Essays in Financial Economics

Essays in Financial Economics

A dissertation presented

by

Gordon Liao

to

The Department of Economics and

The Finance Unit at the Harvard Business School

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Business Economics

Harvard University

Cambridge, Massachusetts

May 2017

Dissertation Committee Co-Chairs:
Professor Robin Greenwood
Professor Samuel G. Hanson

Author:
Gordon Liao

Essays in Financial Economics

My dissertation is composed of three papers in financial economics. In the first essay, "Credit Migration and Covered Interest Rate Parity," I document economically large and persistent discrepancies in the pricing of credit risk between corporate bonds denominated in different currencies. This violation of the Law-of-One-Price (LOOP) in credit risk is closely aligned with violations of covered interest rate parity in the time series and the cross-section of currencies. I explain this phenomenon with a model of market segmentation. Post-crisis regulations and intermediary frictions have severely impaired arbitrage in the exchange rate and credit markets each on their own, but capital flows, either currency-hedged investment or debt issuance, bundle together the two LOOP violations. Limits of arbitrage spill over from one market to another.

The second essay, joint with Robin Greenwood and Sam Hanson, studies theoretically how do large supply shocks in one financial market affect asset prices in other markets. We develop a model in which capital moves quickly within an asset class, but slowly between asset classes. While most investors specialize in a single asset class, a handful of generalists can gradually re-allocate capital across markets. Upon arrival of a supply shock, prices of risk in the impacted asset class become disconnected from those in others. Over the long-run, capital flows between markets and prices of risk become more closely aligned. While prices in the impacted market initially overreact to shocks, under plausible conditions, prices in

related asset classes underreact. Our model suggests that the short-run price impact of a supply shock on different markets may not accurately reveal the long-run impact, which is often of greater interest to policymakers.

The final essay, joint with Robert Barro, develops a new options-pricing formula that applies to far-out-of-the money put options on the overall stock market when disaster risk is the dominant force, the size distribution of disasters follows a power law, and the economy has a representative agent with Epstein-Zin utility. In the applicable region, the elasticity of the put-options price with respect to maturity is close to one. The elasticity with respect to exercise price is greater than one, roughly constant, and depends on the difference between the power-law tail parameter and the coefficient of relative risk aversion, γ. The options-pricing formula conforms to data from 1983 to 2015 on far-out-of-the-money put options on the U.S. S&P 500 and analogous indices for other countries. The analysis uses two types of data—indicative prices on OTC contracts offered by a large financial firm and market data provided by OptionMetrics, Bloomberg, and Berkeley Options Data Base. The options-pricing formula involves a multiplicative term that is proportional to the disaster probability, p. If γ and the size distribution of disasters are fixed, time variations in p can be inferred from time fixed effects. The estimated disaster probability peaks particularly during the recent financial crisis of 2008-09 and the stock-market crash of October 1987.

Contents

Abstract **iii**

Acknowledgments **xi**

1 Credit Migration and Covered Interest Rate Parity[1] **1**

1.1 Introduction 1

1.2 Measuring residualized credit spread 10

1.2.1 Data 10

1.2.2 Matrix pricing of corporate credit 11

1.2.3 Comparison with benchmark credit spreads 14

1.2.4 Robustness in the measurement of the credit spread differential . . . 16

1.3 Alignment of credit differential and CIP violation 17

[1] I am indebted to Robin Greenwood, Sam Hanson, Andrei Shleifer, and Jeremy Stein for their guidance and support. I am also thankful for helpful conversations with Malcom Baker, Robert Barro, Vitaly Bord, May Bunsupha, John Campbell, Jeffrey Frankel, Xavier Gabaix, Gita Gopinath, Ben Hébert, Victoria Ivashina, Derek Kaufman, Owen Lamont, Patrick Luo, Yueran Ma, Matteo Maggiori, Filippo Mezzanotti, Mikkel Plagborg-Møller, Andreas Schaab, David Scharfstein, Jesse Schreger, Emil Siriwardane, Erik Stafford, Sophia Yue Sun (discussant), Adi Sunderam, Boris Vallée, Luis Viceira, Kevin Wang and seminar participants at Harvard, conference participants of the Corporate Finance Conference at Washington University in St. Louis Olin Business School. I also thank the following institutions for insightful conversations with their employees: Bracebridge Capital, Credit Suisse, Deutsche Bank, Harvard Management Company, JP Morgan, Morgan Stanley, Nikko Asset Management, Nippon Life, Nomura, Norinchukin Bank, and UBS. I thank HBS Research Services and HBS Japan Research Center. I am grateful for financial support from the Bradley Foundation Fellowship.

1.4 A model of aligned deviations in credit and currency markets 24
1.4.1 Firm decision 25
1.4.2 Credit markets 26
1.4.3 Currency swap market 28
1.4.4 Summary of equilibrium conditions and predictions 30
1.4.5 Falsifiable alternative 33
1.5 Discussions 34
1.5.1 Source of ε_c and ε_b shocks 34
1.5.2 Limits of arbitrage 38
1.5.3 Firms as natural cross-market arbitrageurs 41
1.6 Additional empirical results 45
1.6.1 Data and definition 45
1.6.2 Prediction 1: Spillover of deviations 48
1.6.3 Prediction 2: Issuance flow and net deviation 52
1.6.4 Prediction 3: Total issuance and deviation alignment 58
1.6.5 Prediction 4: Spillover of Limits to Arbitrage 60
1.7 Conclusion 62

2 Asset Price Dynamics in Partially Segmented Markets[2] 63
2.1 Introduction 63
2.2 Model 69
2.2.1 Single asset model 69
2.2.2 Partially segmented markets 76
2.2.3 Defining market integration 85

[2]This paper was written jointly with Robin Greenwood and Samuel G. Hanson. We thank Daniel Bergstresser, Yueran Ma, Jeremy Stein, Adi Sunderam, and seminar participants at Berkeley Haas, Brandeis, MIT Sloan, NYU Stern, University of Minnesota Carlson, University of North Carolina Kenan-Flager, and University of Texas McCombs for useful feedback. We thank David Biery for research assistance. Greenwood and Hanson gratefully acknowledge funding from the Division of Research at Harvard Business School.

2.3 Market integration following large supply shocks 88
2.3.1 Unanticipated supply shocks 89
2.3.2 Extensions 98
2.4 Discussion and Applications 102
2.4.1 Event studies and changes in the price of risk 102
2.5 Conclusion 105

3 Options-Pricing Formula with Disaster Risk[3] **107**
3.1 Introduction 107
3.2 Baseline Disaster Model and Previous Results 109
3.3 Pricing Stock Options 112
3.3.1 Setup for pricing options 112
3.3.2 Power-law distribution of disaster sizes 114
3.3.3 Options-pricing formula 116
3.3.4 Maturity of the option 117
3.3.5 Diffusion term 120
3.3.6 Stochastic Volatility 121
3.4 Empirical Analysis 123
3.4.1 Data and methodology 123
3.4.2 Basic model fit 126
3.4.3 Estimated disaster probabilities 131
3.4.4 Model robustness 135
3.5 Conclusions 144

Bibliography **146**

[3]This paper was written jointly with Robert J. Barro. We appreciate helpful comments and assistance with data from Josh Coval, Ben Friedman, Xavier Gabaix, Tina Liu, Matteo Maggiori, Greg Mankiw, Robert Merton, Richard Roll, Steve Ross, Emil Siriwardane, Jessica Wachter, and Glen Weyl, and participants in the macroeconomics seminar at Harvard University and finance seminar at MIT.

List of Tables

1.1 Bond data summary 11
1.2 Descriptive regression of Credit Spread Differential on CIP Deviations 23
1.3 Issuance flow and net deviation 56
1.4 Firm-level issuance choice and violations in credit and CIP 59
1.5 Debt issuance amount and deviation alignment 61

2.1 Illustrative model parameters 88
2.2 Model comparative statics 95

3.1 Baseline Regression Estimates 128
3.2 Estimation with At-The-Money Put Options 130
3.3 Means, Standard Deviations, Quantiles of Estimated Disaster Probabilities . 132
3.4 Robustness: Elimination of Lowest Exercise Price, ε 136
3.5 Robustness: Two Ranges of Option Maturity 138
3.6 Robustness: Inclusion of Puts with 1-Year Maturity 139
3.7 Robustness: Alternative Sample Periods 140
3.8 U.S. Regression estimates, 1983-2017 142

List of Figures

1.1 Credit risk price discrepancies and CIP deviations for EURUSD 4
1.2 Residualized foreign currency credit spreads relative to dollar credit spread . 13
1.3 Comparison of residualized credit spread diff. (EU-US) with un-residualized benchmarks 15
1.4 Covered Interest Rate Parity deviations at the 5-year horizon 20
1.5 Credit spread differential and CIP violation relative to USD 22
1.6 Credit spread differential and CIP violation 24
1.7 Schematic of institutional details 35
1.8 Textual analysis of FX-hedged foreign debt issuance for S&P 500 firms . . . 44
1.9 Net deviation 47
1.10 Spillover of deviations: orthogonalized impulse responses of deviations and issuance flow for EURUSD 50
1.11 Spillover of deviations: Panel VAR 51
1.12 Issuance flow and net deviation between Europe and the U.S. 53
1.13 Orthogonalized impulse response of monthly issuance flows to shock to net deviation for EURUSD 57

2.1 Price impact of an unanticipated shock to the supply of asset A 91

2.2 Portfolio adjustments in response to an unanticipated shock to the supply of asset A . . . 92
2.3 Price impact with multiple securities in each market . . . 101
2.4 Event study confidence interval following an unanticipated shock to the supply of asset A . . . 106

3.1 Comparison of Prices with Black-Scholes Predicted Prices . . . 129
3.2 Estimated Disaster Probabilities . . . 133
3.3 Estimated U.S. Disaster Probabilities, 1983-2017 . . . 143

Acknowledgments

First and foremost, I am indebted to the members of my Dissertation Committee, Robert Barro, Robin Greenwood, Sam Hanson, Andrei Shleifer, and Jeremy Stein. This dissertation would not exist without your guidance and advice. Thank you for your support, enthusiasm, and insight. This dissertation reflects how much I've learned from working with each of you—and how I've grown as a researcher.

Robert, thank you for introducing me to the world of research when I was a junior in college. Thank you for always being there as a mentor and friend. Robin, you've coached me on practically everything during my grad school career—not just on how to conduct academic research, but also on creating a positive influence on people around you and how to live a balanced life. I am excited to run many more stadiums together. Sam, thank you for meticulously and patiently teaching me so much over the past four years. You are like a big brother to me. Andrei, thank you for always believing in me, for your infinite wisdom and for making every meeting such a fun one. Jeremy, you've always inspired me to push myself further and served as a role model in all aspects of life. Thank you for your invaluable feedback and constant inspiration.

I am also grateful for a number of other faculty members at Harvard who have advised me over the years, especially, Malcolm Baker, John Campbell, David Laibson, Adi Sunderam and Luis Viceira. I'm so grateful for the HBS/Harvard Econ doctoral community. Thank

you to my student peers whom I've come to consider as family: Cheng Gao, Patrick Luo, Ash Gandhi, David Choi, Vitaly Bord, Yueran Ma, May Bunsupha among others. Cheng, thank you for convincing me to attend graduate school. I could not imagine a better way to spend my twenties learning and growing with awesome people like you. Thank you for being a great friend for the past nine years.

Finally, this dissertation is dedicated to my parents and wife, whom I love dearly. Sarah, thank you for always supporting me and believing in me, even when I had many doubts in myself. Mom and Dad, I am always thankful for my strong upbringing and your unconditional love.

1 Credit Migration and Covered Interest Rate Parity[1]

1.1 Introduction

The finance literature is full of examples in which security markets violate the Law of One Price (LOOP), a cornerstone of finance theory stating that assets with identical payoffs should have identical prices. For instance, closed-end funds, twin shares, and stub pricing are well-documented examples of price discrepancies in securities with similar cashflows[2] (see Lamont and Thaler 2003 for survey). These violations are often studied in isolation and attributed to behavioral and institutional frictions in the particular market. I show, in a novel setting, that LOOP violations in one market can arise as an equilibrium outcome of

[1]I am indebted to Robin Greenwood, Sam Hanson, Andrei Shleifer, and Jeremy Stein for their guidance and support. I am also thankful for helpful conversations with Malcom Baker, Robert Barro, Vitaly Bord, May Bunsupha, John Campbell, Jeffrey Frankel, Xavier Gabaix, Gita Gopinath, Ben Hébert, Victoria Ivashina, Derek Kaufman, Owen Lamont, Patrick Luo, Yueran Ma, Matteo Maggiori, Filippo Mezzanotti, Mikkel Plagborg-Møller, Andreas Schaab, David Scharfstein, Jesse Schreger, Emil Siriwardane, Erik Stafford, Sophia Yue Sun (discussant), Adi Sunderam, Boris Vallée, Luis Viceira, Kevin Wang and seminar participants at Harvard, conference participants of the Corporate Finance Conference at Washington University in St. Louis Olin Business School. I also thank the following institutions for insightful conversations with their employees: Bracebridge Capital, Credit Suisse, Deutsche Bank, Harvard Management Company, JP Morgan, Morgan Stanley, Nikko Asset Management, Nippon Life, Nomura, Norinchukin Bank, and UBS. I thank HBS Research Services and HBS Japan Research Center. I am grateful for financial support from the Bradley Foundation Fellowship.

[2]To be clear, these are LOOP violations in the classical, frictionless sense, if one were to actually construct an arbitrage strategy, the cashflows might very well be different.

arbitrageur actions intended to correct LOOP violations in another market.

I begin by documenting large and persistent differences in the pricing of credit risk for corporate bonds denominated in different currencies. Textbook asset pricing theory predicts that identical claims issued by the same firm but traded in different markets are priced similarly due to arbitrage. I show that persistent discrepancies exist for the entire euro corporate bond market versus the dollar bond market (as well as between other currencies). For example, in November 2014, AT&T, the BBB-rated and U.S.-based telecommunication giant, had a credit spread of 203 basis points on its 15-year U.S. dollar-denominated bond, while its euro-denominated bonds of similar maturity had a credit spread of 129 basis points. Credit risk of AT&T is therefore priced differently in the U.S. and European bond markets.

Generalizing from this example is difficult because no two bonds are perfectly alike. Different terms of maturity, rating, liquidity, and firm-specific characteristics create challenge in the comparison. AT&T, for example, issues more long-term bonds in euro than in dollar. Applying cross-sectional regressions on a large panel of bond credit spreads, I build a measure of currency-specific pricing of credit risk that controls for other characteristics. I interpret the currency fixed effects in the regressions as measures of the price of credit risk associated with different bond denomination currencies. Taking fixed effects normalizes bond characteristics and using credit spread as a price measure removes differences in risk-free funding rates across currencies. Thus, the difference in residualized credit spreads constitutes a difference in the pricing of credit default risks.

The difference in residualized credit spreads between major currencies have dramatically widened since the Global Financial Crisis. From 2004 to 2007, the residualized credit spreads of Australian dollar (AUD), Canadian dollar (CAD), Swiss francs (CHF), Euro (EUR), British Pound Sterling (GBP), and Japanese Yen (JPY) relative to USD maintained a narrow range of 10 bps. Since 2008, however, these spreads have diverged significantly and have been large even in tranquil periods. For instance, the difference between the residualized credit spread of EUR and USD had reached over 70 basis points in 2016. The price discrepancies are substantial in terms of dollar value given the sheer size of the aggregate bond markets (e.g. EUR corporate bond market has $3 trillion of long-term outstanding debt, USD corporate bond market has $10 trillion of outstanding debt[3]). A 70 basis points price discrepancy amounts to $25 billion or represent 84% of net (12% of gross) annual issuance in the euro corporate bond market.

I then show that the LOOP violations in credit market between bonds of different denomination currencies are closely related to deviations from Covered Interest Rate Parity condition, another LOOP violation that has recently attracted attention from a variety of other papers (Sushko, et al. [2016], Du, Tepper, and Verdelhan [2016], Iida, Kimura, and Sudo [2016]). Covered Interest Rate Parity (CIP) condition is a textbook no-arbitrage relation asserting that the forward currency exchange rate must be equal to the spot exchange rate after adjusting for the funding rate differential between two currencies. The CIP condition held tightly prior to 2008. However, large deviations from the CIP relation appeared in the aftermath of the financial crisis and have persisted through 2016. For a detailed documentation and exposition of CIP violations, see Du, Tepper, and Verdelhan (2016).

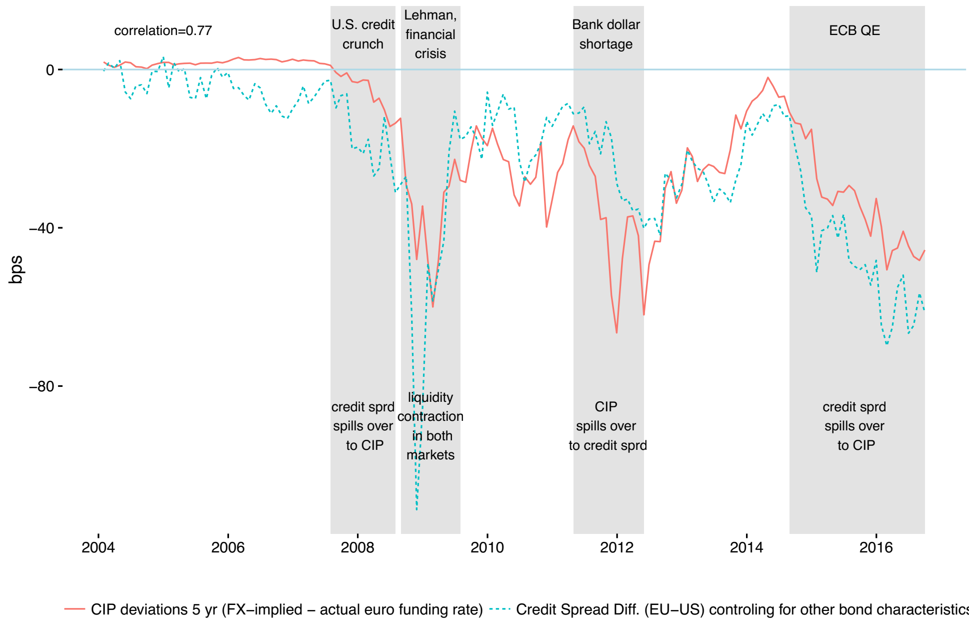
Figure 1.1 shows the time series of price discrepancies in credit risk and deviations from CIP for EUR/USD. Periods when the price of credit risk is lower in euro than in dollar (more negative dashed blue line) tend to coincide with periods with a lower FX-implied euro funding rate relative to actual euro funding rate (more negative CIP deviation as indicated by the red solid line). The two time series share similar magnitude of deviation and are highly correlated (77%). The close alignment of the two LOOP violations is not mechanically driven by interest rate fluctuation, as explained in Section 1.3. This comovement of LOOP violations also holds true in other currencies. In a pooled sample of AUD, CAD, CHF, EUR, GBP, and JPY relative to USD, the correlation between CIP violation and credit

[3] ECB; Federal Reserve Flow of Funds L.213

Figure 1.1: Credit risk price discrepancies and CIP deviations for EURUSD
This figure shows the residualized credit spread differential (dotted blue) and violations of CIP at the 5-year horizon (solid red) for EURUSD. To construct estimates of residualized credit spread, I estimate the following cross-sectional regression at each date t

$$S_{it} = \alpha_{ct} + \beta_{ft} + \gamma_{mt} + \delta_{rt} + \varepsilon_{it}$$

where S_{it} is the yield spread over the swap curve for bond i that is issued in currency c, by firm f, with maturity m and rating r. The residualized credit spread of euro relative to dollar is defined as $\hat{\alpha}_{eur,t} - \hat{\alpha}_{usd,t}$. Details of the measure's construction are provided in Section 1.2.2.



price discrepancies is 81%.

I provide an explanation for the joint determination of credit pricing discrepancies in different currencies and CIP violations based on a model of market segmentation and limited arbitrage. When markets are segmented, prices of risk in one market may be disconnected from those in other markets. The two LOOP deviations reflect two distinct market segmentations – the credit markets are divided by denomination currencies while the CIP violation is a disconnect between spot and forward exchange rates in the FX markets. I develop a model in which the integration of either asset class requires cross-market arbitrageurs to bridge through the other asset class.

To understand the conceptual framework, consider again the AT&T example, the firm finds it cheaper to issue in EUR than in USD when considering the cost of debt payment alone. However, for AT&T to take advantage of the lower credit spread in EUR, it would be exposed to substantial amount of FX volatility[4]. To hedge for this volatility, AT&T would need to buy EUR in the forward market for the future repayment of its debt – in fact, AT&T did exactly this: it issued €800 million ($1 billion) in a 15-year euro-denominated bond and entered into currency derivatives as a hedge. In its 10K statement, AT&T describes the pervasiveness of its FX-hedged global bond issuance,

> "We have entered into multiple cross-currency swaps to hedge our exposure to variability in expected future cash flows that are attributable to foreign currency risk generated from the issuance of our Euro, British pound sterling, Canadian dollar and Swiss Franc denominated debt."

It is therefore natural to think of AT&T as a corporate arbitrageur that not only links together the two credit markets but also connects the FX forward and spot markets through its currency hedges.

There are four players in my model: a FX arbitrageur, two specialized credit investors, and a representative debt-issuing firm. The two specialized credit investors each invest in corporate bonds in their respective home currencies, the euro and the dollar, and they each have a downward sloping demand curves in the credit markets. The FX arbitrageur connects the spot and forward exchange rate markets and also has a downward sloping demand curve because of limited balance sheet capacity to perform the arbitrage.

The firm connects the credit and FX markets by engaging in FX-hedged debt issuance. Its objective is to minimize its overall financing cost by choosing the optimal share of debt to issue in each currency. When the foreign credit spread is low, the firm allocates a greater share of debt to be issued abroad. Issuing in the foreign currency, however, generates FX

[4] A back of envelope calculation suggests that a 10% appreciation of USD could wipe out one-third of AT&T's annual profit if the firm does not hedge its FX exposure on its outstanding foreign currency debt.

exposure, which the firm hedges using currency forwards. To integrate the two downward-sloping demand curves in the bond markets, the firm has to walk down the demand curve in the FX forward market. Conversely, when CIP violations are large, the firm chooses to integrate the forward and spot FX exchange rates instead while walking down the demand curves of the credit markets. The two violations of LOOP are aligned such that the firm's first order condition is satisfied. While cross-market arbitrageurs are modeled in this paper as a debt-issuing firm, they can also be broadly interpreted as global debt investors.

Two types of exogenous demand shocks affect the system. First, there are credit demand shocks (perhaps originating from central bank purchase outside of the model) that raise the relative price of credit for bonds in one currency versus the other. Second, there are CIP shocks originating from other end-users of FX forwards that decouple the forward exchange rates from the spot exchange rate. The shocks are transmitted between the FX and credit markets by firms engaged in currency-hedged foreign debt issuance. Credit demand shocks cause discrepancies in the price of credit risk as well as deviations from CIP. Similarly, CIP shocks also spill over to affect the relative price of credit.

The model generates four key predictions. First, LOOP violation in one market (FX or credit) spills over to the other market. Arbitrage processes are imperfect in both markets, but capital flow ensures that the two LOOP deviations are aligned. Second, the amount of cross-currency issuance, which represents arbitrage position, co-varies with the profitability of the arbitrage. The profit margin is indicated by the difference between credit spread differential and CIP deviation. Third, an exogenous increase in cross-market arbitrage capital in the form of higher total amount of debt issuance aligns the two deviations. Lastly, limits of arbitrage in one market (FX or credit) spill over to the other market and become a constraining friction in the other market.

Empirical analyses lend support to the model predictions. A counterintuitive implication of the model, which also appears in the data, is that the net deviation from LOOP is small even when both deviations in CIP and credit are large individually. When the two deviations are meaningfully large (greater than 20 basis points), the level of net deviation, which represents the amount of arbitragable profit, is only around a quarter of the size of the two individual deviations. Evidence from currency-hedged debt issuance accords with the model. A textual analysis of 10K filings by S&P 500 firms indicates that around 40% of firms have issued currency-hedged foreign debt in recent years. Furthermore, issuance flow at the monthly and quarterly horizon fluctuates with the net deviation. For each one standard deviation increase in the difference between residualized credit spread differential and CIP violation for EURUSD, firms respond by shifting around 5% of the aggregate debt issuance towards the cheaper currency (0.75 standard deviation of issuance flow). Vector Autoregression analyses show that issuance flow responds to shocks in credit and FX markets in the direction predicted by the model. The transmission of shocks is slow moving, which is consistent with theories on slow moving capital (Duffie [2010], Greenwood, Hanson, and Liao [2015]). Firm-level panel regressions confirm the same result as in the aggregate data. In addition, an increase in the overall debt issuance, as instrumented by maturing debt that needs to be rolled over, contributes to the alignment of the two LOOP violations.

Why do the two deviations persist? One way of explaining the co-existence of the two LOOP violations is that each of them serves the role of a short-sell constraint to the other. This joint determination of the two LOOP violations is analogous to heavily-shorted stocks being overvalued at the same time that they have high cost to borrow (Negal [2005], D'Avoli [2002]).

My paper takes the idea of limits of arbitrage a step further. Traditionally, LOOP violations are studied in isolation. Noise trader risks and agency problems pose limits to the amount of arbitrage activities (De Long et al. [1990], Shleifer and Vishny [1997]) in a single market. I provide a conceptual framework and document a clear-cut example in which arbitrage constraints and violations of LOOP spill over from one market to a completely different market. The two LOOP violations are determined jointly in equilibrium.

My paper also contributes to the literature on the determination of foreign exchange rate dynamics. Gabaix and Maggiori (2015) provide a theory of the determination of exchange rates based on capital flows in imperfect financial markets. The study of exchange rate determination typically focuses on *uncovered* interest rate parity. In contrast, I model and provide empirical evidence for the determination of *covered* interest rate parity violations. The two concepts are intimately related. As deviation from CIP becomes large, firms and investors eventually forgo hedging (since CIP deviation is a hedging cost), the unhedged capital flow thus leads to UIP violation. Unlike the risk-bearing financial intermediaries in the Gabaix and Maggiori (2015) model, FX-arbitrageurs in my model face little risk, but CIP arbitrage is capital intensive and therefore costly to implement. Ultimately, the real arbitrageurs of the CIP market are investors and treasuries of firms that must fund the cost of arbitrage through bond markets.

This paper also contributes to previous work showing that corporations behave like arbitrageurs in their financing activities (Baker and Wurgler [2000] and Baker, Foley, and Wurgler [2009], Greenwood, Hanson, and Stein [2010], and Ma [2015]). My paper contributes to the literature on firms as arbitrageurs in two ways. First, this paper shows that firm are advantageous at exploiting LOOP violations in addition to previously documented arbitrage of inexact valuation differences, e.g. between debt and equity and market timing of issuance. These arbitrage strategies of LOOP violations typically require specialized knowledge and capital, and were previously reserved for sophisticated hedge funds. Firms' increasing involvement in specialized arbitrage demonstrates the difficulty of deploying traditional arbitrage capital in the post-crisis financial and regulatory environment. Second, firms are arbitraging multiple markets at the same time – e.g. credit and FX, and they play a role in transforming LOOP violation of one form into that of another form.

A small set of literature has examined short-term CIP violations during the financial crisis

(Baba, Packer, and Nagano [2008], Coffey, Hrung, and Sarkar [2009] Griffoli and Ranaldo [2011], and Levich [2012]). Fletcher and Taylor (1996) document long-term CIP violations of the early 1990s and conclude that these violations have diminished or disappeared over time. While these papers discuss limits to arbitrage that prevent the elimination of CIP violations, their examinations of the root cause of deviation in both crisis and non-crisis periods are limited.

More closely related to my paper are Ivashina, Scharfstein, and Stein (2015), Du, Tepper, and Verdelhan (2016), and Sushko et al. (2016). Ivashina, Scharfstein, and Stein (2015) examine the dollar funding and lending behaviors of European banks during the Eurozone Sovereign Crisis in 2011-2012 and explore how shrinkage of wholesales dollar funding compelled the banks to swap their euro funding into dollar, which in turn generated CIP violations and affected lending. Bräuning and Ivashina (2016) further explore the role of monetary policy in affecting global bank's funding sources and the use of FX hedges. Du, Tepper, and Verdelhan (2016) extensively document persistent deviations from CIP in recent periods and propose explanations based on costly financial intermediation and global imbalances. Sushko et al. (2016) examine the role of hedging demands and costly balance sheet in the determination of CIP violations. Relative to these papers, my contribution is to document and explain the joint determination of both CIP violation and price discrepancies in corporate bonds of different denomination currencies. I show that the two LOOP violations need to be considered together in formulating an explanation of the equilibrium prices and capital flows.

The paper proceeds as follows. Section 1.2 discusses the measurements of residualized credit spread. Section 1.3 presents the stylized fact that residualized credit spread differential and CIP deviation are highly aligned. Section 1.4 provides a model to explain the co-determination of these two violations. This is followed by discussion in Section 1.5. Additional model predictions are tested empirically in Section 1.6.

1.2 Measuring residualized credit spread

In this section, I develop a procedure to measure the price of credit risk in different currencies. The ideal experiment is to find pairs of otherwise identical bonds (same issuer, maturity, etc) in different currencies. This is challenging because no two bonds are perfectly alike. My proposed methodology relies on cross-sectional regression to control for differences in rating, maturity, and firm characteristics. From here on in the paper, I refer to the differential in the residualized credit spread of bonds denominated in different currencies simply as *credit spread differential*.

1.2.1 Data

I utilize a comprehensive sample of individual bond yields from Bloomberg and bond attributes from Financial Securities Data Company (SDC) Platinum Global New Issues data set. The selection of bonds is as exhaustive as possible. I obtain yields of more than 35,000 corporate bonds in seven major funding currencies (USD, EUR, GBP, JPY, AUD, CHF, CAD) from 2004 to 2016. The selection includes all fixed-coupon, bullet corporate bonds with outstanding amount of at least $50 million and original maturity of at least one year available on Bloomberg and in the SDC dataset. These bonds were issued by more than 4,600 entities. The issuing entities also include a number of large supranational (such as the World Bank) and sovereign agencies (such as state-owned banks) that are generally considered a part of the corporate bond market. The total notional of outstanding bonds in the database as of June 2016 is around $10 trillion. These bonds represent the majority of bonds outstanding in the market. I use the yield spread against the swap curve as a measurement of credit spread. Pricing data on swaps are obtained from Bloomberg. Additional bond attributes used for robustness checks are obtained from Moody's Default & Recovery Database. A summary of the bond data is provided in Table 1.1.

Table 1.1: Bond data summary

This table presents a summary of the bond data used in the main analyses. Bond characteristics are from Thompson One SDC Platinum.

		All bonds		Global issuers only	
		Number	Notional $bil	Number	Notional $bil
currency	all	35,204	15,937	24,090	12,294
	usd	12,772	6,443	7,954	4,561
	eur	8,625	5,446	6,653	4,556
	jpy	8,152	1,969	5,316	1,474
	gbp	1,492	766	1,238	678
	cad	1,124	516	700	419
	chf	2,017	478	1,301	304
	aud	1,022	319	928	302
rating	AA- or higher	12,060	7,331	10,528	6,741
	A+ to BBB-	13,732	5,796	8,593	3,782
	HY (BB+ or lower)	1,932	899	1,057	541
	NA	7,480	1,912	3,912	1,230
maturity	<3yrs	1,268	807	1,012	691
	3-7 yrs	14,850	7,173	10,415	5,702
	7-10 yrs	4,755	1,904	3,141	1,396
	10yr+	14,331	6,054	9,522	4,505

1.2.2 Matrix pricing of corporate credit

To assess the impact of denomination currency on the pricing of credit risk, I estimate the following cross-sectional regression at each date t

$$S_{it} = \alpha_{ct} + \beta_{ft} + \gamma_{mt} + \delta_{rt} + \varepsilon_{it} \tag{1.1}$$

where S_{it} is the yield spread over the swap curve for bond i traded in the secondary market at time t. α_{ct}, β_{ft}, γ_{mt}, and δ_{rt} are fixed effect estimates for currency c , firm f, maturity bucket[5] m and rating bucket r respectively at date t. The firm fixed effect is important here since it controls for other characteristics of bonds that are common at the firm level, e.g. industry effect. Furthermore, the data sample is limited to only bonds belonging to multi-

[5] The maturity of the bond at each pricing date t is categorized into four buckets (under 3 years, 3 to 7 years, 7 to 10 years and beyond 10 years). Alternative specification that includes maturity as a linear control is also tested and produce similar results.

currency issuers. As with the AT&T example in the introduction, the idea here is to match bonds of similar characteristics issued by the same firm with the only difference being the currency in which they are denominated. α_{ct} thus measures the *residualized* credit spread controlling for all other observables. This method of attribution is analogous to the standard industry practice of matrix pricing in which a bond with unknown prices is assessed against other bonds with similar maturity and rating.

I use the residualized credit spread differential to measure the LOOP violation of credit risk between currencies. Specifically, the currency fixed effect estimates $\hat{\alpha}_{ct} - \hat{\alpha}_{USDt}$ measures the deviation in the pricing of credit risk in currency c relative to the pricing of credit risk in dollar. The large number of observations for each date t ensures a reasonably tight confidence interval[6].

Figure 1.2 presents time series of the point estimates of $\alpha_{ct} - \alpha_{USDt}$ at each date for currencies EUR, GBP, JPY and AUD. All four credit spread differentials were relatively small from 2004 to 2007. The spreads blew out during the Global Financial Crisis. Yen, sterling, and euro credit all tightened considerably relative to U.S. dollar. In particular, euro and yen credit spread differentials reached deviations beyond -100 basis points during the peak of the crisis. The deviations briefly reversed after the crisis. However, since 2010, the credit spread differentials have widened again. Cross-sectionally, the spread differentials for each market have been persistent. JPY credit (purple long dashed line) has been the most over-priced (negative spread) relative to dollar credit, and AUD credit (solid red) has been under-priced (positive spread) relative to the dollar credit market. EUR credit spread differential (green dots) became more negative since 2014, and reached -70 basis points in 2016.

The dollar magnitude of the deviations is substantial and economically large. As of June 2016, the total amount of outstanding long-term corporate debt in EUR is €3.2 trillion[7].

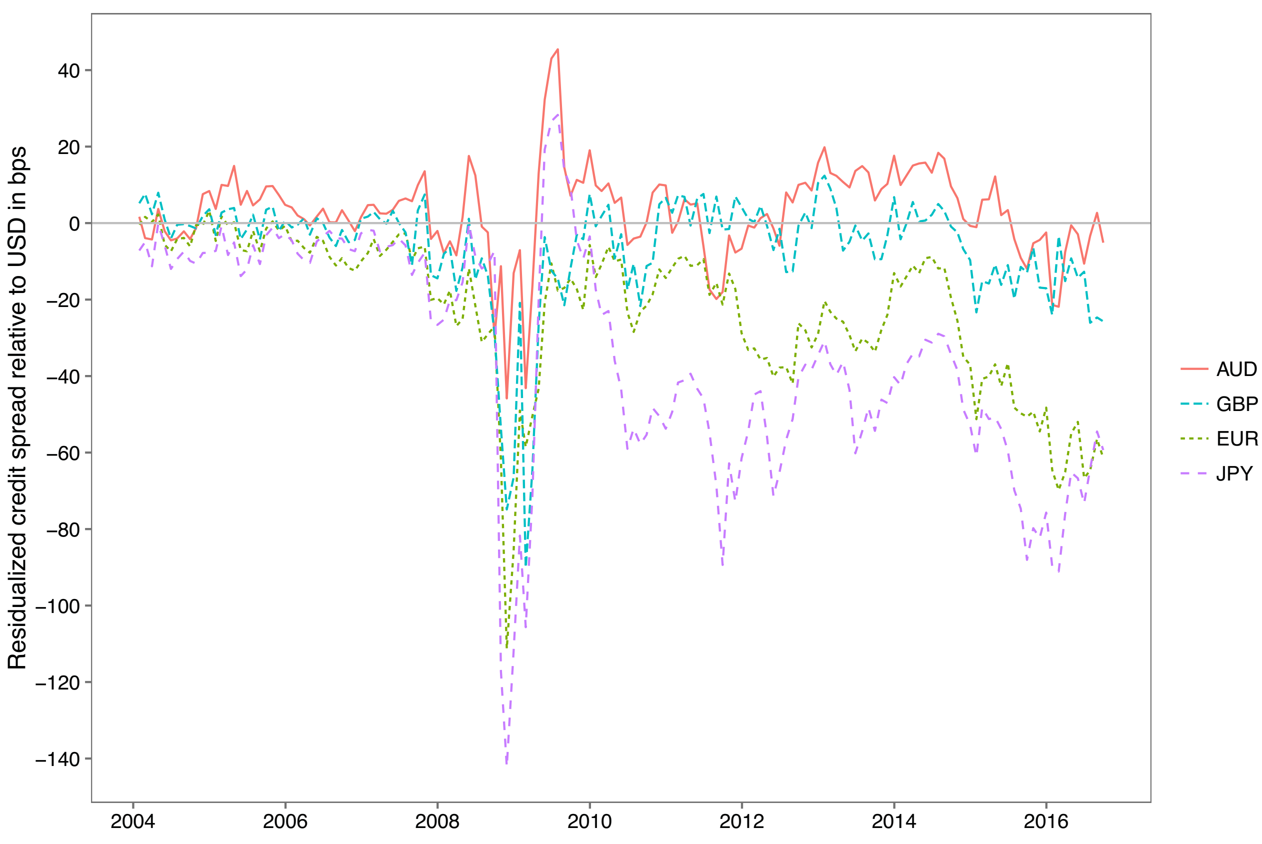
[6]Confidence interval is provided in Figure 1.5

[7]ECB defines long-term debt as debt with original maturity at issuance of greater than one year.

Figure 1.2: Residualized foreign currency credit spreads relative to dollar credit spread
This figure presents the residualized credit spreads in each currency relative to dollar credit spread. To construct this measure, I estimate the following cross-sectional regression at each date t

$$S_{it} = \alpha_{ct} + \beta_{ft} + \gamma_{mt} + \delta_{rt} + \varepsilon_{it}$$

where S_{it} is the yield spread over the swap curve for bond i that is issued in currency c, by firm f, with maturity m and rating r. The residualized credit spread of currency c relative to dollar is defined as $\hat{\alpha}_{c,t} - \hat{\alpha}_{usd,t}$. Details of the measure's construction are provided in Section 1.2.2.



The residualized credit spread differential between EUR and USD in June 2016 is -70 basis points. A back-of-the-envelope calculation suggests that the discrepancy in the pricing of default risk represents a dollar value difference of around $25 billion if all EUR corporate bonds were priced in USD instead. This amount is economically large, representing 84% of the net issuance amount (12% of gross issuance) in EUR by the corporate sector in 2015[8].

1.2.3 Comparison with benchmark credit spreads

The residualization of credit spreads using the above methodology produces time series that offer substantial improvements over un-residualized aggregate credit spreads. I compare the residualized credit spread differential in EURUSD against two un-residualized benchmark indices – the Bank of America Merrill Lynch Corporate Single A index and Barclays Corporate Single A index in Figure 1.3. The residualized and un-residualized spreads are quantitatively and qualitatively different. While the residualized spreads were always negative (indicting tighter euro credit spread than dollar), the unrestricted versions of the spread were positive for a substantial part of the sample and had larger magnitudes. This large difference between the residualized and un-residualized versions is due to compositional differences of the aggregate indices for EUR and USD benchmark bond portfolios provided by Bank of America and Barclays. The regression methodology addresses the compositional difference by controlling for firm and other bond characteristics using individual bond prices.

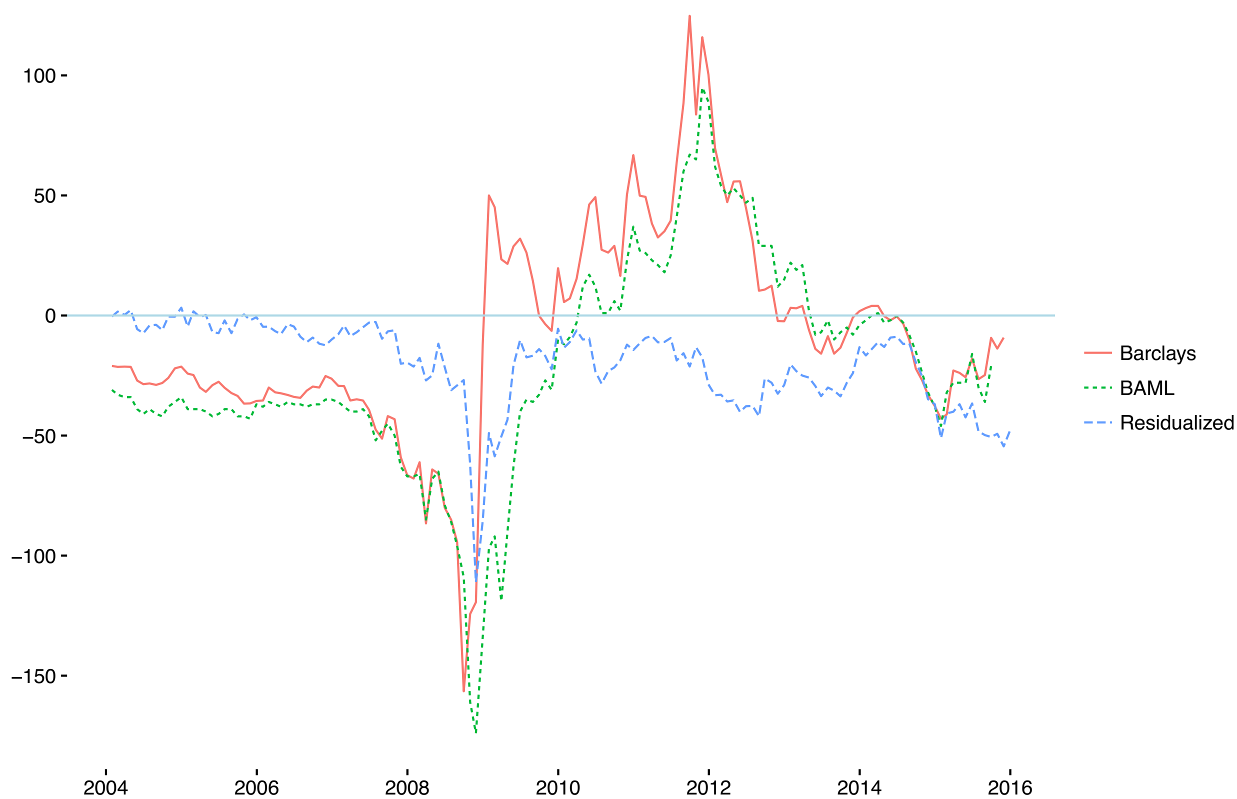
[8]Total net issuance of long-term debt by corporates in 2015 is €26.6 billion and gross issuance is €192.2 billion according to ECB statistics.

Figure 1.3: Comparison of residualized credit spread diff. (EU-US) with un-residualized benchmarks

This figure compares the EU-US residualized credit spread differential (dashed blue) with un-residualized credit spread differentials constructed from Bank of America Merrill Lynch Single A Corporate index (BAML, dotted green) and Barclays Single A Corporate index (solid red). The un-residualized euro minus dollar credit spread differential is constructed by subtracting the dollar-denominated single A aggregate option adjusted spread from euro-denominated single A aggregate option adjusted spread provided by BAML and Barclays. To construct estimates of residualized credit spread, I estimate the following cross-sectional regression at each date t

$$S_{it} = \alpha_{ct} + \beta_{ft} + \gamma_{mt} + \delta_{rt} + \varepsilon_{it}$$

where S_{it} is the yield spread over the swap curve for bond i that is issued in currency c, by firm f, with maturity m and rating r. The residualized credit spread of euro relative to dollar is defined as $\hat{\alpha}_{eur,t} - \hat{\alpha}_{usd,t}$. Details of the measure's construction are provided in Section 1.2.2.



1.2.4 Robustness in the measurement of the credit spread differential

In this section, I conduct a number of robustness checks in the estimation of the residualized credit spread differential.

1.2.4.1 Additional Controls

I augment the regression specification of Equation 1.1 with three additional controls – amount outstanding, age, and seniority. The first two controls serve as liquidity proxies. Larger bond issuance size and newly issued bonds are known to be more liquid. On-the-run bonds, or newly issued bonds, have a premium when compared to off-the-run bonds of similar maturities (Krishnamurthy 2002). To capture this effect, the control for age of the bond is defined as the ratio of remaining maturity to initial maturity of the bond. An additional control for bond seniority (e.g. senior secured, unsecured, subordinate, etc) is obtained from the Moody's Default & Recovery Database and also added to the expanded regression. These controls make little difference on the estimates of the credit spread differentials.

Furthermore, while there might be other idiosyncratic bond attributes not captured in the augmented specification, these additional features should not affect the *aggregate* residualized credit spread differential. As can be seen in Figure 1.2, the residualized credit spread differentials were small prior to the financial crisis. It is unlikely that bond-specific unobservables only begin to vary *systematically* across currencies after the crisis. Therefore, additional unobserved bond features are treated as idiosyncratic noise in the estimation.

1.2.4.2 Heterogeneity for different credit ratings

Another potential concern is that the aggregate credit rating varies significantly across different currency-segmented bond markets. That is, if all euro-denominated bonds have rating of AAA while all dollar-denominated bonds have rating of single-A, then naturally there would be a tighter credit spread for euro-denominated bonds. Under this hypothetical scenario, the residualized credit spread differential would pick up the difference between AAA bonds and single-A bonds rather than a differential due to the denominating currency.

I address this concern in two ways. First, I limit the sample on each date to only bonds that are issued by entities that have debt outstanding in another currency. In this case, controlling for firm fixed-effects alleviate the concern raised above, as bonds issued by the same firm generally have similar credit ratings. Second, a further robustness check is to split the sample for high-grade and low-grade bonds (not shown). When the sample is restricted to low-grade bonds only, the credit spread differentials are larger in magnitude than those of high-grade bonds. This is intuitive since low-grade bonds have higher credit spreads to begin with, the credit spread differential are also intensified.

1.3 Alignment of credit differential and CIP violation

In this section, I define and discuss the measurement of deviation from Covered Interest Rate Parity condition and show the similarities in the time series of CIP deviations and credit spread differentials. Taking the currency pair EUR/USD as an example, the classic text book definition of CIP condition is

$$F_T = S\frac{(1+r_{D,T})^T}{(1+r_{E,T})^T} \tag{1.2}$$

where S is the spot exchange rates expressed in dollars per euro, F_T is the forward exchange rate with maturity T also expressed in dollars per euro, $r_{D,T}$ and $r_{E,T}$ denote the T-period risk-free zero-coupon funding rates in dollar and euro respectively. A violation of CIP occurs when the above equation fails to hold. For expositional purpose, assume that $T = 1$. We can rewrite equation 1.2 as

$$0 = \underbrace{\frac{S}{F}(1 + r_D)}_{\substack{\text{FX-implied} \\ \text{euro funding rate}}} - \underbrace{(1 + r_E)}_{\substack{\text{actual} \\ \text{euro funding rate}}} .$$

In other words, CIP condition states that the FX-implied foreign funding rate is equal to the actual foreign funding rate. A violation of CIP condition can be expressed as a basis b

$$b = \underbrace{\frac{S}{F}(1 + r_D)}_{\substack{\text{FX-implied} \\ \text{euro funding rate}}} - \underbrace{(1 + r_E)}_{\substack{\text{actual} \\ \text{euro funding rate}}} . \tag{1.3}$$

I measure b empirically using the level of cross-currency basis swap, consistent with other concurrent papers[9] studying CIP deviations. A cross-currency basis swap is a market instrument that allows the market participant to simultaneously borrow in one currency and lend in another currency at the respective floating interest rates. The counter party of the swap transaction agrees to take on the reverse position. A currency basis is a market-determined adjustment to the reference floating funding rates. It is analogous to the market pricing of b in Equation 1.3 above. The empirically-relevant funding rates, represented by r_D and r_E in Equation 1.3, are Libor-based swap rates [10]. The details of cross-currency basis swap,

[9] Sushko, et al. [2016], Du, Tepper, and Verdelhan [2016], Iida, Kimura, and Sudo [2016]

[10] Alternative definition using Overnight Index Swap rates based on actual transactions such as Fed Fund Effective Rate or Eonia rate generates similar results. Calculating CIP deviations using FX forward and spot rates also yield similar results.

relation with CIP violation and maturity of CIP deviations are discussed in the appendix[11].

To provide intuition for b, I continue with the earlier example. Suppose AT&T issues in EUR as the euro credit spread is 74 basis points tighter than the dollar credit spread. If there were no CIP deviation, i.e. $b = 0$, AT&T is able to keep the entire 74 basis points by issuing in EUR and swapping EUR into USD. The hedging cost (or benefit) would just be the interest rate differential. If there were a CIP basis $b \neq 0$, the hedging cost would adjust accordingly.

The sign of b is also intuitive. In my example, AT&T issues in EUR and wants to swap EUR to USD. This FX swap transaction can be equivalently stated in two other ways. A FX swap of EUR to USD is equivalent to 1) simultaneously borrowing dollar to lend in euro, and 2) sell euro in the spot market and buy euro in the forward market. Holding the spot exchange rate S and interest rates r_D and r_E fixed in equation 1.3, an increase in F necessitates a decrease in b. Therefore when b is negative, it is expensive to swap from euro to dollar (expensive to buy euro in the forward market), and when b is positive, it is expensive to swap from dollar to euro.

Figure 1.4 shows the deviations from CIP at the 5-year horizon for AUD, EUR, GBP, and JPY relative to USD. This condition had been upheld tightly prior to 2008. However, large deviations from the CIP relation appeared in the aftermath of the financial crisis and persist through 2016.

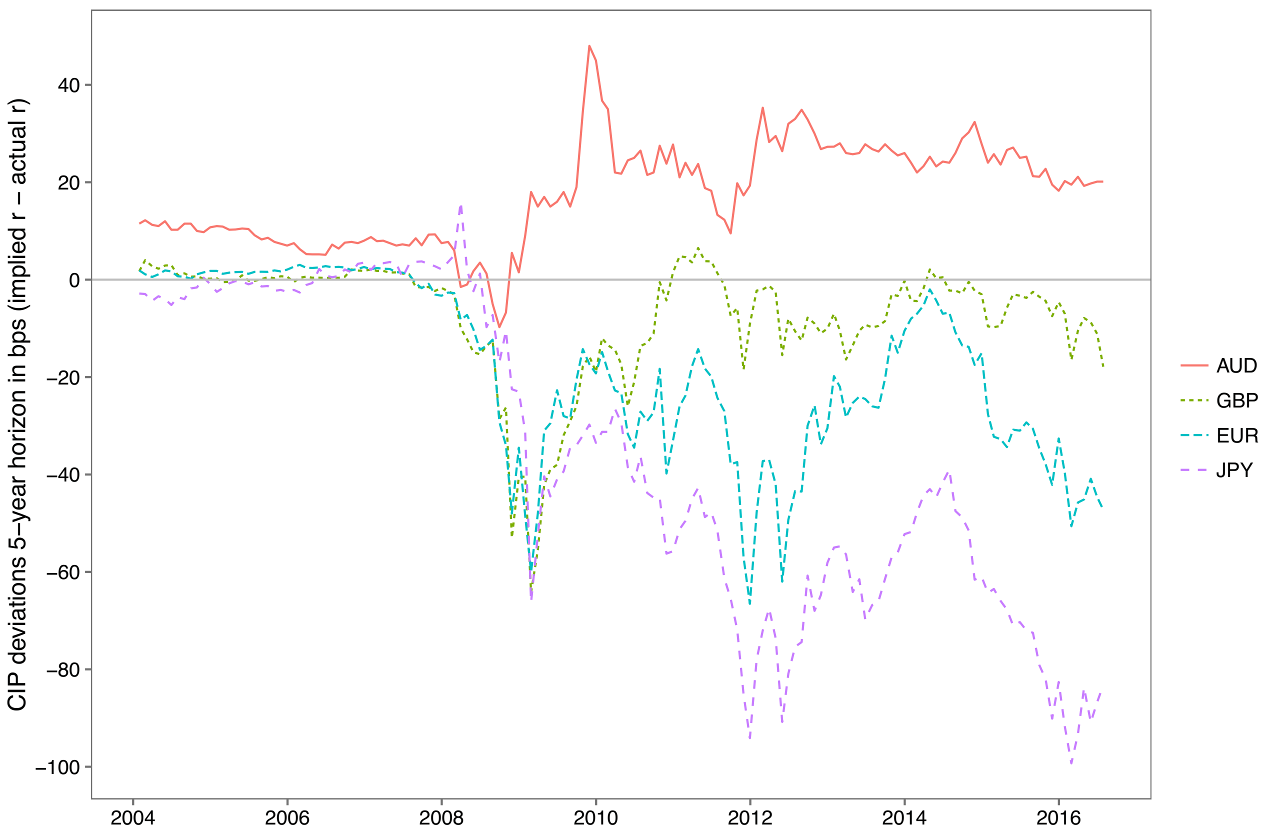
My key finding is that CIP violation and credit spread differential are highly correlated. Figure 1.5 graphs the time series of credit spread differential and CIP deviations at the 5

[11] In the appendix, I show that T-horizon CIP deviation b_T is related to cross-currency basis swap rate B_T by the following approximation:

$$b_T \approx B_T \left[\sum_{t=1}^{T} (1 + Z_t^*)^{-t} \right] \frac{1 + Z_T^*}{T}$$

where Z_t^* denotes the foreign zero-coupon rate with maturity t.

Figure 1.4: Covered Interest Rate Parity deviations at the 5-year horizon
This figure presents the violations of covered interest rate parity at the 5-year horizon between each of the four major free-floating funding currencies - EUR, GBP, JPY, AUD - and USD. Deviations from CIP are measured as the FX-implied local funding rate minus the actual local funding rate. Details of this measure are provided in Section 1.3.



-year horizon for six major funding currencies. The time series of the two violations match closely in magnitude and direction for each currency especially outside of the crisis period. The correlation in the cross-section is also high. Pooling the observations across time and currency, the two violations have a correlation of 81%.

Figure 1.6 shows a scatter plot with credit spread differential on the horizontal axis and deviation from CIP on the vertical axis. This figure highlights both the cross-sectional and time series correlation between the two violations. Japan has negative deviations in both CIP and credit, meaning that yen credit spread is tighter than dollar credit spread for comparable bonds and it is costly to swap yen to dollar. Australia, on the other hand, has both positive deviations, meaning that both its credit spread is wider and it is costly to swap from USD to AUD.

Descriptive regressions also confirm both cross-sectional and time-serial correlation between credit spread differential and CIP deviations. Table 1.2 presents the relationship between the two LOOP violations for the six currencies in panel and individual regressions. The regressions coefficients are highly significant. Most coefficients range from 0.7 to close to 1. Column 2 and 3 present regressions controlling for time and currency fixed effects. While these regressions cannot be interpreted as causal, nonetheless they demonstrate the close alignment of the two LOOP violations. Empirical identification of the impact of one LOOP violation on another is achieved through additional empirical tests of model predictions in subsequent sections.

Figure 1.5: Credit spread differential and CIP violation relative to USD

This figure presents the residualized credit spread differentials ($\alpha_c - \alpha_{USD}$) and CIP deviations ($r_c^{\text{FX implied}} - r_c$) relative to USD for six major funding currencies (c = EUR,GBP,JPY,AUD,CHF,CAD). The CIP deviations are in solid red. Credit spread differentials are in dotted blue. Vertical bars (grey) represent the 95% confidence interval for the estimated credit spread differentials constructed using robust standard errors clustered at the firm level. Details of the measures' construction are provided in Section 1.2.2 and 1.3.

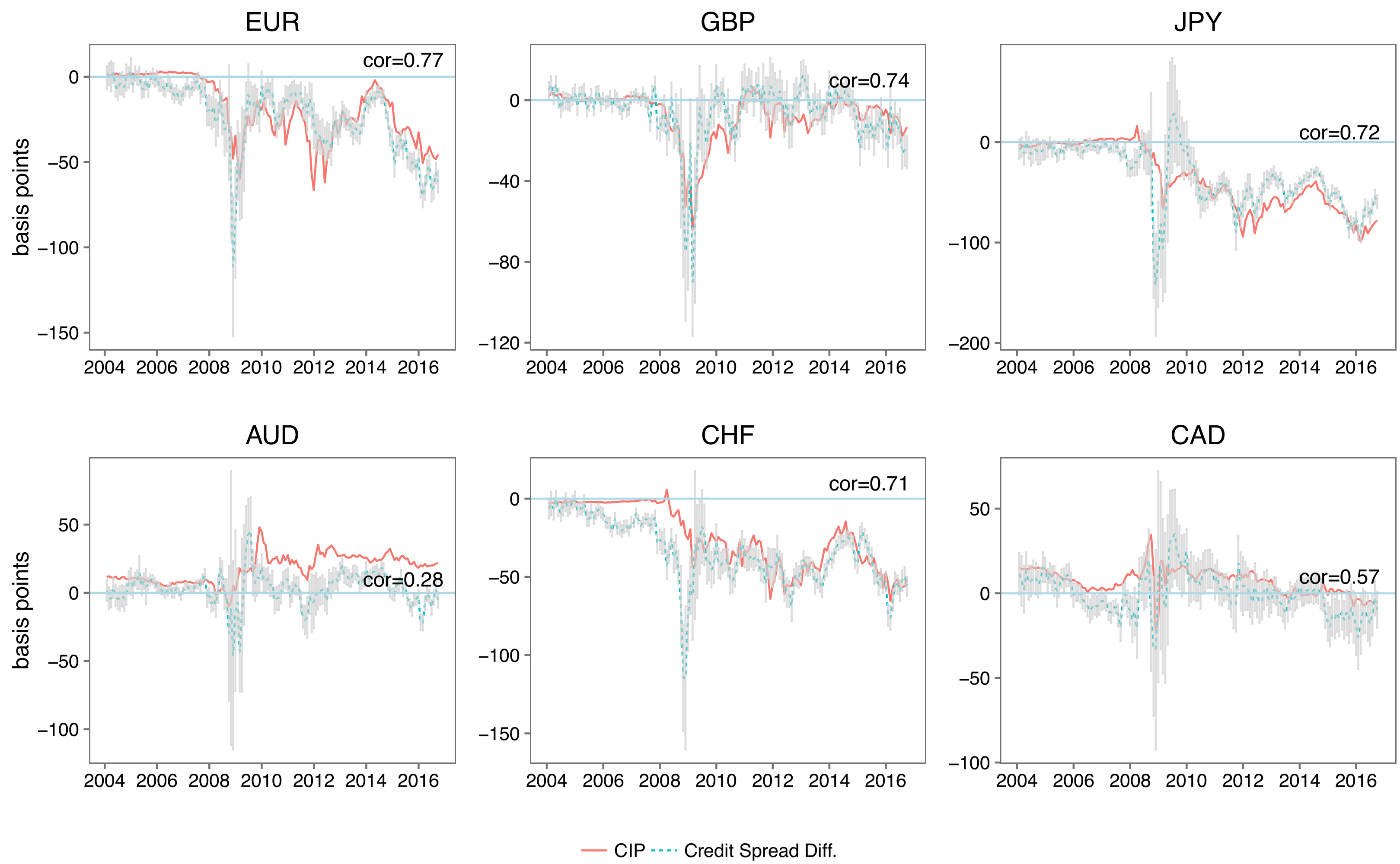


Table 1.2: Descriptive regression of Credit Spread Differential on CIP Deviations

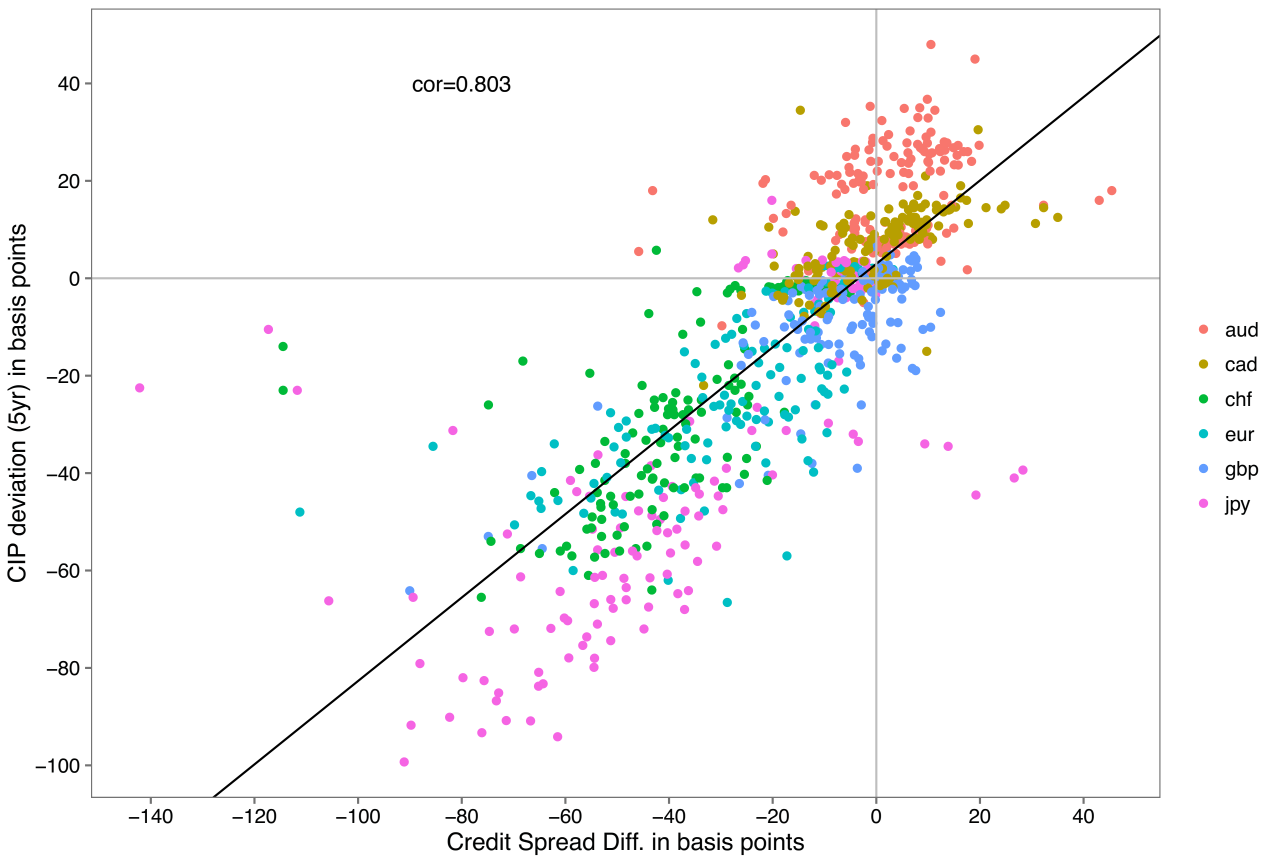
This table presents regressions of credit spread differential on CIP deviations at the 5-year horizon for six major currencies each against the U.S dollar. The sample period is from January 2004 to July 2016 with monthly observation. Column 1 presents the pooled sample regression, columns 2 and 3 present panel regressions with time and currency fixed effects, columns 4 to 9 present regressions for each of the six currencies. In columns 1 to 3, t-statistics in brackets are based on Driscoll and Kraay (1998) standard errors with a maximum lag of 12 months. In columns 4 to 6, t-statistics in brackets are based on Newey-West standard errors with lag selection following Newey-West (1994).

$$crd_{ct} = a + b \cdot cip_{ct} + \varepsilon_{ct} \quad (1)$$
$$crd_{ct} = a_t + b \cdot cip_{ct} + \varepsilon_{ct} \quad (2)$$
$$crd_{ct} = a_c + b \cdot cip_{ct} + \varepsilon_{ct} \quad (3)$$
$$crd_{ct} = a_{ct} + b_{ct} \cdot cip_{ct} + \varepsilon_{ct} \quad (4-9)$$

	Residualized credit spread differential								
	(1) Pooled	(2) Time FE	(3) ccy FE	(4) EUR	(5) GBP	(6) JPY	(7) AUD	(8) CHF	(9) CAD
CIP	0.737	0.683	0.704	0.799	0.821	0.688	0.292	0.705	0.918
	[18.62]	[31.72]	[11.32]	[5.22]	[4.34]	[13.02]	[2.78]	[7.87]	[3.73]
_cons	-7.36			-5.94	0.888	-9.65	-2.4	-14.3	-7.1
	[-5.16]			[-3.67]	[0.64]	[-2.41]	[-0.97]	[-3.79]	[-3.56]
rsq	0.65	0.82	0.48	0.56	0.50	0.53	0.07	0.52	0.36
N	906	906	906	151	151	151	151	151	151

Figure 1.6: Credit spread differential and CIP violation

This figure presents the residualized credit spread differential and CIP violations relative to USD for EUR, GBP, JPY, AUD, CHF and CAD. Details of each measures' construction are provided in Section 1.2.2 and 1.3.



1.4 A model of aligned deviations in credit and currency markets

In this section, I present a model of segmented markets that provide an explanation for the high degree of alignment between the two LOOP violations. In this model, I assume that there are two credit markets, one denominated in euro and another denominated in dollar. These two credit markets are segmented from one another except through capital flow provided by a representative debt-issuing firm. The issuer has funding needs in dollar but issues in both currencies and engages in currency hedging. While the cross-market arbitrageur is modeled as a firm selling debt, it can also be alternatively interpreted as global investors that both purchase and sell across markets. The intuitions and model implications are unchanged when a representative global investor replaces the firm in the model. I use the model to illustrate the transmission of shocks across markets, the alignment of LOOP violations, and the response of issuance capital flow. In addition, the model delivers testable predictions that are examined in Section 1.6.

1.4.1 Firm decision

In this static model, a representative price-taking firm chooses the currency of debt denomination given a fixed debt amount D that needs to be raised. It faces two prices. First, the firm observes a credit spread differential between euro-denominated bonds and dollar-denominated bonds denoted as c. Recall from the earlier example, c is -74 basis points, meaning that AT&T's euro bond credit spread is 74 basis points tighter than the dollar bond spread. If CIP holds, AT&T would save 74 basis points by issuing in EUR and swapping the issuance to USD with currency hedge instead of directly issuing in USD. This is because CIP condition implies that the currency hedging cost is entirely accounted for by the interest rate differential. However, when CIP fails, the firm faces additional hedging cost. It observes a CIP basis, denoted b. As defined earlier in Section 1.3, a negative b means that it is expensive to swap EUR to USD. Suppose $b = -50$, this means that AT&T must pay 50 basis points to swap its euro bond issuance proceeds to dollar. Effectively, AT&T observes a net issuance cost saving of $c - b = 25$ basis points by issuing in EUR instead of USD. Given this cost saving and absent any firm capital structure frictions, AT&T would choose to conduct its entire debt capital raising in EUR instead of USD. That is, the firm chooses dollar issuance share μ to minimize cost

$$\min_{\mu}\left(\underbrace{-c}_{\text{credit spread diff.}}+\underbrace{b}_{\text{CIP/hedging cost}}\right)\mu D$$

where D is the total amount of debt that needs to be raised.

Two predictions emerge immediately from this simple setup. First, if the net deviation (the effective credit spread difference) is negative, $c-b<0$, then the firm chooses $\mu=0$, otherwise it chooses $\mu=1$. More generally stated, issuance capital flow responds to the net deviation of credit and CIP violations. Second, if the total amount of debt D is large, then $c-b$ is driven to zero in general equilibrium. That is, the two deviations are perfectly aligned when the capital available for cross-market arbitrage is large.

In this model, I assume for simplicity that UIP holds (to focus on CIP), firms always currency-hedge when issuing abroad, and that there are no capital structure frictions to prevent firms from issuing all of its debt in one currency versus another. These assumptions can all be relaxed without changing the main results. I provide an extended model in the appendix that provides an interior solution to μ and yield similar predictions. For expositional purpose, I continue with the simple version of the firm's decision.

1.4.2 Credit markets

While the above setup generates simple intuitions for the alignment and elimination of the two types of LOOP violations, understanding how deviation in one market spills over to the other requires endogenizing the two violations. We start with endogenizing c.

There are two credit markets (EUR and USD bond markets), and three main credit market players: active local investors in Europe, active local investor in the U.S. and the representative firm from above that has access to both debt markets.

Local investors U.S. active investors specialize in the investment of corporate bonds denominated in dollars, and European investors only invests in EUR denominated bonds. Investors borrow at the domestic short rate, r_i, and purchase bonds with a promised net yield of Y_i, where $i = EUR$ or USD. The two bonds have identical default probability π, loss-given-default L. The payoff of bonds has a variance of V, which is treated as an exogenous constant in the model for tractability[12]. U.S. and European investors have a mean-variance preference with identical risk tolerance τ and choose investment amount X_i to solve the following

$$\max_{X_i} \left[X_i \left((1-\pi) Y_i - \pi L - r_i \right) - \frac{1}{2\tau} X_i^2 V \right] \tag{1.4}$$

which has the solution $X_i = \frac{\tau}{V} \left((1-\pi) Y_i - \pi L - r_i \right)$ for $i = EUR$ or USD.

Market clearing conditions In addition to active local investors, there are exogenous euro-relative-to-dollar bond demand ε_c, perhaps representing demand shocks that originate from Quantitative Easing or preferred-habitat investors with inelastically demands such as passive pension funds. The sources of exogenous ε_c shocks are discussed in Section 4. Combining the demand with firm debt issuance supply defined earlier, the market clearing conditions for the dollar and euro credit markets are

$$X_U = \mu D \tag{1.5}$$

$$X_E + \epsilon_c = (1-\mu) D. \tag{1.6}$$

We can rewrite the difference between the two promised yields as a credit spread difference and interest rate difference, $Y_E - Y_U \equiv c + (r_E - r_U)$. Combining the investor demands with the market clearing conditions and applying first-order taylor approximation for π around 0, we can express credit spread differential as:

[12] A Bernoulli default distribution with probability π, loss-given-default L and promised yield Y implies that $V = \pi (1-\pi) (Y+L)^2$. The solution to the investors' problem would contain a quadratic root. To keep the model tractable, V is assumed to be an exogenous constant and the same for both EUR- and USD-denominated bonds.

$$\underbrace{c}_{\substack{\text{credit spread}\\ \text{differential}\\ \text{(eu-us)}}} = \underbrace{\frac{V}{\tau}}_{\substack{\text{elasticity of}\\ \text{bond demand}}} \underbrace{\left(\underbrace{(1-2\mu)\,D}_{\text{relative debt issuance}} - \underbrace{\varepsilon_c}_{\text{exog. eur bond demand}} \right)}_{\substack{\text{net bond supply}\\ \text{eur relative to usd}}} \tag{1.7}$$

c represent a LOOP violation in credit since the default probability and loss given default are identical for the two bonds. The intuition is that c is determined by the net supply and demand imbalances between the two markets multiplied by the elasticity of bond demand.

The cross-currency issuer has limited ability to influence the relative credit spread. If it chooses all of its debt to be issued in euro instead of dollar, i.e. $\mu = 0$, then the relative credit spread in euro would widen (c increases) as a result of the additional debt supply. The issuer's impact is limited, however, by the size of its total debt issuance D.

1.4.3 Currency swap market

Next, I endogenize CIP basis b and describe the dynamics of the currency swap market. The intuition is essentially similar to that of credit LOOP violation, but instead of risk preference that determines the slope of demand curve, arbitrage in CIP is limited by intermediary collateral and capital constraints. There are two main players in this market: currency swap traders and issuers.

Currency swap traders Currency swap traders choose amount of capital to devote to either CIP deviations, denoted as b, or alternate investment opportunity with profit of $f(I)$, where I is the amount of investment. b is defined in the same way as in Section 3.

The arbitrageur has to set aside a haircut H when it enters the swap transaction to arbitrage CIP violation. Following Garleanu and Pedersen (2011), the amount of haircut is assumed to be proportional to the size s of the swap position, $H = \gamma|s|$. Therefore, the capital devoted towards alternative investment is $I = W - \gamma|s|$. Swap traders has total wealth W and solve the following

$$\max_{s} bs + f\left(W - \gamma|s|\right)$$

which generates the intuitive result that the expected gain from conducting a unit of additional CIP arbitrage is equal to marginal profitability of the alternative investment, $b = sign[s]\gamma f'\left(W - \gamma|s|\right)$. A simple case is when the alternative investment activity is quadratic, $f\left(I\right) = \phi_0 I - \frac{1}{2}\phi I^2$. In this case, $b = sign[s]\gamma\left(\phi_0 - \phi W + \gamma\phi|s|\right)$.

I make an additional simplifying assumption that CIP deviation b disappears when there is no net demand for swaps, but as soon as there is net demand for swaps, b becomes non-zero. This assumption is equivalent to stating $\frac{\phi_0}{\phi} = W$, which means that arbitrageur has just enough wealth W to take advantage of all positive-NPV investment opportunities in the alternative project $f\left(I\right)$. Simplifying with this assumption remove the constant intercept term in the equation for b, and we obtain that CIP deviation is proportional to swap trader position, $b = \phi\gamma^2 s$. I further normalize $\phi = 1$. This model of swap traders is analogous to that of Ivashina, Scharfstein, and Stein (2015) which models the outside alternative activity of the trader with a log functional form instead of the quadratic form used here.

Equilibrium The representative firm from earlier relies on FX market to hedge its foreign debt issuance. It swaps its euro issuance proceed amount $D\left(1 - \mu\right)$ to dollar. In addition, there are exogenous shocks to CIP basis ε_b that represent other non-issuance-related use of FX-swaps. The sources of shocks are discussed in Section 4.

Market clearing condition of the FX swap market implies that the equilibrium level of CIP deviation satisfies

$$\underbrace{b}_{\text{CIP basis}} = - \underbrace{\gamma^2}_{\substack{\text{haircut} \\ \text{on collateral}}} \underbrace{\left(D\left(1-\mu\right)+\varepsilon_b\right)}_{\substack{\text{net hedging demand} \\ \text{(swap euro to dollar)}}} \tag{1.8}$$

The negative sign arise since the swap trader takes the opposite position of the hedging demand. CIP deviation b is proportional to net hedging demand multiplied by the elasticity of supply, which is determined by the collateral margin. Higher haircut γ amplifies the impact of hedging demand, but without net hedging demand, b does not deviate from zero.

One additional insight on the role of the issuer in the above setup is that debt issuer hedging demand $D\left(1-\mu\right)$ does not have to have the same sign as other exogenous hedging demand ε_b. If ϵ_b has the opposite sign as and larger in magnitude than the issuer demand, the issuer would incur an additional benefit (instead of cost) through hedging. In this case, the firm would contribute to the elimination of CIP deviation and act as a supplier of liquidity in the currency forward market.

An extension of the model with natural hedges hedging using the firm's real asset and cashflows in the foreign currency) and partial hedging is analyzed in the appendix, but it does not alter the main predictions in the model.

1.4.4 Summary of equilibrium conditions and predictions

The three equilibrium conditions are summarized below:

1. Credit spread differential (EU-US):

$$\underbrace{c}_{\text{credit deviation}} = \underbrace{\frac{V}{\tau}}_{\substack{\text{elasticity of} \\ \text{bond demand}}} \underbrace{\left(\left(1-2\mu\right)D-\varepsilon_c\right)}_{\substack{\text{net bond supply} \\ \text{in EUR rel. to USD}}}$$

2. CIP basis (negative means more costly to swap into USD):

$$\underbrace{b}_{\text{CIP basis}} = -\underbrace{\gamma^2}_{\substack{\text{elasticity of}\\ \text{fx swap supply}}} \underbrace{(D(1-\mu)+\varepsilon_b)}_{\substack{\text{net hedging demand}\\ \text{to swap euro to dollar}}}$$

3. Firm choice of dollar issuance ratio:

$$\mu = \begin{cases} 1 & \text{if } c - b > 0 \text{ cheaper to issue in dollar} \\ 0 & \text{if } c - b < 0 \text{ cheaper to issue in euro and swap to dollar} \end{cases}$$

With these equilibrium conditions, we can analyze the transmission of ε_c and ε_b shocks from one market to the other. A positive euro credit demand shock ε_c directly reduces credit spread differential c and net deviation $c - b$. In response to the falling cost of issuing in euro, the firm switches its dollar bond issuance to euro bond issuance, leading to a decrease in the dollar issuance ratio μ. As the firm issues more in euro and swaps the bond proceed back to dollar, the hedging demand then endogenously raises the cost of FX swapping from EUR to USD, resulting in a decrease in b. Thus, a credit demand shock is transformed into a deviation from CIP. c and b both decrease due to a positive ε_c shock.

Conversely, a positive demand shock for dollar liquidity, ε_b, can also spillover to the credit market. An increase in the exogenous demand for swapping euro into dollar directly reduces b, raising the hedging cost of issuing in euro. As the effective cost of euro issuance $c - b$ increases, the firm issues more in dollar, raising μ. This increase in supply in turn widens the credit spread in dollar, reducing c. Therefore, the shock to CIP is transmitted to credit market. As with the ε_c shock, an ε_b shock also induces c and b to commove in the same direction.

While these transitions occur discretely at the boundary when $c-b$ flips sign, a small amount of friction to the firm's capital structure would generate a continuous spillover of deviations as shown in the appendix.

The above analysis can be stated more formally as the following propositions.

Proposition 1. *(Spillover of deviations)* If $\varepsilon_c \uparrow$, then $c \downarrow \Rightarrow \ \mu \downarrow \Rightarrow \ b \downarrow$. If $\varepsilon_b \uparrow$, then $b \downarrow \Rightarrow \ \mu \uparrow \Rightarrow \ c \downarrow$. Shocks to one market are transmitted to the other through capital flows. Credit spread differential c and CIP deviations b respond in the same direction to either credit demand shocks ε_c or FX swap demand shocks ε_b. Dollar issuance share μ responds differentially to the two shocks.

While Proposition 1 has a clear prediction for the signs of c and b, the sign of μ is ambiguous without precisely distinguishing whether the shock originates from ε_c or ε_b. However, the correlation between μ and the net deviation $c - b$ is unambiguous and testable, which leads to the following prediction.

Proposition 2. *(Issuance flow and net deviation)* $(c - b) \downarrow \implies \ \mu \downarrow$ Cheaper net cost of issuance in euro induces more issuance flow in euro and less issuance in dollar.

Another related prediction that follows from the above is that more cross-market arbitrage capital reduces the net deviations and the two deviations are perfectly aligned in the limit.

Proposition 3. *(Arbitrage capital and aligned deviations)* $\frac{\partial |c-b|}{\partial D} < 0$ and $\lim\limits_{D \to \infty} c - b = 0$. An increase in the total amount of debt issuance decreases the absolute value of the net deviation. As the total amount of debt increases towards infinity, the two deviations become identical.

Proposition 4. *(Limits to arbitrage spillover)* Additional comparative statics of the model are summarized in the following table:

	FX haircut $\gamma \uparrow$	Credit investor risk tol. $\tau \uparrow$	bond risk V$\uparrow$
$\lvert c \rvert$	$\uparrow$	$\downarrow$	$\uparrow$
$\lvert b \rvert$	$\uparrow$	$\downarrow$	$\uparrow$

Proposition 4 suggests that limits of arbitrage are carried over from one market to the other. For instance, while the amount of haircut on FX swap trades, γ, directly affects CIP basis b, γ also affects the credit spread differential c indirectly through the cross-market arbitraging firm. Similarly, the risk tolerance of localized bond investors that do not engage in FX swaps also affects the level of CIP deviation through capital flow. Thus, limits of arbitrage can spill over to a completely different market.

On the surface, the prediction of aligned deviation might appear to be similar to implications of intermediary-based asset pricing models that have a single intermediary trading in multiple markets. To distinguish my explanation from those of intermediary-based asset pricing, I discuss the falsifiable alternative below.

1.4.5 Falsifiable alternative

The model developed above is also useful for assessing alternative explanations of the alignment between the two LOOP violations. One alternative hypothesis relies on intermediary-based asset pricing: deviations might be correlated when there are fluctuations in the binding constraints for a common intermediary that operates in both markets. That is, arbitrageurs face the same constraint to arbitrage in credit and CIP, and a shock is delivered to this constraint. An equivalent way of stating this hypothesis in the framework of my model is to set $\gamma^2 = \frac{V}{\tau} \equiv \lambda$ and suppose there is a shock to λ.

There are two reasons for why this alternative hypothesis would not explain the alignment of the credit and CIP violations. First, absent of net demand imbalances in each market, changes in λ would not *cause* deviations to occur; it would only *amplify* the effect of demand imbalances. Second, while the *absolute value of deviations* would be correlated through intermediary capital, i.e. $\frac{\partial |b|}{\partial \lambda} \propto \frac{\partial |c|}{\partial \lambda}$, changes in λ would not explain the high alignment in the *direction* and *magnitude* of the deviations in b and c. Fluctuations in the common constraint

λ are therefore distinct from a spillover of deviation and frictions from one market to the other. Furthermore, one would not expect to observe changes in capital flow as represented by μ under this alternative explanation.

1.5 Discussions

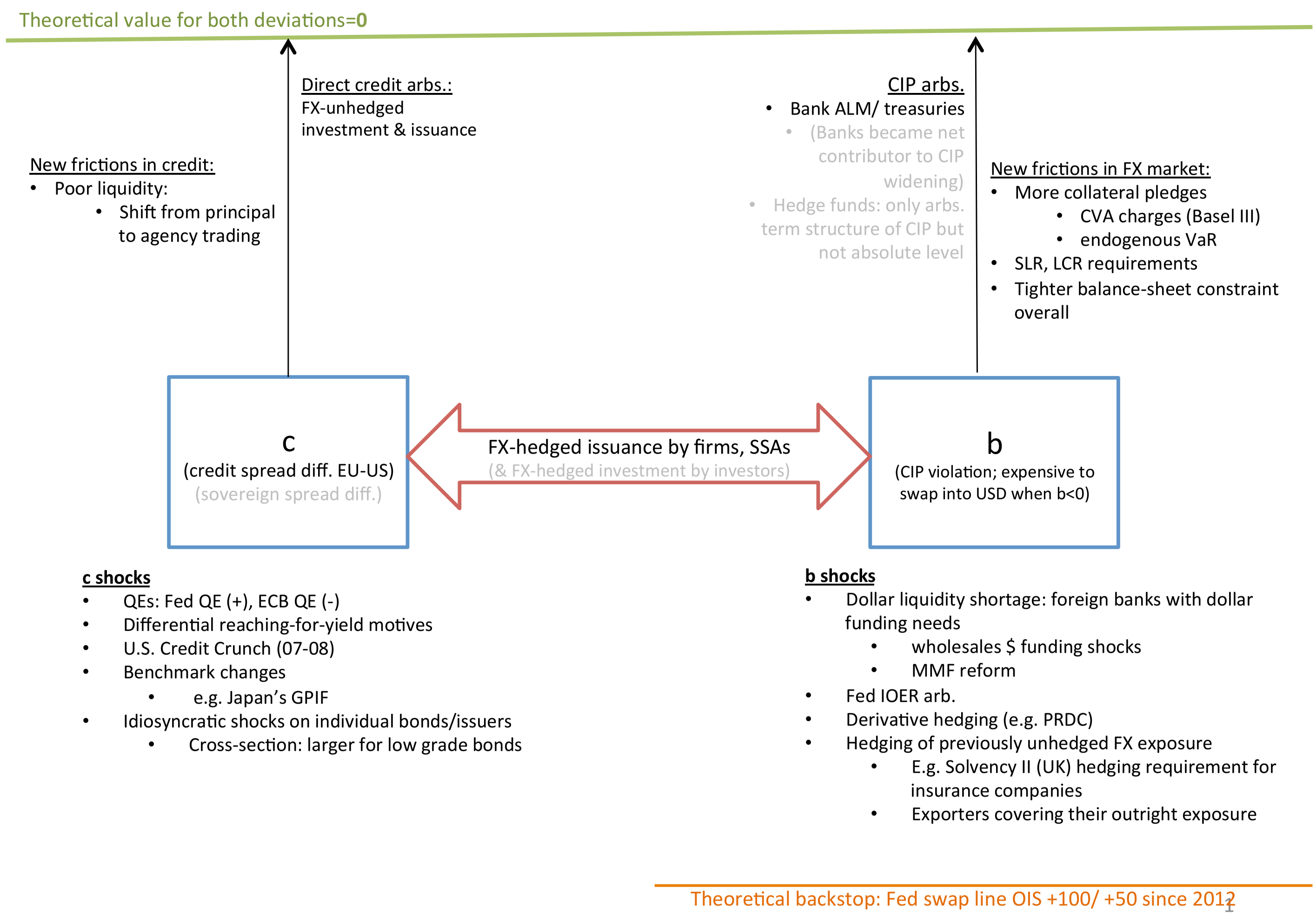
In this section, I discuss the sources of shocks, limits to arbitrage in each market and why firms are natural cross-market arbitrageurs. The schematics in Figure 1.7 summarizes the discussion.

1.5.1 Source of ε_c and ε_b shocks

1.5.1.1 ε_c shocks

- **Central bank QE** Large asset purchasing programs by central banks have contributed to the displacement of traditional government debt investors in search of high-yielding assets such as corporate bonds. The differential timing and sizes of ECB and Fed quantitative easing programs likely changed the relative demand for credits in Europe and the U.S., resulting in changes in ε_c.
- **Passive investor portfolio changes** Shifts to passive institutional investor's benchmarks and portfolios can bring large changes to the demand for assets. Portfolio benchmark changes can be distinct from shifts in the investment of active investors presented in the model due to their slow decision making process and a number of intuitional constraints. For instance, Japan's Government Pension Investment Fund, which holds US$1.2 trillion in asset and serves as the most frequently used portfolio benchmark for other Japanese-based asset managers, decided in October 2014 to reduce its domestic bond holding from 60% to 35% and increase its allocations to stocks

Figure 1.7: Schematic of institutional details



and foreign assets. This large, one-time portfolio shift differs from that of active credit specialists who decide on bond investments based on credit risks at higher frequencies.

- **Regulatory-driven demand shocks** Portfolio shifts can also be driven by regulatory reforms. One such regulatory change occurred in the United Kingdom, where the 2005 Pension Reform Act forced pension funds to mark their liabilities to market by discounting them at the yield on long-term bonds. This reform significantly increased the demand for long-term securities (Greenwood and Vayanos 2010).

- **Credit-market sentiments** A number of papers have analyzed the role of credit sentiment on asset prices and the real economy (López-Salido, Zakrajšek and Stein [2015], Bordalo, Gennaioli, and Shleifer [2016], Greenwood, Hanson, and Jin [2016], Greenwood and Hanson [2014]). A shock to the relative credit demand between bond markets can arise if credit sentiments differentially impact different markets. One such episode occurred around the time of the Bear Stearns collapse, when the *residualized* dollar credit spread widened relative to the euro credit spread as fears of US credit market meltdown heightened.

1.5.1.2 ε_b shocks

- **Dollar liquidity shortage** Since the crisis, non-U.S. banks, in need of short-term dollar funding for their U.S. operations, have become active borrowers of dollar through FX swaps[13]. A particularly striking episode of demand shock for FX swaps into dollar is during the Eurozone Sovereign Crisis in 2011-2012. Dollar money-market funds stopped lending to European banks in of fear of fallouts from the sovereign crisis. The

[13] Banks do not all have dollar liquidity shortage (i.e. ε_b could also be negative). For instance, in Australia, banks need to fund abroad their long term needs as the base of investors lending long-term is small. They borrow in USD or EUR and swap it back in AUD. CIP deviations in AUD indicates that it is more expensive to swap into AUD instead of the other way around (due to the negative ε_b shock). This demand is partially captured in my data on corporate debt issuance since the Australian banks fund both through long-term debt market and short-term money market.

swapping of deposits and wholesale fundings by banks are typically concentrated in short maturities.

- **Money market reform** in the U.S. that took effect in October 2016 has reduced the availability of wholesales dollar funding to foreign banks and increased their reliance on funding via currency swaps (Pozsar and Smith 2016).
- **Structured note issuers** also utilize currency swaps in the hedging of ultra long-dated structured products whose payoff depends on exchange rate at a future date. The convexity embedded in these notes produced enormous hedging needs in FX forwards under certain market conditions for AUD, JPY, and other Asian or Pacific currencies. In particular, the hedging of Power Reverse Dual Currency Notes by issuers had been an important driver of currency basis in AUD, JPY and other Asian currencies.
- **Regulatory-driven hedging demands** New regulatory requirements for the hedging of previously under-hedged exposures also have been a factor driving the CIP basis. Solvency II Directives on E.U. and U.K. insurance companies demanded greater usage of longer-dated cross-currency basis swaps to reduce foreign currency exposure of insurance firm asset holdings[14]. The Solvency II rules started with initial discussions in 2009 and finally took effect in 2016. Regulatory reforms are generally slow and filled with uncertainty during the interim.
- **Central bank policies** European banks with EUR excess liquidity have been able to take advantage of the higher Interest on Excess Reserve (IOER) rate offered by the Fed by lending their EUR through FX swap and use the resulting USD to lend at the IOER. As of September 2016, foreign bank offices in the U.S. have a total excess reserve at the Fed of $766 billion, of which $429 billion[15] are funded through Fed Fund and Repo agreements as a part of the IOER-Fed Fund arbitrage.[16] This leaves the

[14]Previously, insurance firms partially hedged using rolling short-dated FX forwards
[15]Flow of Funds Table L.112
[16]Foreign bank branches can fund at the lower Fed Fund rate and lend at the IOER without paying FDIC

remaining $337 billion as currency-swapped liabilities from abroad. This motive is best described with a quote from an European bank executive:

> In response to the ECB's move to adopt negative rates on bank deposits [...] Rabobank Group, one of Europe's best-capitalized banks, said it has withdrawn a total of €40 billion in recent months and moved it to other large central banks like the Bank of England, the Swiss National Bank and the Federal Reserve."At least there, you don't have to pay to park your money," said Chief Financial Officer Bert Bruggink. (WSJ, August 2014)

The policies at other central banks also had impacts on CIP violations. For example, the termination of ECB's sterilization programs reduced the amount of High Quality Liquid Asset for European banks and were a contributing factor to the widening of the CIP violation in 2014[17].

- **Hedging demand from investors** I do not consider this as an ε_b shock since the issuers in my model can be broadly interpreted as both sellers and buyers of bonds. Another reason why investors are not a major contributor to long-term CIP violations is that they often hedge FX risk using rolling short-dated forwards[18].

1.5.2 Limits of arbitrage

To understand why the credit and CIP violations exist, we must understand who are the arbitrageurs in each market and the constraints that they each face. These constraints

assessment cost since they are uninsured. This is known as the IOER-Fed Fund arbitrage for foreign banks.

[17] ECB's Security Market Program that started in 2010 and the Outright Monetary Transaction program that started in 2012 both were initially sterilized purchasing programs. Sterilization encouraged the use of ECB excess reserved and provided a way for banks to obtain HQLA (High Quality Liquid Asset) needed to fulfill LCR (Liquidity Coverage Ratio) requirements. The end of ECB sterilization in 2014 meant that European banks needed to look for other HQLA to replace around $200 billion of ECB excess reserve. Therefore, these banks had to either invest in Euro assets or swap into other currencies and park their cash at the Fed or other central banks.

[18] Most benchmark indices calculate total returns on foreign sovereign and corporate bonds either as unhedged returns or hedged returns using 1-month rolling FX forwards. Bank of America Merrill Lynch, Barclays, and Citi each state in their index methodology that 1 month rolling forwards are used in the calculation of total returns for currency hedged indices. Longer horizon FX hedges are sometimes used but generate tracking errors from benchmark for investors. Of course, the long- and short- dated CIP basis are integrated to a certain extend as discussed below.

are represented in the model by the elasticity of supply and demand curve, γ^2 and $\frac{V}{\tau}$, but they take on realistic interpretations in practice. The main conclusion from the following discussion is that post-crisis regulatory restrictions and intermediary frictions have severely hindered arbitrage in the FX and credit markets each on their own, but capital flows (from either issuers or investors) bundle together the two deviations.

1.5.2.1 Why CIP deviations cannot be eliminated alone?

Unlike the textbook notion of costless arbitrage, eliminating CIP violations in practice is a very capital-intensive transaction. Suppose one were to arbitrage the CIP violation in EURUSD, when reduced to the simplest form, even deploying the strategy on CIP deviations at the 1-day horizon requires the delivery of large amount of cash in dollar and receiving a large amount of cash in euro today and reversing the transaction tomorrow. The problem is that the arbitrageur needs to 1) fund this large amount of dollar in cash and 2) invest the large amount of euro that is received. If one were able to do (1) and (2) costlessly at either the Libor rate (or the Overnight Index Swap rate), then CIP deviations would easily be eliminated. Below I discuss and rule out possible arbitrageurs:

- **Banks** Traditionally, depository institutions' Asset Liability Management desks eliminated CIP deviations by flexibly lending out their balance sheets as needed. However, few institutions are able to do so today in the post-crisis environment with tightened balance-sheet constraints. On the contrary, as discussed earlier, banks had become a net contributor to CIP violation as they themselves rely on FX swaps to fund in different currencies.

- **Hedge funds** are often mistakenly viewed as a source of arbitrage capital for eliminating CIP violation. In reality, hedge funds only integrate the *term structure* of the currency forwards but provide little mitigation of the *outright level* of deviation from

CIP. This is because outright arbitrage of CIP is a capital-intensive transaction that requires the physical delivery of cash. It is impossible for hedge funds to obtain funding at Libor or OIS rates[19]. The key point is that low-risk, balance sheet intensive activities are costly to conduct. Instead, hedge funds transmit shocks across the maturity curve of CIP deviations by entering into forward starting cross-currency basis swaps that do not have physical exchanges of notional, and they unwinds the trade well-ahead of the actual delivery of cash. This form of term structure integration can be modeled similarly as Vayanos and Vila (2009) and Greenwood and Vayanos (2014).

- **Debt issuers and investors** The ability to borrow and to invest large amount of cash in a deep market is a defining characteristic of the debt capital markets. Therefore, It is natural to expect issuers and investors to play a large role in eliminating CIP violation. This is precisely why CIP violation is linked to corporate credit spread differential (and sovereign spread differentials to some extent[20]).

More stringent regulatory requirements have also raised the cost of arbitraging CIP deviations. In other words, γ has increased. Many of the regulatory change came about because of large losses by certain financial institutions. In this sense, the margin on trades arose endogenously a la Geanakoplos (2010) and further exacerbated the violations. Prior to 2008, many of the FX derivative instruments related to forward exchange rate required little collateral and margining, since then, the trading of these derivatives are much more prohibitive in balance sheet requirements. Specifically, Supplementary Leverage Ratio has increased the cost of holding low-risk positions. Mandatory margining by different local regulator and other Basel III rules has also increased the cost of trading FX swaps. An alphabet soup

[19] Alternatively, using equity capital from investors to arbitrage CIP earns unattractive returns

[20] While the government bond market is more liquid, developed market sovereigns seldomly issue in foreign currencies with the same covenants as their domestic bonds. Sovereigns can also choose to default on foreign bonds without defaulting on domestic bonds. Investors would face different sovereign risk if they were to bundle together the arbitrage of CIP violation with government debt investments. On the other hand, bonds issued by corporates and supranational in multiple currencies have the same underlying credit risks across denominating currencies, therefore, corporate debt is a natural choice for facilitating CIP arbitrage.

of different funding costs has also emerged[21] in response to post-financial-crisis regulatory and market environment. Relatedly, Levich (2012) finds that trading in over-the-counter currency forward has declined in favor of currency futures. In short, there are hefty costs to low-risk, low-return projects.

1.5.2.2 Why credit spread differential cannot be eliminated alone?

With a distortion in CIP, credit spread differential along currency lines cannot be eliminated unless issuers or investors forgo currency hedging. A simple long-short strategy in the bond market alone would incur large amount of currency mismatch. Given the high levels of FX volatility (e.g., EURUSD annualized volatility has averaged 10% since 2004), few investors and issuers would forgo the hedging to earn the credit spread differential. Hedging for the FX exposure, however, requires arbitrageurs to be exposed to CIP violations. All of the constraints in the FX forward market are thus carried over to the credit market.

Furthermore, bond market liquidity conditions have worsened in recent years. The shift from principal-based to agent-based market-making by dealers has increased the cost of transacting in large sizes and lengthened the amount of time it takes to execute large trades. Regulatory rules affecting funding have also contributed to a reduction in market liquidity, as emphasized in Brunnermeier and Pedersen (2009).

1.5.3 Firms as natural cross-market arbitrageurs

Having discussed the constraints and the lack of arbitrageurs in the credit and CIP market each on their own, we turn towards understanding cross-market arbitrageurs between credit

[21] These funding costs include CVA (Credit Valuation Adjustments) that accounts for counter-party default risk, KVA (Capital Valuation Adjustment) imposed by banks on clients to account for the lifetime capital consumption of individual trades, MVA (Margin Valuation Adjustment) that adjusts for interest earned on the initial margin to reflect interest on investments of similar risk elsewhere, and FVA (Funding Valuation Adjustment) that adjusts for differential funding rates associated with derivative collateral posting. Collectively these are known as XVAs.

and CIP. While the cross-market arbitrageurs in the model can be interpreted as global investors as well as firms, I focus my analysis on firms for two reasons. First, bond issuance data is easily obtainable. This data allows the testing of model predictions on capital flow, shock transmissions, and deviation elimination. Second, firms are natural cross-market arbitrageurs that can better withstand noise trader shocks and more easily overcome limits of arbitrage problems raised by Shleifer and Vishny (1997). This point had been argued by previous papers including Baker and Wurgler (2000), Greenwood, Hanson, and Stein (2010), and in particular, Ma (2015) explores the role of firms as cross-market arbitrageurs in their own equity and debt securities.

To observe issuance flow as arbitrage capital, it must be the case that investors are not supplying sufficient arbitrage capital. Why might investors be constrained in performing the arbitrage? While many institutional investors such as pension funds, life insurance companies and endowments have diversified exposure to bonds in different currencies, they often have clear mandates on their benchmarks and currency exposure. The rigidity of their mandates allow for little discretion in their portfolio allocation choice. They are also often limited in their usage of derivatives due to the lack of expertise and regulatory restrictions. Mutual funds and hedge funds in fixed income also typically follow benchmarks. Unrestricted global funds are limited in size. For instance, global retail bond fund holds only a total of €55 billion of EUR corporate bonds[22]. The small number of hedge funds that do engage in the active trading of foreign credit markets face balance-sheet constraints as discussed earlier and high transaction costs in long-short strategy. This is because a long-short strategy requires conducting repo in one market and reverse-repo in the other market to fund the bond positions while also engages in FX hedging. Limits to arbitrage associated with investor redemption and short investment horizon as highlighted in Shleifer and Vishny (1997) pose a challenge to all specialized funds that perform arbitrage. In short, dedicated investors

[22]EPFR data

simply do not have enough capital or risk tolerance to digest large demand shocks.

Firms are natural arbitrageurs to exploit capital-intensive, slow-convergence arbitrage opportunities. They have the ability to bear noise-trader risk, withstand large mark-to-market losses and endure long investment horizons. Because firms have stable cash flows and do not face redemptions, making a one time issuance and hedging decision is equivalent to holding the arbitrage trades to maturity. The standard deviation of monthly issuance flow between the Eurozone and the U.S. is in excess of $6 billion. This is equivalent to the creation of a sizable hedge fund fully dedicated to exploiting the two LOOP violations every month.

1.5.3.1 Evidence from textual analysis of SEC filings

I conduct a textual analysis of SEC filings by S&P 500 firms that is indicative of the pervasive use of currency-hedged debt issuance. Figure 1.8 shows the result of this analysis. I graph the fraction of 10K filings with mentions of words relating to 1) "debt", 2) "exchange rate", 3) "hedging" and 4) "derivatives" in the *same* sentence. The restriction of having all four groups of words to appear in a single sentence likely under-estimates the actual disclosure of currency-hedged issuance since the disclosure could be relayed in multiple sentences. While this proxy might be imperfect, it nonetheless indicates that a substantial fraction of S&P 500 firms had engaged in currency-hedged issuance in recent years. The sharp rise in this proxy from 2007 to 2010 corresponds to the period when deviations in the credit and CIP markets first begin to widen. This analysis of SEC filings shows the pervasiveness of firms acting as cross-market arbitrageurs between the credit market and CIP market in recent periods[23].

[23] Figure 1.8 also shows that a smaller fraction of firms have indicated currency-hedged issuance as early as 2004 even though both the CIP violation and the aggregate credit spread differentials were small prior to 2007. This is possibly explained by issuer-specific idiosyncratic credit spread differentials that did not appear in the aggregate.

Figure 1.8: Textual analysis of FX-hedged foreign debt issuance for S&P 500 firms
This figure presents a textual analysis of SEC filings for S&P500 firms that had indicated cross-currency debt issuance in their annual 10-K filings. Panel A shows three examples of firms that has mentioned in their SEC filings that they engaged in currency-hedged foreign debt issuance. Panel B presents the fraction of SEC 10K filings of S&P500 firms with mentions of words relating to 1) "debt", 2) "exchange rate", 3) "hedging" and 4) "derivative" in the same sentence by year.

Panel A: Examples of SEC filings with mentions of currency-hedged debt issuance

AMGEN

10K: "To hedge our exposure to foreign currency exchange rate risk associated with certain of our long-term notes denominated in foreign currencies, we entered into cross-currency swap contracts, which effectively convert the interest payments and principal repayment of the respective notes from euros/pounds sterling to U.S. dollars."

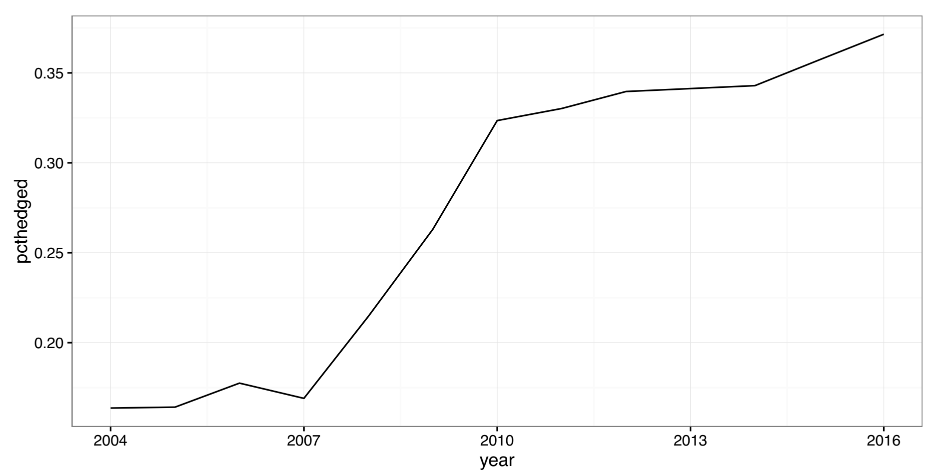
APPLE INC.

10Q: "In the first quarter of 2015, the Company issued €2.8 billion of Euro-denominated long-term debt. To manage foreign currency risk associated with this issuance, the Company entered into currency swaps with an aggregate notional amount of $3.5 billion, which effectively converted the Euro-denominated notes to U.S. dollar-denominated notes."

at&t

10K: "We have entered into multiple cross-currency swaps to hedge our exposure to variability in expected future cash flows that are attributable to foreign currency risk generated from the issuance of our Euro, British pound sterling, Canadian dollar and Swiss Franc denominated debt. "

Panel B: Fraction of 10K filings with mentions of currency-hedged debt issuance



1.6 Additional empirical results

In this section, I take the model to the data. I first describe the issuance data, the measurement of net deviations, and patterns in the misalignment. Then I present supporting evidence for the model predictions.

1.6.1 Data and definition

1.6.1.1 Issuance flow μ

To test the model predictions on cross-currency capital flow, I analyze the amount of corporate debt issued by public firm in the seven free-floating funding currencies. Debt issuance amount and other bond characteristics are obtained from Thompson One SDC Platinum data set. I define the monthly bilateral issuance flow between two currency regions as the amount of debt issuance by foreign firms in dollar minus the amount of debt issuance by U.S. firms in that currency expressed as a percentage of total issuance. For instance, the issuance flow between Europe and the U.S. is expressed as

$$issPct^{EU \to US} = \frac{\text{EU firm issuance in dollar - US firm issuance in euro}}{\text{total issuance in dollar \& euro}}.$$

This measure of issuance flow proxies for μ in my static model.

1.6.1.2 Net deviation $(c - b)$

I define *net deviation* as the difference between the residualized credit spread differential and CIP violation, i.e. $c-b$. The easiest way to construct the net deviation is to directly subtract CIP deviations from the residualized credit spread differential. However, the maturity of FX forward used for hedging each individual bond is different. To construct a measure of the net deviation, I first adjust the swap yield curve by the corresponding CIP deviation maturity curve before linearly interpolating to each individual bond's maturity in calculating the bonds' effective credit spreads. Then I conduct cross-sectional regression as specified in Equation 1.1 using this effective credit spread as the dependent variable. I take the currency fixed effects as estimates of the net deviation that corrects for maturity mismatches between FX forwards and bonds. This procedure produces estimates of $c - b$ that is not too different from directly subtracting the 5-year CIP deviation from the credit spread differential.

Misalignment of LOOP violations The two violations are misaligned when the size of net deviation is large or when their correlation is low. Figure 1.9 shows the net deviation time series for each of the six currency pairs (relative to USD). Apart from the financial crisis period, the net deviation is much smaller in magnitude in comparison to either CIP deviation or credit spread differential alone. This indicates that the two violations in credit and CIP are generally well aligned in magnitude. The misalignment, however, is larger during the financial crisis. This is consistent with the model predictions that larger demand shocks in the FX and credit market, more risk aversion, and less debt issuance lead to larger misalignment between c and b. Credit spread differential had higher spikes during the peak of the crisis than CIP deviation for most currencies. This is in part because CIP deviations were eventually capped when the U.S. Federal Reserve established swap lines with other central banks for the lending of dollar funding to foreign institutions. On the other hand, credit market distortions were exacerbated during the financial crisis by the lack of liquidity in fixed income trading.

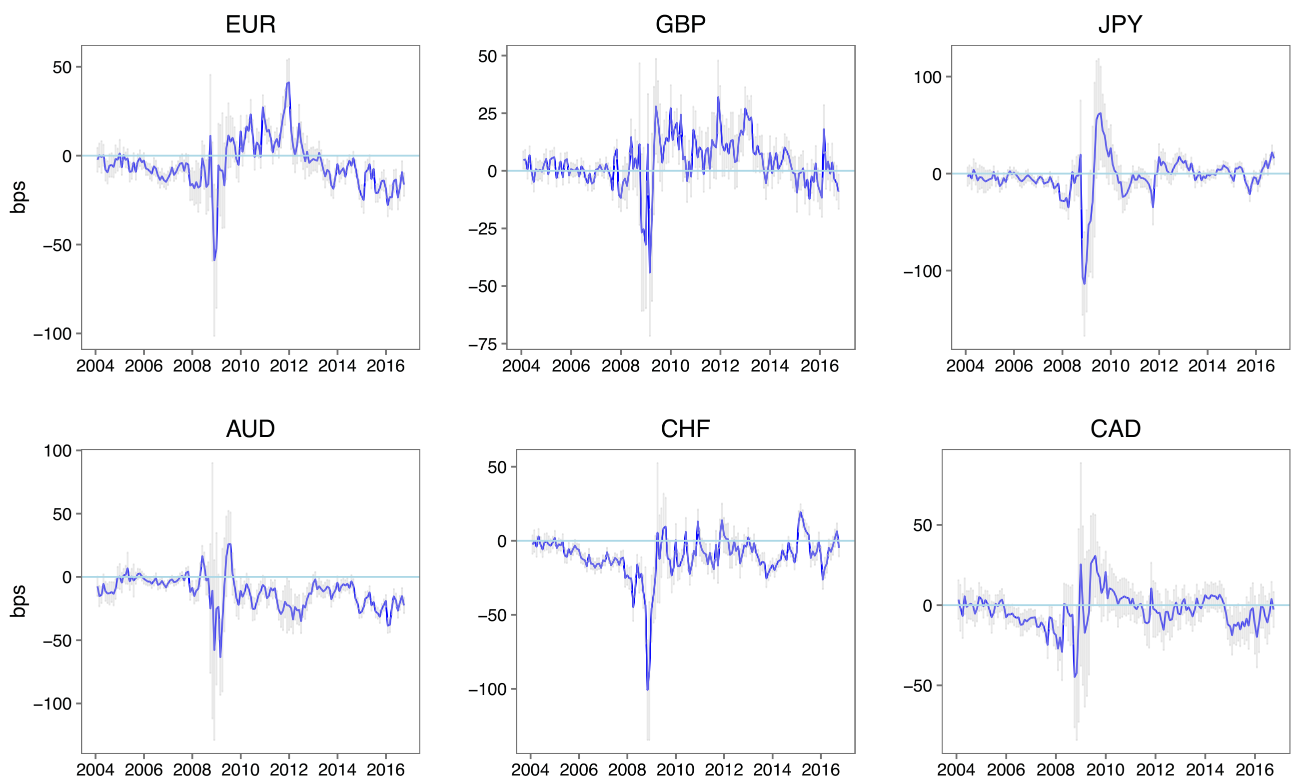
The net deviation represents the effective credit spread differential accounting for the hedging cost that firms observe. Thus, the net deviation time series make it obvious that while yen credit spread is much more compressed related to dollar as presented earlier in Figure 1.5, firms have little net incentive to issue in yen during most of the non-crisis period.

Figure 1.9: Net deviation

This figure presents the net deviation or the effective residualized credit spread (credit spread differentials minus CIP deviations with matching maturities) for EUR, GBP, JPY, AUD, CHF and CAD relative to USD. Vertical bars (grey) represent the 95% confidence interval for the estimated net deviation. To construct the net deviation, I estimate the following cross-sectional regression at each date t

$$S_{it}^{adj} = \alpha_{ct} + \beta_{ft} + \gamma_{mt} + \delta_{rt} + \varepsilon_{it}$$

where S_{it}^{adj} is the yield spread over the CIP-adjusted swap curve for bond i that is issued in currency c, by firm f, with maturity m and rating r. The CIP-adjustment is calculated by subtracting maturity-specific CIP deviation from each bond's yield spread. The net deviation or effective residualized credit spread for currency c relative to dollar credit spread is calculated as $\hat{\alpha}_{c,t} - \hat{\alpha}_{usd,t}$. Details of net deviation's construction are provided in Section 1.6.1.2.



1.6.2 Prediction 1: Spillover of deviations

I test the spillover of deviations through the channel of debt issuance by analyzing the impulse responses of credit spread differential c, CIP violation b, and issuance flow μ to ε_c and ε_b shocks. In addition, I provide interpretation of the time series magnitudes and lead-lags relationships.

1.6.2.1 VAR analysis

VAR analysis is useful in this context since the shocks to credit and CIP can occur simultaneously and transmission could be slow. As discussed in Section 1.5, there are many source of ε_c and ε_b shocks. These shocks can occur concurrently and might be anticipated long before the actual delivery, e.g. gradual regulatory changes. Furthermore, arbitrage capitals provided by non-specialized agents are often slow to react to market distortions due to inattention and institutional impediments to immediate trade (Duffie 2010). In this context, cross-currency issuance transmits the shocks gradually.

Figure 1.10 presents the orthogonalized impulse response functions with shocks to credit and CIP. The impulse response in this figure applies Cholesky Decomposition using a strict ordering of variables. I assume that issuance respond with a lag to both c and b, and b respond with a lag to c. That is, I estimate the following,

$$\begin{bmatrix} 1 & 0 & 0 \\ a_{c\mu} & 1 & 0 \\ a_{b\mu} & a_{bc} & 1 \end{bmatrix} \begin{bmatrix} \mu_t \\ c_t \\ b_t \end{bmatrix} = B \begin{bmatrix} \mu_{t-1} \\ c_{t-1} \\ b_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{\mu,t} \\ \varepsilon_{c,t} \\ \varepsilon_{b,t} \end{bmatrix}.$$

Proposition 1 states that an exogenous increase in the euro credit spread c (less demand of euro credit, $\varepsilon_c \downarrow$) raises dollar debt issuance μ and currency basis b (less FX swapping cost from euro to dollar) as firms avoid the higher credit spread in EUR and issue more in USD. The first row of Figure 1.10 confirms this model prediction. Upon a shock that increases

c(top left), both b (top middle) and μ (top right) are raised. Credit spread differential then gradually declines after the initial shock as do μ and b.

The slow responses of issuance flow μ and CIP deviation b to an ε_c shock are reflective of the slow moving nature of corporate financing decisions. The under-reaction of price movement in the market not directly affected by the shock, FX market in this case, is also a prediction of cross-market price dynamics with slow moving capital in a model developed in Greenwood, Hanson, and Liao (2016).

The bottom row presents the impulse responses with an exogenous increase in b that signals an increase in the cost of swapping dollar to euro. We observe the exact opposite dynamics in the second row as predicted by Proposition 1. Cost of swapping into euro initially is raised then gradually declines over time (bottom middle). The slow moving capital effect is also easily seen. Issuance flow initially shifts towards euro (bottom right) to take advantage of the lower cost of swapping into dollar before gradually normalizing over the next nine months. Credit deviations also increase gradually before plateauing around 6 months after the shock (bottom left).

Since it is ambiguous whether LOOP violation in CIP proceeds violation in credit risk pricing, I also consider an alternate ordering in which issuance respond with a lag to both c and b, and c respond with a lag to b. This alternate specification yields similar results as Figure 1.10 and is presented along with a partial identification approach[24] in the Internet Appendix. Furthermore, I conduct the same analysis on all six currency pairs against the dollar in a panel VAR. The resulting impulse response function is similar to that of EURUSD and is presented in Figure 1.11.

[24]The partial identification approach restricts μ to respond with a lag to c and b but allow the c and b to have contemporaneous effects on each other.

Figure 1.10: Spillover of deviations: orthogonalized impulse responses of deviations and issuance flow for EURUSD

I estimate a first order vector autoregression (VAR) of the form

$$\begin{bmatrix} 1 & 0 & 0 \\ a_{c\mu} & 1 & 0 \\ a_{b\mu} & a_{bc} & 1 \end{bmatrix} \begin{bmatrix} \mu_t \\ c_t \\ b_t \end{bmatrix} = B \begin{bmatrix} \mu_{t-1} \\ c_{t-1} \\ b_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{\mu,t} \\ \varepsilon_{c,t} \\ \varepsilon_{b,t} \end{bmatrix}$$

where μ_t is the bilateral issuance flow (defined in Section 1.6.1.1), c_t is the credit spread differential and b_t is the CIP deviation. I apply Cholesky Decomposition by ordering the variables as μ, c and b. This ordering assumes that issuance responds with a lag to both ε_c and ε_b shocks, and CIP violation respond with a lag to credit shock. The orthogonalized impulse responses to ε_c and ε_b shocks are graphed below. The choice of lag 1 is selected by Bayesian Information Criteria. 95% confidence intervals are shown in gray.

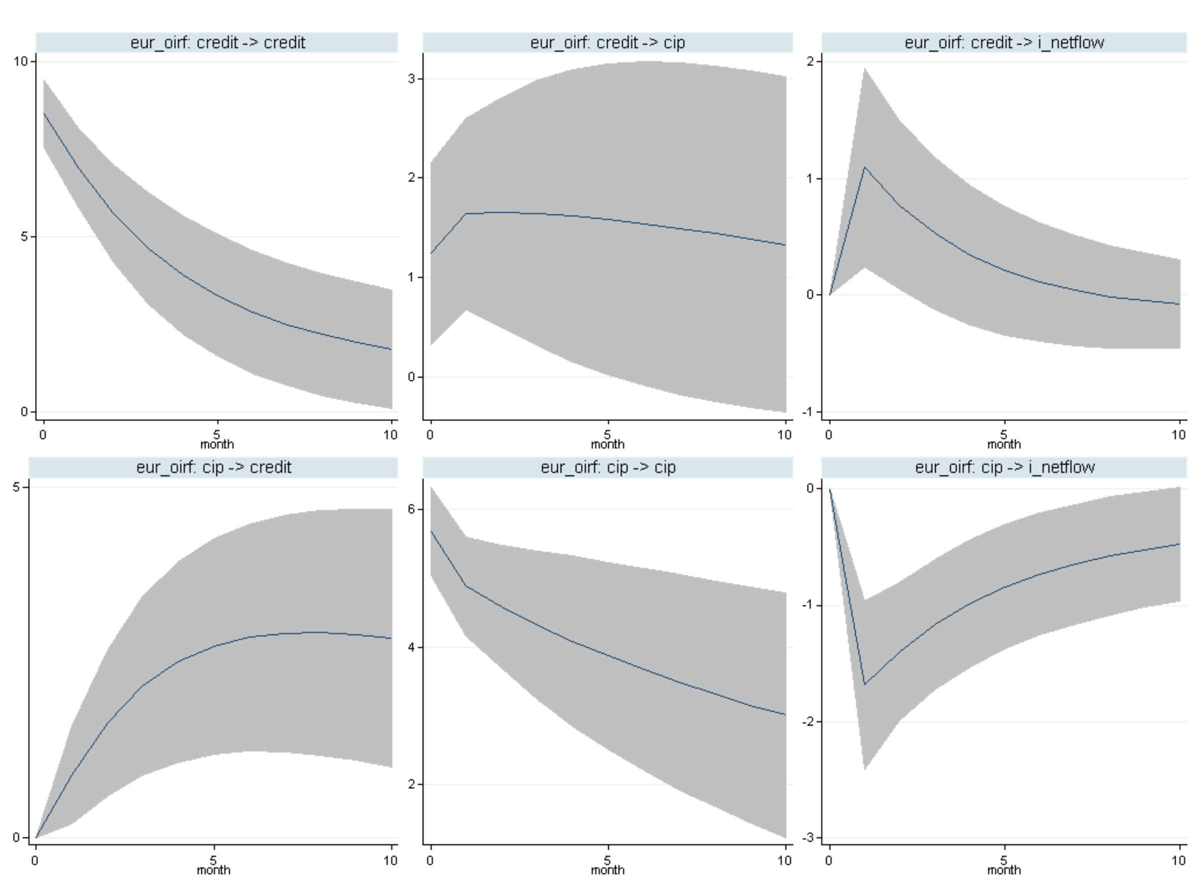
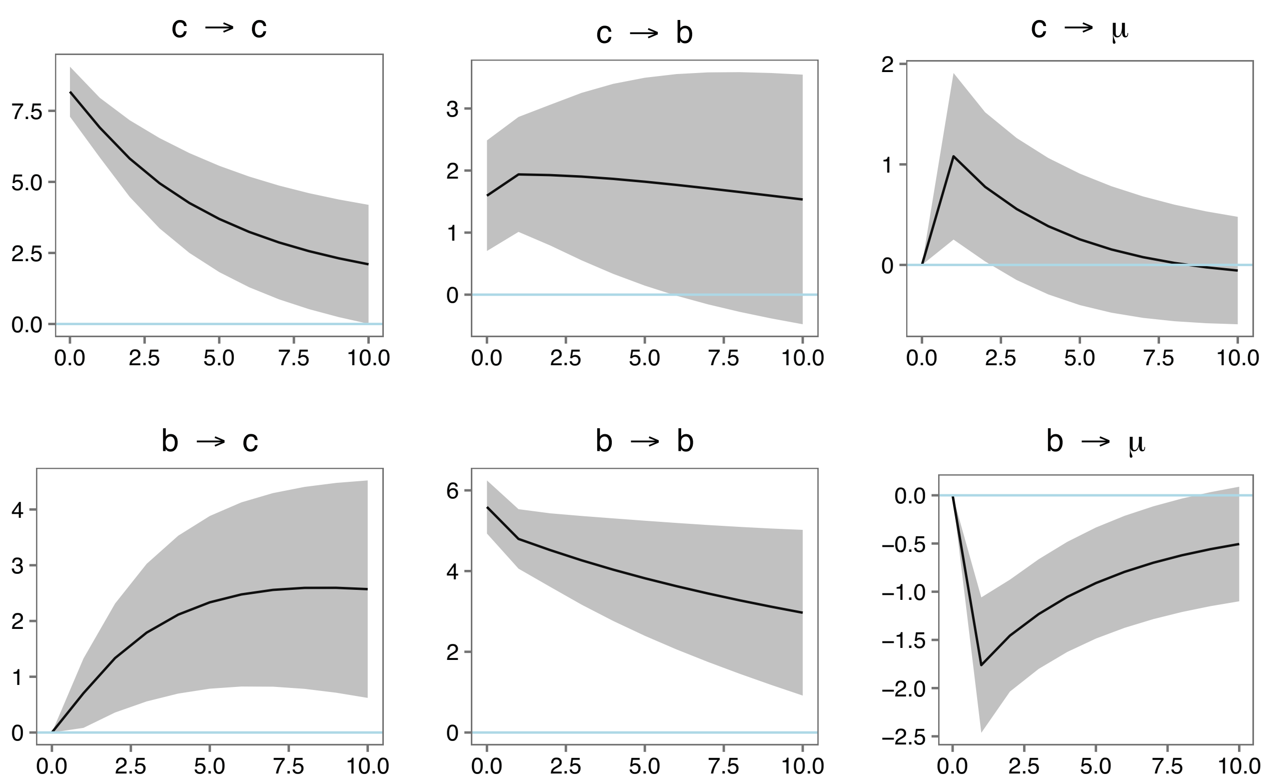


Figure 1.11: Spillover of deviations: Panel VAR

I estimate a first order panel vector autoregression (PVAR) for the six currency pairs (i = EURUSD, GBPUSD, JPYUSD, AUDUSD, CHFUSD, CADUSD)

$$\begin{bmatrix} 1 & 0 & 0 \\ a_{c\mu} & 1 & 0 \\ a_{b\mu} & a_{bc} & 1 \end{bmatrix} \begin{bmatrix} \mu_{i,t} \\ c_{i,t} \\ b_{i,t} \end{bmatrix} = B \begin{bmatrix} \mu_{i,t-1} \\ c_{i,t-1} \\ b_{i,t-1} \end{bmatrix} + \begin{bmatrix} \delta_{i,\mu} \\ \delta_{i,c} \\ \delta_{i,b} \end{bmatrix} + \begin{bmatrix} \varepsilon_{i,\mu,t} \\ \varepsilon_{i,c,t} \\ \varepsilon_{i,b,t} \end{bmatrix}$$

where μ_t is the bilateral issuance flow (defined in Section 1.6.1.1), c_t is the credit spread differential, b_t is the CIP deviation and δ_i is a vector of fixed effects. I apply Cholesky Decomposition by ordering the variables as μ, c and b. This ordering assumes that issuance responds with a lag to both ε_c and ε_b shocks, and CIP violation respond with a lag to credit shock. Confidence intervals at the 95% level using bootstrapped standard errors are shown in gray.



1.6.2.2 Time series

Beyond VAR analysis, the time series of the two LOOP violations are also informative in establishing the direction of spillover. While the ambiguity in the ordering of the LOOP violations poses a challenge to the VAR analysis, the changing lead-lag relationship between c and b, in conjunction with relative magnitude of the two deviations, in different periods provide valuable insights on identifying whether shocks might have originated from credit demand or FX forward demand. As seen in Figure 1.1, CIP deviation appears to have led the credit spread differential both in time and magnitude during the 2011-2012 Eurozone Sovereign Crisis that tightened foreign bank's wholesales dollar funding conditions[25]. In more recent periods, credit spread differential have overtaken CIP deviation in magnitude and time lead, potentially a reflection of credit demand shocks originating from ECB asset purchases.

1.6.3 Prediction 2: Issuance flow and net deviation

Another key prediction from the model is that capital flow fluctuates with net deviation. In the case of corporate arbitrageurs, capital flow is represented by cross-currency issuance.

I focus on bilateral issuance flows with the U.S. since the U.S. corporate bond market is the largest, with over a third of the global corporate debt issuance in the data sample. Figure 1.12 compares the quarterly time series of the issuance flow and net deviation for EURUSD. Consistent with the model prediction on the comovement between μ and $c-b$, issuance flows from Europe to the U.S. when the effective residualized credit spread of euro-denominated debt is high relative to dollar-denominated debt, and vice versa.

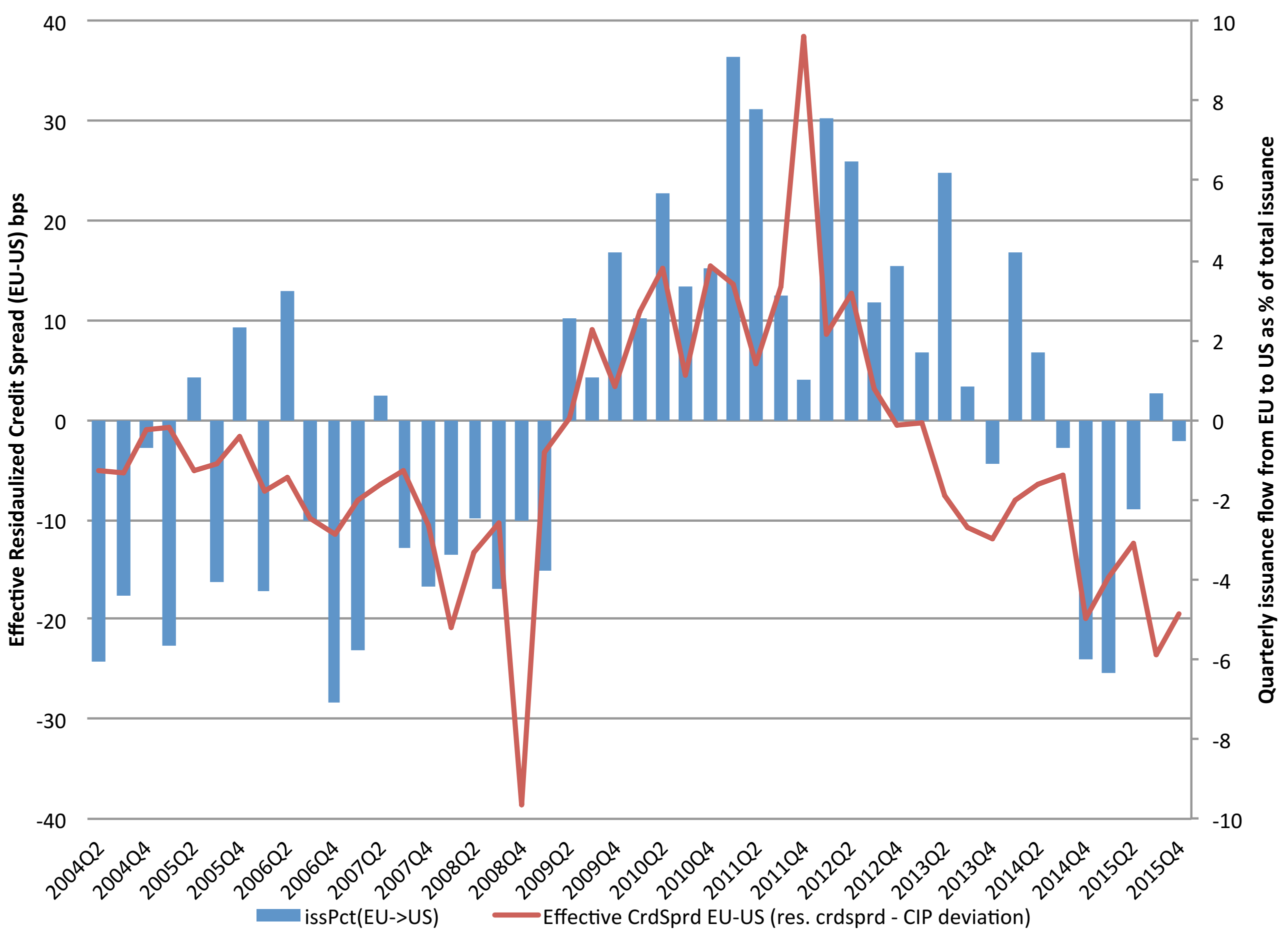
The sign reversals of the issuance flow and net deviation mark distinct time periods in Figure

[25] Chernenko and Sunderam (2014) document that the total money-fund holdings of Eurozone bank paper declined by 37%, from \$453 billion to \$287 billion, between May and August of 2011.

Figure 1.12: Issuance flow and net deviation between Europe and the U.S.
This figure presents issuance flow between the Eurozone and the U.S. and the net deviation (effective residualized credit spread difference) between the euro and the dollar . To construct the net deviation, I estimate the following cross-sectional regression at each date t

$$S_{it}^{adj} = \alpha_{ct} + \beta_{ft} + \gamma_{mt} + \delta_{rt} + \varepsilon_{it}$$

where S_{it}^{adj} is the CIP-adjusted yield spread over the swap curve for bond i that is issued in currency c, by firm f, with maturity m and rating r. The CIP-adjustment is calculated by subtracting maturity-specific CIP deviation from each bond's yield spread. The net deviation or effective residualized credit spread for euro relative to dollar credit spread is calculated as $\hat{\alpha}_{eur,t} - \hat{\alpha}_{usd,t}$. Details of net deviation's construction are provided in Section 1.6.1.2.
Issuance flow is defined as the amount of dollar debt issuance by Eurozone firms minus the amount of euro debt issuance by U.S. firms. I express this measure as a percentage of total issuance between the two countries. Details of the issuance flow's construction are provided in Section 1.6.1.1.



1.12. Prior to the credit crunch in 2007, the net deviation was relatively small and issuance flow oscillated between the two markets with a tilt towards issuance flowing into Europe. The onset of the U.S.-led credit crunch in 2007 reduced the euro credit spread relative to dollar credit spread, which is surprising in itself since the residualized measure suggests that similar bonds issued by the *same* firm are differentially affected by the credit crunch's risk-off sentiment depending on the bond's currency of denomination. This change in net deviation is coupled with several quarters of strong issuance flow from the U.S. to Europe. As the U.S. Federal Reserve begins its quantitative easing (QE) program in late 2008 and early 2009, both the signs for issuance flow and net deviation flipped to the positive side. Even though the asset purchase was in treasury and MBS, QE also indirectly affected the corporate bond market but with lag (Mamaysky 2014, Greenwood, Hanson, and Liao 2015). Foreign issuance in dollar, nicknamed Yankee bond, was popular during this period of Fed QE. In the more recent period since 2014, both time series have reversed sign once again towards the negative. The tapering of Fed QE and the step up of ECB asset-purchasing program arguably led to lower euro-relative-to-dollar credit spread. Reverse-Yankee bonds, or issuance of non-dollar denominated debt by U.S. firms, have picked up and driven the net issuance flow towards Europe.

The comovement of issuance flow and net deviation can also be examined in regression analysis. Table 1.3 presents regression results showing the relation between net deviation (effective credit spread differential) and issuance flow. As seen earlier in the VAR analysis, issuance flow continues for several months after a shock to the credit and CIP violations. Thus, I examine the relation between net deviation at month t and issuance flow averaged over the following six months. The coefficients for the panel regression and for the individual regressions of EUR, GBP, JPY, and CHF are all significant while they are insignificant for AUD, and CAD. One possible interpretation is that while issuance flow is an important source of arbitrage capital in some markets, it is not a dominant force of arbitrage capital for AUD and CAD. Instead, the coefficients on interest rate differential, which represent unhedged carry trade margins, is highly significant for AUD and CAD. This indicates that issuers might be engaged in unhedged issuances in these two currencies for reasons unexplored in this paper. Correspondingly, CIP deviations in AUD and CAD relative to USD are less correlated with their credit spread differentials as can be seen in Figure 1.5 and Table 1.2. While investors-driven hedged capital flows might still be a force that aligns the two deviations, investors generally face more constraints than firms as discussed earlier, therefore, leaving a larger misalignment.

The coefficient on net deviation for EUR-USD issuance flow is the largest and most significant. This is perhaps because the euro and dollar corporate credit markets are highly developed and large in size, issuers are relatively flexible to issue between them. It is also a reflection of the data sample that concentrates on EUR- and USD- denominated bonds.

To explore the dynamics of slow moving capital, I conduct a VAR study on issuance flow and the net deviation as I had done with the individual credit and CIP deviations in earlier section. Figure 1.13 presents the orthogonalized impulse response function of issuance flow upon a shock to the net deviation assuming that issuance respond with a lag to changes in net deviation. The impulse response shows that issuance flow continues to be significant up to 10 months after a shock to the net deviation.

1.6.3.1 Firm-level panel

The aggregate results showing the response of capital flow to the two LOOP violations and to the net deviation can equivalently be tested using a panel of firm-specific credit spread differentials and net deviations. I explore the decision of firm's currency debt choice with a linear probability model in Table 1.4. All of the predictions in the aggregate data are also supported by the firm level regressions with controls for time, currency, and firm fixed

Table 1.3: Issuance flow and net deviation

This table presents forecasting regressions of future issuance flow using effective residualized credit spread differentials (net deviation). $issPct^{Forign \to US}$ is defined as the amount of debt issuance by foreign firms in dollar minus the amount of debt issuance by U.S. firms in the foreign currency expressed as a percentage of total issuance. The sample period is from January 2004 to July 2016 with monthly observation. t-statistics in brackets are based on Newey-West (1987) standard errors with lag selection following Newey-West(1994).

$$issPct^{EU \to US}_{\text{6m.avg.}} = \beta_0 + \beta_1 \text{netdev}_t + \beta_2 \text{ratediff}_t + \varepsilon_{t+1}$$

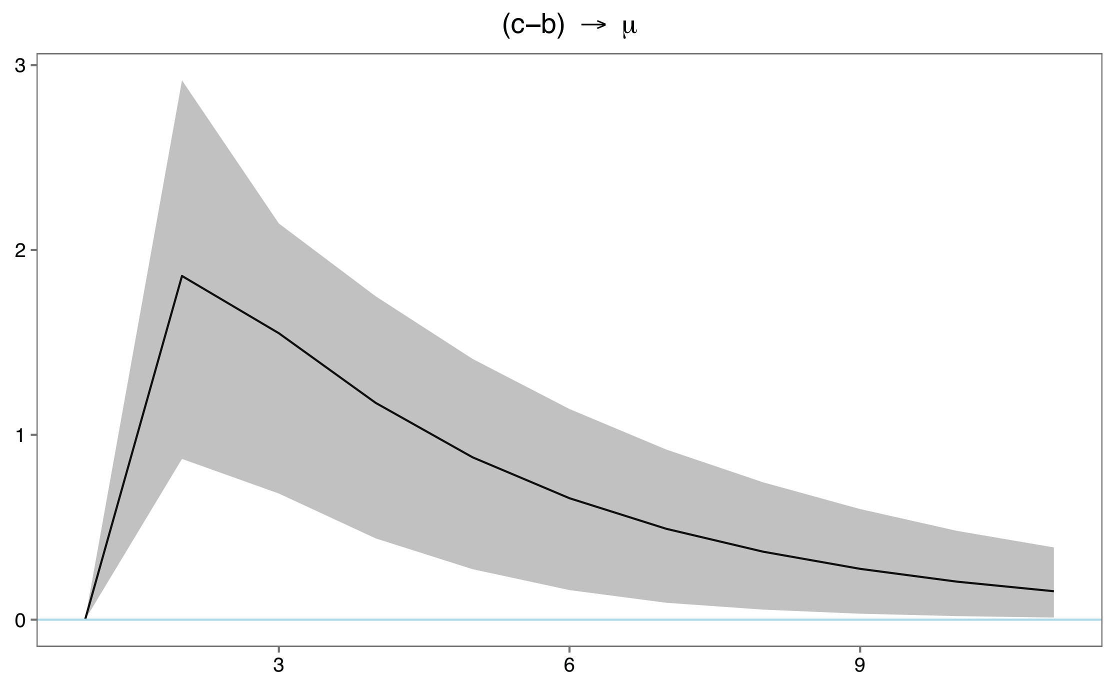
	Net issuance flow (EU→US) /total issuance pct.					
	EUR	GBP	JPY	AUD	CHF	CAD
net dev.	0.247	0.157	0.0353	0.00709	0.119	-0.0534
	[5.08]	[2.11]	[2.10]	[0.07]	[3.47]	[-0.75]
rate diff.	0.0175	-0.0165	0.0256	0.0271	0.00675	0.093
	[1.65]	[-0.77]	[5.50]	[3.52]	[1.14]	[5.32]
_cons	0.984	9.51	5.94	2.26	0.266	7.32
	[0.99]	[4.92]	[4.46]	[1.49]	[0.31]	[6.63]
rsq	0.39	0.13	0.45	0.18	0.29	0.33
n	151	151	151	151	151	151

Figure 1.13: Orthogonalized impulse response of monthly issuance flows to shock to net deviation for EURUSD

I estimate a first order vector autoregression (VAR) of the form

$$\begin{bmatrix} 1 & 0 \\ a_{c-b,\mu} & 1 \end{bmatrix} \begin{bmatrix} \mu_t \\ c_t - b_t \end{bmatrix} = B \begin{bmatrix} \mu_{t-1} \\ c_{t-1} - b_{t-1} \end{bmatrix} + \varepsilon_{\mathbf{t}}$$

where μ_t is the bilateral issuance flow (defined in Section 1.6.1.1), c_t is the credit spread differential and b_t is the CIP deviation. I plot the impulse response of issuance flow μ to shocks to the net deviation $c_t - b_t$. I conduct Cholesky Decomposition by assuming that issuance responds with a lag to shocks to the net deviation. The choice of lag 1 is selected by Bayesian Information Criteria. Confidence intervals at 95% level are shown in gray.



effects. The firm-level panel regressions serve as robustness checks to the aggregate result.

1.6.4 Prediction 3: Total issuance and deviation alignment

Prediction 3 says that an exogenous increase in debt issuance amount D allows firms to deploy more capital and reduces the net deviation. The debt issuance amount D can be seen as the amount of arbitrage capital available to be deployed toward cross-currency credit and CIP arbitrage. As D increases towards infinity, we would expect the net deviation to converge to zero. In this section, I analyze whether large financing needs reduces arbitrageable deviation by first testing in an OLS regression followed by instrumental variable approach that uses the amount of debt maturing to instrument for the need to rollover and refinance through new debt issuance. Specifically, I run a change-on-change regression of the following form

$$\Delta|c-b|_{t,c} = \alpha_c + \beta_1 D_{t,c} + \varepsilon_t$$

where $\Delta|c-b|_{t,c}$ is the monthly change in the absolute value of net deviation and $D_{t,c}$ is the *total* amount of debt issued in both currency c and USD in month t. Note that $D_{t,c}$ is the amount of debt issued, not the outstanding amount of debt.

Conceptually, the analysis relies on the assumption that firms are being opportunistic on the relative allocation of issuance in different currencies rather than being opportunistic on the issuance size in market timing. While the latter motive is important and documented in a number of studies (Baker and Wurgler [2000], Greenwood, Hanson, and Stein [2010], Ma [2015], etc.), it does not preclude the choice analyzed here that focuses on the relative currency denomination conditional on firms having decided the total amount of debt to issue.

To address the potential concerns with endogenous debt issuance decision, I instrument debt issuance amount with maturing debt amount, $M_{t,c}$. Firms frequently issue debt just to rollover existing maturing debt. When deciding to rollover old debt, firms can choose a

Table 1.4: Firm-level issuance choice and violations in credit and CIP

This table presents regressions of firm-level debt denomination choice on credit spread differential and CIP deviation. I estimate the probability that a firm issues debt in currency c conditional on the firm issuing debt in that quarter. I estimate the following specifications in column 1

$$D^{\text{iss}}_{fct} = \beta_0 + \beta_1 \text{Crddiff}_{fct} + \beta_2 \text{CIP}_{ct} + \varepsilon_{fct}$$

where D^{iss}_{fct} is a dummy that equals to 1 if firm f issues in currency c in quarter t, Crddiff_{fct} is the firm-specific residualized credit spread estimated as $\hat{\alpha}_{ct} + \hat{\alpha}_{ct} \cdot \hat{\delta}_{ft}$ in the following cross-sectional regression at each date t

$$S_{it} = \alpha_{ct} + \delta_{ft} + \alpha_{ct} \cdot \delta_{ft} + \varepsilon_{it}$$

where S_{it} is the yield spread over the swap curve for bond i issued in currency c, by firm f. In column 2, I estimate the following regression

$$D^{\text{iss}}_{fct} = \beta_0 + \beta_1 \text{NetDiff}_{fct} + \varepsilon_{fct}$$

where $\text{NetDiff}_{fct} = \text{Crddiff}_{fct} - \text{CIP}_{ct}$. t-statistics in brackets are based on robust standard errors clustered by firm and time.

	probability of issuing in ccy c	
	(1)	(2)
credit dev. c	-0.0727	
	[-5.41]	
cip	0.135	
	[3.19]	
net dev. (c-b)		-0.074
		[-5.53]
firm FE	x	x
time FE	x	x
ccy FE	x	x
rsq	0.18	0.18
n	28726	28726

currency of denomination different from that of the maturing debt. In effect, the amount of debt that needs to be rolled over is capital that corporate arbitrageurs can deploy to take advantage of profitable deviations.

Table 1.5 shows the result of this analysis. AUD and CAD are excluded in this analysis, as issuance is less relevant for the determination of deviations in these two currencies as discussed earlier in Section 1.6.3. For each billion-dollar increase in amount of total debt matured, the net deviation is reduced by roughly 0.1 basis points. While statistically significant, the economic magnitude of this estimation is small, likely because market participants have priced in the effect of large issuance needs from maturing debt given that the debt maturities are easily observable both at the individual and aggregate level.

1.6.5 Prediction 4: Spillover of Limits to Arbitrage

Lastly, I discuss possible tests of the prediction on the spillover of limits to arbitrage. The model suggests that frictions constraining in one market can also be constraining for the other market. These limits to arbitrage frictions can be either directly observable, such as transaction costs, or agency frictions embedded in institutional details. In the model, these constraints are represented by FX swap collateral haircut γ in Equation 1.7, and the ratio of bond risk to risk tolerance $\frac{V}{r}$ in Equation 1.8. The FX haircut is a direct cost while the latter might proxy for indirect agency costs associated with holding an arbitrage position that could become more dislocated before converging as in Shleifer and Vishny (1997).

The empirical measures of these two types of Limits to Arbitrage are difficult to obtain. FX collateral haircut for derivative transactions depends on the currency, maturity and counterparty. The cost of holding LOOP arbitrage positions to maturity are also difficult to quantify. As a rough proxy, I analyze the impact of broker-dealer leverage, proxying for γ, and the VIX index, proxying for $\frac{V}{r}$, on the absolute level of credit spread differential and

Table 1.5: Debt issuance amount and deviation alignment

This table presents regressions of the monthly change in the absolute value of net deviation $(c-b)$ on total debt issuance amount (including both domestic and cross-currency debt) in the same month. The regression is specified as follows

$$\Delta|c-b|_{c,t} = \alpha_c + \beta_1 D_{c,t} + \varepsilon_t,$$

where $D_{c,t}$ is the total amount of debt issued in both currency c and USD expressed in $billions, where $c =$ EUR,GBP,JPY, or CHF. The amount of debt issued is further instrumented by the amount of maturing debt, $M_{c,t}$. Column 1 shows the OLS result with debt issued. Column 2 shows the reduced form regression with maturing debt. Column 3 shows the first stage regression of issued debt on maturing debt. Column 4 shows the IV regression. t-statistics in brackets are based on robust standard errors clustered by time.

	$\Delta\|c-b\|_{c,t}$			
	OLS	Reduced Form	1st stage	IV
$D_{c,t}$ ($\hat{D}_{c,t}$)	-0.080			-0.0939
	[-3.98]			[-2.05]
$M_{c,t}$		-0.0500	0.525	
		[-2.42]	[4.94]	
$\Delta\|c-b\|_{c,t-1}$	-0.089	-0.073		-0.0929
	[-1.44]	[-1.16]		[-1.29]
ccy fe	x	x	x	x
rsq	0.05	0.01	0.63	0.05
n	1180	1180	1198	1180

CIP deviation. The results are in line with Prediction 4. However, for reasons discussed above, the proxies are imprecise and thus relegated to the Internet Appendix.

1.7 Conclusion

This paper examines the connection between violations of covered interest rate parity and price discrepancy of credit risk for bonds of different denominating currencies. I document that these two forms of LOOP violations are substantial and persistent since the financial crisis. Moreover, the two violations are highly aligned in magnitude and direction in both time series and cross section of currencies. I develop a model of market segmentation along two dimensions – in credit market along currency denomination and in FX market between spot and forward exchange rates. Arbitrage processes are imperfect in either markets but capital flow ensures that the two types of LOOP violations are intimately connected.

2 Asset Price Dynamics in Partially Segmented Markets[1]

2.1 Introduction

How do supply shocks in one financial market affect the pricing of assets in other markets? If markets for different asset classes are tightly integrated, then a shock that affects the pricing of a risk factor in one asset class will have a similar effect on other asset classes exposed to the same risk. When markets are more segmented, however, prices of risk in one market may be disconnected from those in other markets. Segmentation arises because institutional and informational frictions lead investors to specialize in a particular asset class or a narrow set of assets (Merton [1987], Grossman and Miller [1988], Shleifer and Vishny [1997]). Although specialization can facilitate arbitrage across securities within an asset class, it can impede arbitrage across asset classes. Specialists' limited willingness to trade across markets may lead the pricing of risk to become disconnected across markets.

The degree of segmentation between different financial markets depends on time horizon.

[1] This paper was written jointly with Robin Greenwood and Samuel G. Hanson. We thank Daniel Bergstresser, Yueran Ma, Jeremy Stein, Adi Sunderam, and seminar participants at Berkeley Haas, Brandeis, MIT Sloan, NYU Stern, University of Minnesota Carlson, University of North Carolina Kenan-Flager, and University of Texas McCombs for useful feedback. We thank David Biery for research assistance. Greenwood and Hanson gratefully acknowledge funding from the Division of Research at Harvard Business School.

Over the long run, the forces of arbitrage ensure that capital will flow from underpriced markets to overpriced markets. However, the process of market integration can be slow, because investors with the flexibility to trade across asset classes do not do so immediately. For example, investment committees at pension funds and endowments—who have the flexibility to allocate capital across asset classes—typically only reallocate capital annually or biannually.

In this paper, we develop a dynamic model of financial markets in which capital moves quickly between securities within a given asset class, but more slowly between different asset classes. Our key contribution is to show how supply shocks in one market are reflected in prices and flows into neighboring markets. In particular, we show how the reaction of neighboring markets depends on time horizon.

Consider two similar long-term risky assets trading in partially segmented markets, such as corporate bonds and Treasury securities. Both assets are exposed to a common fundamental risk factor, making them partial substitutes. This means that, absent frictions, their prices would be tightly linked by cross-market arbitrage. To introduce market segmentation, we assume that there are two sets of risk-averse market specialists, each of whom can flexibly trade one of the risky assets as well as a short-term risk-free asset. Specialists are unable to allocate capital across the two markets. However, markets are partially integrated by risk-averse generalist investors who periodically reevaluate their portfolios and shift between the two risky assets. This setup is similar to Gromb and Vayanos (2002), except that the cross-market arbitrageurs are slow-moving, much like in Duffie (2010). Because of the gradual nature of cross-market arbitrage, markets are more integrated in the long run than the short run.

In the setting we have just described, what happens when there is an unanticipated supply shock in one market? Suppose, for concreteness, that the Federal Reserve announces that it will sell a large portfolio of long-term U.S. Treasury bonds, permanently expanding the amount of interest rate risk that investors need to bear in equilibrium. Treasury market specialists react immediately to the shock, absorbing the increased supply into their inventories. The risk premium on long-term Treasury bonds will rise, lifting their yields. However, Treasury yields will overreact—the short-run price impact will exceed the long-run impact—because the amount of capital that can initially accommodate the shock is limited to specialists and a handful generalists. Over the long-run, generalist investors will allocate additional capital to the Treasury market, muting the price impact of the supply shock at longer horizons. Price dynamics of this sort are similar to those described in Duffie (2010), who shows how prices react to shocks when there is slow-moving capital.

Our key contribution is to characterize how prices evolve in related markets that are not directly impacted by the supply shock. Consider the question of how corporate bond prices (or stock prices for that matter) will react to a shock to the supply of long-term Treasuries. Although this supply shock does not directly impact the corporate bond market, this market is indirectly affected because generalist investors will respond by increasing their holdings of long-term Treasuries and reducing their holdings of long-term corporate bonds. These cross-market capital flows drive down the prices of corporate bonds and push up corporate bond yields. In this way, the trading of generalist asset allocators transmits supply shocks across markets, serving to increase market integration. While yields in the Treasury market initially overreact to the supply shock, under plausible parameter values, we show that corporate bond yields will underreact: the short-run price impact is less than the long-run impact. The overreaction of Treasury yields and the underreaction of corporate yields are both driven by the fact that generalists only reallocate capital slowly. As a result, it takes time for financial markets to fully digest large supply shocks.

If all investors were generalists, the two markets would be fully integrated in the sense that exposures to common risk factors would always have the same prices in the two markets. However, in the more realistic case when markets are partially segmented, risk prices can differ across markets. This occurs because risks cannot be easily unbundled from assets and because markets receive periodic supply shocks, making cross-market arbitrage risky for generalists, as in the model developed by Gromb and Vayanos (2002). For example, interest rate risk may not be priced identically in the corporate bond market and the Treasury market. Following a supply shock, the premia associated with similar risk exposures can differ significantly between the two asset markets. As generalists react to pricing discrepancies across markets, differences in risk premia will gradually narrow. However, the differences will not vanish in the long run because of the permanent risks associated with cross-market arbitrage. Put differently, partial segmentation creates a form of noise trader risk.

The price dynamics in our model depend critically on the fractions of specialists in each market, the number of time periods it takes generalists to fully rebalance their portfolios, and the degree of substitutability between the two asset markets. The fraction of specialists and generalist investors play an especially important role. When there are a small number of slow-moving generalists, the Treasury market overreacts while the corporate market underreacts to the shock to Treasury supply. However, if there are many slow-moving generalists, markets are well-integrated and supply shocks can result in short-run overreaction in both markets.

In describing our model, we have made no distinction between a risky "asset" and the "market" in which it trades. This distinction arises when we introduce multiple risky assets into each market. Individual assets differ in their degree of exposure to common risk factors. For example, the "market" for U.S. Treasury securities contains bonds of many different maturities, which have different exposures to interest rate risk. Extending our model to allow for multiple securities per asset market, we show that a conditional CAPM prices all assets in the first market and that another conditional CAPM—with different prices of risk—prices all assets in the second market. Critically, these two market-specific pricing models are linked over time by the cross-market arbitrage activities of slow-moving asset allocators. For example, the pricing of interest rate risk for 2-year Treasuries is always perfectly consistent with the pricing of interest rate risk for 10-year Treasuries. However, the pricing of interest rate risk in the Treasury market may differ somewhat from that in the corporate bond market. And these cross-market differences will be most pronounced following the arrival of major shocks that take time for slow-moving generalists to digest.

Recognizing the partially segmented nature of capital markets is helpful for understanding fixed income markets in the QE-era. A key question about central bank government bond purchases is whether they impact the prices of financial assets outside of the market for government bonds. The favored methodology for answering this question has been to use event studies of intraday or one-day price changes following central bank policy announcements. A number of these studies have concluded that the effects of quantitative easing are most pronounced in the market in which the central bank is transacting, with only modest spillovers to other related markets (Woodford [2012] and Krishnamurthy and Vissing-Jorgensen [2013]). Others have suggested that at longer horizons, the spillovers are more significant. Mamaysky (2014) suggests that if one expands the measurement window by a few days or weeks, the effects in other markets may be much larger.

Our model suggests that the short-run price impact of a supply shock on different markets may not accurately reveal the long-run impact, which is often of greater interest to policymakers. We illustrate this idea by analyzing the statistical power of short-run event studies within our model. We show that the horizon at which statistical power is maximized is often much shorter than the horizon at which the long-run price impact is achieved. A broad message to emerge from our paper is that while the event study methodology is appropriate for measuring cash flow news, it is less suitable for analyzing policies that impact discount rates.

Our model is closely related to two strands of research in financial economics. The idea that front-line arbitrageurs in financial markets are highly specialized traces back to Merton (1987) and Grossman and Miller (1988), and is a central tenet of the theory of limited arbitrage (De Long et al [1990], Shleifer and Vishny [1997], and Gromb and Vayanos [2002]). A small literature in finance describes asset prices and returns in segmented markets (Stapleton and Subrahmanyam [1977], Errunza and Losq [1985], Merton [1987]). More recently, a number of researchers have demonstrated downward-sloping demand curves for individual financial asset classes, which would be puzzling if markets were fully integrated (Gabaix, Krishnamurthy, and Vigneron [2007], Gârleanu, Pedersen, and Poteshman [2009], Greenwood and Vayanos [2014], and Hanson [2014]). These researchers have often motivated their analysis by positing an extreme form of market segmentation in which a different pricing kernel is used to price the securities in each distinct asset class. Our paper emphasizes how the actions of slow-moving asset allocators serve to link these market-specific pricing kernels together, thereby offering a middle ground between these models positing extreme segmentation and traditional models featuring perfect integration.

Second, our paper is related to research on "slow-moving capital," which is the idea that capital does not flow as quickly towards attractive investment opportunities as textbook theories might suggest (Mitchell, Pedersen, Pulvino [2007], Duffie [2010], Acharya, Shin, and Yorulmazer [2013]). Here, our model draws most heavily from Duffie (2010), who studies the implications of slow moving capital for price dynamics in a single asset market. Compared to his paper, our contribution is to analyze the impact of supply shocks and slow moving capital across multiple asset markets as well as the implication on event study methodology. Duffie and Strulovici (2012) present a model of the movement of capital across two partially segmented markets, but their focus is on the endogenous speed of capital mobility, which we take as exogenous. Our key contribution here is characterizing the dynamics of prices across related asset markets and the patterns of cross-market arbitrage in response to large supply or demand shocks.

Our paper also relates to studies of FOMC news releases on asset prices (Akhtar [1997], Harvey and Huang [2002], Sokolov [2009], and Cieslak, Morse, and Vissing-Jorgensen [2016]) and liquidity and limits to arbitrage in the treasury market, e.g. Hu, Pan, Wang (2013). Relative to these prior studies, our paper emphasizes the long-term effect of asset prices in response to large supply shocks rather than the most immediate effect detectable as event studies.

2.2 Model

We develop the model in two steps. We first develop a tractable, benchmark model for pricing long-term fixed-income assets that are exposed to *both* interest rate risk and default risk. The model builds on the default-free term structure models in Vayanos and Vila (2009) and Greenwood and Vayanos (2014) in which interest rate risk is priced by a set of specialized, risk-averse bond arbitrageurs, leading to a downward-sloping aggregate demand curve for bond risk factors. In this first step, we develop a simple way to incorporate default risk into this class of models. In the second step, we introduce a second asset class and a richer institutional trading environment that contains both generalists and specialists. In this richer environment, we describe how prices and investor positions in both markets evolve following a supply shock that directly impacts only one market.

2.2.1 Single asset model

2.2.1.1 Defaultable perpetuities

Consider a homogenous portfolio of perpetual, defaultable bonds each of which promises to pay a coupon of C each period. Let $P_{L,t}$ denote the the price of each long-term bond at time t. Suppose that a random fraction h_{t+1} of the bonds default at $t+1$ and are worth $(1 - L_{t+1})(P_{L,t+1} + C)$ where $0 \leq L_{t+1} < 1$ is the (possibly random) loss-given-default as a fraction of market value. The remaining fraction $(1-h_{t+1})$ of the bonds do not default and are worth $(P_{L,t+1}+C)$. Thus, the return on the bond portfolio is

$$1+R_{L,t+1}=\frac{(1-Z_{t+1})\,(P_{L,t+1}+C)}{P_{L,t}}, \tag{2.1}$$

where $Z_{t+1}=h_{t+1}L_{t+1}$, satisfying $0\leq Z_{t+1}<1$, is the portfolio default realization at time $t+1$. If $Z_{t+1}\equiv 0$, the bonds are default-free. If Z_{t+1} is stochastic, the bonds are defaultable with high realizations of Z_{t+1} corresponding to larger default losses at time $t+1$. This formulation of default risk follows Duffie and Singleton's (1999) "recovery of market value" assumption which has become standard in the credit risk literature.

To generate a tractable linear model, we use a Campbell-Shiller (1988) log-linear approximation to the return on this portfolio of defaultable perpetuities. Specifically, defining $\theta\equiv 1/\left(1+C\right)<1$, the one-period log return on the bonds is

$$r_{L,t+1}\equiv\ln\left(1+R_{t+1}^{L}\right)\approx\overbrace{\frac{1}{1-\theta}}^{D}y_{L,t}-\overbrace{\frac{\theta}{1-\theta}}^{D-1}y_{L,t+1}-z_{t+1}, \tag{2.2}$$

where $y_{L,t}$ is the log yield-to-maturity at time t,

$$D=\frac{1}{1-\theta}=\frac{C+1}{C} \tag{2.3}$$

is the Macaulay duration when the bonds are trading at par, and $z_t=-\ln\left(1-Z_t\right)$ is the log default loss at time t.[2]

To derive this approximation note that the Campbell-Shiller (1988) approximation of the 1-period log return is

$$\begin{aligned} r_{L,t+1} &= \ln\left(P_{L,t+1}+C\right)-p_{L.t}-z_{t+1} \\ &\approx \kappa+\theta p_{L.t+1}+(1-\theta)\,c-p_{L.t}-z_{t+1} \end{aligned} \tag{2.4}$$

[2] This log-linear approximation for default-free coupon-bearing bonds appears in Chapter 10 of Campbell, Lo, and MacKinlay (1997). Our approximation for defaultable bonds then follows trivially given the assumption that default losses are a (random) fraction of market value.

where $\theta = 1/\left(1+\exp\left(c-\overline{p}_L\right)\right)$ and $\kappa = -\log\left(\theta\right) - \left(1-\theta\right)\log\left(\theta^{-1}-1\right)$ are parameters of the log-linearization. Iterating equation (2.4) forward, we find that the log bond price is

$$p_{L,t} = \left(1-\theta\right)^{-1}\kappa + c - \textstyle\sum_{i=0}^{\infty}\theta^{i}E_t\left[r_{L,t+i+1}+z_{t+i+1}\right]. \tag{2.5}$$

Applying this approximation to *promised cashflows* (i.e., $z_{t+i+1} \equiv 0$ for all $i \geq 0$) and the *yield-to-maturity*, defined as the *constant return* that equates bond price and the discounted value of *promised* cashflows, we obtain

$$p_{L,t} = \left(1-\theta\right)^{-1}\kappa + c - \left(1-\theta\right)^{-1}y_{L,t}. \tag{2.6}$$

Equation (2.2) then follows by substituting the expression for $p_{L,t}$ in equation (2.6) into the Campbell-Shiller return approximation in equation (2.4). Assuming the steady-state price of the bonds is par ($\overline{p}_L = 0$), we have $\theta = 1/\left(1+C\right)$. Thus, bond duration is $D = -\partial p_{L,t}/\partial y_{L,t} = \left(1-\theta\right)^{-1} = \left(1+C\right)/C$. Since $-\partial p_{L,t}/\partial y_{L,t} = -\left(\partial P_{L,t}/\partial Y_{L,t}\right)\left(\left(1+Y_{L,t}\right)/P_{L,t}\right) = (Y_{L,t}^{\cdot}+1)/Y_{L,t}$ this corresponds to Macaulay duration when the bonds are trading at par ($Y_{L,t} = C$).

2.2.1.2 Risk factors

Investors in defaultable long-term bonds are exposed to three different types of risk: interest rate risk, default risk, and supply risk. First, investors are exposed to *interest rate risk*. In our model, investors face an exogenous short-term interest rate that evolves randomly over time and will suffer a capital loss on their bond holdings if short-term rates rise unexpectedly. Second, investors face *default risk*: the future period-by-period default realization is unknown and evolves randomly over time. Finally, investors are exposed to *supply risk*: there are random supply shocks which impact the prices and yields on long-term bonds, holding fixed the expected future path of short-term interest rates and expected future defaults. Thus, using Campbell's (1991) terminology, interest rate risk and default risk are forms of fundamental "cash flow" risk, whereas supply risk is a form of "discount rate" risk.

We make the following concrete assumptions:

- **Short-term interest rates:** The log short-term riskless rate available to investors between time t and $t+1$, denoted r_t, is known at time t. We assume that r_t also follows an exogenous AR(1) process
$$r_{t+1} = \overline{r} + \rho_r (r_t - \overline{r}) + \varepsilon_{r,t+1}, \tag{2.7}$$
where $Var_t[\varepsilon_{r,t+1}] = \sigma_r^2$. One can think of the short-term rate as being determined outside the model either by monetary policy or by a stochastic short-term storage technology that is available in perfectly elastic supply.

- **Default losses:** We assume that the default process z_t follows
$$z_{t+1} = \overline{z} + \rho_z (z_t - \overline{z}) + \varepsilon_{z,t+1} \tag{2.8}$$
where $Var_t[\varepsilon_{z,t+1}] = \sigma_z^2$.

- **Supply:** We assume that the perpetuity is available in an exogenous, time-varying supply s_t. We assume that supply follows an AR(1) process
$$s_{t+1} = \overline{s} + \rho_s (s_t - \overline{s}) + \varepsilon_{s,t+1}, \tag{2.9}$$
where $Var_t[\varepsilon_{s,t+1}] = \sigma_s^2$.

For simplicity, we will assume that $\varepsilon_{s,t+1}$, $\varepsilon_{r,t+1}$, and $\varepsilon_{z,t+1}$ are mutually orthogonal. However, it is straightforward to relax this assumption.

2.2.1.3 Specialist demand and market clearing

There is a unit mass of specialized bond arbitrageurs, each with risk tolerance τ. Specialist arbitrageurs can earn an uncertain future return of $r_{L,t+1}$ from to t to $t+1$ by investing in the defaultable long-term bond. Alternatively, they can earn a certain return of r_t by investing at the short-term interest rate. Specialist arbitrageurs are concerned with their interim wealth.

Formally, we assume that at date t specialist arbitrageurs have mean-variance preferences over their wealth at $t+1$. This means that arbitrageurs choose their holdings of the perpetuity to solve

$$\max_{b_t} \left\{ b_t E_t \left[rx_{L,t+1} \right] - (2\tau)^{-1} (b_t)^2 Var_t \left[rx_{L,t+1} \right] \right\}, \tag{2.10}$$

where $rx_{L,t+1} \equiv r_{L,t+1} - r_t$ is the log excess returns on the defaultable long–term bond over the short-term interest rate between t and $t+1$. Thus, arbitrageur demand for the risky bond is

$$b_t = \tau \frac{E_t \left[rx_{L,t+1} \right]}{Var_t \left[rx_{L,t+1} \right]}. \tag{2.11}$$

Equation (2.11) says that arbitrageurs borrow at the short-term rate and invest in risky long-term bonds when the expected return on perpetuities exceeds that the short rate ($E_t \left[rx_{L,t+1} \right] > 0$). Conversely, arbitrageurs sell short bonds and invest at the short rate when $E_t \left[rx_{L,t+1} \right] < 0$. And they respond more aggressively to these movements in risk premia when they are more risk tolerant and when the variance of excess bond returns is low.

Market clearing ($b_t^* = s_t$) implies that the bond risk premium, $E_t \left[rx_{L,t+1} \right]$ is given by

$$E_t \left[rx_{L,t+1} \right] = \tau^{-1} V_L^{(1)} s_t, \tag{2.12}$$

where $V_L^{(1)} = Var_t \left[(D-1)\, y_{L,t+1} + z_{t+1} \right]$ is the equilibrium variance of 1-period excess returns.

Thus, bond risk-premia are increasing in bond supply, s_t. When a positive supply shock arrives, bond risk premia jump instantaneously. If the shock is almost permanent ($\rho_s \approx 1$), the impact on the risk premium will be long lived. If the shock is transient ($0 < \rho_s \ll 1$), supply will quickly revert to steady-state ($\overline{s}$) and risk premia will revert to their steady-state level, $\tau^{-1} V_L^{(1)} \overline{s}$.

2.2.1.4 Solution and equilibrium yields

To solve the model, we conjecture that equilibrium bond yields take the linear form

$$y_{L,t} = \alpha_0 + \alpha_r \left(r_t - \overline{r}\right) + \alpha_z \left(z_t - \overline{z}\right) + \alpha_s \left(s_t - \overline{s}\right). \tag{2.13}$$

Using this conjecture, in the Internet Appendix we show that a linear equilibrium of this form exists so long as arbitrageurs are sufficiently risk tolerant (i.e., if τ is large enough). We show that the equilibrium variance of 1-period excess bond returns, $V_L^{(1)}$, must satisfy the following quadratic equation

$$V_L^{(1)} = \left(\frac{\theta}{1-\rho_r\theta}\sigma_r\right)^2 + \left(\frac{1}{1-\rho_z\theta}\sigma_z\right)^2 + \left(\tau^{-1}\frac{\theta}{1-\rho_s\theta}\sigma_s\right)^2 \left(V_L^{(1)}\right)^2. \tag{2.14}$$

The total risk premium can be decomposed into compensation for bearing interest rate risk, compensation for bearing credit risk, and compensation for bearing supply risk:

$$\begin{aligned} E_t\left[rx_{L,t+1}\right] = {} & \overbrace{\tau^{-1}\left(\frac{\theta}{1-\rho_r\theta}\sigma_r\right)^2 s_t}^{\text{Interest rate risk premium}} + \overbrace{\tau^{-1}\left(\frac{1}{1-\rho_z\theta}\sigma_z\right)^2 s_t}^{\text{Credit risk premium}} \\ & + \overbrace{\tau^{-1}\left(\tau^{-1}\frac{\theta}{1-\rho_s\theta}\sigma_s\right)^2 \left(V_L^{(1)}\right)^2 s_t}^{\text{Supply risk premium}}. \end{aligned} \tag{2.15}$$

The level of supply (s_t) appears three times on the right hand side of equation (2.15) because all three components of the total risk premium move in lock in our single asset model.

As in Greenwood and Vayanos (2014), when there is supply risk ($\sigma_s^2 > 0$) a linear equilibrium only exists if bond arbitrageurs are sufficiently risk tolerant.[3] In this case, there are two possible solutions to (2.14): one in which yields are highly sensitive to supply shocks and one in which yields are less sensitive. What is the intuition for the multiplicity of equilibria? If yields are highly sensitive to supply shocks, then bonds become highly risky for arbitrageurs. Hence, arbitrageurs absorb supply shocks only if they are compensated by large changes in

[3] If τ is too small and there are supply shocks ($\sigma_s^2 > 0$), no linear equilibrium exits because bonds become extremely risky for arbitrageurs and it is impossible to clear the market.

yields, making the high sensitivity of yields to shocks self-fulfilling. Conversely, if yields are less sensitive to supply shocks, then bonds become less risky for arbitrageurs and arbitrageurs are willingly absorb supply shocks even if they are only compensated by modest changes in yields. Equilibrium multiplicity of this sort is common in overlapping generations models such as ours where arbitrageurs with short investment horizons hold a long-lived asset that is subject to supply shocks (see e.g., DeLong, Shleifer, Summers, and Waldmann [1990]).

[**referee wants more on equilibrum selection, insert here**]

Following Greenwood and Vayanos (2014), we focus on the well-behaved and economically relevant equilibrium in which yields are less sensitive to supply shocks, which corresponds to the smaller root of equation (2.14).[4] It is then straightforward to show that $V_L^{(1)}$ is increasing in σ_r^2, σ_z^2, σ_s^2, ρ_r, ρ_z, ρ_s, and D $[= (1-\theta)^{-1}]$ and decreasing in τ. Thus, for a given level of bond supply, the total risk premium is larger when short-term rates are more volatile, when there is greater uncertainty about future defaults, and when supply shocks are more volatile. Furthermore, the risk premium is larger when each of these three processes is more persistent. Finally, the risk premium is increasing in the duration of the perpetuity and is decreasing in arbitrageur risk tolerance.

Rewriting equation (2.2) as $y_{L,t} = E_t\left[(1-\theta)\left(r_t + rx_{L,t+1} + z_{t+1}\right) + \theta y_{L,t+1}\right]$ and iterating forward, we see that the equilibrium yield on the defaultable perpetuity is a weighted average of expected future short rates, future default losses, and future risk premia

$$y_{L,t} = (1-\theta)\textstyle\sum_{i=0}^{\infty}\theta^i E_t[\overbrace{r_{t+i}}^{\text{Short rate}} + \overbrace{z_{t+i+1}}^{\text{Default loss}} + \overbrace{\tau^{-1}V_L^{(1)}s_{t+i}}^{\text{Risk premium}}]. \tag{2.16}$$

Because of the coupon-bearing nature of the long-term bond, equation (2.16) shows that expected short rates, default losses, and risk premia in the near future have a larger effect on

[4]As $\sigma_s^2 \to 0$, this smaller root converges to the solution for $V_L^{(1)}$ when $\sigma_s^2 = 0$ (i.e., to $\left((\theta\sigma)_r / (1-\rho_r\theta)\right)^2 + \left(\sigma_z / (1-\rho_z\theta)\right)^2$) whereas the larger root diverges to infinity as $\sigma_s^2 \to 0$. All of the relevant comparative statics on $V_L^{(1)}$ have the intuitive signs at the smaller root, but have the opposite signs at the larger root.

bond yields than those in the distant future.[5] Making use of the assumed AR(1) dynamics for r_t, z_t, and s_t, we can express the equilibrium yield as

$$y_{L,t} = \overbrace{\left[\overline{r} + \frac{1-\theta}{1-\rho_r\theta}\left(r_t - \overline{r}\right)\right]}^{\text{Expected future short rates}} + \overbrace{\left[\overline{z} + \frac{1-\theta}{1-\rho_z\theta}\rho_z\left(z_t - \overline{z}\right)\right]}^{\text{Expected future default losses}} \tag{2.17}$$
$$+ \overbrace{\left[\tau^{-1}V_L^{(1)}\overline{s} + \tau^{-1}V_L^{(1)}\frac{1-\theta}{1-\rho_s\theta}\left(s_t - \overline{s}\right)\right]}^{\text{Risk premium}}.$$

Equation (2.17) shows that the perpetuity yield is more sensitive to movements in short rates when the short-rate process is more persistent and when bond duration is shorter (i.e., $\partial^2 y_{L,t}/\partial r_t \partial \rho_r > 0$ and $\partial^2 y_{L,t}/\partial r_t \partial D < 0$). Similarly, the yield is more sensitive to movements in current default losses (z_t) when the default process is more persistent and when bond duration is shorter. Yields are more sensitive to bond supply when short-rates are more volatile or more persistent or when defaults are more volatile or more persistent. Finally, yields are also more sensitive to supply shocks when risk tolerance is low, supply shocks are more volatile, or supply shocks are more persistent.[6]

2.2.2 Partially segmented markets

With this machinery in place, we now introduce a second risky asset and a richer trading environment, to capture the idea that the two assets trade in partially segmented markets. Our goal is to study how shocks to asset supply in one market are transmitted over time to the second market.

[5]This is similar to Campbell and Shiller's (1988) analysis of the price of a dividend-paying stock.

[6]The sign of $\partial^2 y_{L,t}/\partial s_t \partial D$ is ambiguous since $\partial V_L^{(1)}/\partial D > 0$, but $\partial\left[\left(1-\theta\right)/\left(1-\rho_s\theta\right)\right]/\partial D < 0$. This corresponds to the finding in Vayanos and Greenwood (2014) that, depending on the persistence of supply shocks, a current increase in bond supply can have a greater impact on the yields of intermediate or long-dated bonds. Specifically, highly persistent supply shocks have the greatest impact on long-dated yields, while transitory supply shocks have the greatest impact on intermediate-dated yields.

2.2.2.1 Asset markets

Suppose now that there are two portfolios of perpetual risky assets, A and B. A is default–free and exposed only to interest rate risk. Borrowing notation from above, portfolio A pays a coupon of C_A each period, so the gross return on A is $1+R_{A,t+1} = (P_{A,t+1} + C_A)/P_{A,t}$. The log excess return on the A portfolio over the short-term interest rate from time t to $t+1$ is

$$rx_{A,t+1} \approx \frac{1}{1-\theta_A} y_{A,t} - \frac{\theta_A}{1-\theta_A} y_{A,t+1} - r_t, \tag{2.18}$$

where $\theta_A = 1/(1+C_A)$.

The second portfolio, B, is subject to default risk which makes it an imperfect substitute for asset A. Specifically, the B portfolio carries a *promised* coupon payment of C_B each period. However, the gross return on the B portfolio from time t to $t+1$ is $1 + R_{B,t+1} = (1 - Z_{t+1})(P_{B,t+1} + C_B)/P_{B,t}$ where $0 \leq Z_{t+1} \leq 1$ is the default realization at time $t+1$. Therefore, the log excess return on B from time t to $t+1$ is

$$rx_{B,t+1} \approx \frac{1}{1-\theta_B} y_{B,t} - \frac{\theta_B}{1-\theta_B} y_{B,t+1} - z_{t+1} - r_t, \tag{2.19}$$

where $\theta_B = 1/(1+C_B)$. The additional z_{t+1} term in equation (2.19) reflects the time $t+1$ default realization that is specific to the B asset. The variance of z_{t+1} determines, in part, the degree of substitutability between assets A and B.

We assume that the processes for the short rate r_t and for default losses z_t are as in equations (2.7) and (2.8) above. However, we assume that the two asset markets are subject to different supply shocks which also limits their substitutability for investors with shorter horizons. The net supply that investors must hold of asset A evolves according to

$$s_{A,t+1} = \overline{s}_A + \rho_{s_A}(s_{A,t} - \overline{s}_A) + \varepsilon_{s_A,t+1}, \tag{2.20}$$

where $Var_t[\varepsilon_{s_A,t+1}] = \sigma^2_{s_A}$. Similarly, the net supply that investors must hold of asset B

evolves as

$$s_{B,t+1} = \overline{s}_B + \rho_{s_B}\left(s_{B,t} - \overline{s}_B\right) + \varepsilon_{s_B,t+1}, \tag{2.21}$$

where $Var_t\left[\varepsilon_{s_B,t+1}\right] = \sigma^2_{s_B}$.We continue to assume that $\varepsilon_{r,t+1}$, $\varepsilon_{z,t+1}$, $\varepsilon_{s_A,t+1}$, and $\varepsilon_{s_B,t+1}$ are mutually orthogonal.

2.2.2.2 Market participants

There are three types of investors, all with identical risk tolerance τ. Investors are distinguished by their ability to transact in different markets and by the frequency with which they can rebalance their portfolios. Fast-moving A-specialists are free to adjust their holdings of the A asset and the riskless short-term asset each period; however, A-specialists cannot hold the B asset. A-specialists are present in mass q_A and we denote their demand for A by $b_{A,t}$. Analogously, fast-moving B-specialists can freely adjust their holdings of the B asset and the riskless asset each period, but cannot hold the A asset. B-specialists are present in mass q_B and their demand for asset B is $b_{B,t}$.

The third group of investors is a set of slow-moving generalists who can adjust their holdings of A and B asset, as well as the riskless short-term asset, but can do so only every k periods. Generalists are present in mass $1 - q_A - q_B$. Fraction $1/k$ of these generalists investors are active each period and can reallocate their portfolios between the A and B assets. However, they must then maintain this same portfolio allocation for the next k periods. As in Duffie (2010), this is a reduced form way to model the frictions that limit the speed of capital flows across markets.

The market structure we have described here is a natural way to capture the industrial organization of real world asset management. Due to agency and informational problems, savers are only willing to give delegated managers the discretion to adjust their portfolios quickly if the manager accepts a narrow, *specialized* mandate. These same agency and informational frictions also mean that savers are only willing to give managers the discretion to adjust quickly if the manager gives them an open-ended claim (e.g., Stein (2005)). As a result, fast-moving investors often have endogenously *short horizons.* By contrast, most institutions, such as endowments and pensions, that have *longer horizons* and possess *greater flexibility* to re-allocate capital across asset classes are subject to governance mechanisms—themselves a response to informational and agency frictions—that limit the speed of any such capital movement. In combination, we believe that a model with fast-moving specialists and slow-moving generalists is a tractable, reduced-form way to capture real-world arbitrage frictions.

In this paper, we focus on a "medium-run" equilibrium in which the parameters governing market structure (q_A, q_B, and k) are regarded as fixed and exogenously given. However, one could extend the model to endogenize the market structure. In the resulting "very long-run" equilibrium, q_A, q_B, and k would adjust so that A specialists, B specialists, and generalists all have the same expected utility in the long-run.[7]

Fast-moving A-specialists and B-specialists have mean-variance preferences over 1-period portfolio log returns. Thus, their demands are given by

$$b_{A,t} = \tau \frac{E_t\left[rx_{A,t+1}\right]}{Var_t\left[rx_{A,t+1}\right]}, \tag{2.22}$$

and

$$b_{B,t} = \tau \frac{E_t\left[rx_{B,t+1}\right]}{Var_t\left[rx_{B,t+1}\right]}. \tag{2.23}$$

Since they only rebalance their portfolios every k periods, slow-moving generalist investors have mean-variance preferences over their k-period *cumulative* portfolio excess return. Defining $rx_{A,t\to t+k} \equiv \sum_{i=1}^{k} rx_{A,t+i}$ and $rx_{B,t\to t+k} \equiv \sum_{i=1}^{k} rx_{B,t+i}$ as the cumulative k-period returns

[7] For instance, one could assume that A and B specialists must pay a cost to set up a specialized, fast-moving fund and that generalists must pay a cost in order to adjust more quickly. q_A, q_B, and k would then need to adjust so that (i) investors expect to earn the same long-run Sharpe ratio, net of costs, from all three structures and (ii) generalists' marginal benefit from adjusting their portfolios more frequently equals the marginal cost of more frequent adjustment.

from t to $t+k$ on A and B, the k-period portfolio excess return of generalists who are active at t is[8]

$$rx_{d_t,t\to t+k} = d_{A,t} \times rx_{A,t\to t+k} + d_{B,t} \times rx_{B,t\to t+k}. \tag{2.24}$$

Thus, generalist investors who are active at time t choose their holdings of asset A and B, denoted $d_{A,t}$ and $d_{B,t}$, to solve

$$\max_{d_{A,t},d_{B,t}} \left\{ E_t\left[rx_{d_t,t\to t+k}\right] - (2\tau)^{-1}\left(Var_t\left[rx_{d_t,t\to t+k}\right]\right)\right\}. \tag{2.25}$$

This implies that

$$\begin{bmatrix} d_{A,t} \\ d_{B,t} \end{bmatrix} = \frac{\tau}{1-R^{2(k)}_{AB}} \begin{bmatrix} \frac{E_t\left[rx_{A,t\to t+k}\right]}{Var_t[rx_{A,t\to t+k}]} - \beta^{(k)}_{B|A}\frac{E_t\left[rx_{B,t\to t+k}\right]}{Var_t[rx_{B,t\to t+k}]} \\ \frac{E_t\left[rx_{B,t\to t+k}\right]}{Var_t[rx_{B,t\to t+k}]} - \beta^{(k)}_{A|B}\frac{E_t\left[rx_{A,t\to t+k}\right]}{Var_t[rx_{A,t\to t+k}]} \end{bmatrix}, \tag{2.26}$$

where, for example, $\beta^{(k)}_{B|A}$ is the coefficient from a linear regression of $rx_{B,t\to t+k}$ on $rx_{A,t\to t+k}$ and $R^{2(k)}_{AB}$ is the goodness of fit from this regression.[9]

Equation (2.26) says that, all else equal, generalist investors allocate more capital to market A when asset A becomes more attractive from a narrow risk-reward standpoint (i.e., $d_{A,t}$ is increasing in $E_t[\sum_{i=1}^{k} rx_{A,t+i}]/Var_t[rx_{A,t\to t+k}]$). Further, assuming B and A co-move positively ($\beta^{(k)}_{B|A} > 0$), generalists allocate less capital to market A when asset B becomes more attractive from a risk-reward standpoint (i.e., $d_{A,t}$ is decreasing in $E_t[\sum_{i=1}^{k} rx_{B,t+i}]/Var_t[rx_{B,t\to t+k}]$). In this way, the response of generalist investors transmits supply shocks in the B market

[8] Formally, this means we assume that slow-moving generalists re-invest all capital initially allocated to the A market (B market) in the A market (B market) over their k-period investment horizon. Also, our implicit log-linearization of the portfolio return omits the second-order Jensen's inequality adjustments familiar from Campbell and Viceira (2002). However, in the case of low-volatility fixed-income instruments, these adjustments are quantitatively small and do not alter the core economic intuition of the model.

[9] We obtain similar results if we alter equation (2.25) to reflect the fact that the cumulative return from rolling over an investment at the short-rate for k periods, $\sum_{i=0}^{k-1} r_{t+i}$, is unknown at time t. As in Campbell and Viceira (2001), this adds an I-CAPM-like hedging motive for holding long-duration assets that have high excess returns when $\sum_{i=0}^{k-1} r_{t+i}$ turns out to be lower than expected. Formally, this means that generalists solve $\max_{d_{A,t},d_{B,t}} \left\{E_t\left[r_{P,t\to t+k}\right] - \frac{1}{2\tau}\left(Var_t\left[r_{P,t\to t+k}\right]\right)\right\}$ where $r_{P,t\to t+k} = (\sum_{i=0}^{k-1} r_{t+i}) + d_{A,t} \times (\sum_{i=1}^{k} rx_{A,t+i}) + d_{B,t} \times (\sum_{i=1}^{k} rx_{B,t+i})$. The solution takes the same form as (2.26), replacing $E_t[\sum_{i=1}^{k} rx_{A,t+i}]$ with $E_t[\sum_{i=1}^{k} rx_{A,t+i}] - \tau^{-1}\ Cov_t[\sum_{i=1}^{k} rx^{A}_{t+i}, \sum_{i=0}^{k-1} r_{t+i}]$ and simiilarly for asset B.

to the A market, promoting cross-market integration over time. Cross-market capital flows become more responsive to differences in risk-reward between markets when the two assets are closer substitutes (i.e., when $R^{2(k)}_{AB}$ is higher). In the limit as the two assets become perfect substitutes, $R^{2(k)}_{AB}$ approaches 1 and generalist investors become extremely aggressive in exploiting any cross-market pricing differences.

2.2.2.3 Equilibrium yields

In market A at time t, there is a mass q_A of fast-moving specialists, each with demand $b_{A,t}$, and a mass $(1 - q_A - q_B)\,k^{-1}$ of active slow-moving generalists, each with demand $d_{A,t}$. These investors must accommodate the *active supply*, which is the total supply of $s_{A,t}$ less any supply held off the market by inactive generalist investors, $(1 - q_A - q_B)\,k^{-1}\sum_{j=1}^{k-1} d_{A,t-j}$. Thus, the market-clearing condition for asset A is

$$\overbrace{q_A b_{A,t}}^{\substack{\text{Specialist}\\ \text{demand}}} + \overbrace{(1 - q_A - q_B)k^{-1} d_{A,t}}^{\substack{\text{Active generalist}\\ \text{demand}}} = \overbrace{s_{A,t}}^{\substack{\text{Total bond}\\ \text{supply}}} - \overbrace{(1 - q_A - q_B)(k^{-1}\sum_{i=1}^{k-1} d_{A,t-i})}^{\substack{\text{Inactive generalist}\\ \text{holdings}}}. \tag{2.27}$$

The market-clearing condition for asset B is analogous.

We conjecture that equilibrium yields and generalist demands are linear functions of a state vector, $\mathbf{x}_t$, that includes the steady-state deviations of the short-term interest rate, the default realization, the supply of asset A, the supply of asset B, inactive generalist holdings of asset A, and inactive generalist holdings of asset B. Formally, we conjecture that long-term yields in market A and B are

$$y_{A,t} = \alpha_{A0} + \alpha'_{A1}\mathbf{x}_t, \tag{2.28}$$

$$y_{B,t} = \alpha_{B0} + \alpha'_{B1}\mathbf{x}_t, \tag{2.29}$$

and that the demands of slow-moving generalists are

$$d_{A,t} = \delta_{A0} + \delta'_{A1}\mathbf{x}_t, \tag{2.30}$$

$$d_{B,t} = \delta_{B0} + \delta'_{B1}\mathbf{x}_t, \tag{2.31}$$

where the $2(1+k) \times 1$ dimensional state vector, $\mathbf{x}_t$, is given by

$$\mathbf{x}_t = [r_t - \overline{r}, z_t - \overline{z}, s_{A,t} - \overline{s}_A, s_{B,t} - \overline{s}_B, d_{A,t-1} - \delta_{A0}, \cdots, d_{A,t-(k-1)} - \delta_{A0}, d_{B,t-1} - \delta_{B0}, \cdots, d_{B,t-(k-1)} - \delta_{B0}]'. \tag{2.32}$$

These assumptions imply that the state vector follows an AR(1) process

$$\mathbf{x}_{t+1} = \mathbf{\Gamma}\mathbf{x}_t + \epsilon_{t+1}, \tag{2.33}$$

where the transition matrix $\mathbf{\Gamma}$ depends on generalist demands.

As we show in the Internet Appendix, equilibrium yields take the same basic form as in (2.17) with only specialist investors. For market A, the yield is given by

$$\begin{aligned} y_{A,t} = & \overbrace{\left\{\overline{r} + \left(\frac{1-\theta_A}{1-\rho_r\theta_A}\right)(r_t - \overline{r})\right\}}^{\text{Expected future short rates}} \\ & + \overbrace{\left[(q_A\tau)^{-1} V_A^{(1)} \left(\overline{s}_A - (1-q_A-q_B)\,\delta_{A0}\right)\right]}^{\text{Unconditional term premia}} \\ & + \overbrace{\left[(q_A\tau)^{-1} V_A^{(1)} \left(\begin{array}{c} \frac{1-\theta_A}{1-\theta_A\rho_{s_A}}(s_{A,t} - \overline{s}_A) \\ -(1-\theta_A)(1-q_A-q_B)\,k^{-1}\sum_{i=0}^{\infty}\theta_A^i E_t[\sum_{j=0}^{k-1}(d_{A,t+i-j} - \delta_{A0})] \end{array}\right)\right]}^{\text{Conditional term premia}}, \end{aligned} \tag{2.34}$$

where $V_A^{(1)} = Var_t\left[rx_{A,t+1}\right]$ is the equilibrium variance of 1-period excess returns on asset A. The yield for asset B has an extra term relating to expected future defaults, but is otherwise similar

$$y_{B,t} = \overbrace{\left\{\overline{r} + \left(\frac{1-\theta_B}{1-\rho_r\theta_B}\right)(r_t - \overline{r})\right\}}^{\text{Expected future short rates}} + \overbrace{\left\{\overline{z} + \frac{1-\theta}{1-\rho_z\theta}\rho_z(z_t - \overline{z})\right\}}^{\text{Expected future default losses}} \tag{2.35}$$

$$+ \overbrace{\left[(q_B\tau)^{-1} V_B^{(1)} \left(\overline{s}_B - (1-q_A-q_B)\,\delta_{B0}\right)\right]}^{\text{Unconditional term/credit premia}}$$

$$+ \overbrace{(q_B\tau)^{-1} V_B^{(1)} \left[\begin{array}{c} \frac{1-\theta_B}{1-\theta_B\rho_{s_B}}(s_{B,t} - \overline{s}_B) \\ -(1-\theta_B)(1-q_A-q_B)\,k^{-1}\sum_{i=0}^{\infty}\theta_B^i E_t[\sum_{j=0}^{k-1}(d_{B,t+i-j} - \delta_{B0})] \end{array}\right]}^{\text{Conditional term/credit premia}}.$$

Although equations (2.34) and (2.35) show that yields in markets A and B take a similar algebraic form, the risk premia in the two markets will not be the same because of the different risks that market specialists must bear in equilibrium.

As explained further in the Internet Appendix, solving the model involves finding a solution to a system of $8k$ polynomial equations in $8k$ unknowns. Specifically, we need to determine the way that equilibrium yields and active generalist demand in markets A and B respond to shifts in asset supply in A and B: this generates 8 unknowns. We also need to determine how equilibrium yields and active generalist demand in A and B respond to the holdings of inactive generalists: this generates $8(k-1)$ unknowns.

As in the single-asset case, a solution only exists if investors are sufficiently risk tolerant (i.e., for τ sufficiently large). And, there can be a multiplicity of equilibrium solutions. However, as above, there is a unique solution that has well-behaved limiting behavior.

There are three separate forces that give rise to equilibrium multiplicity:

1. Since specialists have short-horizons, a steeply-downward sloping demand curve creates a self-fulfilling form of discount rate risk for specialists, just as in the single-asset model. However, the relevant and well-behaved solution features a smaller equilibrium response of A yields to A supply shocks, and similarly for asset B. As above, the solutions featuring a larger response to supply shocks explode in the limiting case where supply risk vanishes.

2. Although generalists have longer investment horizons than specialists, their investment horizons are still shorter than the maturity (perpetual) of the A and B assets. Since generalists are concerned about the supply risk associated with cross-market arbitrage, the degree of equilibrium segmentation between the A and B can be self-fulfilling. For instance, if yields in market B are insensitive to shocks to the supply of A (and vice versa), cross-market arbitrage becomes very risky for generalists. Hence, generalists will not aggressively integrate markets, making the low sensitivity of B yields to A supply shocks self-fulfilling. Conversely, if generalists behave as if markets are highly integrated, then cross-market arbitrage becomes less risky and, yields in B will be more sensitive to A supply shocks (and vice versa). However, the relevant and well-behaved solution always features more aggressive cross-market arbitrage and, thus, tighter cross-market integration. The solutions with weak cross-market arbitrage explode in the limit where supply risk vanishes: to induce generalists to absorb a A supply shock, the yields in B must drop massively in response to a tiny rise in the supply of A.

3. The final source of multiplicity stems from the way that active generalists and, therefore, bond yields react to the holdings of inactive generalists. In the unique, well-behaved equilibrium, active generalists reduce their holdings less than one-for-one in response to abnormally large holdings of inactive generalists. As a result, large holdings of inactive generalists reduce equilibrium yields. However, there are also solutions in which active generalists "overreact" to the holdings of inactive generalists, reducing their holding more than one-for-one. This can lead to situations where large holdings of inactive generalists actually raises equilibrium yields. This solution behaves oddly in the limit where the number of generalists grows vanishingly small, with a tiny number of active generalists taking extremely large bets.

We solve this system of polynomial equations numerically using the Powell hybrid algorithm. This algorithm performs a quasi-Newton search to find roots of a system of nonlinear equations starting from an initial guess vector. To find all of the roots, we apply this algorithm by sampling over $10,000$ different initial conditions. As discussed above, we restrict attention to those solutions where active generalists reduce their holdings less than one-for-one in response to abnormally large holdings of inactive generalists. Of these, we focus on the single solution where the price of A (B) is less sensitive to shocks to the supply of A (B) and is more sensitive to shocks to the supply of B (A).[10]

2.2.3 Defining market integration

What do we mean by "market integration"? We define markets as being *integrated in the short-run* if, at each date, *conditional risk premia* in both markets reflect the same conditional prices of factor risk. For example, the pricing of interest rate risk is conditionally integrated across markets if, at each date, the expected return per unit of exposure to short-rate shocks is the same in markets A and B. Similarly, we will say that markets are *integrated in the long-run* when average, or *unconditional risk premia* in both markets reflect the same unconditional prices of risk. Unconditional integration is therefore a weaker form of market integration than conditional integration.

Note that, in our model, market integration has nothing to do with the speed by which *fundamental cash flow news* is reflected in asset prices. In our model, fundamental cash flow news is reflected instantaneously in both markets. To see this, consider the terms in curly brackets in equations (2.34) and (2.35) above. Both A and B share exposure to news about changes in future short rates and this news is reflected *identically* in their yields.

In our model, the degree of market integration depends on which investors can bear risk at different horizons and is driven by two parameters: $(1 - q_A - q_B)$ and k. The first parameter,

[10] Specifically, we select solution vectors that satisfy the restrictions $-1 < \sum_{i=1}^{k-1} \delta_{A1}[d_{A,t-i}] < 0$ and $-1 < \sum_{i=1}^{k-1} \delta_{B1}[d_{B,t-i}] < 0$, where $\delta_{A1}[d_{A,t-i}]$ denotes the element of the δ_{A1} solution vector that captures the way that active generalists' demands for A responds to inactive generalists' holdings of A in period $t-i$. We then pick the single solution among the remaining with the smallest value of $\alpha_{A1}[s_A]$ and $\alpha_{B1}[s_B]$.

$(1 - q_A - q_B)$, is the population share of generalists. This parameter determines the degree of long-run integration between markets. For instance, if $(1 - q_A - q_B) \approx 1$, markets will be well integrated in the long-run even if k is large. The second parameter, k, indexes the speed with which generalist capital can flow between markets. Thus, k determines the degree of short-run integration. Markets are perfectly segmented if $(1 - q_A - q_B) = 0$ or $k \to \infty$. If either of these conditions holds, the two markets operate independently of each other.

Formally, collect all of the 1-period returns in a vector $\mathbf{rx}_{t+1}$ and the asset supplies in a vector $\mathbf{s}_t$. Letting $rx_{M_t,t+1} = \mathbf{s}_t' \mathbf{rx}_{t+1}$, markets are *integrated in the short-run* if

$$\begin{aligned} E_t \left[\mathbf{rx}_{t+1} \right] &= \tau^{-1} Var_t \left[\mathbf{rx}_{t+1} \right] \mathbf{s}_t \\ &= \beta_t \left[\mathbf{rx}_{t+1}, rx_{M_t,t+1} \right] E_t \left[rx_{M_t,t+1} \right] \end{aligned} \tag{2.36}$$

where $\beta_t \left[\mathbf{rx}_{t+1}, rx_{M_t,t+1} \right] = Var_t \left[\mathbf{rx}_{t+1} \right] \mathbf{s}_t / \left(\mathbf{s}_t' Var_t \left[\mathbf{rx}_{t \to t+j} \right] \mathbf{s}_t \right)$ and $E_t \left[rx_{M_t,t+1} \right] = \mathbf{s}_t' E_t \left[\mathbf{rx}_{t+1} \right]$. In other words, markets are integrated in the short-run if, at each date, a conditional-CAPM based on the current market portfolio ($rx_{M_t,t+1} = \mathbf{s}_t' \mathbf{rx}_{t+1}$) prices both the A and B assets. In our model, markets are integrated in the short-run if and only if $(1 - q_A - q_B) = 1$ and $k = 1$.[11]

Similarly, markets are *integrated in the long-run* if

$$\begin{aligned} E \left[\mathbf{rx}_{t \to t+k} \right] &= \tau^{-1} Var_t \left[\mathbf{rx}_{t \to t+k} \right] E \left[\mathbf{s}_t \right] \\ &= \beta[\mathbf{rx}_{t \to t+k}, rx_{\overline{M},t \to t+k}] E[rx_{\overline{M},t \to t+k}], \end{aligned} \tag{2.37}$$

where $\beta[\mathbf{rx}_{t \to t+k}, rx_{\overline{M},t \to t+k}] = Var_t \left[\mathbf{rx}_{t \to t+k} \right] E \left[\mathbf{s}_t \right] / \left(E \left[\mathbf{s}_t' \right] Var_t \left[\mathbf{rx}_{t \to t+k} \right] E \left[\mathbf{s}_t \right] \right)$ and $E[rx_{\overline{M},t \to t+k}] = E \left[\mathbf{s}_t' \right] E \left[\mathbf{rx}_{t \to t+k} \right]$. In other words, markets are integrated in the long-run if the same unconditional-

[11] In our setting, a conditional-CAPM holds if and only if the conditional prices of factor risk are the same in both markets at each date. To see this, write $rx_{A,t+1} - E_t \left[rx_{A,t+1} \right] = \phi_A' \varepsilon_{t+1}$ where ε_{t+1} are the (four) factor innovations and ϕ_A are the factor loadings for asset A. Proceeding similarly for market B and stacking these equations, we have $\mathbf{rx}_{t+1} - E_t \left[\mathbf{rx}_{t+1} \right] = \mathbf{\Phi} \varepsilon_{t+1}$ where $\mathbf{\Phi} = [\phi_A \ \phi_B]'$. Therefore, when $(1 - p_A - p_B) = 1$ and $k = 1$, we have $E_t \left[\mathbf{rx}_{t+1} \right] = \tau^{-1} Var_t \left[\mathbf{rx}_{t+1} \right] \mathbf{s}_t = \mathbf{\Phi} \left(\tau^{-1} \mathbf{\Sigma} \mathbf{\Phi}' \mathbf{s}_t \right) = \mathbf{\Phi} \lambda_t$ where $\lambda_t = \left(\tau^{-1} \mathbf{\Sigma} \mathbf{\Phi}' \mathbf{s}_t \right)$ are the (four) conditional prices of factor risk at time t. By contrast, when $(1 - p_A - p_B) \neq 1$ and $k \neq 1$, there is no conditional CAPM that will price the 1-period returns on the A and B assets.

CAPM based on the average market portfolio ($rx_{\overline{M},t\to t+k} = E[\mathbf{s}'_t]\,\mathbf{rx}_{t\to t+k}$) prices both the A and B assets on average. In our model, markets are integrated in the long-run if and only if $(1-q_A-q_B)=1$, irrespective of k.

Economically, the reason markets are not integrated is because cross-market arbitrage is risky for generalists, much like in Gromb and Vayanos (2002). Unless $(1-q_A-q_B)=1$ and $k=1$, short-run integration fails because generalists demand compensation for the risk associated with the short-run trades they place to exploit the cross-market pricing differences that arise following supply shocks. Similarly, unless $(1-q_A-q_B)=1$, long-run integration fails because generalists are engaged in a risky "cross-market arbitrage" trade even in the long run and must be compensated for its risks. Specifically, when $(1-q_A-q_B)<1$, generalists will not hold the market portfolio in the steady-state (i.e., $E[d_{A,t}]\neq E[s_{A,t}]$ and $E[d_{B,t}]\neq E[s_{B,t}]$). Relative to the market portfolio, generalists' portfolio will incorporate a tilt that reflects cross-market pricing differences. And, generalists will demand compensation for bearing the risks stemming from this portfolio tilt.[12] Thus, in the general case where $(1-q_A-q_B)<1$ and $k>1$, we obtain neither short-run nor long-run market integration.

How should one think about the relevant values for $(1-q_A-q_B)$ and k empirically? Clearly, the relevant values of $(1-q_A-q_B)$ and k depend crucially on the two markets being considered. For instance, U.S. Treasury and U.S. Agency bonds are often overseen by the same portfolio manager within a large institution. As a result, we would expect $(1-q_A-q_B)$ to be near 1 and k to be low, so the two markets would be tightly integrated even in the short-run: Treasury supply shocks would be rapidly transmitted to Agency debt markets and vice versa. However, in other cases, such as U.S. Treasury bonds and corporate bonds, or the fixed-income market and the equity market, it is natural to think that $(1-q_A-q_B)$

[12] In the symetric case where $q_A=q_B$ and $\overline{s}_A=\overline{s}_B$, we have $\delta_{B0}>\delta_{A0}$. The reason is that the B asset is riskier than the A asset since the former is exposed to cash-flow risk. As a result, B-specialists will hold less of the B-asset than A-specialists hold of the A-asset. Relative to the market portfolio, this means that generalists will be overweight the B asset and underweight the A asset.

is well below 1 and that $k > 1$. Although different asset classes are often held by the same generalists—e.g., pension funds or endowments, most of these investors are quite slow to reallocate capital.

2.3 Market integration following large supply shocks

How do prices adjust across different asset markets following supply shocks? Here we use our model to explore asset price dynamics following shocks.

Table 2.1 lists the illustrative set of parameter values that we use in these numerical exercises. For the purposes of these illustrations, it may be helpful to think of market A as the U.S. Treasury market and market B as the corporate bond market. We use annualized values so that one period in our numerical exercises corresponds to one year. The total average supply of assets in each market is normalized to be one unit.

Table 2.1: Illustrative model parameters

This table presents the illustrative model parameters that we use throughout our numerical exercises. We use annualized values so that one period corresponds to one year.

Parameter	Description	Value
q_A, q_B	Percentage of investors that are specialists in A and B	45%
k	Number of periods between generalist portfolio rebalancing	4
$\bar{r}$	Average short-term riskless rate	4%
σ_r	Volatility of annual shocks to short-term riskless rate	1.3%
ρ_r	Annual persistence of short-term riskless rate	0.85
$\bar{z}$	Expected default losses per annum on asset B	0.2%
σ_z	Volatility of annual shocks to default losses on asset B	0.7%
ρ_z	Annual persistence of default losses on asset B	0.85
$\bar{s}_A, \bar{s}_B$	Average asset supplies	1
$\sigma_{s_A},\ \sigma_{s_B}$	Volatility of annual supply shocks	0.6
$\rho_{s_A},\ \rho_{s_B}$	Annual persistence of supply shocks	0.999
$D_A,\ D_B$	Macaulay duration in years (implies $\theta_A = \theta_B = 0.8$)	5 years
τ	Investor risk tolerance	50

We have calibrated the parameters in our model to the extent possible to actual data on U.S.

fixed income markets as described in the internet appendix. For parameters that cannot be calibrated, we used reasonable values that are illustrative of the main takeaway.

We begin our analysis by choosing $k = 4$ years and $q_A = q_B = 45\%$, but later show comparative statics for these parameters. Based on these values, our simulations assume that most of the capital in each market is operated by specialists, with 10% being controlled by flexible generalist investors, one-fourth of whom re-allocate their portfolios each year. Our choice of $k = 4$ is somewhat arbitrary, but we think of this as capturing the empirically relevant case of pension funds or endowments who typically review their asset allocations on an annual or biannual basis and, even then, only sluggishly adjust their portfolios towards some evolving target. A recent paper by Bacchetta and van Wincoop (2017) have estimated the portfolio re-allocation frequency to be in line with our assumption.

A simple way to think about $k = 4$ is that upon a shock, one-fourth of the capital is able reallocate in a relative fast manner, while the remaining three-fourth of the total capital reacts sluggishly. Given that relative small share of capital are managed by unrestricted asset managers with the ability to react quickly, e.g. macro hedge funds, even on a levered basis, we think having one-fourth of the total capital being able to react quickly is not an extreme choice.

2.3.1 Unanticipated supply shocks

Baseline example We first consider the impact of an unanticipated supply shock that increases the supply of asset A (Treasuries) by 50% in period 10. To make the intuition as stark as possible, we focus on the case of a near-permanent supply shock and set $\rho_{s_A} = 0.999$. Specifically, Figure 2.1 illustrates the price impact of this shock, plotting the evolution of expected annual returns and bond yields in market A (Treasuries) and market B (corporate bonds). Figure 2.2 shows how specialists and generalist investors adjust their holdings in response to the shock.

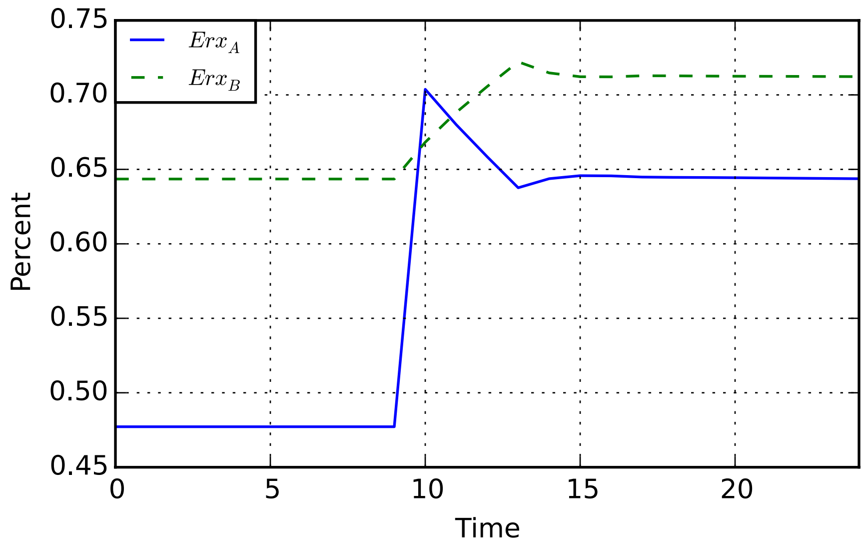
Prior to the supply shock in period 10, Figure 2.1 shows that the risk premium in market B (corporate bonds) is 0.64% per annum versus a risk premium of 0.48% in market A (Treasuries). The additional risk premium of 0.16% obtains because market B (corporate bonds) is subject to default risk, which exposes investors to an additional source of cash flow risk and amplifies the supply risk facing corporate bond holders. The initial yield in market B is 4.84% per annum versus a yield of 4.48% in market A. The steady-state yield in market A equals the average short-term riskfree rate of $\overline{r} = 4.00\%$ plus the steady-state risk premia of 0.48%. The 0.36% steady-state yield spread between the B and A markets equals the difference in steady-state risk premia of 0.16% plus the market B's expected default losses of $\overline{z} = 0.20\%$ per annum.

When the supply shock hits the market A in period 10, expected returns and yields in both markets react immediately. Figure 2.1.A shows that expected returns in market A *overreact* and reach a peak of 0.70% before ultimately falling back to a long-run level of 0.64%. The overreaction of expected returns for asset A illustrates a general property of models that feature slow-moving capital and has been shown in some models, such as Duffie (2010), namely, the relative *steepness* of short-run demand curves and relative *flatness* of long-run demand curves.

In contrast, Figure 2.1.A shows the key novel implication of our model: expected returns in market B actually *underreact* to the shock to the supply of A, rising slowly from 0.64% to a new long-run level of 0.72%. Why does market A overreact to the supply shock while market B underreacts? Figure 2.2.A shows how the positions of different market participants evolve over time. Following the initial supply shock in market A, both specialist demand in A ($b_{A,t}$) and active generalist demand in A ($d_{A,t}$) spike upwards. As a partial hedge against their increased holdings of A, active generalists reduce their holdings in market B. This reduction in generalists' B holdings is motivated by a need to reduce the common short-rate risk (and

Figure 2.1: Price impact of an unanticipated shock to the supply of asset A
This figure shows the impact on annual bond risk premia and bond yields of an unanticipated shock that increases the supply of asset A by 50% in period 10. Panel A shows the evolution of annual bond risk premia in market A, $E_t[rx_{A,t}+1]$, and market B, $E_t[rx_{B,t}+1]$, over time. Panel B shows the evolution of bond yields in market A, $y_{A,t}$, and market B, $y_{B,t}$, over time.

Panel A: Annual bond risk premia



Panel B: Bond yields

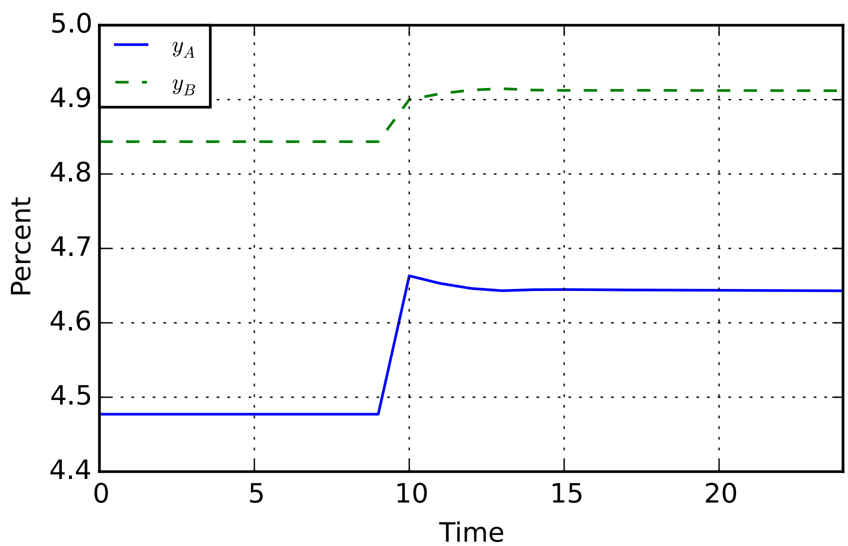
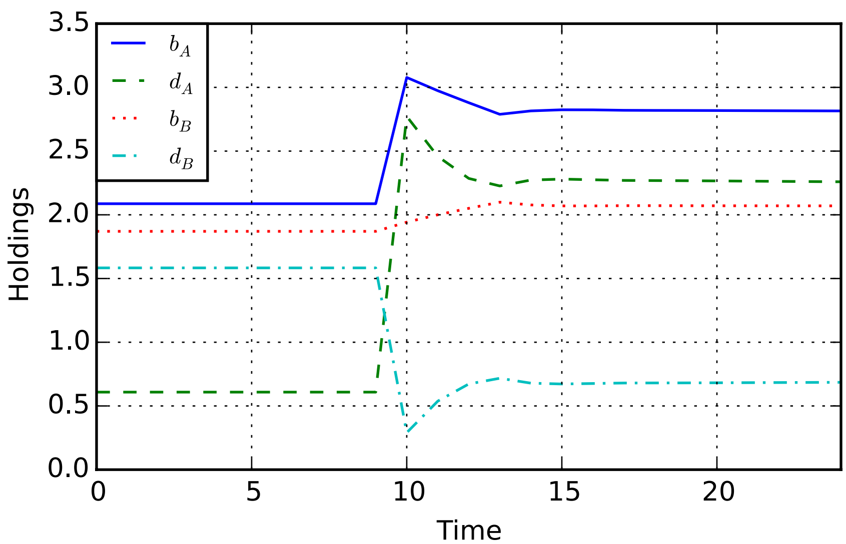


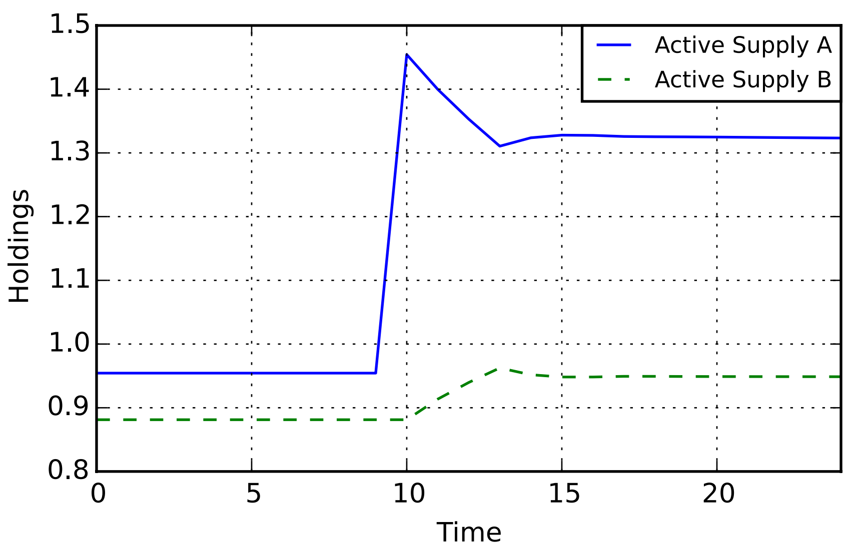
Figure 2.2: Portfolio adjustments in response to an unanticipated shock to the supply of asset A

This figure shows the impact on investor positions and active asset supplies of an unexpected shock that increases the supply of asset A by 50% in period 10. Panel A shows the evolution of specialists holdings in markets A and B ($b_{A,t}$ and $b_{B,t}$) as well as the positions of active generalists ($d_{A,t}$ and $d_{B,t}$). Panel B shows the evolution of the "active supplies" of assets A and B. The active supply of A is $s_{A,t} - (1 - q_A - q_B)k^{-1}) \sum_{i=1}^{k-1} d_{A,t-i}$ and the active supply of B is defined analogously.

Panel A: Specialist holdings and positions of active generalists



Panel B: Active asset supply



supply risk) across their holdings in both markets. To fill the void left by the generalists, specialists in market B must hold more of the B asset. As time passes and more generalists reallocate their portfolios in response to the shock, the active demands for A decline slowly towards their new long-run levels.

In our model, the dynamics of risk premia are tied to the dynamics of the “active supply” of A and B that must be absorbed by active market participants each period. By "active supply" we mean the total supply less the assets that are being held off the market by inactive generalists. The evolution of the active supplies is shown in Figure 2.2.B. The dynamics of active supply mirror those for bond risk premia shown in Figure 2.1.A. The initial supply shock to A in period 10 immediately increases the active supply in A but has no immediate effect on the active supply of B. This is because slow-moving generalists have yet to reduce their holdings in market B. Over the ensuing periods, generalists gradually increase their holdings of A and reduce their holdings of B. Therefore, the active supply in A gradually declines while the active supply in B gradually rises.

Recall that $k = 4$ in this example, so by period 13 all generalist investors have re-allocated their portfolios in response to the supply shock in period 10. However, the gradual adjustment of generalists gives rise to modest echo effects after period 13, generating a series of dampening oscillations that converge to the new long-run equilibrium. As in Duffie (2010), these oscillations arise because generalists who reallocate soon after the supply shock hits take large opportunistic positions. These large positions temporarily reduce the active supply of A and then need to be absorbed in later periods.[13]

Because markets are partially segmented, large supply shocks can have surprising effects on seemingly unrelated risk premia in our model. For example, because it triggers significant cross-market capital flows, the shock to the supply of asset A (Treasuries) actually raises

[13] Using an extension of Duffie’s (2010) model, Bogousslavsky (2016) argues that infrequent rebalancing can explain the echo effects found intra-day and daily returns.

the risk premium that corporate bond investors earn for bearing default risk in market B (corporate bonds), even though Treasury bonds themselves have no exposure to default risk [14].

In Figure 2.1.B, we shift from risk premia to bond yields, which are simply risk premia integrated over time. The overreaction of the A market and the underreaction of the B market is more muted in yield space than in risk premium space. This is natural since bond yields reflect weighted averages of future bond risk premia.[15] Market A yield overreacts by 11% of the total long-run impact and market B yield underreacts by 19% of the total long-run impact.

Comparative statics In Table 2.2, we perform a variety of comparative statics exercises to illustrate how the price dynamics following supply shocks depend on the parameters of our model. We focus on the parameters governing market structure: the population share of generalist investors $(1 - q_A - q_B)$ and the frequency at which generalists can rebalance (k).

For a given set of model parameters, we summarize the impact of the supply shock on both the A and B markets by listing the yields and expected annual returns in (i) the period before the shock arrives (labeled as "pre-shock level"), (ii) the period when the shock arrives (labeled as "short-run Δ"), and (iii) in $2k$ periods after the shock arrives (labeled as "long-run Δ").

We define the degree to which bond yields over- or underreact as the difference between the short-run change and the long-run change, expressed as a percentage of the long-run change[16]

[14] In this way, our model may shed light on the otherwise puzzling finding that central bank purchases of long-term government bonds appear to have reduced credit risk premia (Krishnamurthy and Vissing-Jorgensen [2011]).

[15] Specifically, generalizing (2.16) we have $y_{A,t} = (1-\theta_A)\sum_{i=0}^{\infty}\theta_A^i E_t[r_{t+i} + \tau^{-1}V_A^{(1)} b_{A,t+i}]$ and $y_{B,t} = (1-\theta_B)\sum_{i=0}^{\infty}\theta_B^i E_t[r_{t+i} + z_{t+i+1} + \tau^{-1}V_B^{(1)} b_{B,t+i}]$.

[16] Since our supply shock is not quite permanent, we subtract off the constant $(1-\rho_{S_A}^{2k})/\rho_{S_A}^{2k}$ from *%Over-*

Table 2.2: Model comparative statics

This table shows how the price impact of the same supply shock — an unanticipated shock that increases asset supply by 50% — varies as we change key model parameters one at a time. All other parameters are held constant at the values listed in Table 2.1. For a given set of model parameters, we summarize the impact of the supply shock on both the A and B markets by listing (i) the yields and expected annual returns in the period before the shock arrives (labeled as "pre-shock level"), (ii) the changes in yields and expected annual returns in the period when the shock arrives (labeled as "short-run Δ"), and (iii) in $2k$ periods after the shock arrives (labeled as "long-run Δ"). Finally, we report the degree to which bond yields over- or underreact as the difference between the short-run change and the long-run change, expressed as a percentage of the long-run change

$$\%\textit{Over-Reaction}(y) \equiv \frac{(y_t - y_{t-1}) - (y_{t+2k} - y_{t-1})}{(y_{t+2k} - y_{t-1})}.$$

Our measure of over-reaction for risk premia is defined analogously.

		Market *A*								Market *B*							
		Risk premia, $E_t[rx_{A,t+1}]$				Yields, $y_{A,t}$				Risk premia, $E_t[rx_{B,t+1}]$				Yields, $y_{B,t}$			
		Pre shock level	Short run Δ	Long run Δ	Over-react	Pre shock level	Short run Δ	Long run Δ	Over-react	Pre shock level	Short run Δ	Long run Δ	Over-react	Pre shock level	Short run Δ	Long run Δ	Over-react
Supply shock hits market *A*																	
(1) Base case		0.48	0.23	0.17	35%	4.48	0.19	0.17	11%	0.64	0.02	0.07	-65%	4.84	0.06	0.07	-19%
(2) More risk tolerant	$\tau = 60$	0.39	0.18	0.14	35%	4.39	0.15	0.14	11%	0.52	0.02	0.06	-65%	4.72	0.05	0.06	-19%
(3) No Generalists	$q_A = q_B = 0.5$	0.47	0.24	0.23	0%	4.47	0.23	0.23	0%	0.73	0.00	0.00	NA	4.93	0.00	0.00	NA
(4) More Generalists	$q_A = q_B = 0.2$	0.46	0.35	0.12	180%	4.46	0.18	0.12	47%	0.58	0.13	0.11	20%	4.78	0.12	0.11	14%
(5) More *B* specialists	$q_A = 0.3, q_B = 0.6$	0.59	0.34	0.22	56%	4.59	0.26	0.22	17%	0.56	0.03	0.07	-64%	4.76	0.06	0.07	-18%
(6) Fast-adjusting Generalists	$k = 2$	0.47	0.20	0.17	21%	4.47	0.17	0.17	4%	0.64	0.04	0.07	-39%	4.84	0.06	0.07	-6%
(7) Slow-adjusting Generalists	$k = 6$	0.48	0.24	0.17	40%	4.48	0.20	0.17	17%	0.65	0.02	0.07	-76%	4.85	0.05	0.07	-30%
(8) Larger *A*, smaller *B* market	$\bar{s}_A = 5/3, \bar{s}_B = 1/3,$ $q_A = 0.75, q_B = 0.15$	0.38	0.11	0.10	14%	4.38	0.10	0.10	4%	0.55	0.03	0.07	-60%	4.75	0.06	0.07	-16%

$$\%Over\text{-}Reaction(y) \equiv \frac{(y_t - y_{t-1}) - (y_{t+2k} - y_{t-1})}{(y_{t+2k} - y_{t-1})}.$$

Our measure of over-reaction for risk premia, $\%Over\text{-}Reaction(E[rx])$, is defined analogously. According to this definition, using our baseline set of parameters, yields in market A overreact by approximately 11%, while yields in market B underreact by 19%.

The second row in Table 2.2 shows that, if market participants are more risk tolerant, this reduces the price impact of the supply shock on both market A and market B. Changing investor risk tolerance has a similar impact on the short- and long-run response of yields to shocks. Thus, the degree of overreaction or underreaction in each market is unchanged in percentage terms.

We next change the mix between generalist and specialist investors. q_A and q_B indicate the relative fraction of specialists in market A and B, respectively. In row 3, we set $q_A = q_B = 0.5$ so there are no generalists and the two markets are completely segmented: a supply shock in the market A is not transmitted to the market B and vice versa.

In contrast, in the case of many generalists and few specialists, the markets are well integrated, so that both the A and B markets overreact to a supply shock that directly hits only the A market. In this case, shown in row 4 which sets $q_A = q_B = 0.2$, the two markets behave as essentially one and the result is similar to the single-market case with slow-moving capital studied in Duffie (2010).

We next change the mix between market A specialists and market B specialists, holding fixed the overall mix between generalists and specialists. Row 5 of Table 2.2 shows that if we hold the total number of specialists the same at $q_A + q_B = 0.9$, then as we increase the proportion of specialists in B and decrease the proportion of specialists in A, we get

Reaction to ensure that our measure is zero the case of perfectly conditionally-integrated ($1 - p_A - p_B = k = 1$) or perfectly segmented markets ($1 - p_A - p_B = 0$) in which there is no "over-reaction" but only "reaction." This is because in these limiting cases we have $[(y_t - y_{t-1}) - (y_{t+2k} - y_{t-1})] / [y_{t+2k} - y_{t-1}] = \left[\alpha_{s_A} s_{A,t} - \alpha_{s_A} s_{A,t} \rho_{s_A}^{2k}\right] / \alpha_{s_A} s_{A,t} \rho_{s_A}^{2k} = (1 - \rho_{s_A}^{2k}) / \rho_{s_A}^{2k}$. For $k = 4$ and $\rho_{s_A} = 0.999$, this constant is 0.8%.

more over-reaction in A. The B market is only modestly affected by this change because the supply shock is primarily being absorbed by generalists anyway.

Recall that k is the number of periods it takes for generalists to fully reallocate their portfolios and that $k = 4$ in our base case. In row 6 we instead set $k = 2$, so half the generalists reallocate their portfolio each period, and the other half reallocate in the next period. Naturally, this smaller value of k reduces the over-reaction in market A and the under-reaction in market B. Similarly, when we set $k = 6$ in row 7, there is more over-reaction in market A, and more under-reaction in market B.

Note that k also affects the unconditional risk premium that investors earn over the long run. As we increase k, there are two competing effects on unconditional risk premium. On the one hand, generalists with longer horizons worry less about a fixed amount of transitory discount rate risk, leading to a decline in the unconditional price of discount rate risk.[17] On the other hand, the steady state quantity of discount rate risk that investors must bear actually grows with generalist horizons k.[18] As shown in Table 2.2, the latter effect generally tends to dominate.[19] In summary, as we increase k, *supply shocks* have a larger impact on conditional risk premia and this increase in *supply risk* tends to raise unconditional risk premia.

[17]Since mean-reverting supply shocks generate negative serial correlation in returns, the variance ratio $Var_t[rx_{A,t \to t+k}]/k$ will be decreasing in k holding fixed the endogenous parameters that govern the return generating process. Thus, as in Campbell and Viceira (2002), longer-horizon investors worry less about transitory supply (discount rate) risk, leading them to take larger positions in risky assets.

[18]Formally, as we increase k, the endogenous parameters that govern the return generating process are not held fixed. Since fewer long-horizon investors are active in a given period, the short-term price impact of supply shocks grows, leading to an rise in the quantity of discount rate risk.

[19]Formally, let $E[rx^A_{t+1}] = (p_A \tau)^{-1} \left(\bar{s}^A - (1 - p_A - p_B)\,\delta_{A0}\right) V^{(1)}_A$ denote the unconditional risk premium. We have

$$\frac{\partial E[rx^A_{t+1}]}{\partial k} = (p_A \tau)^{-1} \left[\frac{\partial V^{(1)}_A}{\partial k} \left(\bar{s}^A - (1 - p_A - p_B)\,\delta_{A0}\right) - V^{(1)}_A \left(1 - p_A - p_B\right) \frac{\partial \delta_{A0}}{\partial k} \right]$$

When $\sigma^2_{s_A}, \sigma^2_{s_B} = 0$, $\partial V^{(1)}_A / \partial k = \partial \delta_{A0} / \partial k = 0$ and unconditional risk premia are independent of k. When $\sigma^2_{s_A}, \sigma^2_{s_B} > 0$, we have $\partial V^{(1)}_A / \partial k > 0$ and $\partial \delta_{A0} / \partial k > 0$. In general, we find that the former effect tends to dominate, so that unconditional risk premia are increasing in k.

In row 8, we ask how our results depend on the relative sizes of the A and B markets. Relative to our base case of two equally sized markets, we find that generalists are better able to integrate a small market with a larger market. We keep $\overline{s}_A + \overline{s}_B = 2$ and $q_A + q_B = 0.9$, but we now assume that $\overline{s}_A = 5/3$ and $\overline{s}_B = 1/3$ so average supply in market A is $5\times$ that in market B. We also assume that $q_A = 0.45 \times \overline{s}_A$ and $q_B = 0.45 \times \overline{s}_B$ as in the baseline, which implies $q_A = 0.75$ and $q_B = 0.15$. As in our baseline, we consider a shock that raises the supply of A by 0.5. Row 8 shows that A over-reacts less and that B under-react less to the shock than under our baseline. The explanation is that market B is now much smaller relative to total generalist risk tolerance. As a result, a cross-market arbitrage position of a given size is better able to keep prices in market B close to those in market B.

2.3.2 Extensions

The model is also useful for understanding asset price dynamics following the announcement of a large future change in asset supply. For example, many central bank asset purchase programs occurring 2008 and 2013 were announced weeks or months before the asset purchases actually began. We can use the model to describe how prices and yields react to anticipated supply shocks, but leave the formal details for the Internet Appendix. As we describe below, the long-run impact of supply shocks on risk premia is the same whether or not the shock is pre-announced. But the short run effects can be quite different.

Consider the simple case of the pre-announcement of a one-time, near permanent jump in the supply of asset A. Pre announcing the supply shocks mobilizes slow-moving generalists before the supply of A actually rises. This early mobilization reduces the active supply of asset A that must eventually be absorbed when the shock lands, dampening the overreaction of prices and yields in market A. While this limits overreaction, it also lengthens the amount of time that it takes for the full impact of the shock to be reflected in prices, particularly in market B. The overall impact is that yields in both market A and market B rise more gradually to their new steady-state levels. Compared to the unanticipated shock case, we see less overreaction in market A (and potentially underreaction) and even greater underreaction in market B.

Another extension we have explored is whether the results carry over to a more complex setting in which there are multiple risky assets trading in each market. Details of this analysis are presented in the Internet Appendix. Concretely, in describing the model so far, we have made no distinction between "asset" A and "market" A. The distinction takes on meaning when the market A contains many securities, but flows between all of these securities and market B are limited by the generalist arbitrageurs. For example, in the Treasury market, there are many different securities with different maturities and thus exposure to interest rate risk.

Subject to some mild conditions which guarantee that cross-market arbitrage remains risky, the intuitions from the two risky asset model carry over to a richer setting with more securities[20]. We show that the dynamics of the price of interest rate risk and default risk in markets A and B follow the same path as in the simpler case, but different securities in these markets vary in their exposure to changes in these risk prices. Specifically, a conditional CAPM prices all assets in the first market, and a different conditional CAPM prices all assets in the second market. The two market specific pricing models are linked over time by the cross-market arbitrage activities of the slow-moving asset allocators, who take steps to equalize the prices of risk in the two markets. And, much as before, the degree of market integration depends on the risks faced by cross-market arbitrageurs.

To illustrate the impact of a supply shock on two different markets that each contain multiple assets, we provide the following example. Specifically, we numerically solve the model in the

[20]The key condition is that generalists are unable or unwilling to use the A assets to construct a factor mimicking portfolio that isolates exposure interest rate shocks. If the generalists can do this in both markets, then a riskless arbitrage exists.

case where generalists re-allocate their portfolios every $k = 2$ periods and with $N = 2$ assets in both the A and B markets. The two assets in each market differ solely in their durations. The short-term bonds, denoted $A1$ and $B1$, have a duration of 2 years (i.e., $D_{A_1} = D_{B_1} = 2$). The long-term bonds, denoted $A2$ and $B2$, have duration of 10 years (i.e., $D_{A_2} = D_{B_2} = 10$). As before, the two bonds in market A are default free whereas the two bonds in market B are exposed to default risk.[21]

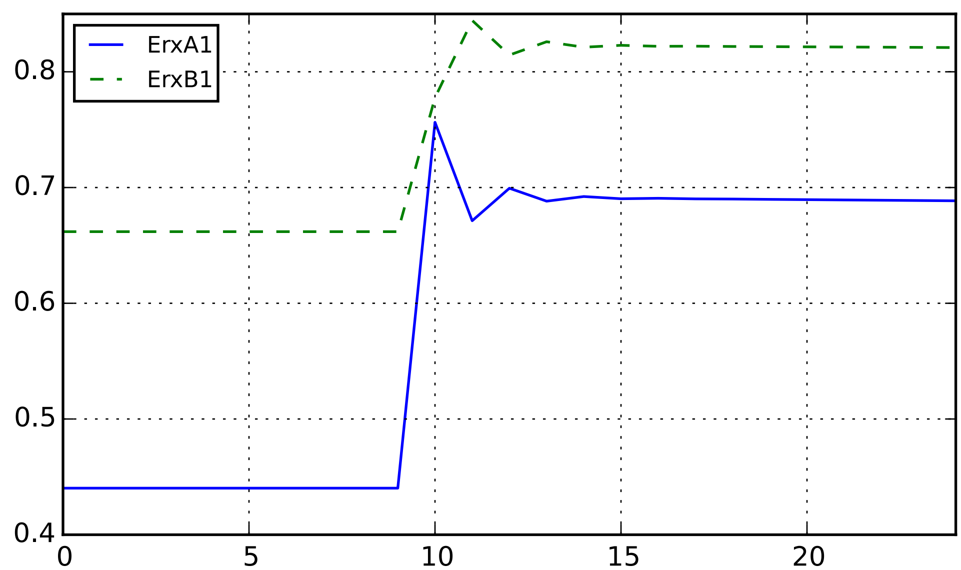
Figure 2.3 shows the evolution of risk premia following an unexpected shock at time 10 which permanently increases the supply of long-term default-free bonds ($A2$) and reduces the supply of short-term default-free bonds ($A1$) by an equal amount. This scenario corresponds to a "reverse Operation Twist" in which the Federal Reserve sells long-term Treasuries and reinvests the proceeds in short-term Treasuries.

Since this supply shock increases the total amount of interest rate risk than investors must bear, Figure 2.3 shows that the risk premia for all four assets rise after impact. Furthermore, this supply shock has a larger impact on the risk premium for long-maturity bonds in each market, leading both the A and B yield curves to steepen (since the shock is permanent). These patterns are consistent with those generated by existing models of bond supply shocks (e.g., Greenwood and Vayanos [2014] and Greenwood, Hanson, and Vayanos [2015]). However, since markets are partially segmented in our example, it takes time for the slow-moving generalists to integrate the A and B markets following the shock, leading the risk premia of the two A assets to initially over-react and the risk premia of the two B assets to initially underreact in Figure 2.3. Furthermore, because cross-market arbitrage is risky for generalists, market integration remains imperfect even in the long run. Specifically, even many period after the supply shock, Figure 2.3 shows that the yield curve in market A has steepened more than the yield curve in market B.

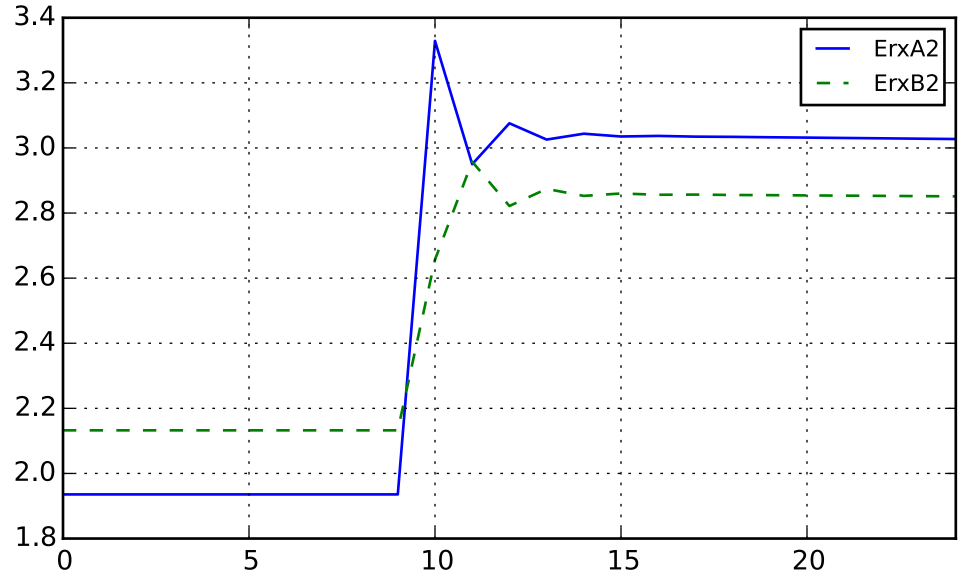
[21] For simplicity, we assume that $B1$ and $B2$ have the same exposure to the common default process, z_t. We also assume that there is no idiosyncratic default risk—i.e., $u_{B_1,t} = u_{B_2,t} \equiv 0$.

Figure 2.3: Price impact with multiple securities in each market
This figure shows the impact on bond risk premia of an unexpected supply shock that permanently increases the supply of long-term default-free bond (A2) and decreases the supply of short-term bond (A1) by an equal amount at time 10. Panel A shows the evolution of risk premia for short-term securities in each market (A1 and B1). Panel B shows the evolution of risk premia for long-term securities in each market (A2 and B2).

Panel A: Risk premia for short-maturity bonds in each market



Panel B: Risk premia for long-maturity bonds in each market



2.4 Discussion and Applications

2.4.1 Event studies and changes in the price of risk

In response to a rapidly evolving financial crisis and worldwide recession, in late 2008 and early 2009, central banks around the world announced their intention to aggressively purchase government bonds and other long-term debt securities. A crucial question in assessing the effectiveness of these asset purchase programs is whether they impacted securities prices beyond government bonds. Suppose, for example, that the impact of asset purchase programs was limited to markets in which the purchases were being made (Treasury bonds and MBS), perhaps because these markets are highly segmented from other financial markets. Such a finding should dampen central bankers' enthusiasm for these programs, and cast doubt that asset purchases could affect broader economic activity.

Our model provides a natural framework for understanding how these asset purchase programs should spill across different financial markets over time. According to our model, the largest short-run effects of these programs should be in the securities being purchased. In the long run, however, changes in risk premia in the market being targeted should spill over to non-targeted markets. Differences between the short-run and long-run price impact should reflect the degree to which the programs were anticipated, the length of time between the announcement date and implementation, and the effective degree of segmentation between different financial markets.

Most empirical studies of these purchase programs have used an event study methodology, focusing on the 1-day or even intraday impact on bond yields following announcements of future asset purchases. In one of the first of these event studies, Gagnon, Raskin, Remache, and Sack (2011) report interest rate changes around a set of Federal Reserve announcement days between November 2008 and January 2010. Cumulating over all announcement dates associated with the Fed's first round of quantitative easing (QE1), they report a 62 basis points decline in 10-year US Treasury yields, a 123 basis points decline in agency MBS yields, and a 74 basis points decline in Baa-rated corporate bond yields. Krishnamurthy and Vissing-Jorgensen (2011) extend this analysis to the Fed's second round of quantitative easing (QE2) and also discuss the impact on other assets, including high yield corporate bonds. After controlling for other factors, Krishnamurthy and Vissing-Jorgensen conclude that the effects of asset purchases were most pronounced among the assets being purchased (MBS and Treasuries in QE1 and Treasuries in QE2), suggesting a high degree of segmentation between different fixed income markets.

At the same time, some researchers have recognized that short horizon announcement returns may not capture the full impact of these asset purchase programs. In their empirical assessment of the Bank of England's quantitative easing program, Joyce, Lasoasa, Stevens and Tong (2010) suggest that it may have impacted corporate bonds and equities. Fratzcher, Lo Duca, and Straub (2013) suggest that the Fed's QE programs triggered portfolio flows that ultimately impacted emerging market asset prices and foreign exchange rates. Mamaysky (2014) suggests that QE might ultimately spill into the asset markets through portfolio allocation, but notes that "it is unlikely that such portfolio flows can take place quickly."

Researchers have used different approaches to measure the long-run effects of QE. Joyce, Lasoasa, Stevens and Tong (2010) report the cumulative change in asset prices for the longer period between March 4, 2009 and May 31, 2010 in addition to 1-day announcement returns. They show that corporate bond yields fall by a cumulative 70 basis points around asset purchase announcements, but by 400 basis points over the longer period. Mamaysky (2014) takes a more tailored approach to each asset market: he chooses an announcement window that maximizes the statistical power of the measured return. Using this approach, he shows that the impact of QE on both equity and high yield bond markets is much larger after 15 days than what one would measure using a 1-day window. But even this approach may significantly understate the long-run effect, because, as we have noted, the full impact of supply shocks may easily take quarters or years to be felt.

Our model clarifies the broader issue at stake: event studies are a useful methodology for detecting short-run price changes, but often lack the statistical power to detect changes in risk premia occuring at longer horizons. The event study methodology was originally developed in the 1970s to tackle questions of *informational efficiency* of stock prices, not changes in risk premia, but has increasingly been used in other settings, such as in event studies assessing QE.

The limitations of event studies are particularly severe when there is noise from "cash flow" news. For instance, consider the effects of σ_z—the volatility of fundamental cash-flow shocks in market B—on our ability to detect the impact on prices in market B stemming from a supply shock that hits market A . When σ_z is large relative to σ_r, there is insufficient power to detect changes in risk premia in market B. The statistical power would increase with the number of events, but power is nonetheless decreasing in σ_z. Thus, our model suggests that short-run event studies may have a hard time detecting spillover effects on markets, such as equities and high yield bonds, where there can be significant confounding news. More generally, our framework suggests that event studies are an inappropriate methodology for measuring cross-market price impact, at least in the short run.

Figure 2.4 illustrates more formally the potential inability of event studies to detect cross-market spillover effects from an unanticipated supply shock in market A. In an environment with low short-rate volatility, a supply shock to market A can have statistically significant impact on market A but not on market B.[22] Even though the shock to the supply of A impacts yields in the B market, the short-term effect is not statistically significant—e.g., the confidence interval for the 1-day change includes zero—and, thus, it would not be detected

[22]The parameter values used to generate Figure 13 reflect the low interest rate volatility during QE when short rates were pinned down at zero and supply risk was low. In addition, arbitrageur risk tolerance (τ) was small during this period.

by conventional event study techniques. In addition, the long-term effect, while economically meaningful, is also statistically insignificant.

In summary, our model suggests that we should be extremely cautious in using event studies to assess the long-run impact of supply shocks on market prices and risk premia. However, measuring the long-run impact of supply shocks across markets is inherently difficult because the full economic impact may occur over such a long time that it is swamped by other factors.

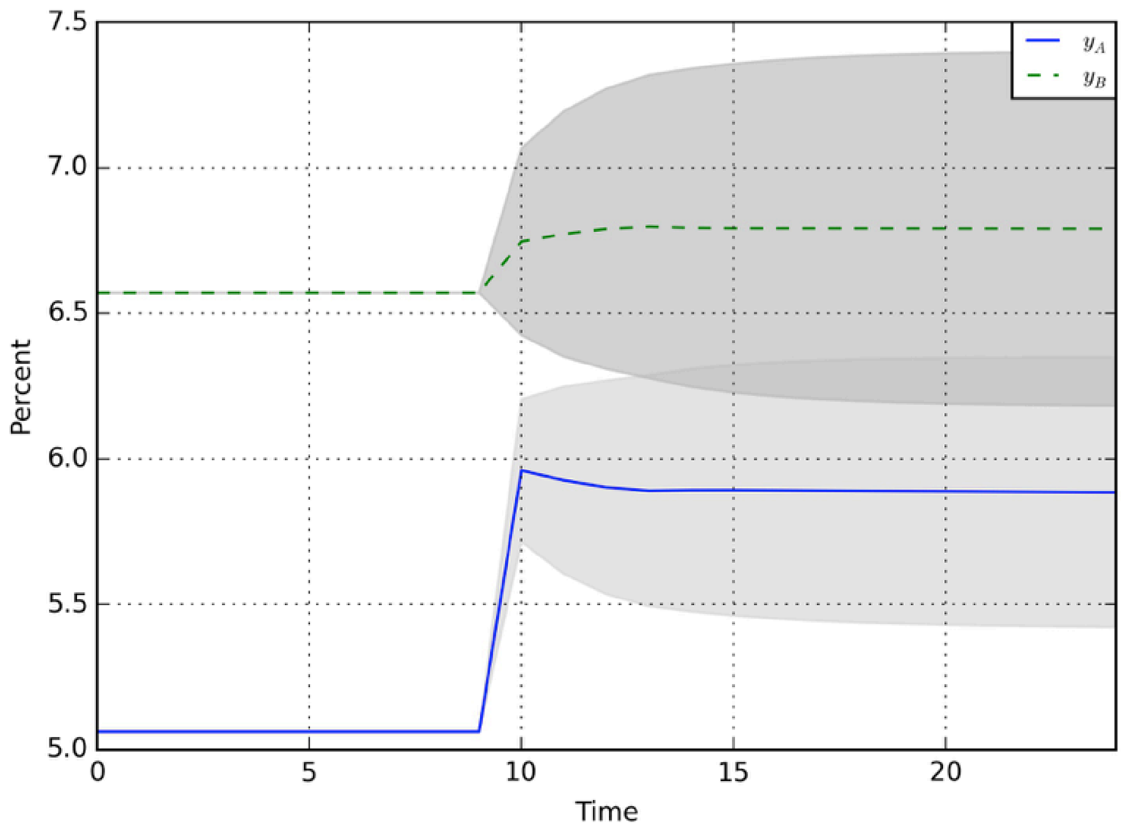
2.5 Conclusion

Modern financial markets are highly specialized. While specialization brings many benefits, the boundaries of securities markets are tested when there are large shocks to the supply of an entire asset class. In this paper, we develop a model to describe securities prices when shocks must draw in arbitrageurs from other related asset markets. We use the model to study the process by which capital flows across markets, and how quickly and by what magnitude prices adjust in different markets. Unlike textbook theories in which asset prices are determined solely by the *stock* of risky assets supplied, our approach suggests that *supply flows*—i.e., the rate at which the supply stock is changing—also matter in the short run. Even when a large amount of capital is mobile in the long run, different asset markets need not be fully integrated because market segmentation creates risks for arbitrageurs.

Our model explores the consequences of specialization when markets are hit with large shocks. However, we have taken the existence of specialists as given. But what determines the boundaries of specialists' expertise and, hence, the fault lines between different asset classes? Answering this question remains an important task for future research.

Figure 2.4: Event study confidence interval following an unanticipated shock to the supply of asset A

The yields of markets A and B and their respective 95% confidence intervals are shown. An unanticipated shock that doubles the supply of asset A is delivered in period 10. The following parameters are used: $\tau = 0.5$, $\sigma_{s_A} = \sigma_{s_B} = 0$, $\sigma_r = \sigma_z = 0.2\%$. All other parameters are the same as those listed in Table 2.1. For period $t > 9$, we compute the model-implied confidence interval for the cumulative changes in yields for market A and B from period 9, $y_{A,t} - y_{A,9}$ and $y_{B,t} - y_{B,9}$, assuming that all shocks are normally distributed. These confidence intervals are shaded in gray.



3 Options-Pricing Formula with Disaster Risk[1]

3.1 Introduction

We derive a new options-pricing formula that applies when disaster risk is the dominant force, when the size distribution of disasters is characterized by a power law, and when the economy has a representative agent with Epstein-Zin utility with a constant coefficient of relative risk aversion. Specifically, we consider far-out-of-the-money put options on the overall stock market, corresponding empirically to the S&P 500 in the United States and analogous indices for other countries. The pricing formula applies when the option is sufficiently far out of the money (operationally, a relative exercise price or moneyness of 0.9 or less) and when the maturity length is not too long (operationally, up to 6 months).

In the prescribed region, the elasticity of the put-options price with respect to maturity is close to one. The elasticity with respect to the exercise price is greater than one, roughly constant, and depends on the difference between the power-law tail parameter, denoted α, and the coefficient of relative risk aversion, γ. (This difference has to be positive for various

[1]This paper was written jointly with Robert J. Barro. We appreciate helpful comments and assistance with data from Josh Coval, Ben Friedman, Xavier Gabaix, Tina Liu, Matteo Maggiori, Greg Mankiw, Robert Merton, Richard Roll, Steve Ross, Emil Siriwardane, Jessica Wachter, and Glen Weyl, and participants in the macroeconomics seminar at Harvard University and finance seminar at MIT.

rates of return not to blow up.)

The options-pricing formula involves a term that is proportional to the disaster probability, p. This term depends also on three other parameters: γ, α, and the threshold disaster size, z_0. If these three parameters are fixed, we can use estimated time fixed effects to gauge the time variations in p. Our analysis modifies the options-pricing formula to allow for potential changes in p. Specifically, sharp increases in p can get far-out-of-the-money put options into the money without the realization of a disaster.

We show that the theoretical formula conforms with data from 1983 to 2017 on far-out-of-the-money put options on the U.S. stock market and analogous indices over shorter periods for other countries. Our analysis relies on two types of data—indicative prices on over-the-counter (OTC) contracts offered to clients by a large financial firm and market data provided by OptionMetrics, Bloomberg, and Berkeley Options Data Base. A key advantage of the OTC source is its provision of a rich array of contracts by exercise price and maturity. In particular, the relative exercise price goes down to 0.5, and the maturity can be 12 months or more. A downside of these data is that the reported prices do not necessarily correspond to actual trades. An advantage of the market data is the correspondence with actual trades, but there are problems with stale prices and sizes of bid-ask spreads. The most serious disadvantage of these data is the limited information on far-out-of-the-money options, which rarely trade. The market data (and trades) are also concentrated on short maturities; for example, about half of the OptionMetrics contracts have maturity of two months or less. In any event, we find that the main results are similar from the two types of data sources.

Extensions of the empirical analysis would allow for second-order terms. These terms involve the possibility of multiple disasters, the presence of a diffusion term, and allowances for discounting and expected growth.

3.2 Baseline Disaster Model and Previous Results

We use a familiar setup based on rare-macroeconomic disasters, as developed in Rietz (1988) and Barro (2006, 2009). The model is set up for convenience in discrete time. Real GDP, Y, is generated from

$$\log Y_{t+1} = \log Y_t + g + v_{t+1} + u_{t+1} \tag{3.1}$$

where, $g \geq 0$ is the deterministic part of growth, u_{t+1} (the diffusion term) is an i.i.d. normal shock with mean 0 and variance σ^2, and v_{t+1} (the jump term) is a disaster shock. Disasters arise from a Poisson process with probability of occurrence p per period. When a disaster occurs, GDP falls by the fraction b, where $0 < b \leq 1$. The distribution of disaster sizes is time invariant. (The baseline model includes disasters but not bonanzas.) This jump-diffusion process for GDP is analogous to the one posited for stock prices in Merton (1976, equations [1] [3])[2].

In the underlying Lucas (1978)-tree model, which assumes a closed economy, no investment, and no government purchases, consumption, Ct, equals GDP, Y_t. The implied expected growth rate of C and Y is given, if the period length is short, by

$$g^* = g - pE\left[b\right] + \frac{1}{2}\sigma_u^2 \tag{3.2}$$

where $E\left[b\right]$ is the mean of b. In this and subsequent formulas, we use an equal sign, rather than approximately equal, when the equality holds as the period length shrinks to zero. The representative agent has Epstein-Zin/Weil utility[3], as in Barro (2009):

$$(1-\gamma) U_t = \left\{ C_t^{1-\theta} + \frac{1}{1+\rho}\left[(1-\gamma) E_t U_{t+1}\right]^{\frac{1-\theta}{1-\gamma}} \right\}^{\frac{1-\gamma}{1-\theta}} \tag{3.3}$$

where $\gamma > 0$ is the coefficient of relative risk aversion, $\theta > 0$ is the reciprocal of the intertemporal-elasticity-of-substitution (IES) for consumption, and $\rho > 0$ is the rate of time preference. As shown in Barro (2009) (based on Giovannini and Weil [1989] and Obstfeld

[2]Related jump-diffusion models appear in Cox and Ross (1976).
[3]Epstein and Zin (1989) and Weil (1990)

[1994]), with i.i.d. shocks and a representative agent, the attained utility ends up satisfying the form:

$$U_t = \Phi \frac{C_t^{1-\gamma}}{1-\gamma} \tag{3.4}$$

where the constant $\Phi > 0$ depends on the parameters of the model. Using equations (3.3) and (3.4), the first-order condition for optimal consumption over time follows from a perturbation argument as

$$\left[E_t \left(\frac{C_{t+1}}{C_t} \right)^{1-\gamma} \right]^{\left(\frac{\gamma-\theta}{\gamma-1}\right)} = \frac{1}{1+\rho} E_t \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} R_{t+1} \right] \tag{3.5}$$

where R_{t+1} is the gross rate of return on any available asset from time t to time $t+1$. When $\gamma = \theta$—the familiar setting with time-separable power utility—the term on the left-hand side of equation (3.5) equals one.

The process for C and Y in equation (3.1) implies, if the period length is negligible:

$$E_t \left(\frac{c_{t+1}}{c_t} \right)^{1-\gamma} = 1 + (1-\gamma) g - p + pE \left[1-b\right]^{1-\gamma} + \frac{1}{2} (1-\gamma)^2 \sigma^2. \tag{3.6}$$

This condition can be used along with equation (3.5) to price various assets, including a risk-free bond and an equity claim on a perpetual flow of consumption (that is, the Lucas tree).

Equations (3.5) and (3.6) imply that the constant risk-free interest rate is given by

$$r^f = \rho + \theta g^* - p \left[E (1-b)^{-\gamma} - \frac{\gamma-\theta}{\gamma-1} + E (1-b)^{1-\gamma} - \theta E b + \left(\frac{1-\theta}{\gamma-1} \right) \right] - \frac{1}{2} \gamma (1+\theta) \sigma^2. \tag{3.7}$$

Let P_t be the price at the start of period t of an unlevered equity claim on the Lucas tree. Let V_t be the dividend-price ratio; that is, the ratio of P_t to C_t. In the present model with i.i.d. shocks, V_t equals a constant, V, so that the growth rate of P_t equals the growth rate of C_t. The reciprocal of V equals the dividend-price ratio and can be determined from equations

(3.5) and (3.6) to be

$$\frac{1}{V} = \rho - (1-\theta)\, g^* + p\left[\left(\frac{1-\theta}{\gamma-1}\right) E\,(1-b)^{1-\gamma} - (1-\theta)\, Eb - \left(\frac{1-\theta}{\gamma-1}\right)\right] + \frac{1}{2}\gamma\,(1-\theta)\,\sigma^2. \quad (3.8)$$

The constant expected rate of return on equity, r^e, is the sum of the dividend yield, $1/V$, and the expected rate of capital gain on equity, which equals $g*$, the expected growth rate of the dividend (consumption). Therefore, r^e is the same as equation (3.8) except for the elimination of the term $g*$.[4] The constant equity premium is given from equations (3.7) and (3.8) by:

$$r^e - r^f = \gamma\sigma^2 + p\left[E\,(1-b)^{-\gamma} - E\,(1-b)^{1-\gamma} - Eb\right] \quad (3.9)$$

The disaster or jump term in equation (3.9) is proportional to the disaster probability, p. The expression in brackets that multiplies p depends on the size distribution of disasters, b, and the coefficient of relative risk aversion, γ. These effects were calibrated in Barro (2006) and Barro and Ursua (2012) by using the long-term history of macroeconomic disasters for 40 countries to pin down p and the distribution of b. The results accord with an observed average unlevered equity premium of 0.04 0.05 per year if γ is around 3-4.

The diffusion term, $\gamma\sigma^2$, in equation (3.9) is analogous to the expression for the equity premium in Mehra and Prescott (1985) and is negligible compared to the observed average equity premium if γ and σ^2 take on empirically reasonable values. For many purposes—including the pricing of far-out-of-the-money stock options—this term can be ignored.

[4]The transversality condition, which ensures that the value of tree equity is positive and finite, is $r^e > g*$.

3.3 Pricing Stock Options

3.3.1 Setup for pricing options

We now discuss the pricing of stock options within our model, which fits into the class of jump-diffusion models. Options pricing within this general class goes back to Merton (1976) and Cox and Ross (1976). The use of prices of far-out-of-the-money put options to infer disaster probabilities was pioneered by Bates (1991). This idea has been applied recently by, among others, Bollerslev and Todorov (2011); Backus, Chernov, and Martin (2011); Seo and Wachter (2016); and Siriwardane (2015).

We derive a pricing solution for far-out-of-the-money put options under the assumption that disaster events (jumps) are the dominant force. Key underlying conditions for the validity of the solution are that the option be sufficiently far out of the money and that the maturity not be too long. Under these conditions, we derive a simple pricing formula that reflects the underlying Poisson nature of disaster events, combined with an assumed power-law distribution for the sizes of disasters. This formula generates testable hypotheses—which we subsequently test—on the relation of the put-options price to maturity and exercise price. The formula also allows for a time-fixed-effects procedure to back out a time series for disaster probability.

Consider a put option on equity in the Lucas tree. To begin, suppose that the option has a maturity of one period and can be exercised only at the end of the period (a European option). The exercise price or strike on the put option is

$$\text{exercise price} = \varepsilon P_t \tag{3.10}$$

where we assume $0 < \varepsilon \leq 1$. We refer to ε, the ratio of the exercise price to the stock price, as the relative exercise price (often described as "moneyness").

The payoff on the put option at the start of period $t+1$ is zero if $P_{t+1} \geq \varepsilon \cdot P_t$. If $P_{t+1} < \varepsilon \cdot P_t$, the payoff is $\varepsilon P_t - P_{t+1}$. If $\varepsilon < 1$, the put option is initially out of the money. We focus empirically on options that are sufficiently far out of the money (ε sufficiently below one) so that the diffusion term, u, in equation (3.1) has a negligible effect on the chance of getting into the money over one period. The value of the put option then hinges on the disaster term, v. Specifically, the value of the put option depends on the probability, p, of experiencing a disaster and the distribution of disaster sizes, b. Further, what will mostly matter is the likelihood of experiencing one disaster. As long as the period (the maturity of the option) is not too long, the chance of two or more disasters has a second-order pricing impact that can be ignored as a good approximation[5].

Let the price of the put option at the start of period t be $\Omega \cdot P_t$. We refer to Ω, the ratio of the options price to the stock price, as the relative options price. The gross rate of return, R^O_{t+1}, on the put option is given by

$$\begin{aligned} R^o_{t+1} &= 0 \text{ if } \frac{P_{t+1}}{P_t} \geq \varepsilon \\ R^o_{t+1} &= \frac{1}{\Omega}\left(\varepsilon - \frac{P_{t+1}}{P_t}\right) \text{ if } \frac{P_{t+1}}{P_t} < \varepsilon. \end{aligned} \tag{3.11}$$

If there is one disaster of size b, the put option is in the money at the start of period $t+1$ if

$$\frac{P_{t+1}}{P_t} = (1+g)(1-b) < \varepsilon.$$

In most cases, the length of the period (maturity of the put option) will be short enough so that, for reasonable growth rates, we can ignore the term g.

When expressed in terms of z, the gross rate of return on the put option is modified from equation (3.11) to:

[5] Similarly, if we allowed for possible bonanzas, we could neglect the chance of a disaster and a bonanza both occurring over the period.

$$R^o_{t+1} = \frac{1}{\Omega}\left(\varepsilon - \frac{1+g}{z}\right) \text{ if 1 disaster occurs and } z > \frac{1+g}{\varepsilon} \quad (3.12)$$
$$R^o_{t+1} = 0 \text{ otherwise.}$$

To determine Ω, we use the first-order condition from equation (3.5), with R_{t+1} given by R^O_{t+1} from equation (3.12). The results depend on the form of the distribution for z, to which we now turn.

3.3.2 Power-law distribution of disaster sizes

Based on the findings for the distribution of observed macroeconomic disaster sizes in Barro and Jin (2011), we assume that the density function for z conforms to a power law:

$$f(z) = Az^{-(1+\alpha)} \quad (3.13)$$

where $A > 0, \alpha > 0,\ z \geq z_0 > 1$.

The general notion of this type of power law was applied by Pareto (1897) to the distribution of high incomes. The power-law distribution has since been applied widely in physics, economics, computer science, and other fields. For surveys, see Mitzenmacher (2003) and Gabaix (2009), who discusses underlying growth forces that can generate power laws. Examples of applications include sizes of cities (Gabaix and Ioannides [2004]), stock-market activity (Gabaix, et al. [2003, 2006]), CEO compensation (Gabaix and Landier [2008]), and firm size (Luttmer [2007]). The power-law distribution has been given many names, including heavy-tail distribution, Pareto distribution, Zipfian distribution, and fractal distribution.

The parameter $z_0 > 1$ in equation (3.13) is the threshold beyond which the power-law density applies. For example, in Barro and Ursua (2012), the floor disaster size of $b_0 = 0.095$

corresponds to $z_0 = 1.105$. We treat z_0 as a constant. The condition that $f(z)$ integrate to one from z_0 to infinity implies $A = \alpha z_0^{\alpha}$. Therefore, the power-law density function in equation (3.13) becomes

$$f(z) = \alpha z_0^{\alpha} z^{-(1+\alpha)}. \tag{3.14}$$

The key parameter in the power-law distribution is the Pareto tail exponent, α, which governs the thickness of the right tail. A smaller α implies a thicker tail.

The probability of drawing a transformed disaster size above z is given by

$$1 - F(z) = \left(\frac{z}{z_0}\right)^{-\alpha}. \tag{3.15}$$

Thus, the probability of seeing an extremely large transformed disaster size, z (expressed as a ratio to the threshold, z_0), declines with z in accordance with the tail exponent $\alpha > 0$.

One issue about the power-law density is that some moments related to the transformed disaster size, z, might be unbounded. For example, in equation (3.7), the risk-free rate depends inversely on the term $E(1-b)^{-\gamma}$. Heuristically (or exactly with time-separable power utility), we can think of this term as representing the expected marginal utility of consumption in a disaster state relative to that in a normal state. When $z \equiv 1/(1-b)$ is distributed according to $f(z)$ from equation (3.14), we can compute

$$\begin{aligned} E(1-b)^{-\gamma} &= E(z^{\gamma}) \\ &= \left(\frac{\alpha}{\alpha-\gamma}\right) z_0^{\gamma} \text{ if } \alpha > \gamma \end{aligned} \tag{3.16}$$

The term on the right side of equation (3.16) is larger when γ is larger (more risk aversion) or α is smaller (fatter tail for disasters). But, if $\alpha \leq \gamma$, the tail is fat enough, relative to the degree of risk aversion, so that the term blows up. In this case, r^f equals minus infinity in equation (3.7), and the equity premium is infinity in equation (3.9). Of course, in the data, the risk-free rate is not minus infinity and the equity premium is not infinity. Therefore, the empirical application of the power-law density in Barro and Jin (2011) restricted γ to a range that avoided unbounded outcomes, given the value of α that was estimated from the observed distribution of disaster sizes. That is, the unknown γ had to satisfy $\gamma<\alpha$ in order for the model to have any chance to accord with observed average rates of return[6]. This condition, which we assume holds, enters into our analysis of far-out-of-the-money put-options prices.

Barro and Jin (2011, Table 1) estimated the power-law tail parameter, α, in single power-law specifications (and also considered double power laws). The estimation was based on macroeconomic disaster events of size 10% or more computed from the long history for many countries of per capita personal consumer expenditure (the available proxy for consumption, C) and per capita GDP, Y. The estimated values of α in the single power laws were 6.3, with a 95% confidence interval of (5.0, 8.1), for C and 6.9, with a 95% confidence interval of (5.6, 8.5), for Y.[7] Thus, the observed macroeconomic disaster sizes suggest a range for α of roughly 5-8.

3.3.3 Options-pricing formula

To get the formula for Ω, the relative options price, we use the first-order condition from equations (3.5) and (3.6), with the gross rate of return, R_{t+1}, corresponding to the return R^O_{t+1} on put options in equation (3.12). We can rewrite this first-order condition as

$$1 + \hat{\rho} = (1+g)^{-\gamma} \, E_t \left(z^{\gamma} R^o_{t+1} \right) \tag{3.17}$$

[6]With constant absolute risk aversion and a power-law distribution of disaster sizes, the relevant term has to blow up. The natural complement to constant absolute risk aversion is an exponential distribution of disaster sizes. In this case, the relevant term is bounded if the parameter in the exponential distribution is larger than the coefficient of absolute risk aversion. With an exponential size distribution and constant relative risk aversion, the relevant term is always finite.

[7]Barro and Jin (2011, Table 1) found that the data could be fit better with a double power law. In these specifications, with a threshold of z0=1.105, the tail parameter, α, was smaller in the part of the distribution with the largest disasters than in the part with the smaller disasters. The cutoff value for the two parts was at a value of z around 1.4.

where $z \equiv 1/(1-b)$ is the transformed disaster size and $1+\hat{\rho}$ is an overall discount term, given from equations (3.5) and (3.6) (when the diffusion term is negligible) by

$$1+\hat{\rho} = 1+\rho-(\gamma-\theta)\,g+p\left(\frac{\gamma-\theta}{\gamma-1}\right)\left[E\,(1-b)^{1-\gamma}-1\right]. \tag{3.18}$$

We can evaluate the right-hand side of equation (3.17) using the density $f(z)$ from equation (3.14) along with the expression for R^{O}_{t+1} from equation (3.12). The result involves integration over the interval $z \geq (1+g)/\varepsilon$ where, conditional on having one disaster, the disaster size is large enough to get the put option into the money. The formula depends also on the probability, p, of having a disaster. Specifically, we have:

$$(1+\hat{\rho})\,(1+g)^{\gamma} = \frac{p}{\Omega}\int_{\frac{1+g}{\varepsilon}}^{\infty}\left\{z^{\gamma}\left[\varepsilon-\frac{1+g}{z}\right]\alpha z_0^{\alpha} z^{-(1+\alpha)}\right\}dz. \tag{3.19}$$

Evaluating the integral (assuming $\gamma<\alpha$ and $\varepsilon<[1+g]/z_0$) leads to a closed-form formula for the relative options price:

$$\Omega = \frac{\alpha z_0^{\alpha}}{(1+\hat{\rho}+\alpha g)}\frac{p\varepsilon^{1+\alpha-\gamma}}{(\alpha-\gamma)\,(1+\alpha-\gamma)}. \tag{3.20}$$

3.3.4 Maturity of the option

Equation (3.20) applies when the maturity of the put option is one "period." We now take account of the maturity of the option. In continuous time, the parameter p, measured per year, is the Poisson hazard rate for the occurrence of a disaster. Let T, in years, be the maturity of the (European) put option. The density, h, for the number of hits (disasters) over T is given by[8]

[8]See Hogg and Craig (1965, p. 88).

$$
\begin{aligned}
h(0) &= e^{-pT} \\
h(1) &= pTe^{-pT} \\
h(x) &= \frac{(pT)^x e^{-pT}}{x!},\ x = 0, 1, ...
\end{aligned}
\tag{3.21}
$$

If pT is much less than 1, the contribution to the options price from two or more disasters will be second-order, relative to that from one disaster. For given p, this condition requires a consideration of maturities, T, that are not "too long." In this range, we can proceed as in our previous analysis to consider just the probability and size of one disaster. Then, in equation (3.20), p will be replaced as a good approximation by pT.

The discount rate, $\hat{\rho}$, and growth rate, g, in equation (3.20) will be replaced (approximately) by $\hat{\rho}T$ and gT. For given $\hat{\rho}$ and g, if T is not "too long," we can neglect these discounting and growth terms. The impacts of these terms are of the same order as the effect from two or more disasters, which we have already neglected.

When T is short enough to neglect multiple disasters and the discounting and growth terms, the formula for the relative options price changes from equation (3.20) to:

$$
\Omega = \frac{\alpha z_0^{\alpha} pT \varepsilon^{1+\alpha-\gamma}}{(\alpha-\gamma)(1+\alpha-\gamma)}. \tag{3.22}
$$

Here are some properties of the options-pricing formula:

- The formula for Ω, the ratio of the options price to the stock price, is well-defined if $\alpha > \gamma$, the condition noted before that ensures the finiteness of various rates of return.
- The exponent on maturity, T, equals 1.
- The exponent on the relative exercise price, ε, equals $1 + \alpha - \gamma$, which is constant and

greater than 1 because $\alpha > \gamma$. We noted before that α ranged empirically between 5 and 8. The corresponding range for γ (needed to replicate an average unlevered equity premium of 0.04-0.05 per year) is between 2.5 and 5.5, with lower γ associating with lower α. The implied range for $\alpha - \gamma$ (taking account of the association between γ and α) is between 2.5 and 4.5, implying a range for the exponent on ε between 3.5 and 5.5.

- For given T and ε, Ω depends on the disaster probability, p; the shape of the power-law density, as defined by the tail coefficient, α, and the threshold, z_0; and the coefficient of relative risk aversion, γ. The expression for Ω is proportional to p.
- For given p and γ, Ω rises with a once-and-for-all shift toward larger disaster sizes; that is, with a reduction in the tail coefficient, α, or an increase in the threshold, z_0.
- For given p, α, and z_0, Ω rises for sure with a once-and-for-all shift in γ if $\varepsilon \leq 1$, which is the range that we are considering for put options. Note that, in contrast, the Black-Scholes options-pricing formula implies that Ω is independent of γ.[9]

We can look at the results in terms of the "risk-neutral probability," p^n, defined as the value of p that would generate a specified relative options price, Ω, when $\gamma = 0$. The formula for the ratio of the risk-neutral to the objective probability, p^n/p, implied by equation (3.22) is:

$$\frac{p^n}{p} = \frac{\alpha(1+\alpha)}{(\alpha-\gamma)(1+\alpha-\gamma)}\varepsilon^{-\gamma}. \tag{3.23}$$

Note that p^n/p depends on the relative exercise price, ε, but not on the maturity, T. If we assume parameter values consistent with the previous discussion—for example, α=7 and γ=3.5—the implied pn/p is 5.1 when ε=0.9, 7.8 when ε=0.8, 12.4 when ε=0.7, 21.3 when ε=0.6, and 40.3 when ε=0.5. Hence, the relative risk-neutral probability associated with far-out-of-the-money put options is sharply above one.

To view it another way, the relative options price, Ω, may seem far too high at low ε, when

[9] See, for example, Hull (2000, pp. 248 ff.). However, this standard result depends on holding fixed the risk-free rate, r_f. Equation (7) shows that r_f depends negatively on γ.

assessed in terms of the (risk-neutral) probability needed to justify this price. Thus, people who are paying these prices to insure against the risk of an enormous disaster may appear to be irrational. In contrast, the people writing these far-out-of-the-money puts may seem to be getting free money by insuring against something that is virtually impossible. Yet the pricing is reasonable if people have roughly constant relative risk aversion with γ around 3.5 (assuming a tail parameter, α, for disaster size around 7). The writers of these options will have a comfortable income almost all the time, but will suffer tremendously during the largest rare disasters (when the marginal utility of consumption is extremely high).

3.3.5 Diffusion term

Recall that the derivation of the formula for Ω, the relative options price, in equation (3.22) neglected the diffusion term, u, in the process for GDP (and consumption and the stock price) in equation (3.1). This omission is satisfactory if the put option is sufficiently far out of the money so that, given a reasonable variance σ^2 of the diffusion term, the chance of getting into the money over the maturity T is negligible. In other words, the tail for the normal process is not fat enough to account by itself for, say, 10% or greater declines in stock prices over periods up to, say, a few months. Operationally, our main empirical analysis applies to options that are at least 10% out of the money ($\varepsilon \leq 0.9$) and to maturities, T, that range up to 6 months.

If we consider put options at or close to the money, the diffusion term would have a first-order impact on the value of the option. If we neglect the disaster (jump) term—which will be satisfactory here—we would be in the standard Black-Scholes world. In this setting (with i.i.d. shocks), a key property of the normal distribution is that the variance of the stock price over interval T is proportional to T, so that the standard deviation is proportional to the square root of T. This property led to the result in Brenner and Subrahmanyam (1988) that the value of an at-the-money put option would be roughly proportional to the square root of the maturity. We, therefore, have two results concerning the impact of maturity, T, on the relative options price, Ω. For put options far out of the money (operationally for $\varepsilon \leq 0.9$), the exponent on T is close to 1. For put options close to the money (operationally for $\varepsilon = 1$), the exponent on T is close to one-half. These predictions turn out to hold empirically for put options on the S&P 500 and on analogous market indices for eight other stock markets.

3.3.6 Stochastic Volatility

The asset-pricing formula in equation (3.22) was derived under the assumption that the disaster probability, p, and the size distribution of disasters (determined by α and z_0) were fixed. We focus here on shifting p, but the results are isomorphic to shifting disaster intensity (reflecting changes in α and z_0). We can rewrite equation (3.22) as

$$\Omega = AT\varepsilon^{1+\alpha-\gamma}, \tag{3.24}$$

where $A > 0$ is constant. We can estimate equation (3.24) with data on Ω for far-out-of-the-money put options on, say, the S&P 500. Given ranges of maturities, T, and relative exercise prices, ε, we can estimate elasticities of Ω with respect to T and ε. We can also test the hypothesis that the coefficient, A, is constant. For example, using monthly data, we estimated a time fixed effect for each month and tested the hypothesis that these fixed effects were all equal. The result, detailed in a later section, is a very strong rejection of the hypothesis that A is constant. Instead, the estimated time fixed effects fluctuate dramatically, including occasional sharp upward movements followed by gradual reversion over a few months toward a baseline value. From the perspective of the model, we interpret these shifts as reflecting variations in the disaster probability, p.

If $\gamma > 1$, as we assume, equation (8) implies that a once-and-for-all rise in disaster probability, p, lowers the price-dividend ratio, V, if $\theta < 1$, so that the intertemporal elasticity of substitution, $1/\theta$, exceeds 1. Bansal and Yaron (2004) focus on IES>1 because it corresponds to the "normal case" where an increase in the expected growth rate, $g*$, raises V. Barro (2009) argues that IES>1 is reasonable empirically and, therefore, also focuses on this case.

Generally, the effects on options pricing depend on θ and other parameters and also on the stochastic process that generates variations in p, including the persistence of these changes. However, for purposes of pricing stock options, we need only consider the volatility of the overall term, A, which appears on the right side of equation (3.24). Our first-round look at the data—that is, the estimated time fixed effects—suggests that this term looks like a disaster process. On rare occasions, this term shifts sharply and temporarily upward and leads, thereby, to a jump in the corresponding term in equation (3.24). We think of this shock as generated by another Poisson probability, q, with a size distribution (for changes in stock prices) involving another power-law distribution, in this case with tail parameter $\alpha^* > \gamma$. If this process (for changing p) is independent of the disaster realizations (which depend on the level of p), then equation (3.22) is modified to

$$\Omega = \frac{\alpha z_0^{\alpha} p_t T \varepsilon^{1+\alpha-\gamma}}{(\alpha-\gamma)(1+\alpha-\gamma)} + \frac{\alpha^* \left(z_0^*\right)^{\alpha^*} q T \varepsilon^{1+\alpha^*-\gamma}}{(\alpha^*-\gamma)(1+\alpha^*-\gamma)}. \tag{3.25}$$

The first term on the right side of equation (3.25) reflects put-option value associated with the potential for realized disasters, and the second term gauges value associated with changing pt and the effects of these changes on stock prices.[10]

From the perspective of equation (3.24), we have the revised specification:

$$\Omega = AT\varepsilon^{1+\alpha-\gamma}\left[p_t + Bq\varepsilon^{\alpha^*-\alpha}\right], \tag{3.26}$$

where $A > 0$ and $B > 0$ are constants[11]. The new term involving $B > 0$ is important for fitting the data on put-options prices. This term implies $\Omega > 0$ if $p_t = 0$ because

[10]The formulation would also encompass effects on stock prices from changing α or γ.

[11]These values are constant if α, α^*, z_0, $(z_0)^*$, γ, and q are all constant. $\alpha - \alpha^*$ is identified because we have sample variation in relative exercise prices, ε.

of the possibility that pt will rise a lot during the life of the option. The preclusion of changing p_t (corresponding to $B = 0$) leads, as emphasized by Seo and Wachter (2016), to overestimation of the average level of p_t in the sample. Moreover, this overstatement of the average p_t associates with an underestimation of disaster sizes; that is, α looks two high. Hence, overall, disasters are gauged to be too frequent and too small.[12] Finally, our hypotheses about elasticities of Ω with respect to T and ε in equation (3.26) accord better with the data when $B > 0$ is admitted.

3.4 Empirical Analysis

The model summarized by equation (3.26) delivers some testable predictions. First, the elasticity of the price of a far out-of-the-money put option with respect to maturity, T—denoted β_T—is close to one. Second, the elasticity of the price of a far-out-of-the-money put option with respect to the relative exercise price, ε—denoted β_ε—is greater than one and is a weighted average of $1 + \alpha - \gamma$ and $1 + \alpha^* - \gamma$. Given a value of γ and the estimated value of $\alpha^* - \alpha$ from equation (3.26), the results can be used to back out estimates of the tail parameters α and α^*. Finally, time fixed effects provide estimates of each period's disaster probability, p_t. We assess these theoretical results empirically by analyzing prices of far-out-of-the-money put options on the U.S. S&P 500 and analogous broad indices for other countries.

3.4.1 Data and methodology

Our primary data source is a broker-dealer with a sizable market-making operation in global equities. We utilize over-the-counter (OTC) options prices for nine equity-market indices for

[12] As Seo and Wachter (2016) argue, these problems appear, for example, in Backus, Chernov, and Martin (2011).

developed and emerging markets—S&P 500 (U.S.), FTSE (U.K.), DAX (Germany), Euro Stoxx 50 (Euro zone), Nikkei (Japan), OMX (Sweden), SMI (Switzerland), Nifty (India), and Bovespa (Brazil). We check the results with OTC data against those with market-based information from OptionMetrics for the United States and from Bloomberg for the United States and other countries. This check is useful because the OTC data do not necessarily correspond to actual trades.

Our primary data derive from implied-volatility surfaces generated by the broker-dealer for the purpose of analysis, pricing, and marking-to-market.[13] These surfaces are constructed from transactions prices of options and OTC derivative contracts.[14] The dealer interpolates these observed values to obtain implied volatilities for strikes ranging from 50% to 150% of spot and for a range of maturities from 15 days to 2 years and more. Even at very low strikes, for which the associated options seldom trade, the estimated implied volatilities need to be accurate for the correct pricing of OTC derivatives such as variance swaps and structured retail products. Institutional-specific factors are unlikely to influence pricing in a significant way because other market participants can profitably pick off pricing discrepancies among dealers. Therefore, sell-side dealers have strong incentives to maintain the accuracy of their implied-volatility surfaces.

As mentioned, the OTC data source is superior to market-based alternatives in the breadth of coverage for exercise prices and maturities. Notably, the market data tend to be unreliable or entirely unavailable for options that are far out of the money and for long maturities. For example, OptionMetrics has very limited information on far out-of-the-money put options prices due to the lack of market transactions and methodological challenges. Specifically,

[13] A common practice in OTC trading is for executable quotes to be given in terms of implied volatility instead of the actual price of an option. Once the implied volatility is agreed on, the options price is determined from the Black-Scholes formula based on the readily observable price of the underlying security. Since the Black-Scholes formula provides a one-to-one link between price and volatility, quotes can be given equivalently in terms of implied volatility or price.

[14] Dealers observe prices through own trades and from indications by inter-dealer brokers. It is also a common practice for dealers to ask clients how their prices compare to other market makers in OTC transactions.

their volatility surface is mainly limited to 20-delta options volatilities at the extreme, which correspond to options that are close to the money,[15] whereas the OTC data contain implied volatilities for 5 delta and even 1-delta options.

The broad range of strikes in the broker-dealer data is important for our analysis because it is the prices of far-out-of-the-money put options that will mainly reflect disaster risk. In practice, we focus on put options with exercise prices of 50%, 60%, 70%, 80%, and 90% of spot; that is, we exclude options within 10% of spot.

For maturities, we focus on a range between 30 and 180 days; specifically, for 30 days, 60 days, 90 days, and 180 days.[16] Our main analysis excludes options with maturities greater than six months because the prices in this range may be influenced significantly by the possibility of multiple disaster realizations and also by discounting. However, in practice, the results for 1-year maturity accord reasonably well with those for shorter maturities.

Using the data on implied volatilities, we re-construct options prices from the standard Black-Scholes formula, assuming a zero discount rate and no dividend payouts. We should emphasize that the use of the Black-Scholes *formula* to translate implied volatilities into options prices does not bind us to the Black-Scholes *model* of options prices. The formula is used only to convert the available data expressed as implied volatilities into options prices. Our calculated options prices are comparable to directly quoted prices (subject to approximations related to discounting and dividend payouts).

[15] A 20-delta option has a price that changes by 0.20% for a 1% change in the underlying security price. OptionMetrics Volatility Surface uses interpolation to generate the implied volatility for each security on each day, based on a kernel-smoothing algorithm. The lower bound of this volatility surface is 20-delta. In our use of the OptionMetrics data, we expand on the range of option strikes by applying linear interpolation whenever there are two or more observations for a single trading date. This procedure enlarges the volatility surface.

[16] We omit 15-day options because we think measurement error is particularly serious in this region in pinning down the precise maturity. Even the VIX index, which measures short-dated implied volatility, does not track options with maturity less than 23 days.

3.4.2 Basic model fit

We estimate the model based on equation (3.26) with non-linear least-squares regression. In this form, we think of the error term as additive with a constant variance. Log-linearization with a constant-variance error term (that is, a shock proportional to price) is problematic because it understates the typical error in extremely far-out-of-the-money put prices, which are close to zero. That is, this specification gives undue weight to puts with extremely low exercise prices.

In the non-linear regression, we allow for time fixed effects to capture the unobserved time-varying probability of disaster, p_t, in equation (3.26). We allow the estimated p_t to differ across countries; that is, we estimate country-time fixed effects. Note that, for a given country and date, these effects are the same for each observed maturity, T, and relative exercise price,ε.

We sample the data at monthly frequency, selecting only month-end dates, to allow for ease of computation with a non-linear solver. The selection of mid-month dates yields similar results. The sample period for the United States in our main analysis is August 1994-February 2017. Because of lesser data availability, the samples for the other stock-market indices are shorter. Subsequently, we expand the U.S. sample back to 1983, particularly to assess pricing behavior before and after the global stock-market crash of 1987. However, we do not use this longer sample in our main analysis because the data quality before 1994 is substantially poorer.

Table 3.1 shows the model estimation using non-linear least-squares regression. The estimated elasticities with respect to maturity, β_T, are close to one. For example, the estimated coefficient for the United States is 0.986 (s.e.=0.036) and that for all nine indices jointly is 0.944 (0.038). These results indicate that far-out-of-the-money prices of put options on broad market indices are roughly proportional to maturity, in accordance with our rare-disasters model. This nearly proportional relationship between options price and maturity for far-out-of-the-money put options is a newly documented fact that cannot be explained under the Black-Scholes model. To our knowledge, other theoretical models of options prices also do not predict this behavior.

The results with respect to maturity can be visualized in Figure 3.1, Panel A, which plots ratios of put prices to spot prices against maturity, assuming an exercise price of 80% of spot. The blue curve corresponds to the historical data that underlie Table 3.1. The red curve shows values generated by the Black-Scholes model, assuming a log-normal distribution of shocks and a constant volatility of 30% (chosen to accord with the average observed level of put prices). Most importantly, the Black-Scholes model predicts that these far-out-of-the-money put prices will have a convex relationship with maturity. This pattern deviates from the nearly linear relationship shown by the historical data.

In contrast, as discussed in Brenner and Subrahmanyam (1988), prices of at-the-money put options in the Black-Scholes model are roughly proportional to the square root of the maturity. This result arises because, with a diffusion process driven by i.i.d. normal shocks, the variance of the log of the stock price is proportional to time and, therefore, the standard deviation is proportional to the square root of time. This pattern implies the concave relation between put price and maturity as shown by the red curve in Figure 3.1, Panel B. In this case, the Black-Scholes prediction accords with the historical data, shown by the blue curve in Panel B.

Table 3.2 provides detailed regression estimates for the nine indices for at-the-money put prices. The estimated coefficient on maturity is 0.521 (s.e.=0.007) for the United States and 0.497 (0.006) for the nine indices jointly. Hence, as predicted by Black-Scholes, these coefficients are close to 0.5. For exercise prices between 80% and 100% of spot, the Black-Scholes prediction for the relation between put price and maturity shifts from convex to concave at around 90% of spot (with the exact shift point depending on the underlying

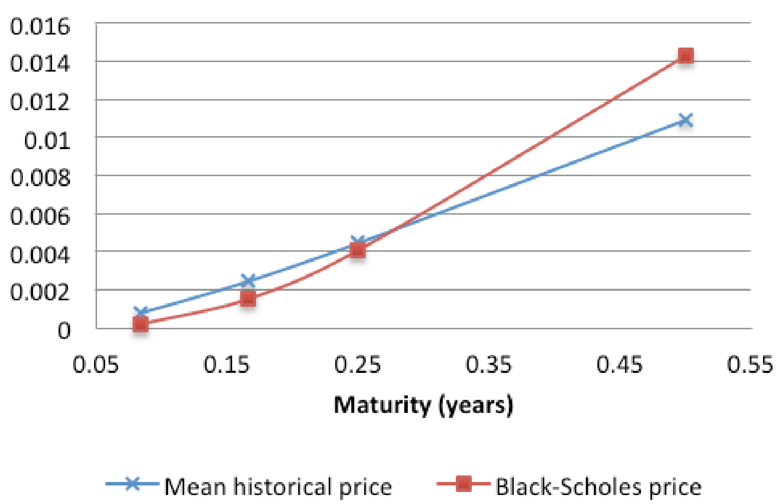
Table 3.1: Baseline Regression Estimates

This table presents non-linear least-squares regression estimates of our main model with variable disaster probability. We use OTC put-option prices for nine countries or areas with maturity, T, ranging from 30 to 180 days and relative exercise price, ε, ranging from 0.5 to 0.9. The estimation for relative options prices, Ω, corresponds to equation (3.26): $\Omega = AT\varepsilon^{1+\alpha-\gamma} \cdot [p_t + B\varepsilon^{\alpha*-\alpha}]$, where p_t is a country's disaster probability, α is the tail parameter for disaster sizes, α^* is the tail parameter for stock-price changes induced by changes in p_t, and γ is the coefficient of relative risk aversion. We apply country-time monthly fixed effects to capture each country's variations in pt, as shown in Figure 3.2. The estimated exponent on T, denoted β_T, should equal 1. The estimated exponent on the first ε term, denoted β_ε, should equal $1+\alpha-\gamma$. Time-clustered standard errors are in parentheses.

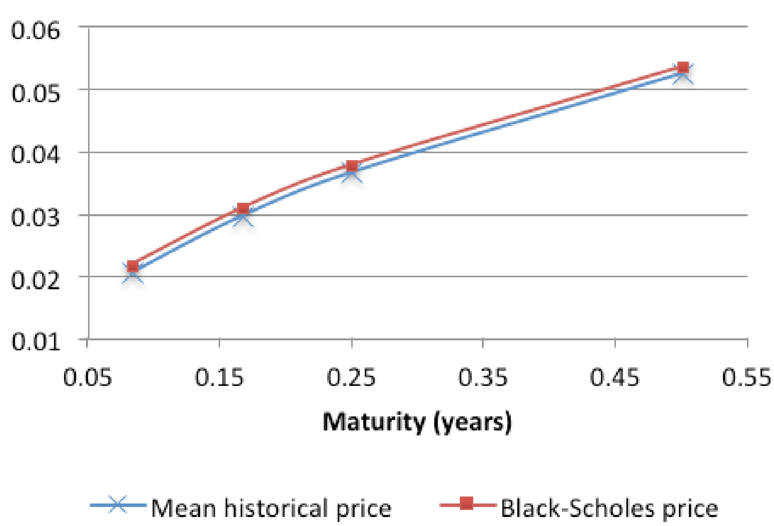
Index	spx	ftse	estx50	dax	nky	omx	swx	bovespa	nifty	all
	US	UK	EURO	GER	JAP	SWE	SWZ	BRAZ	INDIA	
	Aug94 - Feb17	Jan98 - Feb17	Jun98 - Feb17	Jan00 - Feb17	Sep97 - Feb17	Jan98 - Feb17	Jan98 - Feb17	Jan08 - Feb17	Aug06 - Feb17	Aug94 - Feb17
β_T	0.986	0.993	0.939	0.933	0.877	0.917	1.000	0.851	0.914	0.944
	(0.036)	(0.031)	(0.028)	(0.033)	(0.066)	(0.028)	(0.032)	(0.064)	(0.057)	(0.038)
β_ϵ	5.31	4.94	4.45	4.70	3.99	4.46	4.97	8.60	4.45	4.67
	(0.12)	(0.09)	(0.14)	(0.11)	(0.40)	(0.21)	(0.28)	(0.50)	(0.31)	(0.18)
α*-α	9.69	7.21	4.97	6.05	6.05	4.26	6.84	-4.24	4.72	5.80
	(1.49)	(0.73)	(0.21)	(0.44)	(0.22)	(0.32)	(0.88)	(0.60)	(0.19)	(0.20)
B	0.59	0.64	1.17	0.84	1.74	1.50	0.84	1.29	1.00	0.96
	(0.06)	(0.11)	(0.16)	(0.15)	(0.27)	(0.41)	(0.22)	(0.43)	(0.20)	(0.10)

Figure 3.1: Comparison of Prices with Black-Scholes Predicted Prices
These figures compare the mean of observed put prices across maturities with predicted prices from the Black-Scholes model. Black-Scholes prices are generated assuming flat volatility across maturities (30% for out-of-the-money options and 19% for at-the-money options). For ease of comparison, the volatilities are chosen so that the prices scale appropriately to historical prices. Panel A graphs relative put prices on the S&P 500 with strike of 80% of spot. Panel B graphs relative put prices with strike of 100% of spot. We introduced spacing between the two curves in Panel B solely for visual comparison.

Panel A: Out-of-the-Money Put Prices



Panel B: At-the-Money Put Prices



volatility). The predicted relation turns out to be nearly linear for an exercise price around 90% of spot. In contrast, as implied by Table 3.1 and Figure 3.1, Panel A, the relation in the data is roughly linear in maturity for a broad range of exercise prices below 90%—down to at least 50%. These results accord with the rare-disasters model but not with Black-Scholes.

Table 3.2: Estimation with At-The-Money Put Options

This table presents non-linear least-squares regression estimates for at-the-money put options only for the nine stock-market indices used in Table 3.1. We use put options prices with maturity ranging from 30 to 180 days and strike equal to spot price. The estimation includes a country-time multiplicative fixed effect. Time-clustered standard errors are in parentheses.

Index	spx	ftse	estx50	dax	nky	omx	swx	bovespa	nifty	all
	US	UK	EURO	GER	JAP	SWE	SWZ	BRAZ	INDIA	
	Aug94 – Feb17	Jan98 - Feb17	Jun98 - Feb17	Jan98 - Feb17	Sep97 - Feb17	Jan98 - Feb17	Jan98 - Feb17	Jan08 - Feb17	Aug06 - Feb17	Aug94- Feb17
β_T	0.521	0.513	0.488	0.492	0.470	0.478	0.511	0.475	0.504	0.497
	(0.0071)	(0.0070)	(0.0069)	(0.0082)	(0.0095)	(0.0068)	(0.0072)	(0.0163)	(0.0132)	(0.0063)

To summarize, the fit of the Black-Scholes model is good for at-the-money put options but poor for put options with exercise prices at 90% or less of spot. These patterns arise because the diffusion component of shocks dominates pricing of at-the-money put options, whereas disaster (jump) risk, not captured in Black-Scholes, dominates the pricing of far-out-of-the-money put options. As discussed earlier in the modeling section, the roughly proportional relationship between far-out-of-the-money put prices and maturity arises because, in a Poisson context, the probability of a disaster is proportional to maturity. The resulting formula is only approximate because it neglects the potential for multiple disasters within the time frame of an option's maturity, omits a diffusion term, and also ignores discounting. However, for options that are not "too long," these approximations will be reasonably accurate.

Table 3.1 also shows estimates of the elasticity with respect to the relative exercise price, β_ε. This coefficient corresponds in the model to $1 + \alpha - \gamma$, where α is the tail coefficient for disaster sizes and γ is the coefficient of relative risk aversion. The estimates are all positive and greater than one, as predicted by the model. The estimated coefficients are similar across indices, except for Brazil. For the other indices, the estimated values fall in a range from 3.99 (s.e.=0.40) for Japan to 5.31 (0.12) for the United States. The joint estimate across the nine countries is 4.67 (0.18).

Rare-disasters research with macroeconomic data, such as Barro and Ursua (2008) and Barro and Jin (2011), suggested that a γ of 3-4 would accord with observed average (unlevered) equity premia. With this range for γ, the estimated values of $\beta_\varepsilon = 1 - \alpha - \gamma$ from Table 3.1 (aside from Brazil) imply tail coefficients, α, between 6 and 8. This finding compares with a direct estimate for α based on macroeconomic data on consumption in Barro and Jin (2011, Table 1) of 6.3 (s.e.=0.8). Hence, the estimates of α coefficients implied by Table 3.1 accord roughly with those found from observation of the size distribution of macroeconomic disasters (based on GDP or consumption).

3.4.3 Estimated disaster probabilities

We can use the estimated monthly fixed effects for each country from the regressions in Table 3.1, along with equations (3.25) and (3.26), to construct time series of (objective) disaster probabilities, p_{jt}, where j now denotes the country. The assumption here is that the other parameters that multiply p_{jt} in equation (3.25) are constant over time for country j. In that case, the estimated p_{jt} will be proportional to the time fixed effect for country j.

To get a ballpark idea of the level of p_{jt}, we assume that, in each country, the threshold for disaster sizes is fixed at z_0 =1.1 (as in Barro and Jin [2011]) and that the coefficient of relative risk aversion is γ=3. We allow the tail coefficient, α_j, to differ across countries; that is, we allow countries to differ with respect to the size distribution of potential disasters. We use the estimated coefficients from Table 3.1 for β_ε (which equals $1 + \alpha - \gamma$ in the model) to back out the implied tail coefficient, α_j, for country j. (Aside from Brazil, these values range from 6.0 to 7.3.) Figure 3.2, Panel A, presents the resulting time series of disaster probabilities for each of the nine stock-market indices. For clarity, Panel B presents the results just for the United States. Table 3.3 provides summary statistics for the estimated disaster probabilities. Note that the levels of the series, but not the time patterns, depend on our assumed parameter values.

Table 3.3: Means, Standard Deviations, Quantiles of Estimated Disaster Probabilities

This table presents summary statistics on estimated disaster probabilities and the implied estimated α (tail) coefficients from the regressions in Table 3.1. The disaster probabilities are calculated as indicated in the notes to Figure 3.2.

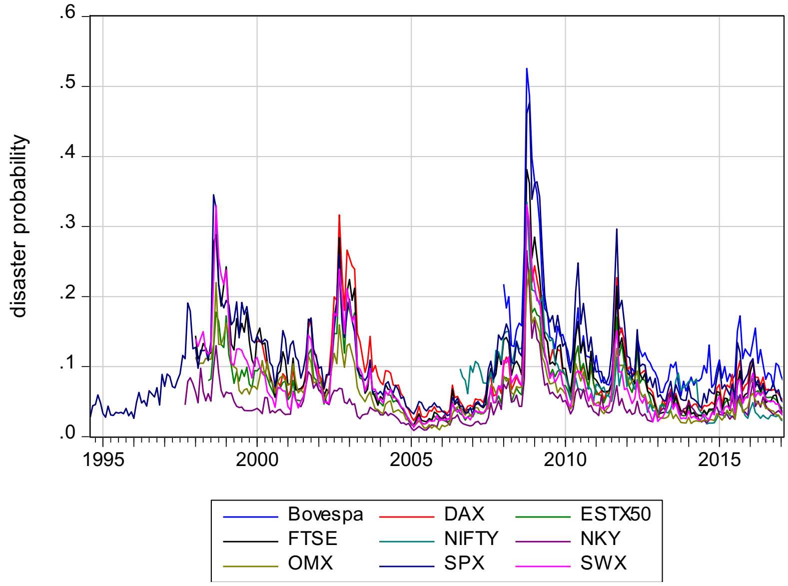
index	# obs	Start date	mean	sd	min	1%	10%	25%	50%	75%	99%	max	α
spx (US)	271	Aug-94	0.105	0.071	0.026	0.029	0.041	0.056	0.086	0.136	0.362	0.477	7.31
FTSE (UK)	230	Jan-98	0.093	0.063	0.017	0.021	0.032	0.046	0.076	0.121	0.287	0.381	6.94
ESTX50 (EURO)	225	Jun-98	0.073	0.043	0.015	0.018	0.030	0.044	0.063	0.087	0.220	0.256	6.45
DAX (GER)	206	Jan-00	0.089	0.055	0.023	0.029	0.038	0.054	0.074	0.102	0.295	0.329	6.70
nky (JAP)	234	Sep-97	0.046	0.030	0.009	0.010	0.020	0.031	0.039	0.056	0.158	0.266	5.99
omx (SWE)	230	Jan-98	0.064	0.040	0.010	0.012	0.023	0.033	0.058	0.084	0.200	0.238	6.46
swi (SWZ)	230	Jan-98	0.077	0.058	0.013	0.019	0.026	0.039	0.058	0.094	0.289	0.331	6.97
bovespa (BRAZ)	110	Jan-08	0.131	0.078	0.058	0.063	0.075	0.087	0.110	0.143	0.479	0.526	10.60
nifty (INDIA)	127	Aug-06	0.077	0.052	0.018	0.019	0.028	0.036	0.068	0.094	0.247	0.335	6.45

The disaster probabilities shown in Figure 3.2, Panel A, have high correlations among the countries, with an average pair-wide correlation of 0.88. For pairs that include the United States (SPX), the correlations are even higher: 0.98 with the United Kingdom (FTSE), 0.97 with the Euro area (ESTX50), 0.96 with Germany (DAX), 0.90 with Japan (NKY), 0.94 with Sweden (OMX), 0.96 with Switzerland (SWX), 0.93 with Brazil (Bovespa), and 0.86 with India (NIFTY). These high correlations indicate that the main part of the inferred

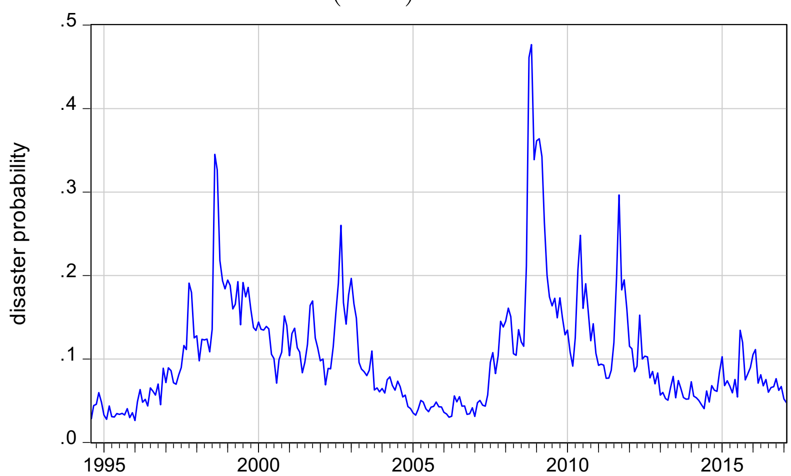
Figure 3.2: Estimated Disaster Probabilities

Panel A graphs the estimated disaster probabilities for nine stock-market indices associated with the regressions in Table 3.1. The annualized disaster probability, p_{jt} for index j, is calculated from the time fixed-effect coefficients in the form of equation (3.26), assuming in equation (3.25) that $z_0 = 1.1$, $\gamma = 3$, and $\beta_\varepsilon = 1 + \alpha_j - \gamma$. Panel B is for the United States only (SPX).

Panel A: 9 Stock-Market Indices



Panel B: United States (SPX)



disaster probability reflects the chance of a common (global) disaster.

The median estimated disaster probability is 8.6% per year for the S&P 500. For the other indices, the medians range from 3.9% for Japan to 7.6% for the U.K. and 11.0% for Brazil (see Table 3.3). These estimates can be compared with average disaster probabilities of 3-4% per year estimated from macroeconomic data on rare disasters—see, for example, Barro and Ursua [2008]). However, this earlier analysis assumed that disaster probabilities were constant across countries and over time.

The estimated disaster probabilities in Figure 3.2, Panel A, are volatile and right-skewed, with spikes during crisis periods. The U.S. disaster probability hit a peak of 48% per year in November 2008. Other countries had their highest disaster probabilities in the range of 21% to 53% in October and November 2008. These patterns mirror the options-derived U.S. equity premia in Martin (2015) and the U.S. disaster probabilities found by Siriwardane (2015). Another sharp peak occurred around the time of the Russian and LTCM crises in August-September 1998. In this case, the estimated U.S. disaster probability reached 30% in August 1998.

Figure 3.2, Panel A, suggests a lower bound on disaster probability around 1-2% per year (except for Brazil, which has a minimum of 6%). These lower bounds compare with the (constant) disaster probability of 3-4% per year found by Barro and Ursua (2008).

The estimated first-order AR(1) coefficient for the estimated U.S. disaster probability in Figure 3.2, Panel B, is 0.88 (s.e.=0.03), applying at a monthly frequency. This coefficient implies that rare-disaster shocks have a half-life of 6 months. The persistence of disaster probabilities for the other countries (Figure 3.2, Panel A) is similar to that for the United States, with the estimated AR(1) coefficients ranging from 0.85 to 0.90, except for Japan, which is at 0.81.

Although we attributed the time pattern shown in Figure 3.2 to variable disaster probability,

p_{jt}, the variations in the time fixed effects may also reflect changes in the other parameters contained in the term that multiplies p_{jt} in equation (25), $\alpha z_0^{\alpha}/[(\alpha-\gamma)(1+\alpha-\gamma)]$.[17] For example, outward shifts in the size distribution of disasters, generated by reductions in the tail parameter, α, or increases in the threshold disaster size, z_0, work like increases in p. Similarly, increases in the coefficient of relative risk aversion, γ, would raise the overall term. This kind of change in risk preference, possibly due to habit formation, has been stressed by Campbell and Cochrane (1999). Separation of changes in the parameters of the disaster distribution from those in risk aversion require simultaneous consideration of asset-pricing effects (reflected in Figure 3.2) with information on the actual incidence and size of disasters (based, for example, on movements of macroeconomic variables).

3.4.4 Model robustness

Tables 3.4-3.7 explore the empirical robustness of the baseline model from Table 3.1 under various scenarios. Table 3.4 shows the effects from eliminating the lowest relative exercise price, ε= 0.5. This change has minor effects on the estimates. In particular, the estimated coefficients β_T and β_ε are similar to those reported in Table 3.1. For example, for all indices jointly, β_T in Table 3.4 is 0.943 (0.038), compared with 0.944 (0.038) in Table 3.1. For βε, the value for all indices jointly is 4.62 (0.18) in Table 3.4, compared with 4.67 (0.18) in Table 3.1.

Table 3.5 does the estimation separately for two ranges of maturity, T—30 and 60 days (Panel A) versus 90 and 180 days (Panel B). In this case, the estimated β_T is higher when the maturity values are lower. For example, for all indices jointly, β_T is 1.07 (0.06) in the

[17] In the model with i.i.d. shocks, this term does not depend on the intertemporal elasticity of substitution for consumption, 1/ϑ, or the rate of time preference, ρ. Kelly and Jiang (2014, p. 2842) assume a power-law density for returns on individual securities. Their power law depends on a cross-sectional parameter and also on aggregate parameters that shift over time. In contrast to our analysis, they assume time variation in the economy-wide values of the tail parameter, analogous to our α, and the threshold, analogous to our z0. (Their threshold corresponds to the fifth percentile of observed monthly returns.)

Table 3.4: Robustness: Elimination of Lowest Exercise Price, ε

This table modifies Table 3.1 to consider the range of relative exercise prices, ε, from 0.6 to 0.9; that is, the lowest value, $\varepsilon = 0.5$, is omitted.

Index	spx	ftse	estx50	dax	nky	omx	swx	bovespa	nifty	all
	US	UK	EURO	GER	JAP	SWE	SWZ	BRAZ	INDIA	
	Aug94 - Feb17	Jan98 - Feb17	Jun98 - Feb17	Jan00 - Feb17	Sep97 - Feb17	Jan98 - Feb17	Jan98 - Feb17	Jan08 - Feb17	Aug06 - Feb17	Aug94 - Feb17
$\boldsymbol{\beta}_T$	0.986	0.992	0.938	0.932	0.876	0.916	0.999	0.85	0.913	0.943
	(0.036)	(0.031)	(0.028)	(0.033)	(0.067)	(0.029)	(0.033)	(0.065)	(0.057)	(0.038)
$\boldsymbol{\beta}_\epsilon$	5.09	4.8	4.49	4.5	4.21	4.09	4.91	4.44	4.54	4.62
	(0.129)	(0.098)	(0.150)	(0.102)	(0.399)	(0.319)	(0.280)	(0.292)	(0.301)	(0.178)
α^*-α	7.23	6.3	5.05	5.35	6.5	4.05	6.52	4.48	4.97	5.63
	(0.483)	(0.399)	(0.190)	(0.241)	(0.177)	(0.185)	(0.851)	(0.535)	(0.210)	(0.187)
$\boldsymbol{B}$	0.7	0.747	1.11	1.06	1.45	2.29	0.895	0.685	0.886	1.02
	(0.100)	(0.127)	(0.115)	(0.157)	(0.199)	(0.830)	(0.253)	(0.178)	(0.166)	(0.099)

lower range and 0.90 (0.03) in the upper range (compared with 0.94 [0.04] for the full range in Table 3.1). These results likely reflect the approximations in the model that relate to possibilities for multiple disasters, discounting, and the omission of a diffusion term. These approximations are best at short maturities, such as the range considered in Table 3.5, Panel A.

Table 3.6 carries out a related analysis in which the structure of included maturities goes out to one year. In this case, the estimated β_T is similar to that found in Table 3.5, Panel B, which applied to maturities of 90 and 180 days.

Table 3.7 explores the stability of the baseline results from Table 1 with regard to sample period. Results apply to the pre-financial-crisis period before 2008 (1994-2007 for the United States), the crisis period, 2008-2010, and the post-crisis period, 2011-2017. With respect to the estimated coefficients β_T and β_ε, the main finding is the somewhat lower values for the crisis interval of 2008-2010. For example, for the United States, the estimates of β_T are 1.03 (s.e.=0.04), 0.89 (0.05), and 1.15 (0.07), respectively, for the three periods. For all countries jointly, the corresponding estimates are 0.98 (0.03), 0.85 (0.06), and 1.10 (0.04). Possibly the low estimated β_T during the crisis period can be explained by a disaster probability,p_{jt}, that was high in the short term but projected to fall more quickly than usual in the near future. This interpretation accords with the low estimated values in the 2008-2010 period for the B coefficient (which picks up the potential for a rise in p_{jt} in the future).

The estimates in Table 3.7 of β_ε for the three samples for the United States are 5.76 (s.e.=0.16), 4.86 (0.10), and 5.06 (0.28), whereas those for all countries jointly are 5.21 (0.12), 4.33 (0.17), and 4.86 (0.17). Possibly the low estimated values of β_ε during the crisis interval reflect a perceived fatter tail than usual for bad outcomes (represented by a low tail exponent α and a correspondingly low value of β_ε, which equals $1 + \alpha - \gamma$).

A lot of analysis of options pricing, starting with Bates (1991), suggests that the nature of

Table 3.5: Robustness: Two Ranges of Option Maturity

This table modifies Table 3.1 to present regression estimates for two ranges of maturity, T. Panel A uses $T = 30$ and 60 days. Panel B uses $T = 90$ and 180 days.

Panel A: T=30 and 60 days

Index	spx	ftse	estx50	dax	nky	omx	swx	bovespa	nifty	all
	US	UK	EURO	GER	JAP	SWE	SWZ	BRAZ	INDIA	
	Aug94 - Feb17	Jan98 - Feb17	Jun98 - Feb17	Jan00 - Feb17	Sep97 - Feb17	Jan98 - Feb17	Jan98 - Feb17	Jan08 - Feb17	Aug06 - Feb17	Aug94 - Feb17
$\boldsymbol{\beta_T}$	1.12	1.15	1.08	1.05	0.945	1.06	1.18	0.895	0.982	1.07
	(0.054)	(0.043)	(0.037)	(0.043)	(0.137)	(0.037)	(0.051)	(0.093)	(0.080)	(0.061)
$\boldsymbol{\beta_\epsilon}$	7.34	7.09	5.91	6.18	4.95	5.25	6.34	6.02	5.65	6.05
	(0.24)	(0.16)	(0.19)	(0.14)	(0.60)	(0.40)	(0.48)	(0.45)	(0.67)	(0.35)
α*-α	34.4	10.4	5.93	6.99	8.38	6.12	6.57	6.98	5.61	7.28
	(126.0)	(2.7)	(0.42)	(0.72)	(0.82)	(0.28)	(0.75)	(2.70)	(0.46)	(0.54)
B	5.5	0.90	2.21	1.41	1.80	3.57	1.98	0.59	1.62	1.53
	(69.0)	(0.17)	(0.56)	(0.45)	(0.70)	(1.46)	(0.59)	(0.38)	(0.38)	(0.37)

Panel B: T=90 and 180 days

Index	spx	ftse	estx50	dax	nky	omx	swx	bovespa	nifty	all
	US	UK	EURO	GER	JAP	SWE	SWZ	BRAZ	INDIA	
	Aug94 - Feb17	Jan98 - Feb17	Jun98 - Feb17	Jan00 - Feb17	Sep97 - Feb17	Jan98 - Feb17	Jan98 - Feb17	Jan08 - Feb17	Aug06 - Feb17	Aug94 - Feb17
$\boldsymbol{\beta_T}$	0.940	0.941	0.889	0.888	0.837	0.873	0.944	0.825	0.879	0.899
	(0.034)	(0.031)	(0.027)	(0.032)	(0.054)	(0.028)	(0.031)	(0.058)	(0.055)	(0.034)
$\boldsymbol{\beta_\epsilon}$	5.24	4.87	4.34	4.62	3.87	4.21	4.94	4.2	4.35	4.59
	(0.135)	(0.093)	(0.145)	(0.112)	(0.419)	(0.250)	(0.290)	(0.232)	(0.327)	(0.179)
α*-α	9.39	6.83	4.31	5.50	5.49	3.43	6.60	3.32	4.12	5.22
	(1.64)	(0.88)	(0.25)	(0.53)	(0.24)	(0.28)	(1.14)	(0.73)	(0.23)	(0.24)
B	0.490	0.525	1.10	0.714	1.77	1.85	0.688	0.844	0.942	0.854
	(0.054)	(0.103)	(0.187)	(0.153)	(0.339)	(0.850)	(0.220)	(0.472)	(0.258)	(0.102)

Table 3.6: Robustness: Inclusion of Puts with 1-Year Maturity

This table corresponds to Table 3.1 except that the maturity range, T, is from 30 days to 1 year. That is, T=360 days is added.

Index	spx	ftse	estx50	dax	nky	omx	swx	bovespa	nifty	all
	US	UK	EURO	GER	JAP	SWE	SWZ	BRAZ	INDIA	
	Aug94 - Feb17	Jan98 - Feb17	Jun98 - Feb17	Jan98 - Feb17	Sep97 - Feb17	Jan98 - Feb17	Jan98 - Feb17	Jan08 - Feb17	Aug06 - Feb17	Aug94 - Feb17
$\boldsymbol{\beta_T}$	0.940	0.935	0.880	0.885	0.844	0.883	0.935	0.851	0.886	0.899
	(0.026)	(0.024)	(0.021)	(0.026)	(0.042)	(0.022)	(0.026)	(0.046)	(0.040)	(0.027)
$\boldsymbol{\beta_\epsilon}$	4.35	4.00	3.76	3.96	3.55	4.12	4.32	3.58	3.66	3.96
	(0.091)	(0.067)	(0.088)	(0.083)	(0.253)	(0.119)	(0.146)	(0.200)	(0.128)	(0.109)
α*-α	7.36	6.34	4.40	5.18	4.62	4.61	7.30	3.97	4.13	5.15
	(0.74)	(0.42)	(0.16)	(0.24)	(0.17)	(0.41)	(1.10)	(0.23)	(0.15)	(0.13)
$\boldsymbol{B}$	0.386	0.432	0.708	0.535	1.360	0.429	0.443	0.489	0.710	0.580
	(0.041)	(0.054)	(0.084)	(0.062)	(0.212)	(0.062)	(0.079)	(0.058)	(0.090)	(0.047)

Table 3.7: Robustness: Alternative Sample Periods

This table corresponds to Table 3.1 except for the sample periods.

		spx	ftse	estx50	dax	nky	omx	swx	bovespa	nifty	all
		US	UK	EURO	GER	JAP	SWE	SWZ	BRAZ	INDIA	
		Aug94 – Feb17	Jan98 - Feb17	Jun98 - Feb17	Jan00 - Feb17	Sep97 - Feb17	Jan98 - Feb17	Jan98 - Feb17	Jan08 - Feb17	Aug06 - Feb17	Aug94 - Feb17
Pre-2008	$\boldsymbol{\beta_T}$	1.03	0.994	0.928	0.894	1.00	0.916	1.020	--	1.03	0.978
		(0.037)	(0.033)	(0.040)	(0.047)	(0.029)	(0.032)	(0.035)		(0.043)	(0.030)
	$\boldsymbol{\beta_\epsilon}$	5.76	4.96	4.94	4.83	6.76	3.96	5.42	--	5.28	5.21
		(0.164)	(0.098)	(0.166)	(0.122)	(0.284)	(0.354)	(0.196)		(0.161)	(0.115)
	α*-α	9.32	7.16	4.56	5.95	8.86	3.95	9.64	--	3.21	5.77
		(1.767)	(0.713)	(0.353)	(0.693)	(1.754)	(0.220)	(3.410)		(0.159)	(0.270)
	B	0.745	0.688	0.943	0.851	0.561	3.79	0.538	--	3.1	0.958
		(0.092)	(0.120)	(0.229)	(0.249)	(0.102)	(1.359)	(0.137)		(1.001)	(0.144)
Jan. 08-Dec.10	$\boldsymbol{\beta_T}$	0.885	0.935	0.903	0.925	0.718	0.856	0.931	0.762	0.812	0.850
		(0.054)	(0.060)	(0.050)	(0.061)	(0.096)	(0.059)	(0.067)	(0.070)	(0.061)	(0.063)
	$\boldsymbol{\beta_\epsilon}$	4.86	4.92	4.00	4.48	3.60	4.61	3.93	3.67	4.07	4.33
		(0.100)	(0.172)	(0.147)	(0.150)	(0.306)	(0.206)	(0.201)	(0.741)	(0.316)	(0.168)
	α*-α	6.82	7.97	4.63	5.98	5.13	5.89	5.04	2.10	3.52	5.23
		(4.845)	(3.197)	(0.461)	(1.360)	(0.264)	(1.962)	(0.295)	(0.281)	(0.249)	(0.665)
	B	0.258	0.352	0.904	0.501	0.662	0.416	1.46	1.69	0.899	0.538
		(0.175)	(0.140)	(0.282)	(0.200)	(0.155)	(0.161)	(0.515)	(3.542)	(0.240)	(0.158)
Jan.11-Feb.17	$\boldsymbol{\beta_T}$	1.15	1.15	1.03	1.05	1.03	1.06	1.11	1.06	1.21	1.10
		(0.072)	(0.060)	(0.039)	(0.046)	(0.032)	(0.042)	(0.052)	(0.031)	(0.034)	(0.038)
	$\boldsymbol{\beta_\epsilon}$	5.06	4.88	4.36	4.85	3.83	5.05	6.00	5.42	5.50	4.86
		(0.283)	(0.310)	(0.175)	(0.279)	(0.087)	(0.157)	(0.360)	(0.152)	(0.085)	(0.166)
	α*-α	6.45	6.60	6.09	6.16	7.54	6.13	7.18	4.41	6.08	6.21
		(0.173)	(0.199)	(0.127)	(0.123)	(0.140)	(0.151)	(0.232)	(0.134)	(0.202)	(0.074)
	B	1.11	1.50	1.87	1.36	3.25	1.17	1.56	1.38	1.16	1.71
		(0.092)	(0.088)	(0.144)	(0.131)	(0.321)	(0.107)	(0.150)	(0.080)	(0.098)	(0.117)

pricing changed in character following the 1987 stock-market crash. In particular, a "smile" in graphs of implied volatility against exercise price is thought to apply only post-1987. To examine this idea, we expanded our analysis to the period June 1983 to July 1994, using quotes on S&P 100 index options from the Berkeley Options Data Base.[18] These data derive from CBOE's Market Data Retrieval tapes. Because of the limited number of quotes on far-out-of-the-money options in this data base, we form our monthly panel by aggregating quotes from the last five trading days of each month.

Table 3.8, Panel A presents the regression estimates for the United States for 1983-2017 in the context of our baseline model. In this estimation, the data from Berkeley Options Data Base related to the S&P 100 for June 1983 to July 1994 are treated as comparable to the OTC data related to the S&P 500 for August 1994-February 2017. The estimate for β_T is 0.985 (s.e.=0.034) and that for β_ε is 5.14 (0.54). These results are close to those in Table 3.1 with U.S. OTC data on the S&P 500 for 1994-2017.

As before, we back out a time series for estimated disaster probability, p_t, based on monthly fixed effects, assuming that the parameters other than pt in the term- for options prices in equation (3.25) are fixed. We also use levels for these other parameters as specified before. Figure 3.3 graphs the time series of estimated disaster probability. Readily apparent is the dramatic jump in p_t at the time of the October 1987 crash, in which the S&P 500 declined by 20.5% in a single day. The estimated pt reached 119% per year but fell rapidly thereafter.[19] The Persian Gulf War of 1990 1991 caused another rise in disaster probability to 19-20%.

Table 3.8, Panel B, shows statistics associated with the time series in Figure 3.3. A comparison pre-crash (June 1983-Sept 1987) and post-crash (Oct 1988-July 1994), based on the data related to the S&P 100 from Berkeley Options Data Base, shows an increase in the

[18] Direct access to this database has been discontinued. We thank Josh Coval for sharing his version of the data. We have the data from Berkeley Options Data Base through December 1995.

[19] Note that a disaster probability above 100% per year is well defined in the context of a continuous-time Poisson specification.

Table 3.8: U.S. Regression estimates, 1983-2017

This table presents regression estimates of our baseline model and summary statistics on the estimated disaster probabilities for the United States from June 1983 to February 2017. The data from June 1983 to July 1994 are based on the S&P 100 index, from the Berkeley Options Database. We form monthly panels of put-options prices by aggregating quotes from the last five trading days of each month. Consistent with the methodology used to analyze OptionMetrics data, we apply a bivariate linear interpolation on the implied volatility surface to obtain put prices with granular strikes at every 5% moneyness interval and maturities ranging from one to six months. The U.S. data from August 1994 to February 2017 are those used in Table 3.1. The monthly fixed effects capture the variations in disaster probability, which are shown in Figure 3.3 and summarized in Panel B. Because of missing data, many months before August 1994 do not appear in the figure. Time-clustered standard errors are in parentheses.

Panel A: Coefficient estimates

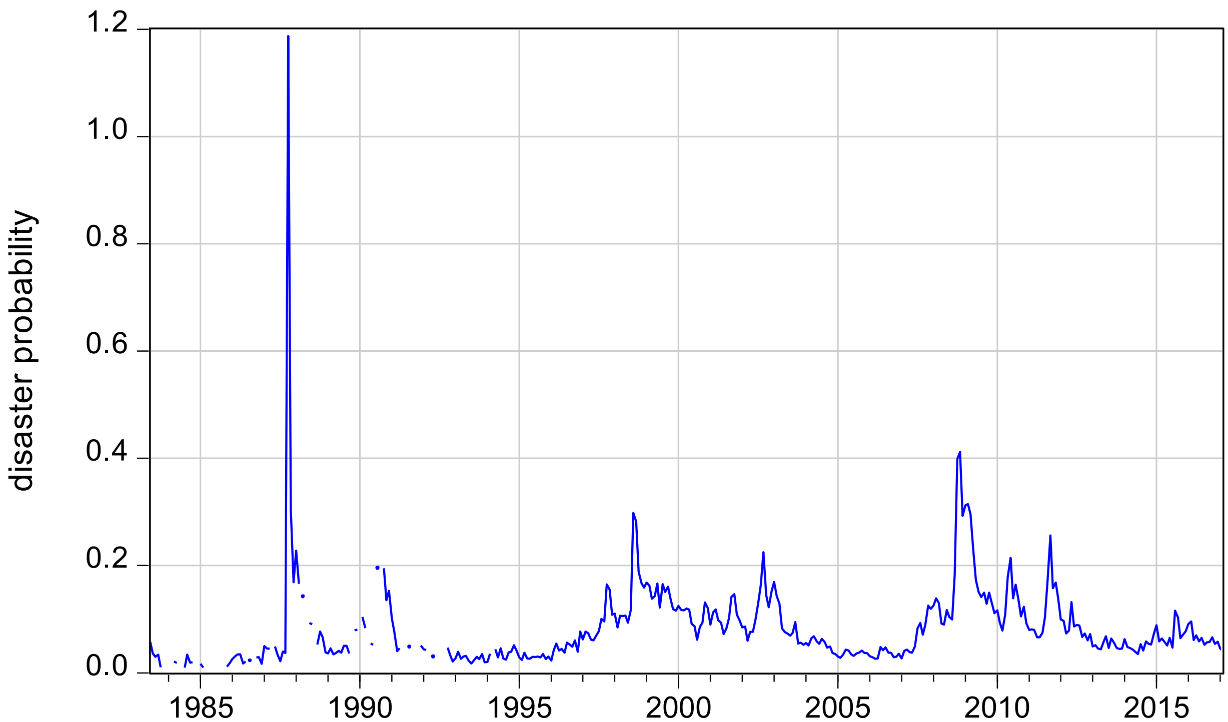
	SPX (US)
	June 1983-Feb 2017
β_T	0.985
	(0.034)
β_ϵ	5.14
	(0.541)
α*-α	6.72
	(6.037)
B	0.669
	(0.419)

Panel B: Disaster probabilities, summary statistics

Period start	Period end	min	25%	median	mean	75%	max
Jun-83	Sep-87	0.010	0.019	0.029	0.028	0.034	0.057
Oct-87	Sep-88	0.053	0.093	0.166	0.270	0.228	1.188
Oct-88	Jul-94	0.017	0.030	0.041	0.054	0.053	0.195
Aug-94	Feb-17	0.023	0.048	0.075	0.091	0.117	0.412

Figure 3.3: Estimated U.S. Disaster Probabilities, 1983-2017

This figure presents the estimated U.S. disaster probabilities, pt, associated with the regression in Table 3.8. The underlying data from June 1983 to July 1994 associate with the S&P 100 and are from Berkeley Options Data Base. The data from August 1994 to February 2017 are the OTC data based on the S&P 500 and are the same as those used in Figure 3.2. The methodology corresponds to that used in Figure 3.2.



typical size and volatility of the estimated disaster probability, pt. The pre-crash mean and median are 0.028 and 0.029, respectively, whereas the post-crash values are 0.054 and 0.041. Moreover, the minimum value pre-crash, 0.010, is lower than that, 0.017, post-crash. These patterns continue in the period August 1994-February 2017, using the OTC data related to the S&P 500. Thus, the overall suggestion is that the October 1987 crash "permanently" raised the average disaster probability and also increased the minimum level to which the disaster probability tended to revert. These changes likely account for the introduction of a smile (or at least an intensified smile) into the graph of implied volatility against exercise price following the October 1987 stock-market crash.

3.5 Conclusions

Options prices contain rich information on market perceptions of rare disaster risks. We develop a new options-pricing formula that applies when disaster risk is the dominant force, the size distribution of disasters follows a power law, and the economy has a representative agent with Epstein-Zin utility. The formula is simple but its main implications about maturity and exercise price accord with U.S. and other data from 1983 to 2017 on far-out-of-the-money put options on broad stock-market indices.

If the coefficient of relative risk aversion and the size distribution of disasters are fixed, the regression estimates of time fixed effects provide information on the evolution of disaster probability. The estimated disaster probability is highly correlated across nine major stock-market indices, applicable to the United States, United Kingdom, Euro area, Germany, Japan, Sweden, Switzerland, Brazil, and India. All of these series show a sharp peak during the financial crisis of 2008-09. Using U.S. data, the peak in the estimated disaster probability is much more dramatic in the stock-market crash of October 1987. This market-based assessment of disaster risk should be a valuable indicator of aggregate economic shocks for practitioners, macroeconomists, and policymakers. For example, in February 2017, the estimated U.S. disaster probability is only 4.7% per year, compared to the median of 8.6% from 1994 to 2017 and the peak of 48% in November 2008.

Bibliography

[1] Acharya, V., H. S. Shin, and T. Yorulmazer, 2013, “Fire-sale FDI,” Korean Economic Review 27, 163-202.

[2] Adrian, T., Etula, E., and Muir, T. (2014). Financial intermediaries and the cross-section of asset returns. The Journal of Finance, 69(6):2557–2596.

[3] Akhtar, M., 1997, “Understanding Open Market Operations”, Federal Reserve Bank of New York,Public Information Department.

[4] Avdjiev, S., Du, W., Koch, C., and Shin, H. S. (2016). The dollar, bank leverage and the deviation from covered interest parity.

[5] Baba, N., Packer, F., and Nagano, T. (2008). The spillover of money market turbulence to fx swap and cross-currency swap markets. BIS Quarterly Review, March.

[6] Bacchetta, P., van Wincoop, E., 2017, “Gradual Portfolio Adjustment: Implications for Global Equity Portfolios and Returns”, NBER Working Paper 23363.

[7] Backus, David, Mikhail Chernov, and Ian Martin. 2011. “Disasters Implied by Equity Index Options,” Journal of Finance, 66(6), 1969–2012.

[8] Baker, M. (2009). Capital market-driven corporate finance. The Annual Review of Financial Economics, 1:181–205.

[9] Baker, M. and Wurgler, J. (2000). The equity share in new issues and aggregate stock returns. The Journal of Finance, 55(5):2219–2257.

[10] Baker, M., Foley, C. F., and Wurgler, J. (2009). Multinationals as arbitrageurs: The effect of stock market valuations on foreign direct investment. Review of Financial Studies, 22(1):337–369.

[11] Bansal, Ravi and Amir Yaron. 2004. “Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles,” Journal of Finance, 59(4), 1481–1509.

[12] Barro, Robert J. 2006. “Rare Disasters and Asset Markets in the Twentieth Century,” Quarterly Journal of Economics, 121(3), 823–866.

[13] Barro, Robert J. 2009. “Rare Disasters, Asset Prices, and Welfare Costs,” American Economic Review, 99(1), 243–264.

[14] Barro, Robert J. and Jose F. Ursua. 2008. “Macroeconomic Crises since 1870,” Brookings Papers on Economic Activity, 255–335.

[15] Barro, Robert J. and Tao Jin. 2011. “On the Size Distribution of Macroeconomic Disasters,” Econometrica, 79(5), 1567–1589.

[16] Barro, Robert. J. and Jose F. Ursúa. 2012. "Rare Macroeconomic Disasters," Annual Review of Economics, 4, 83–109.

[17] Bates, David S. 1991. "The Crash of '87: Was It Expected? The Evidence from Options Markets," Journal of Finance, 46(3), 1009–1044.

[18] Bollerslev, Tim and Viktor Todorov. 2011. "Estimation of Jump Tails," Econometrica, 79 (6), 1727–1783.

[19] Bollerslev, Tim and Viktor Todorov. 2011. "Tails, Fears, and Risk Premia," Journal of Finance, 66(6), 2165–2211.

[20] Bordalo, P., Gennaioli, N., and Shleifer, A. (2010). Salience theory of choice under risk. Technical report, National Bureau of Economic Research.

[21] Brenner, Menachem and Marti G. Subrahmanyam. 1988. "A Simple Formula to Compute the Implied Standard Deviation," Financial Analysts Journal, 44, 80–83.

[22] Brunnermeier, M. K. and Pedersen, L. H. (2009). Market liquidity and funding liquidity. Review of Financial studies, 22(6):2201–2238.

[23] Bräuning, F. and Ivashina, V. (2016). Monetary policy and global banking. Working paper.

[24] Campbell, J. Y. and L. Viceira, 2002, Strategic Asset Allocation: Portfolio Choice for Long-Term Investors, Clarendon Lectures in Economics, Oxford University Press.

[25] Campbell, J. Y. and R. J. Shiller, 1988, "Stock Prices, Earnings, and Expected Dividends,", Journal of Finance 43, 661-76.

[26] Campbell, J. Y., 1991, "A Variance Decomposition for Stock Returns", Economic Journal 101:157–179.

[27] Campbell, John Y. and John H. Cochrane. 1999. "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," Journal of Political Economy, 107(2), 205-251.

[28] Chernenko, S. and Sunderam, A. (2014). Frictions in shadow banking: Evidence from the lending behavior of money market mutual funds. Review of Financial Studies, 27(6):1717–1750.

[29] Cieslak, A., Morse, A. and Vissing-Jorgensen, A., 2016, "Stock Returns over the FOMC Cycle ", Working Paper.

[30] Coffey, N., Hrung, W. B., and Sarkar, A. (2009). Capital constraints, counterparty risk, and deviations from covered interest rate parity. Staff Reports. Federal Reserve Bank of New York.

[31] Corradin, S. and Rodriguez-Moreno, M. (2016). Violating the law of one price: the role of non-conventional monetary policy.

[32] Cox, John C. and Stephen A. Ross. 1976. "The Valuation of Options for Alternative Stochastic Processes," Journal of Financial Economics, 3(1-2), 145-166

[33] De Long, J. B., Shleifer, A., Summers, L. H., and Waldmann, R. J. (1990). Noise trader risk in financial markets. Journal of political Economy, pages 703–738.

[34] DeLong, J. B., A. Shleifer, L. H. Summers, and R. J.Waldmann, 1990, “Noise Trader Risk in Financial Markets,” Journal of Political Economy 98, 703-738.

[35] Du, W., Tepper, A., and Verdelhan, A. (2016). Covered interest rate parity deviations in the post-crisis world. Working paper.

[36] Duffie, D. (2010). Presidential address: Asset price dynamics with slow-moving capital. The Journal of finance, 65(4):1237–1267.

[37] Duffie, D., and B. Strulovici, 2012, “Capital Mobility and Asset Pricing”, Econometrica 80: 2469-2509.

[38] Duffie, D., and K. Singleton, 1999, "Modeling Term Structures of Defaultable Bonds" , Review of Financial Studies 12: 687-720.

[39] Duffie, D., “Asset Price Dynamics with Slow-Moving Capital”, Journal of Finance 2010, 65: 1238-1268.

[40] D’avolio, G. (2002). The market for borrowing stock. Journal of financial economics, 66(2):271–306.

[41] Epstein, Larry G. and Stanley E. Zin. 1989. “Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework,” Econometrica, 57(4): 937–69.

[42] Errunza, V., and E. Losq, 1985, “International Asset Pricing under Mild Segmentation: Theory and Test” Journal of Finance 40, 105-124.

[43] Fletcher, D. J. and Taylor, L. W. (1996). " swap" covered interest parity in long-date capital markets. The Review of Economics and Statistics, pages 530–538.

[44] Fratzscher, M., M. Lo Duca, and R. Straub, 2013, “On the International Spillovers of US Quantitative Easing,” European Central Bank Working Paper.

[45] Gabaix, Xavier and Augustin Landier. 2008. “Why Has CEO Pay Increased So Much?,” Quarterly Journal of Economics, 123(1), 49-100.

[46] Gabaix, Xavier and Yannis Ioannides. 2004. “The Evolution of City Size Distributions,” in V. Henderson and J.F. Thisse, eds., Handbook of Regional and Urban Economics, v. 4, Amsterdam, North-Holland.

[47] Gabaix, Xavier, Parameswaran Gopikrishnan, Vasiliki Plerou, and H. Eugene Stanley. 2003. “A Theory of Power Law Distributions in Financial Market Fluctuations,” Nature, 423, 267 270.

[48] Gabaix, Xavier. 2009. “Power Laws in Economics and Finance,” Annual Review of Economics, 1, 255–293.

[49] Gabaix, Xavier. 2012. “Variable Rare Disasters: an Exactly Solved Framework for Ten Puzzles in Macro-Finance,” Quarterly Journal of Economics, 127(2), 645–700.

[50] Gagnon, J., M. Raskin, J. Remache, and B. Sack, 2011, “The Financial Market Effects of the Federal Reserve’s Large-scale Asset Purchases,” International Journal of Central Banking 7 , 3–43.

[51] Garleanu, N. and Pedersen, L. H. (2011). Margin-based asset pricing and deviations from the law of one price. Review of Financial Studies, 24(6):1980–2022.

[52] Garleanu, N., L. H. Pedersen, and A.M. Poteshman, 2009, “Demand-Based Option Pricing,” The Review of Financial Studies 22, 4259-4299.

[53] Geanakoplos, J. (2010). The leverage cycle. In NBER Macroeconomics Annual 2009, Volume 24, pages 1–65. University of Chicago Press.

[54] Giovannini, Alberto, and Philippe Weil. 1989. “Risk Aversion and Intertemporal Substitution in the Capital Asset Pricing Model,” National Bureau of Economic Research, Working Paper 2824.

[55] Greenwood, R. and D. Vayanos, 2014, “Bond Supply and Excess Bond Returns,” Review of Financial Studies 27, 663-713.

[56] Greenwood, R. and Hanson, S. G. (2013). Issuer quality and corporate bond returns. Review of Financial Studies, 26(6):1483–1525.

[57] Greenwood, R. and Vayanos, D. (2010). Price pressure in the government bond market. The American Economic Review, 100(2):585–590. He, Z. and Krishnamurthy, A. (2013). Intermediary asset pricing. American Economic Review, 103(2):732–70.

[58] Greenwood, R., Hanson, S. G., and Jin, L. J. A model of credit market sentiment. Working Paper.

[59] Greenwood, R., Hanson, S. G., and Liao, G. Y. (2015). Price dynamics in partially segmented markets. Working paper.

[60] Greenwood, R., Hanson, S., and Stein, J. C. (2010). A gap-filling theory of corporate debt maturity choice. The Journal of Finance, 65(3):993–1028.

[61] Greenwood, R., S. G. Hanson, and D. Vayanos, 2015, “Forward Guidance in the Yield Curve: Short Rates versus Bond Supply,” Working Paper.

[62] Greenwood, R., S. G. Hanson, J. Rudoplh, and L. H. Summers, 2015, “The Optimal Maturity of Government Debt,” in The $13 Trillion Question: How America Manages Its Debt, Brookings Institution Press.

[63] Greenwood, R., S.G. Hanson and J.C. Stein, 2010, “A Gap-Filling Theory of Corporate Debt Maturity” Journal of Finance 65, 993-1028.

[64] Gromb, D. and D. Vayanos, 2002, “Equilibrium and Welfare in Markets with Financially Constrained Arbitrageurs,”Journal of Financial Economics 66, 361-407.

[65] Gromb, Denis and Dimitri Vayanos, 2015, “The Dynamics of Financially Constrained Arbitrage,” Working paper.

[66] Grossman, Sanford J & Miller, Merton H, 1988. “Liquidity and Market Structure,” Journal of Finance 43, 617-37.

[67] Hanson, Samuel G., 2014, “Mortgage Convexity,” Journal of Financial Economics 113(2), 270-299.

[68] Harvey, C., and Huang, R., 2002, “The Impact of the Federal Reserve Bank’s Open Market Operations”, Journal of Financial Markets, 5, 223-257.

[69] He, Z. and Krishnamurthy, A. (2013). Intermediary asset pricing. American Economic Review, 103(2):732–70.

[70] Hogg, Robert V and Allen T. Craig. 1965. Introduction to Mathematical Statistics, 2nd Edition, Macmillan, New York, 103-104.

[71] Hu, X., Pan, J., and Wang, J., 2013, "Noise as Information for Illiquidity", Journal of Finance, 68, 2223-2772.

[72] Hull, John C. (2000). Options, Futures, & Other Derivatives, 4th ed., Upper Saddle River NJ, Prentice Hall.

[73] Iida, T, T. K. and Sudo, N. (2016). An upsurge in a cip deviation during the noncrisis period and the role of divergence in monetary policy. Bank of Japan Working Paper Series.

[74] Ivashina, V., Scharfstein, D. S., and Stein, J. C. (2015). Dollar funding and the lending behavior of global banks. The Quarterly Journal of Economics, 1241:1281. 36

[75] Joyce, M. A. S., A. Lasaosa, I. Stevens, and M. Tong, 2011, "The Financial Market Impact of Quantitative Easing in the United Kingdom," International Journal of Central Banking 7, 113-161.

[76] Kelly, Bryan and Hao Jiang. 2014. "Tail Risk and Asset Prices," Review of Financial Studies, 27(10), 2841–2871.

[77] Kim, Y. C. and Stulz, R. (1988). The eurobond market and corporate financial policy: A test of the clientele hypothesis. Journal of Financial Economics, 22(2):189–205.

[78] Kou, S.G. (2002). "A Jump-Diffusion Model for Option Pricing," Management Science, 48(8), 1086-1101.

[79] Krishnamurthy, A. (2002). The bond/old-bond spread. Journal of Financial Economics, 66(2):463–506.

[80] Krishnamurthy, A. and A. Vissing-Jorgensen, 2011, "The Effects of Quantitative Easing on Interest Rates: Channels and Implications for Policy," Brookings Papers on Economic Activity, Fall 2011, 215-265.

[81] Krishnamurthy, A. and A. Vissing-Jorgensen, 2012, "The Aggregate Demand for Treasury Debt," Journal of Political Economy, 120, 233–267.

[82] Krishnamurthy, A., O. Vigneron and X. Gabaix, 2007, "Limits of Arbitrage: Theory and Evidence from the Mortgage-Backed Securities Market", Journal of Finance, 62(2), 557-595.

[83] Lamont, O. A. and Thaler, R. H. (2003). Anomalies: The law of one price in financial markets. The Journal of Economic Perspectives, 17(4):191–202.

[84] Levich, R. M. (2012). Fx counterparty risk and trading activity in currency forward and futures markets. Review of Financial Economics, 21(3):102–110.

[85] López-Salido, D., Stein, J. C., and Zakrajšek, E. (2016). Credit-market sentiment and the business cycle. Technical report, National Bureau of Economic Research.

[86] Lucas, Robert E. 1978. "Asset Prices in an Exchange Economy," Econometrica, 46, 1429–1445.

[87] Luttmer, Erzo G.J. 2007. "Selection, Growth, and the Size Distribution of Firms," Quarterly Journal of Economics, 122(3), 1103-1144.

[88] López-Salido, D., Stein, J. C., and Zakrajšek, E. (2016). Credit-market sentiment and the business cycle. Technical report, National Bureau of Economic Research.

[89] Ma, Y. (2015). Non-financial firms as cross-market arbitrageurs. Working paper.

[90] Maggiori, M. and Gabaix, X. (2015). International liquidity and exchange rate dynamics. Quarterly Journal of Economics, 130(3).

[91] Mamaysky, H. (2014). The time horizon of price responses to quantitative easing. Working paper.

[92] Mamaysky, Harry, 2014, "The Time Horizon of Price Responses to Quantitative Easing", Working Paper.

[93] Martin, Ian. 2015. "What is the Expected Return on the Market?" working paper, London School of Economics, June.

[94] McBrady, M. R. and Schill, M. J. (2007). Foreign currency-denominated borrowing in the absence of operating incentives. Journal of Financial Economics, 86(1):145–177.

[95] Mehra, Rajnish, and Edward C. Prescott. 1985. "The Equity Premium: A Puzzle," Journal of Monetary Economics, 15, 145–161.

[96] Merton, Robert C. 1976. "Option Pricing when Underlying Stock Returns Are Discontinuous," Journal of Financial Economics, 3, 125–144.

[97] Mitzenmacher, Michael. 2003. "A Brief History of Generative Models for Power Law and Lognormal Distributions," Internet Mathematics, 1(2), 226-251.

[98] Nagel, S. (2005). Short sales, institutional investors and the cross-section of stock returns. Journal of Financial Economics, 78(2):277–309.

[99] Obstfeld, Maurice. 1994. "Evaluating Risky Consumption Paths: The Role of Intertemporal Substitutability," European Economic Review, 38(7), 1471–1486.

[100] Pareto, Vilfredo. 1897. Cours d'Economie Politique, v.2, Paris, F. Pichou.

[101] Pedersen, L.H., M. Mitchell , and T. Pulvino, 2007. "Slow Moving Capital," American Economic Review 97, 215-220.

[102] Plerou, Vasiliki, Parameswaran Gopikrishnan, Xavier Gabaix, H. Eugene Stanley. 2004. "On the Origins of Power Law Fluctuations in Stock Prices," Quantitative Finance, 4, C11-C15.

[103] policy. Working Paper.

[104] Pozsar, Z. and Smith, S. (2016). Global Money Notes #5. Credit Suisse Fixed Income Research, (April 13).

[105] Rietz, Thomas A. 1988. "The Equity Risk Premium: A Solution," Journal of Monetary Economics, 22(1), 117-131.

[106] Ross, S. A., 1976, "The Arbitrage Theory of Capital Asset Pricing," Journal of Economic Theory 13, 341-360.

[107] Seo, Sang Byung and Jessica A. Wachter. 2016. “Option Prices in a Model with Stochastic Disaster Risk,” working paper, University of Pennsylvania.

[108] Shleifer, A. and Vishny, R. W. (1997). The limits of arbitrage. The Journal of Finance, 52(1):35–55.

[109] Shleifer, A., and R. W. Vishny, 1997. “The Limits of Arbitrage,” Journal of Finance 52, 35-55.

[110] Siriwardane, Emil. 2015. "The Probability of Rare Disasters: Estimation and Implications," Harvard Business School, working paper 16-061.

[111] Sokolov, V., 2009, “The Impact of Open Market Operations on the Government Bond Market: Microstructure Evidence”, Working Paper, Higher School of Economics.

[112] Stapleton, R.C., and M. G.. Subrahmanyam, 1977, “Market Imperfections, Capital Market Equilibrium and Corporation Finance,” Journal of Finance 32, 307-319.

[113] Stein, J. C. and Sunderam, A. (2016). The fed, the bond market, and gradualism in monetary

[114] Stein, J. C., 2005, “Why Are Most Funds Open-End? Competition and the Limits of Arbitrage,” Quarterly Journal of Economics 120, 247-272.

[115] Sushko, V., Borio, C., Mccauley, R., and Mcguire, P. (2016). Whatever happened to covered interest parity ? Hedging demand meets limits to arbitrage. Working paper

[116] Tuckman, B. and Porfirio, P. (2003). Interest rate parity, money market basis swaps, and cross-currency basis swaps. Fixed income liquid markets research, Lehman Brothers.

[117] Vayanos, D. and J. Vila, 2009, “A Preferred-Habitat Model of the Term Structure of Interest Rates,” NBER Working Paper No. 15487.

[118] Vayanos, D. and Vila, J.-L. (2009). A preferred-habitat model of the term structure of interest rates. NBER Working paper.

[119] Weil, Philippe. 1990. “Nonexpected Utility in Macroeconomics,” Quarterly Journal of Economics, 105(1), 29–42.

[120] Woodford, M., 2012, “Methods of Policy Accommodation at the Interest-Rate Lower Bound,” in The Changing Policy Landscape, Federal Reserve Bank of Kansas City.

[121] Wurgler, J. and E. Zhuravskaya, 2002, “Does Arbitrage Flatten Demand Curves for Stocks?” The Journal of Business 75, 583–608.