetric) cumulants of the global factor innovation, $L^g_{t+1}$, and across its moments (variance, skewness, etc.). For each quantity, we report its time-series mean, volatility, and the 5th and 95th percentiles of its distribution. Results are reported for each of the four specifications reported in Table I.

### Panel A: Alternative Set of Test Assets

<table>
<thead>
<tr>
<th>Specification</th>
<th>Loadsings $\xi_{long}$</th>
<th>$\xi_{short}$</th>
<th>Risk premium $\lambda_{HML}$</th>
<th>By Cumulant Even</th>
<th>Odd</th>
<th>By Moment Variance</th>
<th>Skewness</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>II (All pairs)</td>
<td>Mean</td>
<td>0.93</td>
<td>1.08</td>
<td>3.85</td>
<td>88.20</td>
<td>11.80</td>
<td>82.18</td>
<td>8.39</td>
</tr>
<tr>
<td>Volatility</td>
<td>0.13</td>
<td>0.11</td>
<td>6.66</td>
<td>19.02</td>
<td>19.02</td>
<td>31.04</td>
<td>12.83</td>
<td>20.19</td>
</tr>
<tr>
<td>5%</td>
<td>0.75</td>
<td>0.90</td>
<td>0.04</td>
<td>76.37</td>
<td>2.30</td>
<td>60.93</td>
<td>2.27</td>
<td>0.18</td>
</tr>
<tr>
<td>95%</td>
<td>1.13</td>
<td>1.26</td>
<td>13.37</td>
<td>97.70</td>
<td>23.63</td>
<td>97.52</td>
<td>15.44</td>
<td>26.27</td>
</tr>
</tbody>
</table>

### Panel B: $HML_{FX}$ Factor Replicating Portfolio Formed Using Global Factor Loadings, $\xi_i^t$

<table>
<thead>
<tr>
<th>Specification</th>
<th>Loadsings $\xi_{long}$</th>
<th>$\xi_{short}$</th>
<th>Risk premium $\lambda_{HML}$</th>
<th>By Cumulant Even</th>
<th>Odd</th>
<th>By Moment Variance</th>
<th>Skewness</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>I (HLX + X/USD pairs)</td>
<td>Mean</td>
<td>0.80</td>
<td>1.09</td>
<td>6.41</td>
<td>91.25</td>
<td>8.75</td>
<td>87.79</td>
<td>7.12</td>
</tr>
<tr>
<td>Volatility</td>
<td>0.07</td>
<td>0.08</td>
<td>9.86</td>
<td>6.68</td>
<td>6.68</td>
<td>10.69</td>
<td>4.80</td>
<td>6.67</td>
</tr>
<tr>
<td>5%</td>
<td>0.68</td>
<td>0.98</td>
<td>1.03</td>
<td>78.66</td>
<td>0.69</td>
<td>66.37</td>
<td>0.67</td>
<td>0.04</td>
</tr>
<tr>
<td>95%</td>
<td>0.92</td>
<td>1.23</td>
<td>18.48</td>
<td>99.31</td>
<td>21.35</td>
<td>99.26</td>
<td>15.97</td>
<td>19.50</td>
</tr>
<tr>
<td>II (All pairs)</td>
<td>Mean</td>
<td>0.88</td>
<td>1.14</td>
<td>6.68</td>
<td>87.95</td>
<td>12.05</td>
<td>81.86</td>
<td>8.63</td>
</tr>
<tr>
<td>Volatility</td>
<td>0.10</td>
<td>0.13</td>
<td>9.09</td>
<td>6.58</td>
<td>6.58</td>
<td>11.43</td>
<td>4.16</td>
<td>8.62</td>
</tr>
<tr>
<td>5%</td>
<td>0.71</td>
<td>0.95</td>
<td>1.22</td>
<td>76.50</td>
<td>2.55</td>
<td>61.37</td>
<td>2.48</td>
<td>0.24</td>
</tr>
<tr>
<td>95%</td>
<td>1.03</td>
<td>1.35</td>
<td>18.11</td>
<td>97.45</td>
<td>23.50</td>
<td>97.18</td>
<td>15.08</td>
<td>25.59</td>
</tr>
</tbody>
</table>
Appendix B

Time-changed Lévy Model

Parametrization and Identification

B.1 Model Parametrization

In order to flexibly incorporate time-variation in moments, we model the pricing kernel innovations, $L^g_{t+1}$ and $L^i_{t+1}$, using the time-change framework of Carr and Wu (2004). Unlike in a multiplicative scaling, a time-change affects not only volatility, but also the higher order moments of shocks, enabling the model to better match the empirical features of foreign exchange option data (Bakshi, et al. (2008)). We rely on the cumulant generating functions (CGF) of these time-changed innovations to express key quantities of interest within the model, such as currency risk premia, and option prices. The CGF of a random variable, $\varepsilon_{t+1}$, is given by: $k_{\varepsilon_{t+1}}[u] = \ln E_t[\exp (u \cdot \varepsilon)]$, and provides a complete characterization of its distribution. This section introduces the time-change framework, the details of our shock parametrization, and derives the CGF of the time-changed increments.
B.1.1 Modeling time-varying moments via time-change

To derive the cumulant generating function of a time-changed random variable we rely on Theorem 1 in Carr and Wu (2004), which states that for a generic time change, \( T_{t+1} \), the cumulant generating function of the time-changed Lévy process, \( L_{t+1} \equiv L_{T_{t+1}} \), is given by

\[
k_{T_{t+1}}[u] = k_{\tilde{L}_{t+1}}[u] \cdot \Psi_t \quad \text{(B.1.1a)}
\]

The role of the time-change is therefore to scale up all of the cumulants of the non-time-changed process, \( k_{\tilde{L}_{t+1}}[u] \), affecting all moments of the distribution. For example, the variance, skewness, and kurtosis of a random variable are defined in terms of its cumulants, as follows: \( \mathcal{V} = \kappa_2 \), \( \mathcal{S} = \kappa_3 \cdot (\kappa_2)^{\frac{3}{2}} \), and \( \mathcal{K} = \kappa_4 \cdot (\kappa_2)^{-2} \). Consequently,
the variance, skewness, and kurtosis of the time-changed increments are:

\[
V_{L_{t+1}} = \kappa_{L_{t+1},2} \cdot \Psi_t = V_{\tilde{L}_{t+1}} \cdot \Psi_t \\
S_{L_{t+1}} = \kappa_{L_{t+1},3} \cdot \left( \kappa_{\tilde{L}_{t+1},2} \right)^{-\frac{3}{2}} \cdot \Psi_t^{-\frac{1}{2}} = S_{\tilde{L}_{t+1}} \cdot \Psi_t^{-\frac{1}{2}} \\
K_{L_{t+1}} = \kappa_{L_{t+1},4} \cdot \left( \kappa_{\tilde{L}_{t+1},2} \right)^{-2} \cdot \Psi_t^{-1} = K_{\tilde{L}_{t+1}} \cdot \Psi_t^{-1}
\]

Therefore, the time-change acts both as a channel for inducing time-varying volatility – similar to a stochastic volatility model – but also variation in higher-order moments \((n > 2)\). To facilitate an intuitive interpretation of the state variable, \(\Psi_t\), as a measure of risk, we standardize the non-time-changed increments, such that \(V_{\tilde{L}_{t+1}} = 1\).

**B.1.2 Pricing kernel increments**

To derive the CGF of the time-changed increment, \(L_{t+1}\), we begin by parametrizing their non-time-changed counterpart, \(\tilde{L}_{t+1}\). The variance of the non-time-changed increment is normalized to one, such that the variance of the time-changed increment is controlled by the level of the state variable, \(\Psi_t = \{Z_t, Y^i_t\}\). The increment, \(\tilde{L}_{t+1}\), is modeled as a the sum of two independent components:

\[
\tilde{L}_{t+1} = W_{t+1} + \text{sign} (\eta_t) \cdot X_{t+1}
\]

where \(W_{t+1}\) is a Gaussian innovation with variance \((1 - |\eta_t|)\) and \(X_{t+1}\) is a non-Gaussian innovation with variance \(|\eta_t|\), and \(\eta_t \in [-1, 1]\). In our empirical implementation, the \(X_{t+1}\) shock has a CGMY distribution \((\text{Carr, et al. (2002)})\), which: (a) nests the effects of compound Poisson jumps, infinite-activity jumps with finite variation, as well as, infinite-activity jumps with infinite variation; and, (b) can capture both positively and negatively skewed innovations. The \(\text{sign} (\eta_t)\) factor, along with the asymmetry of \(X_{t+1}\), determines the overall skewness of the \(\tilde{L}_{t+1}\). For example, if
$X_{t+1}$ is negatively skewed, the skewness of the the non-time-changed increment will be negative (positive) if $\eta_t > 0$ ($\eta_t < 0$). We exploit this feature when parameterizing the country-specific innovations by allowing $\eta_t$ to be country-specific, while forcing the parameters governing the distribution of $X_{t+1}$ to be common across-countries.

Given the independence of $W_{t+1}$ and $X_{t+1}$, the CGF of the non-time-changed increment is given by the sum of the CGFs of the two components:

$$k_{L_{t+1}}[u] = k_{W_{t+1}}[u] + k_{X_{t+1}}[\text{sign}(\eta_t) \cdot u]$$ (B.1.7)

The cumulant generating function of the Gaussian innovation is given by:

$$k_{W_{t+1}}[u] = \frac{1 - |\eta_t|}{2} \cdot u^2$$ (B.1.8)

and the corresponding cumulant generating function of the CGMY random variable is given by:

$$k_{X_{t+1}}[u] = \begin{cases} 
C \cdot \Gamma[-Y] \cdot \left((\mathcal{M} - u)^Y - \mathcal{M}^Y + (\mathcal{G} + u)^Y - \mathcal{G}^Y\right) & Y \neq \{0, 1\} \\
- C \cdot \left(\ln \left(1 - \frac{u}{\mathcal{M}}\right) + \ln \left(1 + \frac{u}{\mathcal{G}}\right)\right) & Y = 0 \\
C \cdot \left((\mathcal{M} - u) \cdot \ln \left(1 - \frac{u}{\mathcal{M}}\right) + (\mathcal{G} + u) \cdot \ln \left(1 + \frac{u}{\mathcal{G}}\right)\right) & Y = 1 
\end{cases}$$

where $\Gamma[\cdot]$ is the gamma function. In general, the parameters of the CGMY process, $\{C, \mathcal{G}, \mathcal{M}, Y\}_t$, are allowed to be time-varying our model. The $C^j_t$ parameter is a scaling factor, which is set such that the variance of $X_{t+1}$ is $|\eta^j_t|$; whereas, $\mathcal{G}^j_t$ and $\mathcal{M}^j_t$ determine the exponential dampening of the distribution for negative and positive shocks, respectively. The relative magnitudes of the dampening parameters determine the skewness of the non-Gaussian component.

Putting these results together, the cumulant generating function for the time-changed Lévy increments, $k_{L_{t+1}}[u]$ – for the empirically relevant case when $Y \neq \{0, 1\}$.
is given by:

\[ k_{L_{t+1}}[u] = \left( |\eta_t| \cdot \frac{(M_t - \text{sign} \left( \eta_t \right) \cdot u)^{Y_t} - M_t^{Y_t} - (G_t + \text{sign} \left( \eta_t \right) \cdot u)^{Y_t} - G_t^{Y_t})}{Y_t \cdot (Y_t - 1) \cdot (M_t^{Y_t-2} + G_t^{Y_t-2})} \right) + (1 - |\eta_t|) \cdot \frac{u^2}{2} \cdot \Psi_t \]  

(B.1.9)

Specifically, in the case when \( Y \neq \{0, 1\} \), the second and third cumulants are given by:

\[
\kappa_{X_{t+1},2} = C_t \cdot \Gamma[-Y_t] \cdot Y_t \cdot (Y_t - 1) \cdot (M_t^{Y_t-2} + G_t^{Y_t-2}) = |\eta_t| \]

(B.1.10a)

\[
\kappa_{X_{t+1},3} = C_t \cdot \Gamma[-Y_t] \cdot Y_t \cdot (Y_t - 1) \cdot (Y_t - 2) \cdot (-M_t^{Y_t-3} + G_t^{Y_t-3}) \]

(B.1.10b)

To simplify expressions, in the ensuing derivations we assume \( \eta_t > 0 \); the corresponding expressions for the case when \( \eta_t < 0 \) can be obtained analogously.

**B.2 Distributions under the pricing measure, \( \mathbb{P}^i \)**

In this section, we derive the distribution of the pricing kernel increments, \( \{L^g, L^i, L^j\}_{t+1} \), under the risk-forward measure, \( \mathbb{F}^i \). The risk-forward measure for an investor from country \( I \), is associated with a zero-coupon bond maturing at time \( t + 1 \), and is defined as follows:

\[
\frac{d\mathbb{F}^i}{d\mathbb{P}} = \frac{M^i_{t+1}}{M^i_t} \cdot \exp \left( y^i_{t+1} \cdot \Delta \right) \]

(B.2.11)
where \( \mathbb{P} \) denotes the objective (historical) measure. Recall that the pricing kernel must price the one-period bond, such that we have:

\[
m_{t+1}^i - m_t^i = -\alpha_t^i - \xi_{t+1}^i \cdot L_{t+1}^g - L_t^i
\]

\[
= -y_{t,t+1}^i \cdot \Delta - k_{L_{t+1}^g}^{i} \cdot [-\xi_t^i] - k_{L_{t+1}^g}^{i} [-1] \quad \text{(B.2.12)}
\]

Proceeding from the definition of the cumulant generating function and the measure change, we have:

1. For the global pricing kernel innovation, \( L_{t+1}^g \):

\[
k_{L_{t+1}^g}^{i} [u] = \ln E_t^{p} \left[ \exp \left( u \cdot L_{t+1}^g \right) \right] = \ln E_t^{p} \left[ \exp \left( y_{t,t+1}^i \cdot \Delta + (m_{t+1}^i - m_t^i) + u \cdot L_t^g \right) \right]
\]

\[
= \ln E_t^{p} \left[ \exp \left( -k_{L_{t+1}^g}^{i} \cdot [-\xi_t^i] - k_{L_{t+1}^g}^{i} [-1] + (u - \xi_t^i) \cdot L_{t+1}^g \right) \right]
\]

\[
= k_{L_{t+1}^g}^{i} \cdot [-\xi_t^i] - k_{L_{t+1}^g}^{i} [-1] = \left( k_{L_{t+1}^g}^{i} \cdot [-\xi_t^i] - k_{L_{t+1}^g}^{i} [-1] \right) \cdot Z_t \quad \text{(B.2.13)}
\]

2. For the country-specific innovation in the reference (home) country, \( L_{t+1}^i \):

\[
k_{L_{t+1}^i}^{i} [u] = \ln E_t^{p} \left[ \exp \left( u \cdot L_{t+1}^i \right) \right] = \ln E_t^{p} \left[ \exp \left( y_{t,t+1}^i \cdot \Delta + (m_{t+1}^i - m_t^i) + u \cdot L_t^i \right) \right]
\]

\[
= \ln E_t^{p} \left[ \exp \left( -k_{L_{t+1}^i}^{i} \cdot [-\xi_t^i] - k_{L_{t+1}^i}^{i} [-1] - \xi_t^i \cdot L_{t+1}^i + (u - 1) \cdot L_{t+1}^i \right) \right]
\]

\[
= k_{L_{t+1}^i}^{i} \cdot [-\xi_t^i] - k_{L_{t+1}^i}^{i} [-1] = \left( k_{L_{t+1}^i}^{i} \cdot [-\xi_t^i] - k_{L_{t+1}^i}^{i} [-1] \right) \cdot Y_t^i \quad \text{(B.2.14)}
\]

3. For the country-specific innovation in the foreign country, \( L_{t+1}^j \):

\[
k_{L_{t+1}^j}^{i} [u] = \ln E_t^{p} \left[ \exp \left( u \cdot L_{t+1}^j \right) \right] = \ln E_t^{p} \left[ \exp \left( y_{t,t+1}^j \cdot \Delta + (m_{t+1}^j - m_t^j) + u \cdot L_t^j \right) \right]
\]

\[
= \ln E_t^{p} \left[ \exp \left( -k_{L_{t+1}^j}^{i} \cdot [-\xi_t^j] - k_{L_{t+1}^j}^{i} [-1] - \xi_t^j \cdot L_{t+1}^j - L_{t+1}^i + u \cdot L_t^j \right) \right]
\]

\[
= k_{L_{t+1}^j}^{i} [u] = k_{L_{t+1}^j}^{i} [u] \cdot Y_t^j \quad \text{(B.2.15)}
\]
Intuitively, since the foreign country-specific innovation is independent of the innovations in the pricing kernel of investor’s from country $I$, the measure change leaves its distribution unaltered. Consequently, the CGF is given by (B.1.9).

B.3 Model Identification

This section discusses the identification of the model parameters using cross-sectional asset price data. We begin by highlighting the role of cross-rate options in pinning down the currency return factor structure under the pricing measure, and then detail the procedure by which we recover $HML_{FX}$ and short reference risk premia.

B.3.1 The role of cross-rate options

Cross-rate options carry information about the correlation structure of currency returns, and play an important role in the identification of the global and country-specific components of the exchange rate distribution under the pricing measure. To illustrate this, we consider the situation where one only has access to options on exchange rates quoted against a single reference currency, $I$ (e.g. exchange rates measured relative to the USD), and demonstrate that these data alone cannot reject a model whereby there is no global component in currency returns.

First, recall that the log currency return under the baseline model specification is given by:

$$s_{ji}^{i+1} - s_{ji}^i = -\alpha_j^i + \alpha_i^i - (\xi_t^j - \xi_t^i) \cdot L_{t+1}^j - L_{t+1}^j + L_{t+1}^i$$  \hfill (B.3.16)
Since each of the innovations driving the exchange rate is independent in our model, the CGF of the log exchange rate under the pricing measure is:

$$k^{F_i}_{L_{t+1}}[u] = (s_{t+i}^j - \alpha_{t+i}^j + \alpha_{t+i}^j) \cdot u + k^{F_i}_{L_{t+1}}[(\xi_{t+i}^j - \xi_{t+i}^j) \cdot u] + k^{F_i}_{L_{t+1}}[u] + k^{F_i}_{L_{t+1}}[-u]$$

$$= (s_{t+i}^j - \alpha_{t+i}^j + \alpha_{t+i}^j) \cdot u + \left( k^{F_i}_{L_{t+1}}[(\xi_{t+i}^j - \xi_{t+i}^j) \cdot u - \xi_{t+i}^j] - k^{F_i}_{L_{t+1}}[-\xi_{t+i}^j] \right)$$

$$+ \left( k^{F_i}_{L_{t+1}}[u - 1] - k^{F_i}_{L_{t+1}}[-1] \right) + k^{F_i}_{L_{t+1}}[-u] \quad (B.3.17)$$

This CGF uniquely pins down the model-implied exchange rate distribution of $s_{t+i}$ under the $F_i$-measure, and therefore exchange rate option prices. Any alternative model specification delivering the same CGF, cannot be distinguished using option prices alone. For example, consider an alternative model which has no global component, such that the global ($HML_{FX}$) risk premium is counterfactually absent:

$$s_{t+i}^j - s_{t+i}^j = -\alpha_{t+i}^j + \alpha_{t+i}^j - \hat{L}_{t+1}^j + L_{t+1}^j \quad (B.3.18)$$

The shocks, $\hat{L}_{t+1}^j$, are: (a) assumed to be independent with respect to $L_{t+1}^j$; and, (b) equal in distribution to $(\xi_{t+i}^j - \xi_{t+i}^j) \cdot L_{t+1}^j - L_{t+1}^j$, such that their CGF is given by:

$$k^{F_i}_{L_{t+1}}[u] = k^{F_i}_{L_{t+1}}[(\xi_{t+i}^j - \xi_{t+i}^j) \cdot u] + k^{F_i}_{L_{t+1}}[-u]$$

$$= \left( k^{F_i}_{L_{t+1}}[(\xi_{t+i}^j - \xi_{t+i}^j) \cdot u - \xi_{t+i}^j] - k^{F_i}_{L_{t+1}}[-\xi_{t+i}^j] \right) + k^{F_i}_{L_{t+1}}[-u] \quad (B.3.19)$$

Since the shock, $\hat{L}_{t+1}^j$, is idiosyncratic with respect to investor $I$’s pricing kernel, we additionally have: $k^{F_i}_{L_{t+1}}[u] = k^{F_i}_{L_{t+1}}[u]$. Substituting, $(B.3.18)$, into the definition of the cumulant generating function, one immediately recovers $(B.3.17)$, such that the alternative model produces the same prices for options on $J/I$ exchange rates as the baseline specification. Consequently, option data on exchange rates measured relative to a single currency are insufficient to identify the global risk factor.
While these two models generate identical option prices for options on J/I exchange rates, their implications for cross-rate prices are distinct. To see this formally, consider the CGF of the exchange rate J/K under the alternative model:

\[
\hat{k}^{i^k}_{s^k_{j^k+1}}[u] = \left( s^j_k - \alpha^j_t + \alpha^k_t \right) \cdot u + k^{\hat{L}^q_k}_{\hat{L}^q_{j^k+1}}[u] + k^{\hat{L}^q_i}_{\hat{L}^q_{i^k+1}}[-u] \\
= \left( s^j_k - \alpha^j_t + \alpha^k_t \right) \cdot u + \left( k^{\hat{L}^q_k}_{\hat{L}^q_{j^k+1}}[u - 1] - k^{\hat{L}^q_i}_{\hat{L}^q_{i^k+1}}[-1] \right) + k^{\hat{L}^q_i}_{\hat{L}^q_{i^k+1}}[-u] \\
= \left( s^j_k - \alpha^j_t + \alpha^k_t \right) \cdot u + \left( k^{\hat{L}^q_i}_{\hat{L}^q_{i^k+1}}\left[ (\xi^i_t - \xi^k_t) \cdot (u - 1) - \xi^i_t \right] - k^{\hat{L}^q_i}_{\hat{L}^q_{i^k+1}}[-\xi^i_t] \right) \\
+ k^{\hat{L}^q_i}_{\hat{L}^q_{i^k+1}}[-(u - 1)] - \left( k^{\hat{L}^q_i}_{\hat{L}^q_{i^k+1}}\left[ (\xi^i_t - \xi^k_t) - \xi^i_t \right] - k^{\hat{L}^q_i}_{\hat{L}^q_{i^k+1}}[-\xi^i_t] \right) - k^{\hat{L}^q_i}_{\hat{L}^q_{i^k+1}}[1] \\
+ \left( k^{\hat{L}^q_i}_{\hat{L}^q_{i^k+1}}\left[ -(\xi^i_t - \xi^i_t) \cdot u - \xi^i_t \right] - k^{\hat{L}^q_i}_{\hat{L}^q_{i^k+1}}[-\xi^i_t] \right) + k^{\hat{L}^q_i}_{\hat{L}^q_{i^k+1}}[u] \tag{B.3.20}
\]

where the second transition follows from the properties of the change of measure for a country-specific shock (Appendix B), and the final transition follows from the definition of the CGF of the alternative country-specific shocks, \( \hat{L}^j_{i^k+1} \). For comparison, the CGF of \( s^j_{i^k+1} \) under the baseline model specification is:

\[
k^{\hat{L}^q_i}_{s^j_{i^k+1}}[u] = \left( s^j_k - \alpha^j_t + \alpha^k_t \right) \cdot u + \left( k^{\hat{L}^q_i}_{\hat{L}^q_{i^k+1}}\left[ (\xi^i_t - \xi^k_t) \cdot u - \xi^i_t \right] - k^{\hat{L}^q_i}_{\hat{L}^q_{i^k+1}}[-\xi^i_t] \right) \\
+ \left( k^{\hat{L}^q_i}_{\hat{L}^q_{i^k+1}}[u - 1] - k^{\hat{L}^q_i}_{\hat{L}^q_{i^k+1}}[-1] \right) + k^{\hat{L}^q_i}_{\hat{L}^q_{i^k+1}}[-u] \tag{B.3.21}
\]

which is clearly distinct from the expression for the alternative model in that it is independent of \( \xi^i_t \). Consequently, our calibration can be understood as relying on cross-rate options to establish the role of the global pricing kernel component in the option-implied exchange rate distributions.

**B.3.2 Recovering risk premia**

In order to recover currency risk premia we need to identify the global factor loadings, \( \xi^i_t \), the parameters governing the distribution of the global pricing kernel innovation, \( L^q_{i^k+1} \), and the parameters governing the distributions of the country-specific
pricing kernel innovations, $L_{t+1}^i$. In the case of the $HML_{FX}$ risk premium, this requires incorporating information from the objective measure (e.g. forecasts of conditional variance), analogous to the approach adopted for equity risk premia (Pan (2002), Santa-Clara and Yan (2010), and Andersen, et al. (2013)). Interestingly, we demonstrate that the short reference component of the risk premium within our model can be identified using only exchange rate option data. This result can be traced to our ability to extract both the objective and risk-forward distributions of $L_{t+1}^i$ from options on $I/K$ and $J/I$ exchange rates, respectively. The first set of options – where currency $I$ is the investment (long) currency – is priced under measure $\mathbb{F}^k$, such that the idiosyncratic innovation in country $I$ is unpriced, thus revealing its objective distribution. The second set of options – where currency $I$ is the funding (short) currency – is priced under measure $\mathbb{F}^i$ revealing the risk-forward distribution of $L_{t+1}^i$. We show that the CGF of the innovation under either measure is sufficient to pin down the short reference risk premium for currency $I$.

The global factor risk premium

Exchange rate option data alone are insufficient to separably pin down the global pricing kernel loadings and the distribution of the global innovation. To see this, define the global component of the terminal exchange rate distribution $s_{t+1}^{ji}$ as follows: $g_{t+1}^{ji} = (\xi_t^j - \xi_t^i) \cdot L_{t+1}^g$. We demonstrate the existence of a parameter re-scaling which leaves the global components of exchange rate unaltered, and therefore has no effect on exchange rate option prices, while shifting the model-implied risk premia. Importantly, each parameter set also has distinct implications for conditional exchange rate moments under the objective measure, yielding a simple approach to resolving the non-identification present when only option data are available.

The CGF of the global component of the exchange rate distribution, $s_{t+1}^{ji}$, under the pricing measure $\mathbb{F}^i$ is given by $k_{t+1}^{ji} \left[ (\xi_t^i - \xi_t^j) \cdot u \right]$, and has the following cumu-
where the expressions for the cumulants, $\kappa_{L^g_{t+1},n}$, are obtained by taking successive derivatives of the global CGF (Appendix A) with respect to $u$, and evaluating at zero:

\[
\kappa_{L^g_{t+1},2} = \left( (1 - \eta^g_t) + \eta^g_t \cdot \left( \frac{G^q_t - \xi^i_t}{G^q_t} \right)^{\gamma^q_t-2} \right) \cdot Z_t \tag{B.3.23a}
\]

\[
\kappa_{L^g_{t+1},n} = \eta^g_t \cdot \left( \prod_{i=3}^{n} (\gamma^q_t - i + 1) \right) \cdot \frac{(G^q_t - \xi^i_t)^{\gamma^g_t-n}}{(G^q_t)^{\gamma^q_t-2}} \cdot Z_t \quad \text{for} \quad n \geq 3 \tag{B.3.23b}
\]

We show that all the cumulants of the global component remain unaffected (i.e. the distribution of the global component is unchanged), when the parameters are shifted as follows:

\[
\hat{Z}_t(\delta) = \left( (1 - \eta^g_t) + \eta^g_t \cdot \left( \frac{G^q_t + \delta}{G^q_t} \right)^{\gamma^q_t-2} \right) \cdot (1 + \delta)^2 \cdot Z_t \tag{B.3.24a}
\]

\[
\hat{\eta}_t(\delta) = \frac{\eta^g_t \cdot \left( \frac{G^q_t + \delta}{G^q_t} \right)^{\gamma^q_t-2}}{(1 - \eta^g_t) + \eta^g_t \cdot \left( \frac{G^q_t + \delta}{G^q_t} \right)^{\gamma^q_t-2}} \tag{B.3.24b}
\]

\[
\hat{G}^q_t(\delta) = \frac{G^q_t + \delta}{1 + \delta} \tag{B.3.24c}
\]

\[
\hat{\xi}^i_t(\delta) = \frac{\xi^i_t + \delta}{1 + \delta} \tag{B.3.24d}
\]

where $\delta$ is an arbitrary shift parameter ($\delta > -\min_i \xi^i_t$). This parametrization preserves: (a) the interpretation of $\hat{\eta}_t(\delta)$ as the share of variance due to the CGMY component; and, (b) the normalization that the U.S. loading is fixed at one. We show that the risk-forward cumulants of the baseline variable specification, $L^g_{t+1}(Z_t, \eta^g_t, G^q_t, \xi^i_t)$,
are identical to the cumulants of the shifted shocks structure, \( \hat{L}_t^g \left( \hat{Z}_t, \hat{\eta}_t^g, \hat{G}_t^g, \hat{\xi}_t^i \right) \), and therefore, produce identical exchange rate option prices. Note that \( \delta \) itself could be a function of time, though we suppress its time subscript for parsimony.

First, consider the second cumulant of, \( \hat{g}_{t+1}^{ji} = \left( \hat{\xi}_t^i - \hat{\xi}_t^i \right) \cdot \hat{L}_t^g \), under the risk-forward measure:

\[
k_{g_{t+1},2}^{F_i} = \left( \hat{\xi}_t^i - \hat{\xi}_t^i \right)^2 \cdot \left( 1 - \hat{\eta}_t \right) \cdot \hat{\eta}_t \cdot \left( \frac{\hat{G}_t^g - \hat{\xi}_t^i}{\hat{G}_t^g} \right)^{Y_t^g - 2} \cdot \hat{Z}_t(\delta)
\]

\[
= \left( \hat{\xi}_t^i - \hat{\xi}_t^i \right)^2 \cdot \left( 1 - \hat{\eta}_t \right) \cdot \hat{\eta}_t \cdot \left( \frac{G_t^g - \xi_t^i}{G_t^g + \delta} \right)^{Y_t^g - 2} \cdot \left( 1 - \eta_t^g \right) + \hat{\eta}_t \cdot \left( \frac{G_t^g}{G_t^g + \delta} \right)^{Y_t^g - 2} \cdot Z_t
\]

\[
= \left( \hat{\xi}_t^i - \hat{\xi}_t^i \right)^2 \cdot \left( 1 - \hat{\eta}_t \right) + \eta_t \cdot \left( \frac{G_t^g - \xi_t^i}{G_t^g} \right)^{Y_t^g - 2} \cdot Z_t \equiv k_{g_{t+1},2}^{F_i} \quad \text{(B.3.25)}
\]

Proceeding similarly, for the higher-order \( n \geq 3 \) cumulants under the risk-forward measure we have:

\[
k_{g_{t+1},n}^{F_i} = \left( \hat{\xi}_t^i - \hat{\xi}_t^i \right)^n \cdot \hat{\eta}_t^n \cdot \left( \prod_{m=3}^{n} (\mathcal{Y}_t^g - m + 1) \right) \cdot \left( \frac{\hat{G}_t^g - \xi_t^i}{\hat{G}_t^g} \right)^{Y_t^g - n} \cdot \hat{Z}_t
\]

\[
= \left( \hat{\xi}_t^i - \hat{\xi}_t^i \right)^n \cdot \left( 1 + \delta \right)^{-n} \cdot \hat{\eta}_t^n \cdot \left( \prod_{m=3}^{n} (\mathcal{Y}_t^g - m + 1) \right) \cdot (1 + \delta)^{-n - 2} \cdot \left( \frac{G_t^g - \xi_t^i}{G_t^g + \delta} \right)^{Y_t^g - n} \cdot \hat{Z}_t
\]

\[
= \left( \hat{\xi}_t^i - \hat{\xi}_t^i \right)^n \cdot \eta_t^n \cdot \left( \prod_{m=3}^{n} (\mathcal{Y}_t^g - m + 1) \right) \cdot \left( \frac{G_t^g - \xi_t^i}{G_t^g} \right)^{Y_t^g - n} \cdot \hat{Z}_t \equiv k_{g_{t+1},n}^{F_i} \quad \text{(B.3.26)}
\]

Since the two parameterizations produce identical risk-forward cumulants independent of the choice of \( \delta \), they generate identical risk-forward distributions, and therefore, exchange rate option prices. Consequently, while cross-rate options pin down the contribution of the global component, \( \hat{g}_{t+1}^{ji} \), to the option-implied exchange rate distribution, they do not permit the separable identification of the global factor loadings, \( \xi_t^i \), and distributional parameters of the global innovation, \( L_t^g \).
To illustrate the effect of the parameter shift on risk premia, consider the $HML_{FX}$ component of the currency risk premium for exchange rate $J/K$ demanded by an investor in country $I$:

$$\lambda_{t}^{j,k,i} = k_{L_{t+1}}^{g} [\xi_{t}^{i} - \xi_{t}^{j}] - k_{L_{t+1}}^{g} [\xi_{t}^{i} - \xi_{t}^{k}] + k_{L_{t+1}}^{g} [-\xi_{t}^{i}] - k_{L_{t+1}}^{g} [-\xi_{t}^{j}]$$  \hspace{1cm} (B.3.27)

The risk-forward CGF of the global innovation, $k_{L_{t+1}}^{F_{i}} [u]$, is related to its $\mathbb{P}$-measure counterpart (Appendix B), as follows:

$$k_{L_{t+1}}^{F_{i}} [u] = k_{L_{t+1}}^{g} [u - \xi_{t}^{i}] - k_{L_{t+1}}^{g} [-\xi_{t}^{i}]$$  \hspace{1cm} (B.3.28)

Equivalently, we can write:

$$k_{L_{t+1}}^{g} [u] = k_{L_{t+1}}^{F_{i}} [u + \xi_{t}^{i}] + k_{L_{t+1}}^{g} [-\xi_{t}^{i}]$$  \hspace{1cm} (B.3.29)

We also have: $k_{L_{t+1}}^{F_{i}} [\xi_{t}^{i}] = -k_{L_{t+1}}^{g} [-\xi_{t}^{i}]$, since – by definition – any CGF evaluated at zero is equal to zero. Hence the $\mathbb{P}$-measure CGF is given by:

$$k_{L_{t+1}}^{g} [u] = k_{L_{t+1}}^{F_{i}} [u + \xi_{t}^{i}] - k_{L_{t+1}}^{F_{i}} [\xi_{t}^{i}]$$  \hspace{1cm} (B.3.30)

Re-expressing the global risk premium in terms of the $F_{i}$-measure CGF we obtain:

$$\lambda_{t}^{j,k,i} = k_{L_{t+1}}^{F_{i}} [2 \cdot \xi_{t}^{i} - \xi_{t}^{j}] - k_{L_{t+1}}^{F_{i}} [2 \cdot \xi_{t}^{i} - \xi_{t}^{k}] + k_{L_{t+1}}^{F_{i}} [\xi_{t}^{i} - \xi_{t}^{k}] - k_{L_{t+1}}^{F_{i}} [\xi_{t}^{i} - \xi_{t}^{j}]$$  \hspace{1cm} (B.3.31)

However, exchange rate option data only allow us to pin down the CGF of the global component of the currency return distribution, $g_{t+1}^{j,k} = (\xi_{t}^{i} - \xi_{t}^{k}) \cdot L_{t+1}^{g}$, but not the CGF of the global shock itself. If we re-write the risk premium expression in terms of the CGFs of the global components of the currency return distribution for the pairs
\[
\lambda_{jk,i}^t = k_{\xi^i}^\pi \left( \xi^i_t - \xi^j_t \right) \cdot L_{t+1}^g \left[ \frac{2 \cdot \xi^i_t - \xi^j_t}{\xi^j_t - \xi^i_t} \right] - k_{\xi^k}^\pi \left( \xi^k_t - \xi^j_t \right) \cdot L_{t+1}^g \left[ \frac{2 \cdot \xi^k_t - \xi^j_t}{\xi^j_t - \xi^k_t} \right] \\
+ k_{\xi^k}^{\pi} \left( \xi^k_t - \xi^i_t \right) \cdot L_{t+1}^g [-1] - k_{\xi^k}^{\pi} \left( \xi^k_t - \xi^i_t \right) \cdot L_{t+1}^g [-1] \\
= k_{\xi^j}^{\pi} \left[ \frac{2 \cdot \xi^i_t - \xi^j_t}{\xi^j_t - \xi^i_t} \right] - k_{\xi^k}^{\pi} \left[ \frac{2 \cdot \xi^k_t - \xi^j_t}{\xi^j_t - \xi^k_t} \right] \\
+ k_{\xi^i}^{\pi} [-1] - k_{\xi^j}^{\pi} [-1] \\
\] (B.3.32)

Although the last two terms can be obtained from cross-sectional option data alone, the first two cannot, since they additionally depend on the magnitudes of the global factor loadings. Equivalently, the parameter shift \( \delta \) affects risk premia:

\[
\hat{\lambda}_{jk,i}^t (\delta) = k_{\xi^i}^{\pi} \left[ \frac{2 \cdot \hat{\xi}^i_t - \hat{\xi}^j_t}{\hat{\xi}^j_t - \hat{\xi}^i_t} \right] - k_{\xi^k}^{\pi} \left[ \frac{2 \cdot \hat{\xi}^k_t - \hat{\xi}^j_t}{\hat{\xi}^j_t - \hat{\xi}^k_t} \right] \\
+ k_{\xi^i}^{\pi} [-1] - k_{\xi^j}^{\pi} [-1] \\
= k_{\xi^j}^{\pi} \left[ \frac{2 \cdot \xi^i_t - \xi^j_t + \delta}{\xi^j_t - \xi^i_t} \right] - k_{\xi^k}^{\pi} \left[ \frac{2 \cdot \xi^k_t - \xi^j_t + \delta}{\xi^j_t - \xi^k_t} \right] \\
+ k_{\xi^i}^{\pi} [-1] - k_{\xi^j}^{\pi} [-1] \\
\] (B.3.33)

In order to pin down \( \delta \), and therefore the \( HML_{FX} \) risk premium for individual currency pairs, we exploit the fact that \( \delta \) is reflected in the conditional moments of the log currency returns under the objective measure. In particular, the conditional variance of the log currency return is given by:

\[
Var^\pi_t \left[ s_{t+1}^{ji} - s_t^{ji} \right] = \left( \hat{\xi}^i_t - \hat{\xi}^j_t \right)^2 \cdot \hat{Z}_t + Y^i_t + Y^j_t \\
= \left( \xi^i_t - \xi^j_t \right)^2 \cdot \left( \left( 1 - \eta^g_t \right) + \eta^g_t \cdot \left( \frac{G^g_t + \delta}{G^g_t} \right)^{y^g_t - 2} \right) \cdot \hat{Z}_t \\
+ Y^i_t + Y^j_t \\
\] (B.3.33)
Consequently, to pin down the value of $\delta$ and therefore the risk premia, we augment cross-sectional exchange rate option data with empirical variance forecasts (e.g. GARCH, realized variance).

**The country-specific risk premium**

The country-specific component of the currency risk premium for pair $J/I$ is determined by the $\mathbb{P}$-measure cumulant generating function of the country-specific shock, $L_{t+1}^i$, evaluated at $u = 1$ and $u = -1$:

$$
\lambda_{\text{refFX},t}^{j,i} = k_{L_{t+1}^i}[1] + k_{L_{t+1}^i}[-1]
$$

(B.3.34)

This CGF can be identified from options on exchange rates where currency I is the investment (long) currency, e.g. $I/K$. Recall that the payoffs of these options are priced under measure $\mathbb{F}^k$, which assigns a zero price of risk to innovations $L_{t+1}^i$. As a result, the $\mathbb{F}^k$-measure distribution of the shock is equivalent to the $\mathbb{P}$-measure distribution, which determines the short reference risk premium for currency $I$.

Alternatively, we can show that risk premium expression can be written in terms of the $\mathbb{F}^i$ measure CGF of the country-specific shock, $k_{L_{t+1}^i}^\mathbb{F}^i[u]$, which can be obtained from options on exchange rates where currency I is the funding (short) currency, e.g. $J/I$. Recall that the risk-forward and objective CGFs (Appendix B) are related via:

$$
k_{L_{t+1}^i}^\mathbb{F}^i[u] = k_{L_{t+1}^i}[u] - k_{L_{t+1}^i}[u - 1] - k_{L_{t+1}^i}[-1]
$$

(B.3.35)

Equivalently:

$$
k_{L_{t+1}^i}[u] = k_{L_{t+1}^i}^\mathbb{F}^i[u + 1] + k_{L_{t+1}^i}[-1]
$$

(B.3.36)
Moreover, note that – by definition of the cumulant generating function – \( k_{L_{i+1}} \{0\} = 0 \), such that \( k_{L_{i+1}} \{1\} = -k_{L_{i+1}} \{-1\} \). This allows us to express the \( \mathbb{P} \)-measure CGF entirely in terms of the \( \mathbb{F} \)-measure CGF, \( k_{L_{i+1}} \{u\} \):

\[
  k_{L_{i+1}} \{u\} = k_{L_{i+1}} \{u + 1\} - k_{L_{i+1}} \{1\} \quad (B.3.37)
\]

Substituting this result into the definition of the short reference risk premium, \( \lambda_{\text{refFX},t} \):

\[
  \lambda_{\text{refFX},t} = k_{L_{i+1}} \{2\} + k_{L_{i+1}} \{0\} - 2 \cdot k_{L_{i+1}} \{1\} = k_{L_{i+1}} \{2\} - 2 \cdot k_{L_{i+1}} \{1\} \quad (B.3.38)
\]

Since the risk-forward cumulant generating function, \( k_{L_{i+1}} \{u\} \), is pinned down exclusively from option data, so is the short-reference risk premium. In summary, cross-rate options contain sufficient information to identify the distribution of each country-specific innovation, \( L_{i+1} \), both under the objective measure, \( \mathbb{P} \), as well as, the pricing measure, \( \mathbb{F} \).
Appendix C

CoCos with stock price triggers

Market triggers mostly exist in the academic literature. Their advantage is their high transparency. But they have the drawback that they are prone to manipulation. Here we shortly discuss how they fit into the general framework developed in this paper. Let us consider a CoCo which converts into equity if the stock price \( S_t \) breaches the barrier \( S_\star < S_0 \). That is, the conversion time is of the form

\[
\tau = \inf \{ t \geq 0 : S_t \leq S_\star \}.
\]

In the benchmark case where the interest rate is equal to a constant \( r \) and the stock price evolves like

\[
S_t = S_0 \exp \left\{ \left( r - q - \frac{1}{2} \sigma^2 \right) t + \sigma W_t \right\}
\]

for a constant volatility \( \sigma \) and a \( \mathbb{Q} \)-Brownian motion \( W \), expressions for the quantities (4.1.13) can be derived from well-known results on first-passage times of Brownian motion. Conditioned on \( t < \tau \), one has

\[
\tau = \inf \{ s \geq t : W_s = -c + ds \}
\]
for 
\[ c = \frac{\log S_t - \log S_\tau}{\sigma} \quad \text{and} \quad d = \frac{q + \sigma^2/2 - r}{\sigma}. \]

So the distribution of \( \tau \) under \( Q_t \) is given by 
\[ Q_t[\tau \leq s] = \int_s^\tau f(v)dv \]

where 
\[ f(v) = \frac{c}{\sqrt{2\pi v^3}} \exp \left( -\frac{1}{2v} (c - dv)^2 \right); \]

see for instance, Steele (2001). With this in hand, it is easy to calculate the quantities 
\[ Q_t[\tau > t_i], \quad E^Q_t [P(t, \tau)1_{\{\tau \leq t_i\}}] \quad \text{and} \quad E^Q_t [P(t, \tau)\tau 1_{\{\tau \leq t_i\}}]. \]

Under the distorted measure \( Q^* = \tilde{S}_T/S_0 \cdot Q = \exp(\sigma W_T - \sigma^2 T/2) \cdot Q \), \( W_t^* = W_t - \sigma t \) is a Brownian motion. Therefore, one has 
\[ Q_t^*[\tau \leq s] = \int_s^\tau f^*(v)dv \]

where 
\[ f^*(v) = \frac{c}{\sqrt{2\pi v^3}} \exp \left( -\frac{1}{2v} (c - d^*u)^2 \right), \quad \text{where} \quad d^* = d - \sigma. \]

This allows to compute the expectation 
\[ E^Q_t^*[e^{-q\tau}1_{\{\tau \leq T\}}]. \]

If \((r_t)\) and \((S_t)\) follow more general diffusion dynamics, CoCos with stock price triggers can be priced by solving PDEs like in Section 4.3. Calibration and hedging also works like in Section 4.3. Only now equity shares and options are more closely related to the trigger event than CDS’s. However, equity options often only exist with short maturities and strikes around the money, while CDS contracts are traded with longer maturities; compare to the discussion in Corcuera et al. (2012).
Appendix D

Proofs and Derivations for CoCo Design

D.1 Hitting time for geometric Brownian Motions

If we define the hitting time \( \tau \) to be the hitting time of a level \( l \) for a geometric Brownian Motion \( V_t \) starting from \( V_0 = v \geq l \) with dynamic

\[
\frac{dV_t}{V_t} = \mu dt + \sigma dW_t,
\]

then the Laplace transform of \( \tau \) is

\[
L^\tau (u) = \mathbb{E} \left[ \exp(u \tau) \right] = \left( \frac{l}{v} \right)^{\frac{m+\sqrt{m^2-2uv^2}}{\sigma^2}}
\]

with \( m := \mu - \frac{1}{2} \sigma^2 \). Furthermore, the probability of \( \tau < \infty \) is \( \left( \frac{l}{v} \right)^{\frac{2m}{\sigma^2}} \). Another quantity of interest is \( \mathbb{E} \left[ \int_0^\tau e^{-rs} V_s ds \right] \), which can be computed by applying Itô’s formula to
\[ e^{-rt}V_t \] as follow:

\[
\mathbb{E}\left[ e^{-rt}V_t \right] - v = \mathbb{E}\left[ \int_0^T (\mu - r)e^{-rt}V_t dt \right] + \mathbb{E}\left[ \int_0^T e^{-rt}V_t \sigma dW_t \right]
\]

\[
= (\mu - r)\mathbb{E}\left[ \int_0^T e^{-rt}V_t dt \right]
\]

Since \( V_T = l \),

\[
\mathbb{E}\left[ \int_0^T e^{-rt}V_t dt \right] = \frac{v - L^r(-r)l}{r - \mu}.
\]

### D.2 Proofs

#### Proof of Proposition [6.2.1]

The expression for equity value with constant volatility [6.2.10] is an immediate consequence of the results in Appendix [D.1]. The value function’s convexity is determined by the sign of \((1 - \theta)\mathcal{V}(V_c, b + c + d) - E^{ac}(V_c)\). Assume this quantity to be positive, the value function is concave, and simplifying the condition leads to

\[
(1 - \theta)\mathcal{V}(V_c, b + c + d) - E^{ac}(V_c)
\]

\[
= \frac{1 - \theta}{1 + n} \cdot \left[ \mathcal{V}(V_b, b + d)\mathcal{D}_{\gamma(\sigma^{ac})}(V_c, V_b) - \left( \frac{(1 + n)c}{r} - n\mathcal{V}(V_c, b + d) \right) \right].
\]

Hence, the value function is concave if and only if

\[
\mathcal{V}(V_b, b + d) > \mathcal{D}_{\gamma(\sigma^{ac})}(V_b, V_c) \left( \frac{(1 + n)c}{r} - n\mathcal{V}(V_c, b + d) \right).
\]

If \( n \) is chosen such that the conversion is at par,

\[
\frac{c}{r} = nE^{ac}(V_c).
\]
Similarly, the previous condition becomes \( \mathcal{V}(V_b, b + d) > 0 \), which reduces to exactly
the same condition as a bank without CoCos.

\[ \square \]

**Proof of Theorem 6.2.3**

For the deposits, since they are guaranteed, the value is equal to the discounted future
coupon payments. The result is thus straightforward.

For the CoCos, the value equals to the sum of 1) expectation of discounted coupon
payments \( c \) until conversion 2) expected discounted value of equity received upon con-
version. Both terms can be easily computed using the formula for Laplace transform
of hitting time in \([D.1]\).

For the debts, we need to calculate the Laplace transform of \( \tau_b \):

\[
\mathbb{E}_Q^{V_0 = v}[e^{ut_b}] = \mathbb{E}_Q^{V_0 = v} [\mathbb{E}_Q^{V_0 = V_c} [e^{u\tau_c} e^{u(\tau_b - \tau_c)} | \mathcal{F}_{\tau_c}]]
\]

\[
= \mathbb{E}_Q^{V_0 = v}[e^{u\tau_c}] \mathbb{E}_Q^{V_0 = V_c}[e^{u\tau_b}] = \left( \frac{V}{V_c} \right)^{-\gamma(\sigma^b c)} \left( \frac{V_c}{V_b} \right)^{-\gamma(\sigma^a c)}
\]

The value of debt is equal to the sum of 1) expectation of discounted coupon payments
\( b \) until default 2) expected discounted value of bail-out. Both terms can be evaluated
using the above formula.

Finally, the total value of the bank can be decomposed into:

1. Asset value,
2. Cost of bankruptcy,
3. Tax shield benefits (vary depending on whether CoCo coupons are tax-
deductible),
4. Bail-out value provided by the regulator,
5. Deposit guarantee provided by the regulator, and
6. Cost of premium paid to the regulator for deposit guarantee.

Only the cost of bankruptcy and cost of premium paid to the regulator for deposit guarantee are negative, the rest of them are all positive.

**Proof of Theorem 6.3.1**

The value function $V_b \mapsto R^{noCoCo}(v; V_b)$ is a continuous function of $V_b$ except on the critical point $b + \frac{d}{r}$. It is also differentiable except on $b + \frac{d}{r}$ and $\frac{d}{(1-\alpha)r}$.

For $V_b < \frac{d}{(1-\alpha)r}$, $\phi > 0$, choosing $V_b$ to maximize $R(v)$ is equivalent to maximizing

$$-\left(\theta V_b - \frac{(\theta - p)b + d}{r}\right)\left(\frac{v}{V_b}\right)^{-\gamma(\sigma)}.$$

Taking derivative w.r.t. $V_b$, we obtain:

$$\left(-\theta(\gamma(\sigma) + 1) + \gamma(\sigma)\frac{(\theta - p)b + d}{rV_b}\right)\left(\frac{v}{V_b}\right)^{-\gamma(\sigma)},$$

which is negative for any $V_b \geq \frac{\gamma(\sigma)}{\gamma(\sigma) + 1} \left((1 - \frac{\phi}{\theta})\frac{b}{r} + \frac{d}{r}\right)$. For $V_b \geq \frac{d}{(1-\alpha)r}$, $\phi = 0$, choosing $V_b$ to maximize $R(v)$ is equivalent to maximizing

$$-\left((\theta - (1 - \alpha)p)V_b - (\theta - p)\frac{b + d}{r}\right)\left(\frac{v}{V_b}\right)^{-\gamma(\sigma)}.$$

Taking derivative w.r.t. $V_b$, we obtain:

$$\left(-(\theta - (1 - \alpha)p)(\gamma(\sigma) + 1) + \gamma(\sigma)(\theta - p)\frac{b + d}{rV_b}\right)\left(\frac{v}{V_b}\right)^{-\gamma(\sigma)}.$$

As $\theta > (1 - \alpha)p$, this derivative is negative for any $V_b \geq \frac{\gamma(\sigma)}{\gamma(\sigma) + 1} \frac{\theta - p}{\theta - p + \alpha p} \frac{b + d}{r}$.

As $\frac{\gamma(\sigma)}{\gamma(\sigma) + 1} \left((1 - \frac{\phi}{\theta})\frac{b}{r} + \frac{d}{r}\right) < V_{endo}$ and $\frac{\gamma(\sigma)}{\gamma(\sigma) + 1} \frac{\theta - p}{\theta - p + \alpha p} \frac{b + d}{r} < V_{endo}$, the regulator’s value has negative derivative on the intervals $[V_{endo}, \frac{b + d}{r})$ and $(\frac{b + d}{r}, \infty)$, it is decreasing on both of these intervals.
Therefore, to maximize $R^{\text{noCoCo}}(v; V_b)$, we only need to choose the maximum over $R^{\text{noCoCo}}(v; V_{\text{endo}})$ and $R^{\text{noCoCo}}(v; \frac{b+d}{r})$.

Proof of Proposition [6.3.2]

This result is obvious as $\sigma = \sigma^l$ when $V_b = \frac{b+d}{r}$ and $\sigma = \sigma^h$ when $V_b = V_{\text{endo}}$. $\sigma \mapsto \gamma(\sigma)$ is a decreasing function, the result follows naturally from the monotonicity.
Bibliography


173


[58] O’Kane, D., S. Turnbull, 2003, Valuation of credit default swaps, Lehman Brothers, Fixed Income Quantitative Credit Research.


