

Modeling International Asset Markets

by

Ron Alquist

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in the University of Michigan
2008

Doctoral Committee:

Professor Lutz Kilian, Co-Chair
Professor Linda L. Tesar, Co-Chair
Assistant Professor Benjamin Remy Chabot
Assistant Professor Kathy Z. Yuan

© Ron Alquist

2008

To EAG

Acknowledgments

I owe many people my gratitude for their assistance and support on this long, strange trip. I would like to thank the members of my dissertation committee: Lutz Kilian, Linda L. Tesar, Benjamin R. Chabot, and Kathy Z. Yuan. Lutz, who provided a sterling example of the economist as researcher, taught me not only how to pose but how to answer economic questions in ways that expand our fundamental knowledge of the world. Linda's constructive criticism forced me to think through my assumptions and has commensurately improved the quality of my work. Ben taught me all that I know about asset pricing, while patiently answering my persistent and often naïve questions. Kathy's comments and suggestions helped to focus my work toward the right audience. My dissertation committee possessed complementary strengths reflected in the final product, and for that, I am grateful to each member. I thank Ben and Lutz for allowing me to draw on joint for parts of my dissertation.

Two other professors at Michigan also contributed immensely to my development as an economist. William James Adams was a warm and thoughtful mentor who helped me discover how rewarding teaching can be. Stephen Salant's deep wisdom about microeconomic modeling impressed upon me the power of this type of reasoning; he has encouraged me to cultivate my skills in this area. As senior members of the field, both Jim and Steve provided me with a broad perspective on the evolution of economics as an academic discipline.

When I was an undergraduate, Menzie Chinn's encouragement and thoughtful mentoring pushed me to consider pursuing a PhD. Without his initial guidance, I would not be here, and so I am grateful for his efforts at that early juncture.

Not only have my parents been consistently encouraging during my time at Michigan, but their strong work ethic, essential curiosity, and voracious appetite for books supplied an excellent model, which served me well while I was there. I could not

have finished the dissertation without the important lessons they imparted to me. Their love of learning had a huge influence on the person that I have become.

My aunt and sister were essential for me to keep it real. I cannot recall how many times I called them anxious about an exam or a facet of my research, or wondering if I would ever finish my dissertation. They invariably provided an iconoclastic and humorous perspective on whatever ailed me at that moment. They worked hard to keep me sane, and they were – for the most part – successful.

My wife, Elizabeth, was supremely patient, understanding, and helpful at all stages of this process, even while her own career made serious demands on her time. She always had my love and respect, but now she also has my gratitude for the innumerable ways in which she made this process easier for me. I dedicate this dissertation to her.

For shepherding the final manuscript to completion, I would like to thank Adrienne Janney. Adrienne did a yeoman's job of formatting the final draft of the dissertation on short notice and without complaint. She is a star.

Table of Contents

Dedication	ii
Acknowledgements	iii
List of Figures	vi
List of Tables	vii
Chapter 1: Introduction	1
Chapter 2: What Do We Learn From the Price of Crude Oil Futures?	3
Chapter 3: Did Adhering to the Gold Standard Reduce the Cost of Capital?	55
Chapter 4: How Important is Liquidity Risk for Sovereign Bond Risk Premia?	85
Chapter 5: Conclusion	128

List of Figures

Figure

2-1: Prices of Crude Oil Futures Contracts and Spot Price of Oil, 1983.3-2007.2	44
2-2: Volume of NYMEX Oil Futures Contracts	45
2-3: Oil Futures Spread and Foreign Exchange Futures Spread, 3-Month Horizon	46
2-4a: The Effect of an Increase in Uncertainty on the Marginal Convenience Yield	47
2-4b: The Effect of an Increase in Uncertainty on the Demand for Oil	47
2-5: Negative Spread by Horizon, 1989.1-2007.2	48
2-6: 12-Month Negative Spread and VAR-based Estimate of Precautionary Demand Component of Spot Price, 1989.1-2007.2	49
2-7: Response of the Real Price of Oil to a Positive Precautionary Demand Shock	50
4-1: Market Liquidity in the London Sovereign Bond Market, 1871.1-1907.12	119
4-2: Market Liquidity and Major Defaults, 1871.1-1907.12	120
4-3: Changes in Market Liquidity and Major Defaults, 1871.1-1907.12	121

List of Tables

Table

2-1: 1-Month Ahead Recursive Forecast Error Diagnostics	36
2-2: 3-Month Ahead Recursive Forecast Error Diagnostics	37
2-3: 6-Month Ahead Recursive Forecast Error Diagnostics	38
2-4: 9-Month Ahead Recursive Forecast Error Diagnostics	39
2-5: 12-Month Ahead Recursive Forecast Error Diagnostics	40
2-6: Asymptotic p -values for Forecast Efficiency Regressions	41
2-7: Times Series Features of the Spread (Percent)	42
2-8: Contemporaneous Correlation of the Negative Spread and the VAR Estimate of Precautionary Demand Component of the Real Spot Price of Crude Oil (Percent)	43
3-1: 28-Day Excess Return Regressions: Off-Gold Portfolio-On-Gold Portfolio	71
3-2: 28-Day Excess Return Regressions: On- and Off-Gold Portfolios	73
4-1: Developing Country Bond Issues, Equity Issues, and Loan Commitments, 1990- 2007	107
4-2: Time-Series Features of Bond Market Liquidity	109
4-3: Value-Weighted Sovereign Bond Portfolios Sorted by Spread	110
4-4: Value-Weighted Sovereign Bond Portfolios Sorted by Spread	111
4-5: Equally Weighted Sovereign Bond Portfolios Sorted by Spread	112
4-6: Equally Weighted Sovereign Bond Portfolios Sorted by Spread	113
4-7: Value-Weighted Sovereign Bond Portfolios Sorted by Size	114
4-8: Value-Weighted Sovereign Bond Portfolios Sorted by Size	115
4-9: Equally Weighted Sovereign Bond Portfolios Sorted by Size	116
4-10: Equally Weighted Sovereign Bond Portfolios Sorted by Size	117
4-11: Risk Premia for Value-Weighted Sovereign Bond Portfolios	118

Chapter 1

Introduction

Since the late 1970s, the major financial centers have relaxed interest-rate and foreign-exchange controls and introduced more liberal regulatory regimes. In conjunction, these policies have facilitated the gradual expansion of international asset markets and warrant a detailed study of these increasingly important financial markets. The three papers that comprise my dissertation study international asset markets in order to understand better the information contained in the price of crude oil futures contracts; whether the exchange-rate regime reduces sovereign and private borrowing costs; and the relationship between sovereign borrowing costs and liquidity risk.

The first paper “What Do We Learn from Price of Crude Oil Futures?” assesses the forecasting ability of crude oil futures contracts traded on the New York Mercantile Exchange. Despite their widespread use among policymakers and financial analysts as predictors of the future spot price of oil, my paper establishes that the best forecast of the future spot price is the no-change forecast. This finding does not imply that there is no useful information contained in crude oil futures contracts, however. Using a model of the spot and futures market for crude, the paper shows that one can interpret the spread between the current spot and futures price as an indicator of the precautionary demand for crude oil. The paper documents several independent pieces of evidence that show the data are consistent with this interpretation of the spread between the current spot and futures price. Although crude oil futures prices are poor predictors of the future spot price of oil, they do contain information about the demand for crude oil that stems from increased uncertainty about the future supply of crude oil.

The second paper “Did Adhering to the Gold Standard Reduce the Cost of Capital?” uses a unique data set of all of the stocks and sovereign bonds traded on the London Stock Exchange between 1870 and 1907 to study whether adhering to a fixed

exchange rate regime like the gold standard reduces the borrowing costs. The findings of this paper are striking. In contrast to other papers that have investigated this question, this paper shows that, conditional on British business-cycle risk, there is no evidence that a portfolio of assets issued by countries off gold earned higher excess returns than a portfolio of assets issued by countries on gold. Overall, the returns to both stocks and bonds issued by countries on and off gold are statistically identical. These results are very robust to alternative empirical specifications. More broadly, this paper finds that the exchange rate regime mattered less for borrowing costs than previously thought.

The third paper “How Important is Liquidity Risk for Sovereign Bond Risk Premia?” uses the same data set that I use in the gold standard paper to study the influence of liquidity risk on the sovereign bond risk premia. Limitations on the availability of observable bid and ask prices for internationally traded sovereign debt from the contemporary period warrant studying this earlier period to understand the influence of liquidity risk on the price of sovereign debt. The paper shows that liquidity is an economically important and statistically significant risk that affects sovereign borrowing costs. On average, illiquid sovereign bonds yield higher returns per year than liquid sovereign bonds and the contribution of liquidity risk to the sovereign bond risk premium is economically large. This evidence suggests an independent role for unexpected changes in liquidity as a pervasive risk in determining sovereign bond risk premia, in addition to either business-cycle risk or interest-rate risk.

These three papers teach us about what oil futures prices can and cannot tell us; the inability of the exchange-rate regime to affect borrowing costs; and the key role that liquidity risk plays as a determinant of sovereign borrowing costs. By answering these questions, this dissertation has widened the scope of our understanding and deepened our knowledge of salient features of the international financial landscape.

Chapter 2

What Do We Learn from the Price of Crude Oil Futures?

2.1. Introduction

The surge in the price of crude oil since 2002 has renewed interest in the question of what determines the spot and futures price of crude oil and has highlighted the importance of being able to predict as accurately as possible the evolution of the spot price of oil (see, for example, Greenspan 2004a,b, 2005; Bernanke 2004, 2006; Gramlich 2004; Davies 2007; Kohn 2007). In this paper, I use insights provided by a theoretical model of the spot and futures market for crude oil in conjunction with empirical analysis to shed light on the relationship between the spot price of crude oil, expectations of future oil prices, the price of crude oil futures, and the oil futures spread (defined as the percent deviation of the oil futures price from the spot price of oil).

The paper is organized as follows. In section 2, I document the use of oil futures prices as predictors of spot prices at central banks and international organizations. Futures-based forecasts of the price of crude oil inform monetary policy decisions and affect financial markets' perceptions of the risks to price stability and sustainable growth. It is widely believed that oil futures prices can be viewed as effective long-term supply prices (see, for example, Greenspan 2004a) or as the expected price of oil (see, for example, Bernanke 2004). I put this common practice to the test. Using a newly constructed data set of oil futures prices and oil spot prices that includes data up to February 2007, I assess empirically whether forecasts based on the price of oil futures are more accurate than forecasts from alternative models excluding futures prices. I show that forecasts based on oil futures prices and forecasts based on the oil futures spread tend to be less accurate than forecasts from alternative easy-to-use models such as the no-

change forecast under standard loss functions including the mean-squared prediction error (MSPE). They also are more biased than the no-change forecast.

The result that futures prices are neither unbiased predictors nor the best possible predictors in the MSPE sense is new and surprising because it contradicts widely held views among policymakers and financial analysts. It also differs from some earlier empirical results in the academic literature based on shorter samples. Moreover, it contrasts with related results in the foreign exchange literature. Although the no-change forecast has been shown to work well in other contexts such as exchange rate forecasting, there are important differences between the foreign exchange market and the crude oil market. Forecast efficiency regressions for oil markets generate the expected signs and magnitudes for all coefficients, whereas similar regressions for foreign exchange markets generate coefficients of the wrong sign and magnitude (see, for example, Froot and Thaler 1990). Thus, the superiority of the random walk predictor of oil prices compared with futures prices is by no means expected.

In section 3, I conduct a systematic evaluation of the out-of-sample predictive accuracy of a broader set of oil price forecasting approaches based on the forecast evaluation period 1991.1-2007.2. A robust finding across all horizons from 1 month to 12 months is that the no-change forecast tends to be more accurate than forecasts based on other econometric models and much more accurate than professional survey forecasts of the price of crude oil. This makes the no-change forecast a natural benchmark.

In section 4, I show that the cause of the large mean-squared prediction error (MSPE) of futures-based forecasts relative to the no change forecast is not so much that these forecasts are so different on average, but rather the variability of the futures price about the spot price, as captured by the spread of oil futures. I document that there are large and persistent fluctuations in the oil futures spread that are unlike the fluctuations observed in the spread of foreign exchange futures (see, for example, Taylor 1989).

In section 5, I show that these differences can be linked to the existence of a marginal convenience yield for crude oil that is absent in foreign exchange markets. Oil inventories, unlike inventories of many financial assets, may serve to avoid interruptions of the production process or to meet unexpected shifts in demand. This option value is reflected in a convenience yield (see, for example, Brennan 1991; Pindyck 1994, 2001). I

study the implications of the marginal convenience yield for the oil futures spread in the context of a multi-period, two-country general equilibrium model of the spot and futures markets for crude oil. I show that shifts in the uncertainty about futures oil supply shortfalls may explain the excess variability of oil futures prices relative to the spot price that is responsible for their poor predictive accuracy.

In the model, an oil-producing country exports oil to an oil-consuming country that uses oil in producing a final good to be traded for oil or consumed domestically. Oil importers may insure against uncertainty about stochastic oil endowments by holding above-ground oil inventories or buying oil futures. Oil producers may sell oil futures to protect against endowment uncertainty. The model abstracts from oil below the ground. The spot and futures prices of oil are determined endogenously and simultaneously. Using comparative statics, I establish that under plausible conditions increased uncertainty about future oil supply shortfalls causes the oil futures spread to fall. Such uncertainty shifts also raise the current real spot price of oil, as precautionary demand for oil inventories increases in response to increased uncertainty. Increased uncertainty about future oil supply shortfalls in the model will cause the real price of oil to overshoot initially with no response of oil inventories on impact, followed by a gradual decline of the real price of oil, as inventories are gradually accumulated gradually over time. The model implies that the oil futures spread declines, as the component of the real spot price of oil driven by precautionary demand for crude oil increases. Hence, the negative of the oil futures spread may be viewed as an *indicator* of fluctuations in the spot price of crude oil driven by shifts in precautionary demand for oil.

In section 6, I evaluate these predictions of my model empirically. First, I show that the proposed indicator moves as expected during events such as the Persian Gulf War that a priori should be associated with large shifts in precautionary demand for crude oil. I also find evidence of shifts in the spread associated with the Asian financial crisis, with 9/11 and with the 2003 Iraq War, for example. My findings corroborate earlier results in the literature based on regression dummies as well as historical decompositions derived from structural vector autoregressive (VAR) models. Second, my indicator is highly correlated with an independent estimate of the precautionary demand component of the spot price of crude oil proposed in Kilian (2007a,b). That alternative estimate is based on

a structural VAR model of the global crude oil market and does not rely on data from the oil futures market. I show that the VAR-based measure and the futures-based measure have a correlation as high as 79 percent during 1989.1-2003.12. Third, I show that the overshooting pattern of the response of the real price of oil to a precautionary demand shock in the Kilian (2007a) VAR model is consistent with the predictions of my theoretical model. The concluding remarks are in section 7.

2.2. Do Oil Futures Prices Help Predict the Spot Price of Oil?

It is commonplace in policy institutions, including many central banks and the International Monetary Fund (IMF), to use the price of NYMEX oil futures as a proxy for the market's expectation of the spot price of crude oil.¹ A widespread view is that prices of NYMEX futures contracts are not only good proxies for the expected spot price of oil, but also better predictors of oil prices than econometric forecasts. Forecasts of the spot price of oil are used as inputs in the macroeconomic forecasting exercises that these institutions produce. For example, the European Central Bank (ECB) employs oil futures prices in constructing the inflation and output-gap forecasts that guide monetary policy (see Svensson 2005). Likewise the IMF relies on futures prices as a predictor of future spot prices (see, for example, International Monetary Fund 2005, p. 67; 2007, p. 42). Futures-based forecasts of the price of oil also play a role in policy discussions at the Federal Reserve Board (see, for example, Greenspan 2004a,b; Bernanke 2004; Gramlich 2004). This is not to say that forecasters do not recognize the potential limitations of futures-based forecasts of the price of oil. Nevertheless, the perception is that oil futures prices, imperfect as they may be, are the best available forecasts of the spot price of oil.

¹ Futures contracts are financial instruments that allow traders to lock in today a price at which to buy or sell a fixed quantity of the commodity in a predetermined date in the future. Futures contracts can be retraded between inception and maturity on a futures exchange such as the New York Mercantile Exchange (NYMEX). The NYMEX offers institutional features that allow traders to transact anonymously. These features reduce individual default risk and ensure homogeneity of the traded commodity, making the futures market a low-cost and liquid mechanism for hedging against and for speculating on oil price risks. The NYMEX light sweet crude contract is the most liquid and largest volume market for crude oil trading (NYMEX 2007a).

There are subtle differences in how oil futures prices are interpreted by policymakers. In its strongest form, the price of oil futures is viewed as the best predictor of the spot price of oil. This interpretation is exemplified by Greenspan's (2004a) remark that "... oil futures prices ... can be viewed as effective long-term supply prices." A weaker interpretation is that oil futures prices represent the expected spot price of oil. That view is illustrated by Bernanke's (2004) statement that "... futures prices of \$20 a barrel suggest that traders expect oil prices to remain at that level". Before studying the theoretical support for these statements, in this section I examine their empirical support. I formally evaluate the predictive power of oil futures prices for the spot price of oil since the creation of oil futures markets in the 1980s.

2.2.1. Forecasting Models

The Benchmark Model

Let $F_t^{(h)}$ denote the current nominal price of the futures contract that matures in h periods, S_t the current spot price of oil, and $E_t[S_{t+h}]$ the expected future spot price at date $t+h$ conditional on information available at t . A natural benchmark for forecasts based on the price of oil futures is provided by the random walk model without drift. This model implies that changes in the spot price are unpredictable, so the best forecast of the future spot price of crude oil is simply the current spot price:

$$(1) \quad \hat{S}_{t+h|t} = S_t \quad h=1, 3, 6, 9, 12$$

Below I consider two types of forecasting models based on the price of oil futures. The first model simply treats the current level of futures prices as the predictor; the second model is based on the futures spread.

Futures Prices as Future Spot Prices

The Greenspan (2004a) quote of the introduction implies the forecasting model:

$$(2) \quad \hat{S}_{t+h|t} = F_t^{(h)} \quad h=1, 3, 6, 9, 12.$$

Forecasts Based on the Futures Spread

An alternative approach to forecasting the spot price of oil is to use the spread between the spot price and the futures price as an indicator of whether the price of oil is likely to go up or down (see, for example, Gramlich 2004). If the futures price equals the expected spot price, as stated by Bernanke (2004), the spread should be an indicator of the expected change in spot prices, although not necessarily an accurate predictor of the change in spot prices in the MSPE sense. The rationale for this approach is clear from dividing $F_t^{(h)} = E_t[S_{t+h}]$ by S_t , which results in $E_t[S_{t+h}]/S_t = F_t^{(h)}/S_t$. I explore the forecasting accuracy of the spread based on several alternative forecasting models. The baseline model is:

$$(3) \quad \hat{S}_{t+h|t} = S_t \left(1 + \ln(F_t^{(h)} / S_t)\right), \quad h=1, 3, 6, 9, 12$$

To allow for the possibility that the spread may be a biased predictor, it is common to relax the assumption of a zero intercept:

$$(4) \quad \hat{S}_{t+h|t} = S_t \left(1 + \hat{\alpha} + \ln(F_t^{(h)} / S_t)\right), \quad h=1, 3, 6, 9, 12$$

Alternatively, one can relax the proportionality restriction:

$$(5) \quad \hat{S}_{t+h|t} = S_t \left(1 + \hat{\beta} \ln(F_t^{(h)} / S_t)\right), \quad h=1, 3, 6, 9, 12$$

Finally, I can relax both the unbiasedness and proportionality restrictions:

$$(6) \quad \hat{S}_{t+h|t} = S_t \left(1 + \hat{\alpha} + \hat{\beta} \ln(F_t^{(h)} / S_t)\right), \quad h=1, 3, 6, 9, 12.$$

2.2.2. Data Description and Timing Conventions

Data Construction

In section 2.4, I will compare the real-time forecast accuracy of models (1)-(6). My empirical analysis is based on daily prices of crude oil futures traded on the NYMEX from the commercial provider *Price-Data.com*. The time series begins in March 30, 1983, when crude oil futures were first traded on the NYMEX, and extends through February 28, 2007. Crude oil futures can have maturities as long as 7 years. Contracts are for delivery at Cushing, OK. Trading ends four days prior to the 25th calendar day preceding the delivery month. If the 25th is not a business day, trading ends on the fourth

business day prior to the last business day before the 25th calendar day (NYMEX 2007b). A common problem in constructing monthly futures prices of a given maturity is that an h -month contract may not trade on a given day. I identify the h -month futures contract trading closest to the last trading day of the month and use the price associated with that contract as the end-of-month value. For all horizons, I obtain a continuous monthly time series based on a backward-looking window of at most five days. For maturities up to three months, the backward-looking window is at most three days. My approach is motivated by the objective of computing in a consistent manner end-of-month time series of futures prices for different maturities. This allows us to match up end-of-month spot prices and futures prices as closely as possible.² The daily spot price data are obtained from *Datastream* and refer to the price of West Texas Intermediate crude oil available for delivery at Cushing, OK. Figure 1 plots the monthly prices of oil futures contracts for maturities of 1 through 12 months and the spot price of crude oil starting in 1983.1. Note that contracts of longer maturities only gradually became available over the course of the sample period.

The Choice of Maturities in the Empirical Analysis

The perception that futures prices contain information about future spot prices implicitly relies on the assumption that futures contracts are actively traded at the relevant horizons. In this subsection I investigate how liquid futures markets are at each maturity h . This question is important because one would not expect $F_t^{(h)}$ to have predictive content for future spot prices, unless the market is sufficiently liquid at the relevant horizon.

Policymakers and the public widely believe that the oil futures market provides effective insurance against risks associated with crude oil production shortfalls and conveys the market's assessment of the evolution of future supply and demand conditions in the crude oil market. If the market were effectively pricing the possibility of, say, a shutdown of the Iranian oil fields or the demise of the Saudi monarchy within the next five years, one would expect active trading at such long horizons. The evidence below,

² My approach differs from that in Chernenko, Schwarz, and Wright (2004). Their approach is to treat futures prices from a window in the middle of the month as a proxy for the futures price in a given month. Yet another approach is to substitute the price of a j -month contract for a given day for the missing price of the h -month contract on that day where $j \neq h$, (see Bailey and Chan 1993).

however, suggests otherwise. Figure 2 shows the monthly trading volume corresponding to a futures contract with a fixed horizon that is closest to the last trading day of the month. *Volume* refers to the number of contracts traded in a given month.³ As illustrated in Figure 2, over the past 25 years, trading volume in the futures market has grown significantly, particularly at the 1-month and 3-month horizon, and to a lesser extent at the 6-month horizon. In 1989, the NYMEX introduced for the first time contracts exceeding twelve months and in 1999, a 7-year contract was first introduced. Although such contracts are available, the market remains illiquid at horizons beyond one year even in recent years. Trading volumes fall sharply at longer maturities.

This observation is important for my forecast evaluation because one would not expect forecasts based on futures with long maturities to provide accurate predictions, when only a handful of contracts are trading. Given the evidence in Figure 2, I therefore will restrict myself to futures contracts of up to one year in the empirical analysis below. In addition, the evidence in Figure 2 suggests that the public and policymakers have overestimated the ability of oil futures markets to provide insurance against long-term risks such as political instability in the Middle East or the development of oil resources in the Caspian Sea. Policymakers routinely rely on futures prices for long maturities in predicting future oil prices. For example, Greenspan (2004a) explicitly referred to the 6-year oil futures contract in assessing effective long-term supply prices. For similar statements also see Greenspan (2004b), Gramlich (2004) and Bernanke (2004). As my volume data in Figure 2 show, there is very little information contained in futures prices beyond one year, making it inadvisable to rely on such data. This conclusion is also consistent with prior studies of the crude oil futures market between 1983 and 1994 (see Neuberger 1999) and with perceptions of industry experts.⁴

³ In contrast to open interest, volume measures the total number of contracts, including those in a position that a trader closes or that reach delivery, and thus gives a good sense of the overall activity in the futures market. My method of data construction is consistent with the conventions used in constructing the monthly futures prices.

⁴ According to sources within the oil industry who wish to remain anonymous, oil companies are fully aware of how thin the market is at longer horizons and do not rely on futures price data for such maturities. The perception is that one trader signing a couple of contracts with a medium-term horizon may easily move the futures price by several dollars on a given day.

2.2.3. Real-Time Forecast Accuracy of Futures-Based Forecasting Models

Tables 1 through 5 assess the predictive accuracy of various forecasting models against the benchmark of a random walk without drift for horizons of 1, 3, 6, 9, and 12 months. The forecast evaluation period is 1991.1-2007.2. The assessment of which forecasting model is best may depend on the loss function of the forecaster (see Elliott and Timmermann 2007). I present results for the MSPE and the mean absolute prediction error (MAPE). I also report the bias of the forecasts, and I report the number of times that a forecast correctly predicts the sign of the change of the spot price based on the success ratio statistic of Pesaran and Timmermann (1992). In addition to ranking forecasting models by each loss function, I formally test the null that a given candidate forecasting model is as accurate as the random walk without drift. Suitably constructed p -values are shown in parentheses.

Oil Futures as Predictors of Oil Spot Prices

The first two rows of Tables 1 through 5 document that the no-change forecast has lower MSPE than the futures forecast at the 1-month, 6-month, 9-month and 12-month horizon. Only at the 3-month horizon is the futures forecast more accurate, but the improvement in accuracy is not statistically significant. Moreover, based on the MAPE metric, the random walk forecast is more accurate at *all* horizons. In all cases, the random walk forecast is less biased than the futures forecast. Nor do futures forecasts have important advantages when it comes to predicting the sign of the change in oil prices. Only at the 9-month and 12-month horizons is the success ratio significant at the 10 percent level and 5 percent level, respectively, but the improvement is only 2.6 and 3.6 percentage points. The observation that futures prices are worse predictors of the price of oil than simple no-change forecasts is important because it contradicts commonly held views that current futures prices are a good guide to the evolution of future spot prices, as exemplified by the Greenspan (2004a) and Bernanke (2004) quotations.

Oil Future Spreads as Predictors of Future Spot Prices

Rows 3-6 in Tables 1-5 document that the no-change forecast has lower MSPE than spread-based forecasts at horizons of 6, 9 and 12 months. At horizons 1 and 3 in some cases the spread models has lower MSPE, but the improvement is never statistically

significant and no one spread model performs well systematically. Based on the MAPE rankings, the no-change forecast is superior at all horizons. These results are broadly consistent with the earlier evidence for the futures forecasts. Finally, rows 3-6 reveal some evidence that spread models may help predict the direction of change at horizons of 9 and 12 months. The gains in accuracy are statistically significant, but quite moderate. There is no such evidence at shorter horizons.⁵

2.5.3. Relationship with Forecast Efficiency Regressions

It is useful to compare my results for the spread model in Tables 1 through 5 to the closely related literature on forecast efficiency regressions (see, for example, Chernenko et al. 2004; Chinn, LeBlanc, and Coibion 2005). Consider the full-sample regression model:

$$\Delta s_{t+h} = \alpha + \beta (f_t^{(h)} - s_t) + u_{t+h}, \quad h=1,3,6,9,12$$

where lower-case letters denote variables in logs and u_{t+h} denotes the error term. Forecast efficiency in the context of the oil futures spread means that the hypothesis $H_0 : \alpha = 0, \beta = 1$ holds. A rejection of these restrictions is interpreted as evidence of the existence of a time-varying risk premium (see, for example, Fama and French 1987, 1988; Chernenko et al. 2004).⁶ Chernenko et al. report that the hypothesis of forecast efficiency cannot be rejected at conventional significance levels. It is important to bear in mind that such evidence does not necessarily mean that oil prices are forecastable based on the spread in practice. First, non-rejection of a null hypothesis does not imply that the null model is true. In fact, I showed that the forecasting model (3) that imposes this null does not dominate the no-change forecasts in out-of-sample forecasts. Second, as my forecasting results show, relaxing one or more of the restrictions implied by forecast efficiency may either improve or worsen the forecast accuracy of the spread model, depending on the bias-variance trade-off. In particular, such models require the

⁵ Motivated by term-structure models, I also experimented with models including a weighted average of spreads at different horizons. These models consistently performed so poorly that no results will be reported.

⁶ Such tests implicitly postulate that the trader's loss function coincides with the econometrician's quadratic loss function. If that is not the case, forecast efficiency tests tend to be biased in favor of the alternative hypothesis (see Elliott, Komunjer, and Timmermann 2005).

estimation of additional parameters compared with the no-change forecast, and the resulting loss in forecast precision may outweigh the benefits from reduced forecast bias. Thus, there is no contradiction between my results and the forecast efficiency results in the literature.

In addition, it can be shown that the results in Chernenko et al. are not robust to updating the sample. Despite differences in the timing conventions used in constructing the monthly futures price data, I am able to replicate their results qualitatively using my data, but their sample period. For the full sample, however, I do reject the hypothesis of forecast efficiency at horizons 6 and 12 (see Table 6). This pattern is consistent with the earlier forecasting results. This rejection of forecast efficiency occurs despite the fact that $\hat{\alpha}$ is close to zero and $\hat{\beta}$ fairly close to 1, as suggested by theory, and very much unlike in the foreign exchange literature (see, for example, Froot and Thaler 1990).

2.3. What Is the Best Predictor of the Spot Price of Oil?

The preceding section demonstrated that simple no-change forecasts of the price of oil tend to be more accurate in the MSPE sense than forecasts based on oil futures prices. This does not mean that the no-change forecast is necessarily the best predictor of the spot price. The first part of this section broadens the scope of forecasting methods to include other predictors. One alternative approach to measuring market expectations is the use of survey data. While economists have used survey data extensively in measuring the risk premium embedded in foreign exchange futures (see Chinn and Frankel 1995), this approach has not been applied to oil futures, with the exception of recent work by Wu and McCallum (2005). Yet another approach is the use of more sophisticated econometric forecasting models of the spot price of crude oil.

2.3.1. Other Candidate Forecasting Models

Survey Forecasts

Given the significance of crude oil to the international economy, it is surprising that there are few organizations that produce monthly forecasts of spot prices. In the oil industry, where the spot price of oil is critical to investment decisions, oil firms tend to make annual forecasts of future spot prices for horizons as long as 15-20 years, but these are

not publicly available. The U.S. Department of Energy's International Energy Agency (IEA) uses a structural econometric model of crude oil supply and demand to produce quarterly forecasts of the spot price of oil, but these forecasts are available only beginning in late 2004. The Economist Intelligence Unit has produced annual forecasts since the 1990s for horizons of up to 5 years. None of these sources provides monthly forecasts.

A standard source of monthly forecasts of the price of crude oil is *Consensus Economics Inc.*, a U.K.-based company that compiles private sector forecasts in a variety of countries. Initially, the sample consisted of more than 100 private firms; it now contains about 70 firms. Of interest to us are the survey expectations for the 3- and 12-month ahead spot price of West Texas Intermediate crude oil, which corresponds to the type and grade delivered under the NYMEX futures contract. The survey provides the arithmetic average, the minimum, the maximum, and the standard deviation for each survey month beginning in October 1989 and ending in February 2007. I use the arithmetic mean at the relevant horizon:

$$(7) \quad \hat{S}_{t+h|t} = S_{t,h}^{CF} \quad h = 3, 12$$

Econometric Forecasts

An alternative to modeling expectations of spot prices for crude oil is based on econometric models. One example of such econometric models is the random walk model without drift introduced earlier. In this section, I introduce the random walk with drift and the Hotelling model as additional competitors. Given that oil prices have been persistently trending upward (or downward) at times, it is natural to consider a random walk model with drift. One possibility is to estimate this drift recursively, resulting in the forecasting model:

$$(8) \quad \hat{S}_{t+h|t} = S_t (1 + \alpha) \quad h = 1, 3, 6, 9, 12$$

Alternatively, a local drift term may be estimated using rolling regressions:

$$(9) \quad \hat{S}_{t+h|t} = S_t (1 + \Delta \bar{s}_t^{(l)}) \quad h = 1, 3, 6, 9, 12, \quad l = 1, 3, 6, 9, 12$$

where $\hat{S}_{t+h|t}$ is the forecast of the spot price at $t+h$; and $1 + \Delta \bar{s}_t^{(l)}$ is the geometric average of the monthly percent change for the preceding l months, i.e., the percent change in the spot price between t and $t-l+1$. This model postulates that traders extrapolate from the

spot price's recent behavior when they form expectations about the future spot price. The local drift model is appealing in that it may capture "short-term forecastability" that arises from local trends in the oil price data.

An alternative approach is motivated by Hotelling's (1931) model, which predicts that the price of an exhaustible resource such as oil appreciates at the risk free rate of interest:

$$(10) \quad \hat{S}_{t+h|t} = S_t(1 + i_{t,h}) \quad h=3, 6, 12$$

where $i_{t,h}$ refers to the interest rate at the relevant maturity h .⁷ Although the Hotelling model seems too stylized to generate realistic predictions, I include this method given recent evidence that the Hotelling model does well in forecasting the future spot price of oil (see Wu and McCallum 2005). I use the Treasury bill rate as a proxy for the risk free rate.⁸

2.3.2. Real-Time Forecast Accuracy of Other Forecasting Approaches

In this subsection, I compare the real time forecast accuracy of models (7)-(10) to that of the no-change forecast in (2). Section 2.3 established that the no-change forecast tends to be more accurate than models based on the price of oil futures. An obvious question is whether the no-change forecast can be improved upon, for example, by using information on interest rates.

Hotelling Model

Row 7 in Tables 2, 3, and 5 shows that the random walk model has lower MSPE than the Hotelling model at horizons of 3 and 6 months, whereas at the 12-month horizon the ranking is reversed. This reversal is not statistically significant, however. Based on the MAPE, the no-change forecast is superior at all three horizons. The Hotelling forecasting model has systematically lower bias at all three horizons than the no-change forecast. It also is systematically better at predicting the sign of the change in oil prices than futures

⁷ Assuming perfect competition, no arbitrage, and no uncertainty, oil companies extract oil at a rate that equates: (1) the value today of selling the oil less the costs of extraction; (2) and the present value of owning the oil, which, given the model's assumptions, is discounted at the risk free rate. In competitive equilibrium, oil companies extract crude oil at the socially optimal rate.

⁸ Specifically, we use the 3-month, 6-month, and 12-month constant-maturity Treasury bill rates from the Federal Reserve Board's website <http://federalreserve.gov/releases/H15/data.htm>

forecasts, although I cannot assess the statistical significance of the improvement, given that there is no variability at all in the sign forecast.

Random Walk Models with Drift

The next six rows in Tables 1-5 document that allowing for a drift in no case significantly lowers the MSPE of the random walk model, when the drift is estimated based on rolling regressions, and only in one case, when the drift is estimated recursively. Allowing for a drift lowers the MAPE at some horizons and for some models, but the gains are not systematic and different models work well for different horizons. Again, the Clark and West (2005) test rejects the null of no predictability in several cases (mainly at the nine-month horizon). As discussed earlier, that rejection does not necessarily translate into accuracy gains in real time forecasting, as evidenced by the MAPE rankings. In some cases, allowing for a drift also improves significantly the ability to predict the sign of the change of the oil price at longer horizons, but only when the drift is estimated recursively. In general, the results for the random walk with drift are quite sensitive to the model specification and forecast horizon, and they do not account for the “specification mining” implicit in considering a large number of alternative models with drift (see Inoue and Kilian (2004) and the references therein). There is no evidence that such models dominate the no-change forecast.

Professional Survey Forecasts

The last row in Tables 2 and 5 shows that the consensus survey forecast has much higher MSPE than the no-change forecast at both the 3-month and 12-month horizons. It also has a larger bias and higher MAPE and there is no statistically significant evidence that it is better at predicting signs than a coin flip. The survey forecast is also inferior to the futures forecasts, suggesting that survey respondents do not rely on futures price data alone in forming their expectations.

2.3.3. Why the No-Change Forecast Is a Plausible Measure of the Expected Spot Price

The central result of section 3.2 is that no-change forecasts of the price of oil tend to be more accurate than forecasts based on econometric models and more accurate than survey

forecasts.⁹ This result is consistent with views among oil experts. For example, Peter Davies, chief economist of British Petroleum, has noted that “we cannot forecast oil prices with any degree of accuracy over any period whether short or long” (see Davies 2007). The favorable forecasting performance of the no-change forecast also is consistent with the observed high persistence of the nominal spot price of oil (see, for example, Diebold and Kilian 2000). The high autocorrelation of commodity prices in general has been widely recognized in the literature (see, for example, Deaton and Laroque 1992, 1996). Finally, it is important to stress that the superior forecast accuracy of the random walk model without drift does not contradict theoretical results in the literature that oil prices are endogenous with respect to global macroeconomic conditions (see, for example, Barsky and Kilian 2002). The first point to keep in mind is that macroeconomic determinants such as U.S. interest rates, U.S. inflation, or global economic growth are but one of many determinants of the price of oil. For example, many of the major oil price increases in recent decades have been associated with unforeseen political disturbances in the Middle East and rising concerns about future oil supply shortfalls. Hence, one would not expect forecasting models based on macroeconomic fundamentals alone to be successful in practice. The second point to bear in mind is that predictability that exists in population may be difficult to exploit in real time in finite samples. The link from macroeconomic fundamentals to the price of oil is complicated and likely to be nonlinear. Even if the spot price of crude oil does not truly follow a random walk, random walk forecasts tend to be attractive in terms of their mean-squared prediction error (MSPE) since the reduction in variance from excluding other predictors in small samples will typically more than offset the omitted variable bias. Thus, the superior forecast accuracy of the no-change forecast does not invalidate economic models of the crude oil market.

⁹ This result differs from at least some earlier studies. For example, Chernenko et al. (2004) report evidence that futures-based forecasts have marginally lower MSPE than the no-change forecast at horizons of 3, 6 and 12 months. In related work, Wu and McCallum (2005) find that futures prices are generally inferior to the no-change forecast, but report that spread regressions have lower MSPE than the no-change forecast at short horizons. These findings do not contradict my results. The differences in MSPE rankings can be traced mainly to differences in the sample period. The sample period considered in my paper is longer than in any previous study. Further sensitivity analysis suggests that evidence of accuracy gains, sometimes obtained in samples shorter than mine, tends to vanish when the full sample is examined.

2.4. Why Is the MSPE of Oil-Futures Prices So Large Relative to the No-Change Forecast?

The preceding section demonstrated that under the MSPE metric the best predictor of the nominal price of oil is the no-change forecast. This section examines in greater detail the differences between the no-change forecast and the forecast based on oil futures prices. A formal analysis of what precisely goes wrong with the oil futures forecast will help motivate the theoretical analysis of the oil spot and futures markets in the next section. For this purpose it is convenient to express the deviation of the futures price from the no-change forecast in percentage terms as $f_t^{(h)} - S_t$.

There are two possible reasons for the higher MSPE of $F_t^{(h)}$ relative to S_t . One is higher forecast bias; the other is a higher forecast variance. In Table 7, I first evaluate the possibility that $F_t^{(h)}$ is different on average from S_t . For expository purposes, I focus on the 3-month and 12-month horizons. My sample period is 1989.1-2007.2, as a contiguous time series for the 12-month spread becomes available only starting in 1989.1. Using heteroskedasticity and autocorrelation consistent (HAC) standard errors, on average both the 3-month and 12-month spread are statistically different from zero at the 1% level. Although the rejection is decisive, Table 7 shows that the mean deviation is comparatively small in magnitude (about 1% at the 3-month horizon and below 5% at the 12 month horizon).

Whereas the bias of futures prices relative to the no-change forecast may seem small, the variability about the no-change forecast is not. As Table 7 shows, at any point in time, the discrepancy between the futures price and the spot price may be very large and go in either direction. It is this variability of the deviation of futures prices from spot prices rather than the differences in mean that drives the larger MSPE of futures-based forecasts and that makes the use of such oil price forecasts inadvisable. The time-variation in the spread is not only large, but highly persistent. In Table 7, I measure this persistence by modeling the spread as an autoregression with the lag order selected by the Akaike Information Criterion. Based on the fitted autoregressive models, I compute the sum of the autoregressive coefficients as a measure of persistence as suggested by

Andrews and Chen (1994). The estimated persistence for the 3-month spread in the first column is 0.74, whereas that for the 12-month spread is 0.81.

The evidence in Table 7 is important because it suggests that the key to understanding the poor predictive accuracy of oil futures prices relative to the no-change forecast is to understand the causes of the excess variability of oil futures prices relative to the spot price of oil. The existence of such large fluctuations in the oil futures spread may seem surprising at first, considering the much lower variability and persistence of the futures spread in the widely studied foreign exchange futures market. The spread of foreign exchange futures prices over the spot exchange rate is well explained by the bilateral interest rate differential because the spread captures the opportunity cost of holding assets in one currency as opposed to another. This covered interest rate parity result has been documented, for example, by Taylor (1989). Considering the typical size of interest rate differentials, the spread in major foreign exchange markets tends to be small. This point is illustrated in Figure 3. The oil futures spread is far more variable than the U.S.-U.K. foreign exchange futures spread.

In the next section I propose a theoretical explanation of this discrepancy. I observe that the difference between the oil futures price and the expected spot price of oil is not accounted for by the interest rate alone, but that it also reflects the value that firms derive from having ready access to oil, a fact commonly referred to as the convenience yield. The presence of this convenience yield makes the analysis of oil futures markets fundamentally different from the analysis of the market for foreign exchange futures. I propose a theoretical model that explains the persistent and large fluctuations in the spread in terms of fluctuations in the marginal convenience yield. The model implies that fluctuations in the marginal convenience yield can be directly linked to shifting fundamentals in the form of expectation shifts about future oil supply shortfalls. Whereas concerns about future supply shortfalls may in principle arise in any commodity market, there is reason to believe that such concerns historically have been particularly relevant in the crude oil market and may explain both large and sharp fluctuations in the spread over time.

2.5. A Two-Country General Equilibrium Model of the Oil Futures and Oil Spot Markets

2.5.1. Model Description

The model in this section can be viewed as a generalization of the analysis in Pindyck (1994, 2001). There are two countries, the United States and Saudi Arabia. Saudi Arabia trades its oil endowment with the United States in exchange for a consumption good that the United States produces from oil to be delivered at the end of the period. The United States consumes some of the final consumption good and sells the rest to Saudi Arabia. Saudi Arabia is treated as an endowment economy in recognition of the fact that capacity constraints have been binding in global crude oil production in recent years (see Kilian 2008). The existence of capacity constraints implies that extracting less oil today does not permit more oil to be extracted in the future. Each period, Saudi Arabia receives a random oil endowment $\tilde{\omega}_t$. The oil endowment in period t is $\tilde{\omega}_t = \omega + \varepsilon_t$ with probability θ ; and $\tilde{\omega}_t = \omega - \hat{\varepsilon}_t$ with probability $1 - \theta$ and $\hat{\varepsilon}_t = \theta \varepsilon_t / (1 - \theta)$ such that $E(\tilde{\omega}_t) = \omega$. The variance of the oil endowment is σ_ε^2 .

In each period, the United States chooses: (1) next period's above-ground inventory holdings of oil (I_t); (2) the number of oil futures contracts that deliver one barrel of oil next period; (3) the number of risk-free one-period bonds that yield $(1 + r_{t,t+1})$, and (4) the quantity of oil to use in the production of the consumption good. Saudi Arabia chooses the number of oil futures contracts and the number of risk-free bonds it wishes to hold. The price of the consumption good in period t is P_t and the spot price of oil is S_t . The price of the consumption good is the numeraire.

2.5.2. The United States' Demand for Oil

The United States chooses the amount of oil to use in the production of the consumption good; and the amount of oil to store as above-ground inventory. Imported oil can be transformed into the consumption good using the production function $F(Z_t)$, where Z_t is the quantity of oil the United States uses in producing the consumption good. I

postulate that $F'(Z_t) > 0$, $F''(Z_t) < 0$, $F'''(Z_t) > 0$, and $\lim_{Z_t \rightarrow 0} F'(Z_t) = \infty$. The United States chooses Z_t such that the marginal product of oil equals the real price of oil in terms of the consumption good

$$(11) \quad S_t/P_t = F'(Z_t),$$

which implies the demand schedule:

$$Z(S_t, P_t) \equiv F'^{-1}(S_t/P_t).$$

The resource constraint for crude oil is given by the identity

$$\Delta I_t \equiv \tilde{\omega}_t - Z(S_t, P_t).$$

Re-interpreting equation (11) as a demand function in ΔI_t , I obtain the inverse net demand function expressed as a function of the random Saudi oil endowment and the change in inventories:

$$\frac{S_t}{P_t} = F'(\tilde{\omega}_t - \Delta I_t) \equiv D(\tilde{\omega}_t, \Delta I_t).$$

If S_t/P_t is drawn on the vertical axis and ΔI_t on the horizontal axis, $D(\tilde{\omega}_t, \Delta I_t)$ is upward-sloping in ΔI_t .

2.5.3. No-Arbitrage Condition 1: The Oil Futures Market

If I am willing to impose, in addition, that both the United States and Saudi Arabia are risk neutral, as Bernanke (2004) explicitly assumed, then by the no-arbitrage condition that the expected return from holding inventories must equal the real price of oil today, it follows that

$$E_t[F_t/P_{t+1}] = E_t[S_{t+1}/P_{t+1}].$$

Using a linear Taylor series approximation, I obtain that

$$F_t \approx E_t[S_{t+1}]$$

Thus, the futures price will be an approximately unbiased predictor of the spot price.

2.5.4. No-Arbitrage Condition 2: The Bond Market

Under risk neutrality, the real value of a bond today must equal the discounted real present value of a bond tomorrow:

$$(12) \quad \frac{1}{P_t} = \beta(1+r_{t,t+1})E_t \left[\frac{1}{P_{t+1}} \right] \Leftrightarrow \frac{1}{\beta(1+r_{t,t+1})P_t} = E_t \left[\frac{1}{P_{t+1}} \right].$$

A linear Taylor series approximation implies that:

$$(12') \quad \beta(1+r_{t,t+1}) \approx 1 \Leftrightarrow r_{t,t+1} \approx \beta - 1$$

2.5.5. No-Arbitrage Condition 3: The Market for Storage

The distinguishing feature of my model is the existence of a market for storage. Storage takes the form of holding above-ground oil inventories. The term *convenience yield* in the literature refers to the benefits arising from access to crude oil in the form of inventories such as the ability to avoid disruptions of the production process or the ability to meet unexpected demand for the final good. The convenience yield is a commonly used modeling device (see, for example, Brennan 1991; Fama and French 1988; Gibson and Schwartz 1990; Pindyck 1994). Its microeconomic foundations have been discussed in Ramey (1989), Litzenberger and Rabinowitz (1995), Considine (1997), among others. I denote the convenience yield by $g = g(I_t, \sigma_\varepsilon^2)$. Let $g_1 = g_1(I_t, \sigma_\varepsilon^2)$ denote the marginal convenience yield associated with holding additional above-ground inventories between t and $t+1$. Following the commodity pricing literature, I impose that $g_1 > 0$, $g_{11} < 0$ and $g_{12} > 0$, where g_i denotes the derivative of g with respect to its i^{th} argument and g_{ij} the cross-partial derivative of g with respect to the arguments i and j . As increases in the variance make production shortfalls more likely, the marginal convenience yield from holding inventories is increasing in the variance. Throughout the paper I also postulate that the Inada condition

$$\lim_{I_t \rightarrow 0} g_1(I_t, \sigma_\varepsilon^2) = \infty$$

holds, which ensures that the U.S. holds strictly positive inventories. With $g_1(I_t, \sigma_\varepsilon^2)$ on the vertical axis and above-ground inventory holdings on the horizontal axis, the

intersection of the $g_1(I_t, \sigma_\varepsilon^2)$ curve and inventory holdings I_t describes the equilibrium in the market for storage.

Abstracting from costs of storage, no arbitrage implies that storing a barrel of oil above ground for one period and simultaneously selling short a one-period futures contract is a risk-free strategy:

$$\underbrace{(1+r_{t,t+1})g_1 + E_t \left[\frac{S_{t+1}}{P_{t+1}} \right] - \frac{S_t}{P_t}}_{\text{Return from storing oil}} + \underbrace{E_t \left[\frac{F_t}{P_{t+1}} \right] - E_t \left[\frac{S_{t+1}}{P_{t+1}} \right]}_{\text{Return from selling futures contract}} = (1+r_{t,t+1})g_1 - \frac{S_t}{P_t} + E_t \left[\frac{F_t}{P_{t+1}} \right].$$

By no arbitrage, the returns to this investment must equal the return on investing the same amount at the risk-free rate:

$$r_{t,t+1} \frac{S_t}{P_t} = (1+r_{t,t+1})g_1 - \frac{S_t}{P_t} + E_t \left[\frac{F_t}{P_{t+1}} \right].$$

Since $E_t \left[\frac{F_t}{P_{t+1}} \right] \approx \frac{F_t}{P_t}$ given equation (12), I obtain:

$$(13) \quad (1+r_{t,t+1}) \frac{S_t}{P_t} - \frac{F_t}{P_t} \approx (1+r_{t,t+1})g_1$$

Equation (13) shows that the difference between the capitalized real spot price and the real futures price is equal to the capitalized marginal convenience yield.

2.5.6. A Permanent Mean-Preserving Spread of Oil Endowments

In this subsection, I derive two comparative statics results under risk neutrality. The first result is that an increase in uncertainty about the future oil supply shortfalls immediately raises the real spot price of oil; the second result is that under plausible assumptions this increase in uncertainty lowers the oil futures spread. I model the increase in uncertainty as a mean-preserving increase in the spread of the oil endowment shock. The thought experiment is an increase in ε_t .

The mean preserving spread helps us abstracts from changes in the conditional mean of oil supplies and focus on changes in the conditional variance. The motivation for this modeling choice is best seen by focusing on the example of the Persian Gulf War. Events such as the invasion of Kuwait in August of 1990 have two distinct effects. First, they cause a reduction in expected oil supply. This oil supply shock represents a change

in the conditional mean of oil supplies. It has been documented in the literature that such a shock indeed occurred in 1990, but that this supply shock fails to explain the bulk of the movements in the real price of oil in 1990/91. Second, there is an increase in uncertainty about future oil supply shortfalls. Indirect evidence that the price spike of 1990/91 was driven by increased uncertainty about future oil supply shortfalls has been presented in Kilian (2008). To keep the model tractable, I model this increased uncertainty as an increase in the conditional variance of oil supplies, implicitly abstracting from the global business cycle or any other change in the conditional mean.

Result 1: An Increase in Uncertainty Increases the Real Spot Price

I solve the no-arbitrage condition (13) for S_t/P_t and substitute for $1/((1+r_{t,t+1})P_t)$ from equation (12) to obtain

$$\frac{S_t}{P_t} = \beta E_t \left[\frac{F_t}{P_{t+1}} \right] + g_1(I_t, \sigma_\varepsilon^2)$$

$E_t[F_t/P_{t+1}] = E_t[S_{t+1}/P_t]$ by the no-arbitrage condition in the futures market. Using equation (12) to substitute for the real price of oil in terms of the marginal product, I arrive at:

$$F(\tilde{\omega}_t - \Delta_t) = \beta E_t[F(\tilde{\omega}_{t+1} - \Delta_{t+1})] + g_1(I_t, \sigma_\varepsilon^2),$$

implying that the United States equates the marginal benefits and marginal costs of these inventory holdings. The mean-preserving spread drives a wedge between the left-hand and right-hand side of this intertemporal marginal efficiency condition. Because $F(\cdot)$ is convex, the mean-preserving spread increases $E_t[F(\tilde{\omega}_{t+1} - \Delta_{t+1})]$ by Jensen's inequality (Hirshleifer and Riley 1992). It also increases the marginal willingness to pay for inventories, given by $g_1(I_t, \sigma_\varepsilon^2)$. To re-establish intertemporal marginal efficiency, the United States must increase its inventory holdings such that equality is re-established.

Figure 4 illustrates the dynamic adjustment process of the real price of oil and of U.S. oil inventories in response to an exogenous increase in uncertainty about future oil supply shortfalls. Figure 4a plots the marginal convenience yield. Figure 4b shows the corresponding inverse U.S. demand function for oil. In the model, date t inventory

holdings are determined by the quantity of inventories the U.S. decided to hold at time $t-1$. Suppose that we are at point A in Figure 4a at the beginning of the period. When there is a mean-preserving increase in the endowment spread, the marginal convenience yield schedule shifts upwards instantaneously, because the U.S. values each unit of inventory more than it did prior to the increase in uncertainty. We move along the inventory schedule from point A to point B . Consequently, by the concavity of its production function, the United States finds it optimal to increase its future inventory holdings relative to last period's inventory holdings. Thus $I_{t-1} \neq I_t^*$ and

$\Delta I_t = I_t^* - I_{t-1} > 0$. This implies a decrease in the real price of oil over time, starting from point B , as the United States moves along the marginal convenience yield schedule towards point C . The accumulation of additional inventories is associated with a decline in the real price of oil, as the marginal convenience yield falls. The real price of oil in the new long-run equilibrium will be higher than its level at $t-1$, but lower than its impact level. To summarize, I expect the real price of oil to overshoot in response to increased uncertainty about future oil supply shortfalls, whereas inventories will be accumulated only gradually over time. The overshooting result for the real price of oil is analogous to the overshooting of the exchange rate in the Dornbusch (1976) model. It is driven by the assumption that inventories are predetermined and will not adjust fully to an increase in uncertainty on impact.

Result 2: An Increase in Uncertainty Decreases the Oil Futures Spread

By rearranging equation (13), I obtain an expression for the spread:

$$(14) \quad \frac{F_t - S_t}{S_t} = r_{t,t+1} - (1 + r_{t,t+1}) \frac{g_1(I_t, \sigma_\varepsilon^2)}{S_t/P_t}.$$

A sufficient condition for the oil futures spread to decrease in response to a mean-preserving spread is that

$$\frac{dr_{t,t+1}}{d\varepsilon_t} - \frac{d(1 + r_{t,t+1})}{d\varepsilon_t} \left[\frac{g_1(I_t, \sigma_\varepsilon^2)}{S_t/P_t} \right] - (1 + r_{t,t+1}) \left\{ \frac{1}{F'} \left[g_{11}(I_t, \sigma_\varepsilon^2) \frac{d\Delta I_t}{d\varepsilon_t} + g_{12}(I_t, \sigma_\varepsilon^2) \frac{d\sigma_\varepsilon^2}{d\varepsilon_t} \right] + \frac{F''}{F'^2} g_1(I_t, \sigma_\varepsilon^2) \frac{d\Delta I_t}{d\varepsilon_t} \right\} < 0$$

Since $dr_{t,t+1}/d\varepsilon_t \approx 0$, the first two terms in this expression are zero. The sign of the expression depends on the relative magnitudes of (1) the decrease in the marginal convenience yield associated with the increase in inventories triggered by the shock to the endowment distribution; and (2) the increase in the marginal convenience yield associated with the increase in σ_ε^2 triggered by the same shock. The spread declines if and only if

$$(15) \quad \frac{d\sigma_\varepsilon^2}{d\varepsilon_t} > -\frac{1}{g_{12}} \left[g_{11} + g_1 \frac{F''}{F'} \right] \frac{d\Delta I_t}{d\varepsilon_t}.$$

I can express both $d\sigma_\varepsilon^2/d\varepsilon_t$ and $d\Delta I_t/d\varepsilon_t$ in terms of the model's parameters and show that expression (15) is equivalent to:

$$(15') \quad g_{12} > \frac{\lambda(1-\theta)B}{2\theta\varepsilon_t(1-\lambda)},$$

where $\lambda \equiv (g_{11}/g_1 + F''/F')g_1/(A - g_{11})$, $0 < \lambda < 1$; and

$$A = -F''(\tilde{\omega}_t - \Delta I_t) - \beta E_t[F''(\tilde{\omega}_{t+1} - \Delta I_{t+1})] > 0$$

$$B = \beta\theta[F''(\omega_t + \varepsilon_t - \Delta I_{t+1}) - F''(\omega_t + \hat{\varepsilon}_t - \Delta I_{t+1})] > 0$$

Hence, for a given stock of inventories and increase in ε_t , the spread will decline, provided g_{12} is large enough. The term g_{12} measures the shift in the marginal convenience yield induced by the mean-preserving spread. It represents the sensitivity of the marginal value of inventories in response to an increase in uncertainty. The shift of g_1 reflects the fact that following an increase in uncertainty each unit of inventory has greater value as insurance against supply shortfalls. In other words, the oil futures spread will decline if agents' willingness to pay for an extra barrel of oil to be used as insurance against oil supply shortfalls increases sufficiently in response to an unanticipated shift in uncertainty. It is well-documented that during past uncertainty shocks in the crude oil market, traders were willing to pay exorbitant prices to procure extra stocks of oil (see, for example, Penrose 1976; Terzian 1985). Thus, large values of g_{12} seem empirically plausible. Uncertainty shocks driven by exogenous events provide an economic

explanation for the large and persistent fluctuations in the spread that undermine the forecasting accuracy of oil futures prices.¹⁰

2.6. Model Evaluation

2.6.1. Test 1: Can the Model Explain the Poor Forecast Accuracy of Oil Futures Prices?

The theoretical model predicts that under risk neutrality $F_t^{(h)} \approx E_t[S_{t+h}]$, which is approximately the result asserted by Bernanke (2004). Given this result, it may seem puzzling at first that the forecast accuracy of oil futures prices is poor in practice. This result follows naturally from the model, however. Sections 2 and 3 established that the best proxy for $E_t[S_{t+h}]$ is the no-change forecast S_t . There is no presumption in the theoretical model that $E_t[S_{t+h}] = S_t$. In fact, equation (14) implies that in equilibrium $F_t^{(h)}$ may be larger than, smaller than or equal to S_t . Thus, the evidence in Table 7 that on average over my sample period $F_t^{(h)}$ is slightly smaller than S_t is fully consistent with the theoretical model. Nevertheless, on average the (approximate) model expectation $F_t^{(h)}$ is fairly close to the econometric expectation S_t .

The reason that $F_t^{(h)}$ is a poor predictor is not so much that it is different on average from S_t , but that it fluctuates widely relative to S_t . At any point in time, the discrepancy between the futures price and the spot price may be very large and go in either direction. Taking the spot price of crude oil to be \$65, about its level in late March 2007, for example, the minimum and maximum value of the 12-month spread implies that the futures price may differ from the best predictor by as much as \$20 in one

¹⁰ Earlier I documented that the oil futures spread is highly persistent, but mean reverting (see Table 7). I also documented that the no-change forecast is the best predictor of the nominal spot price of oil. The conclusion that under plausible conditions the mean-reverting spread is associated with changes in the precautionary demand component of the spot price may seem to contradict the random walk result. This is not the case. First, the result about the forecast accuracy refers to the nominal price of oil, whereas the comparative statics result is for the real price of oil. Second, the forecasting results are for total spot price of oil, whereas the results of this section are only for one of the components of the real price of oil. Third, as Diebold and Kilian (2000) demonstrate, for autoregressive processes with degrees of persistence in the range documented in Table 7 an incorrectly specified random walk model will tend to have lower MSPE than the correct mean-reverting model in small samples.

direction or by as much as \$18 in the other (see Table 7). Thus, policymakers relying on oil futures prices are likely to overestimate or underestimate the expected price of oil substantially at any given point in time, and the fact that these mistakes largely average out in the long run is of little consolation. Put differently, it is not that Bernanke's (2004) assertion that oil futures prices can be viewed as expected spot prices is necessarily wrong, but that it is of limited use in practice given the large fluctuations in the futures price relative to the best predictor. My theoretical model provides an explanation of this excess variability. In the model, fluctuations in the spread arise naturally from shifts in uncertainty about future oil supply shortfalls and will be indicative of fluctuations in the spot price of oil driven by precautionary demand for crude oil, provided g_{12} is large enough.¹¹ Hence, the theoretical model helps us understand the poor forecast accuracy of oil futures prices. Whether this explanation is empirically plausible is a question that I turn to next.

2.6.2. Test 2: Does the Proposed Indicator Move as Expected During Known Episodes of Uncertainty Shifts?

One way of judging the empirical content of the model is to verify that the spread moves in the expected direction at times of major unforeseen events such as the outbreak of the wars. In Figure 5, I focus on several clearly defined events in recent history that should have been associated with shifts in the market's uncertainty about future oil supply shortfalls such as the Persian Gulf War and the 2003 Iraq War (which should have caused the spread to fall) and the Asian financial crisis and 9/11 (which should have caused the spread to increase as world demand for crude oil fell, making a shortfall less likely). Clearly, expectations shifts of the type embodied in my theoretical model are not the only possible reason for shifts in the spread, but arguably they are the most important reason.

Figure 5 plots the negative of the spread for 1989.1-2007.2 by horizon. This normalization allows us to interpret positive spikes as increases in the precautionary

¹¹ Strictly speaking, this link holds if and only if a change in demand for oil inventories is confronted with an inelastic supply of oil. In the model, this inelasticity is represented in the form of an endowment structure. While this assumption may be unrealistic for the early 1980s, throughout much of the sample that I consider below this is a reasonable assumption. Kilian (2008) documents that capacity constraints in world crude oil production have been binding since the early 1990s.

demand component of the real spot price. The plot confirms the conclusion in Kilian (2008) that the sharp spike in oil prices during the Persian Gulf War was driven by expectations shifts reflected first in higher precautionary demand, as Iraq invaded Kuwait, and then in lower precautionary demand, as the U.S. troop presence in the region increased (also see Kilian 2007a). Likewise, the spike after mid-2002 in the period leading up to the 2003 Iraq War is as expected, given that the Iraq War was anticipated by the market starting in the summer of 2002 (Barsky and Kilian 2004). The plot also indicates that the temporary decline in oil prices following the Asian crisis (and its reversal after 1999) reflected fluctuations in precautionary demand. There is a similar but smaller temporary decline following the adverse demand shock associated with 9/11. Anecdotal evidence suggests that the spike in 1996 was associated with concerns about tight oil supplies and the spike in 2000 with concerns arising from strong demand for crude oil. In addition, the plot suggests a persistent decline in precautionary demand in recent years. Such a decline seems highly implausible on a priori grounds, given that recent years have been characterized by widespread concerns about future oil supply shortfalls, a point to which I will return below.

2.6.3. Test 3: How does the Proposed Indicator Compare to Alternative Measures of Precautionary Demand Shifts?

The indicator of expectations-driven oil price increases proposed in this paper is not the only possible measure. Recently, an alternative measure of the component of the spot price of crude oil that is driven by shocks to precautionary demand has been proposed by Kilian (2007a,b) based on different data and a different methodology. Unlike the measure developed in this paper, that estimate was based on a structural VAR decomposition of the real price of crude oil. The structural representation of the underlying trivariate autoregressive model is

$$A_0 z_t = \alpha + \sum_{i=1}^p A_i z_{t-i} + \varepsilon_t ,$$

Where p denotes the lag order, ε_t is the vector of serially and mutually uncorrelated structural innovations and z_t a vector variable including the percent change in global crude oil production, a (suitably detrended) index of global real economic activity that

captures fluctuations in the global demand for all industrial commodities (including crude oil), and the real price of oil (in that order), measured at monthly frequency.

Let e_t denote the reduced form VAR innovations such that $e_t = A_0^{-1}\varepsilon_t$. The structural innovations are derived from the reduced form innovations by imposing exclusion restrictions on A_0^{-1} . The identifying assumptions are that (1) crude oil supply will not respond to oil demand shocks within the month, given the costs of adjusting oil production and the uncertainty about the state of the crude oil market; that (2) increases in the real price of oil driven by shocks that are specific to the oil market will not lower global real economic activity within the month. In this model, innovations to the real price of oil that cannot be explained by oil supply shocks or demand shocks that are common to all industrial commodities by construction must be demand shocks that are specific to the oil market. The latter oil-specific demand shock by construction captures fluctuations in precautionary demand for oil driven by fears about the availability of future oil supplies. Kilian (2007a) makes the case that this shock effectively can be interpreted as a precautionary demand shock, given the absence of plausible alternative interpretations and given the time path of this shock during specific historical episodes, during which I would expect precautionary demand to shift.

The structural VAR model postulates a vertical short-run supply curve for crude oil and a downward sloping short-run demand curve that is being shifted by innovations to the business cycle in global industrial commodity markets as well as shifts in the demand for oil that are specific to the oil market such as shifts in the precautionary demand for crude oil. Given these assumptions, one can use the structural moving average decomposition of the VAR model to construct a time series of the component of the real price of oil that can be attributed to shifts in the precautionary demand for crude oil in response to changes in the uncertainty about future oil supply shortfalls. While it is not possible to compare this VAR-based measure of the precautionary demand component of the spot price to the futures-based measure for the full sample period of 1973-2006 considered in Kilian (2007a), given the limited availability of oil futures price data, I may compare these two measures for the period 1989.1-2006.12, which includes several major oil price spikes. Since the oil futures-based measure is essentially an index

and the VAR-based measure is not, the appropriate metric of comparison is their contemporaneous correlation.

Table 8 shows that the two measures in general are highly correlated despite the differences in their method of construction. For the sample period of 1989.1 through 2006.12, the correlation ranges from 39% at the 3-month horizon to 61% at the 12-month horizon. The fit improves monotonically with the horizon, consistent with the view that shifts in precautionary demand are primarily concerned with expectations beyond the short run. Thus, I focus on the 12-month spread. A correlation of 61% between two independently constructed measures of the fluctuations in the spot price of oil driven by precautionary demand is remarkably high. The correlation is even higher if I exclude the last three years of data, for which the spread seems implausibly high, as discussed above. Table 8 shows that, excluding the last three years, the correlation of the two measures rises to 79% at the 12 month horizon. A correlation of near 80% for most of the sample is evidence both of the predictive power of my theoretical model of the oil futures and spot markets and of the realism of the identifying assumptions underlying the VAR-based measure.

Not only does the correlation weaken after 2003.12, but the spread data and the VAR-based measure of the precautionary demand component of the spot price of oil paint a somewhat different picture (see Figure 6). Whereas the VAR-based measure on average remains at a high level after 2003.12, consistent with the perception of sustained uncertainty about future oil supply shortfalls, the futures-based measure systematically declines. This evidence casts further doubt on the credibility of the negative of the spread as an indicator of fluctuations in the precautionary demand component of the spot price over the last three years of the sample. These observations suggest that a structural change may have occurred around 2003.12 that is beyond the scope of the theoretical model in section 5. Indeed, it has been suggested in the financial press that the nature of the oil futures markets has changed in recent years, as hedge funds and other investors with no ties to the oil industry attempted to capitalize on rising oil prices. Data from the Commodity Futures Trading Commission (not shown to conserve space) shed light on the share of speculators among oil futures traders since 1989 and reveal an unprecedented increase in speculative activities after 2003.12. To the extent that increased speculative

trading tends to raise the price of oil futures more than the spot price (and hence increases the spread), this fact might provide an explanation for the weakening of the correlations at the end of the sample. Establishing such a link is left for future research.

2.6.4. Test 4: Does the VAR Response of the Real Price of Oil Match the Model Predictions?

Another implication of the theoretical model is that the real price of oil will overshoot in response to a mean-preserving spread, while inventory holdings will increase only gradually. If the Kilian (2007a) measure of the precautionary demand shock is valid, the response of the real price of oil in that VAR should exhibit overshooting, as predicted by the theoretical model. Figure 7 confirms that the response to an oil-specific demand shock indeed displays overshooting, suggesting that the interpretation of this shock as a precautionary demand shock is justified and indirectly supporting the interpretation of the futures-based indicator as a measure of expectations shifts. There is no evidence of such a pattern in response to other oil demand or oil supply shocks.

2.6.5. Implications for Crude Oil Inventories

Whereas I have focused on the empirical relationship between increased concerns about future oil supply shortfalls and the precautionary demand component of the real spot price of oil, the model also has implications for the behavior of inventories in response to increased uncertainty. Testing these implications is not straightforward, given that inventories move for many reasons other than shifts in uncertainty about future oil supply shortfalls. First, whereas for the real price of oil I was able to use a VAR decomposition to focus specifically on the precautionary demand component of the real price, no similar measure of the precautionary demand component in oil inventories exists, making it impossible to identify the consequences of precautionary demand shocks for inventories. Second, inventory data are trending, and measures of the comovement between the precautionary demand component of the spot price and inventories tend to be sensitive to the method of detrending.

There is, however, anecdotal evidence from oil industry experts documenting that shifts in precautionary demand coincide with a strong motive for inventory accumulation.

This situation has been aptly described by Terzian (1985) in the context of the 1979 oil price shock:

Spot deals became more and more infrequent. The independent refineries, with no access to direct supply from producers, began to look desperately for oil on the so-called 'free market'. But from the beginning of November, most of the big oil companies invoked *force majeure* and reduced their oil deliveries to third parties by 10% to 30%, when they did not cut them off altogether. Everybody was anxious to hang on to as much of their own oil as possible, until the situation had become clearer. The shortage was purely psychological, or 'precautionary' as one dealer put it. (p. 260)

Penrose (1976, p. 46) describes a similar hoarding phenomenon in the period leading up to the 1973 oil price shock, as oil companies became concerned with the possibility of being expropriated. In her words, "the major oil companies became increasingly cautious about outside sales as uncertainty increased". These accounts are consistent with the implications of my theoretical model.

2.7. Conclusion

I introduced a two-country, multi-period general equilibrium model of both the spot market and the futures market for crude oil to provide fresh insights about the interpretation of oil futures prices and related statistics such as the oil futures spread. The key insights can be summarized as follows: First, it is widely believed that prices of oil futures are accurate predictors of forecast spot prices in the MSPE sense. Using observations up to February of 2007, I showed that the price of crude oil futures is not an accurate predictor of the spot price of crude oil. Many users of oil futures-based forecasts are aware of this caveat and understand that futures-based forecasts may be poor, but still believe that they provide the *best* available forecast of spot prices of crude oil. I showed this not to be the case. Futures-based forecasts are inferior to simple and easy-to-use forecasting methods such as the no-change forecast. No-change forecasts are also more accurate than commercial survey-based forecasts.

Second, I showed that the large MSPE of oil futures-based forecasts is driven not by the bias, but by the variability of the futures price about the spot price. I documented large and time-varying deviations of oil futures prices from the spot price of oil, as measured by the oil futures spread. For example, given a spot price of \$65, the 12-month

futures price may deviate as much as \$20 from the expected spot price or as little as \$0, which helps explain the poor forecasting accuracy of oil futures prices.

Third, my analysis demonstrates that fluctuations in the oil futures spread are larger and more persistent than fluctuations in the spread of foreign exchange futures. I showed that this anomaly is linked to the presence of a marginal convenience yield in the oil futures market that is absent in the foreign exchange futures market. I proposed a theoretical model of the oil spot market and oil futures market that incorporates this marginal convenience yield. The model implies that the oil futures spread is directly linked to shifts in oil market fundamentals. I showed that shifts in the uncertainty about future oil supply shortfalls cause fluctuations in the oil futures spread not found in models of the foreign exchange futures market. My model explains the excess variability of oil futures prices relative to the no-change forecast and the resulting poor forecast accuracy of oil futures prices.

Fourth, I showed that, under plausible conditions, the oil futures spread will decline, as the precautionary demand component of the real spot price of crude oil increases. Thus, the negative of the spread may be viewed as an indicator of fluctuations in the real price of crude oil driven by precautionary demand for oil. The time path of the oil futures spread since 1989 suggested major shifts in precautionary demand for oil during the Persian Gulf War and following the Asian crisis, for example. These results provided independent evidence of how shifts in market expectations about future oil supply shortfalls affect the spot price of crude oil. Such expectation shifts have been difficult to quantify, yet play an important role in explaining oil price fluctuations (Kilian 2008). In addition, I documented that my measure of oil price movements driven by uncertainty shifts matches up well with independently obtained VAR based measures, and that my model predicts the overshooting of the real price of oil found in VAR analysis. My analysis also is consistent with anecdotal evidence of hoarding in oil inventory markets during times of crisis.

Notes: The forecast evaluation period is 1991.1-2007.2. The initial estimation window is 1983.4-1990.12. For regressions based on 6-month futures prices the estimation window begins in 1983.10; for the 9-month futures price in 1986.12; for the 12-month futures price in 1989.1. $F_t^{(h)}$ is the futures price that matures in h periods; $i_{t,m}$ is the m month interest rate; and $\Delta \bar{s}_t^{(l)}$ denotes the trailing geometric average of the monthly percent change for l months. p -values are in parentheses. All p -values refer to pairwise tests of the null of a random walk without drift. Comparisons of nonnested models without estimated parameters are based on the *DM*-test of Diebold and Mariano (2005) with $N(0,1)$ critical values. Nested model comparisons with estimated parameters are based on Clark and West (2006). For the rolling regression estimates of the random walk with drift I use $N(0,1)$ critical values under quadratic loss; for recursive estimates under quadratic loss and for all estimates under absolute loss I use bootstrap critical values as described in Clark and West. The sign test in the last column is based on Pesaran and Timmermann (1992).

Table 2-1: 1-Month Ahead Recursive Forecast Error Diagnostics

$\hat{S}_{t+1 t}$	MSPE (<i>p</i> -value)	Bias	MAPE (<i>p</i> -value)	Success Ratio (<i>p</i> -value)
S_t	6.998	0.172	1.941	N.A.
$F_t^{(1)}$	7.106 (0.809)	0.210	1.949 (0.770)	0.443 (0.898)
$S_t(1 + \hat{\alpha} + \hat{\beta} \ln(F_t^{(1)} / S_t))$	6.994 (0.175)	0.200	1.954 (0.580)	0.479 (0.529)
$S_t(1 + \hat{\beta} \ln(F_t^{(1)} / S_t))$	6.975 (0.104)	0.156	1.948 (0.462)	0.423 (0.984)
$S_t(1 + \hat{\alpha} + \ln(F_t^{(1)} / S_t))$	7.138 (0.799)	0.162	1.948 (0.439)	0.526 (0.257)
$S_t(1 + \ln(F_t^{(1)} / S_t))$	7.106 (0.807)	0.212	1.949 (0.676)	0.443 (0.898)
$S_t(1 + \hat{\alpha})$	7.013 (0.384)	0.186	1.945 (0.522)	0.479 (0.497)
$S_t(1 + \Delta \bar{s}_t^{(3)})$	10.044 (0.717)	0.015	2.235 (0.151)	0.521 (0.294)
$S_t(1 + \Delta \bar{s}_t^{(6)})$	8.293 (0.835)	0.005	2.050 (0.087)	0.495 (0.567)
$S_t(1 + \Delta \bar{s}_t^{(9)})$	8.155 (0.932)	-0.016	2.057 (0.806)	0.495 (0.567)
$S_t(1 + \Delta \bar{s}_t^{(12)})$	7.405 (0.305)	-0.023	1.943 (0.521)	0.505 (0.443)

Table 2-2: 3-Month Ahead Recursive Forecast Error Diagnostics

$\hat{S}_{t+3 t}$	MSPE (<i>p</i> -value)	Bias	MAPE (<i>p</i> -value)	Success Ratio (<i>p</i> -value)
S_t	19.560	0.435	3.099	N.A.
$F_t^{(3)}$	19.038 (0.347)	0.631	3.172 (0.920)	0.479 (0.648)
$S_t(1 + \hat{\alpha} + \hat{\beta} \ln(F_t^{(3)} / S_t))$	24.217 (0.870)	0.253	3.610 (0.990)	0.407 (0.996)
$S_t(1 + \hat{\beta} \ln(F_t^{(3)} / S_t))$	22.826 (0.983)	0.804	3.541 (0.998)	0.407 (0.992)
$S_t(1 + \hat{\alpha} + \ln(F_t^{(3)} / S_t))$	22.090 (0.747)	0.315	3.365 (0.965)	0.397 (0.998)
$S_t(1 + \ln(F_t^{(3)} / S_t))$	19.036 (0.348)	0.649	3.176 (0.920)	0.479 (0.648)
$S_t(1 + i_{t,3})$	19.811 (0.715)	0.167	3.111 (0.632)	0.541 N.A.
$S_t(1 + \hat{\alpha})$	19.699 (0.351)	0.484	3.114 (0.345)	0.485 (0.413)
$S_t(1 + \Delta \bar{s}_t^{(3)})$	24.702 (0.961)	0.238	3.461 (0.707)	0.500 (0.524)
$S_t(1 + \Delta \bar{s}_t^{(6)})$	22.098 (0.893)	0.213	3.231 (0.315)	0.485 (0.685)
$S_t(1 + \Delta \bar{s}_t^{(9)})$	20.242 (0.531)	0.224	3.105 (0.023)	0.557 (0.061)
$S_t(1 + \Delta \bar{s}_t^{(12)})$	20.013 (0.454)	0.223	3.071 (0.005)	0.546 (0.101)
$S_{t,3}^{CF}$	30.726 (0.997)	-1.905	4.148 (0.999)	0.500 (0.338)

Notes: See Table 2-1.

Table 2-3: 6-Month Ahead Recursive Forecast Error Diagnostics

\hat{S}_{t+6t}	MSPE (<i>p</i> -value)	Bias	MAPE (<i>p</i> -value)	Success Ratio (<i>p</i> -value)
S_t	34.058	0.937	4.466	N.A.
$F_t^{(6)}$	36.334 (0.716)	1.615	4.608 (0.906)	0.485 (0.483)
$S_t(1 + \hat{\alpha} + \hat{\beta} \ln(F_t^{(6)} / S_t))$	51.809 (0.738)	1.012	5.315 (0.794)	0.485 (0.613)
$S_t(1 + \hat{\beta} \ln(F_t^{(6)} / S_t))$	47.143 (0.917)	1.959	5.200 (0.904)	0.464 (0.703)
$S_t(1 + \hat{\alpha} + \ln(F_t^{(6)} / S_t))$	40.640 (0.710)	1.074	4.692 (0.528)	0.485 (0.576)
$S_t(1 + \ln(F_t^{(6)} / S_t))$	36.475 (0.721)	1.684	4.621 (0.910)	0.485 (0.483)
$S_t(1 + i_{t,6})$	34.906 (0.713)	0.382	4.509 (0.708)	0.557 N.A.
$S_t(1 + \hat{\alpha})$	33.942 (0.132)	1.093	4.678 (0.155)	0.515 (0.021)
$S_t(1 + \Delta \bar{s}_t^{(3)})$	41.100 (0.874)	0.605	4.738 (0.571)	0.479 (0.762)
$S_t(1 + \Delta \bar{s}_t^{(6)})$	35.936 (0.691)	0.671	4.531 (0.170)	0.510 (0.424)
$S_t(1 + \Delta \bar{s}_t^{(9)})$	33.812 (0.293)	0.585	4.372 (0.988)	0.557 (0.091)
$S_t(1 + \Delta \bar{s}_t^{(12)})$	34.379 (0.437)	0.708	4.465 (0.085)	0.510 (0.411)

Notes: See Table 2-1.

Table 2-4: 9-Month Ahead Recursive Forecast Error Diagnostics

$\hat{S}_{t+9 t}$	MSPE (<i>p</i> -value)	Bias	MAPE (<i>p</i> -value)	Success Ratio (<i>p</i> -value)
S_t	46.574	1.791	5.161	N.A.
$F_t^{(9)}$	53.798 (0.887)	2.892	5.370 (0.926)	0.526 (0.080)
$S_t(1 + \hat{\alpha} + \hat{\beta} \ln(F_t^{(9)} / S_t))$	54.225 (0.471)	2.515	5.406 (0.296)	0.546 (0.035)
$S_t(1 + \hat{\beta} \ln(F_t^{(9)} / S_t))$	54.939 (0.632)	3.163	5.411 (0.452)	0.536 (0.026)
$S_t(1 + \hat{\alpha} + \ln(F_t^{(9)} / S_t))$	55.042 (0.725)	2.502	5.313 (0.361)	0.546 (0.025)
$S_t(1 + \ln(F_t^{(9)} / S_t))$	54.642 (0.898)	3.017	5.403 (0.948)	0.526 (0.080)
$S_t(1 + \hat{\alpha})$	46.107 (0.111)	2.090	5.150 (0.130)	0.557 (0.000)
$S_t(1 + \Delta \bar{s}_t^{(3)})$	51.025 (0.658)	1.492	5.258 (0.245)	0.510 (0.431)
$S_t(1 + \Delta \bar{s}_t^{(6)})$	46.300 (0.303)	1.556	5.116 (0.092)	0.595 (0.581)
$S_t(1 + \Delta \bar{s}_t^{(9)})$	45.428 (0.168)	1.581	5.082 (0.048)	0.510 (0.401)
$S_t(1 + \Delta \bar{s}_t^{(12)})$	46.229 (0.315)	1.578	5.139 (0.109)	0.500 (0.516)

Notes: See Table 2-1.

Table 2-5: 12-Month Ahead Recursive Forecast Error Diagnostics

$\hat{S}_{t+12 t}$	MSPE (<i>p</i> -value)	Bias	MAPE (<i>p</i> -value)	Success Ratio (<i>p</i> -value)
S_t	65.978	2.540	5.885	N.A.
$F_t^{(12)}$	77.204 (0.898)	4.009	6.212 (0.767)	0.536 (0.021)
$S_t(1 + \hat{\alpha} + \hat{\beta} \ln(F_t^{(12)} / S_t))$	78.414 (0.523)	3.874	6.272 (0.362)	0.526 (0.032)
$S_t(1 + \hat{\beta} \ln(F_t^{(12)} / S_t))$	84.275 (0.768)	4.352	6.411 (0.623)	0.541 (0.004)
$S_t(1 + \hat{\alpha} + \ln(F_t^{(12)} / S_t))$	76.682 (0.710)	3.839	6.138 (0.427)	0.515 (0.028)
$S_t(1 + \ln(F_t^{(12)} / S_t))$	79.007 (0.916)	4.189	6.279 (0.789)	0.536 (0.021)
$S_t(1 + i_{t,12})$	65.285 (0.480)	1.439	6.018 (0.804)	0.582 N.A.
$S_t(1 + \hat{\alpha})$	64.709 (0.108)	3.200	5.968 (0.269)	0.552 (0.001)
$S_t(1 + \Delta \bar{s}_t^{(3)})$	68.673 (0.484)	2.268	6.056 (0.478)	0.490 (0.668)
$S_t(1 + \Delta \bar{s}_t^{(6)})$	65.632 (0.314)	2.321	5.964 (0.355)	0.438 (0.966)
$S_t(1 + \Delta \bar{s}_t^{(9)})$	64.931 (0.234)	2.340	5.929 (0.274)	0.469 (0.816)
$S_t(1 + \Delta \bar{s}_t^{(12)})$	64.986 (0.238)	2.346	5.906 (0.199)	0.479 (0.728)
$S_{t,12}^{CF}$	107.866 (0.979)	-4.808	6.957 (0.954)	0.515 (0.122)

Notes: See Table 2-1.

Table 2-6: Asymptotic p -Values for Forecast Efficiency Regressions

Horizon	$\hat{\alpha}$	$\hat{\beta}$	$H_0: \alpha=0$	$H_0: \beta=1$	$H_0: \alpha = 0, \beta = 1$
3-month	0.029	1.160	0.063	0.398	0.247
6-month	0.057	0.766	0.037	0.685	0.037
12-month	0.111	0.731	0.008	0.777	0.004

Notes: For the 3- and 6-month regressions, the sample period is 1989.4-2007.2. For the 12-month regression, the sample is 1990.1-2007.2. All t -and $Wald$ -tests have been computed based on HAC standard errors.

**Table 2-7: Time Series Features of the Spread
(Percent)**

	3 Month	12 Month
Mean	-1.12	-4.88
(<i>p</i> -value)	(0.00)	(0.00)
Mean Abs. Deviation	2.72	8.89
Max	12.3	30.1
Min	-10.1	-27.7
Persistence	0.74	0.81

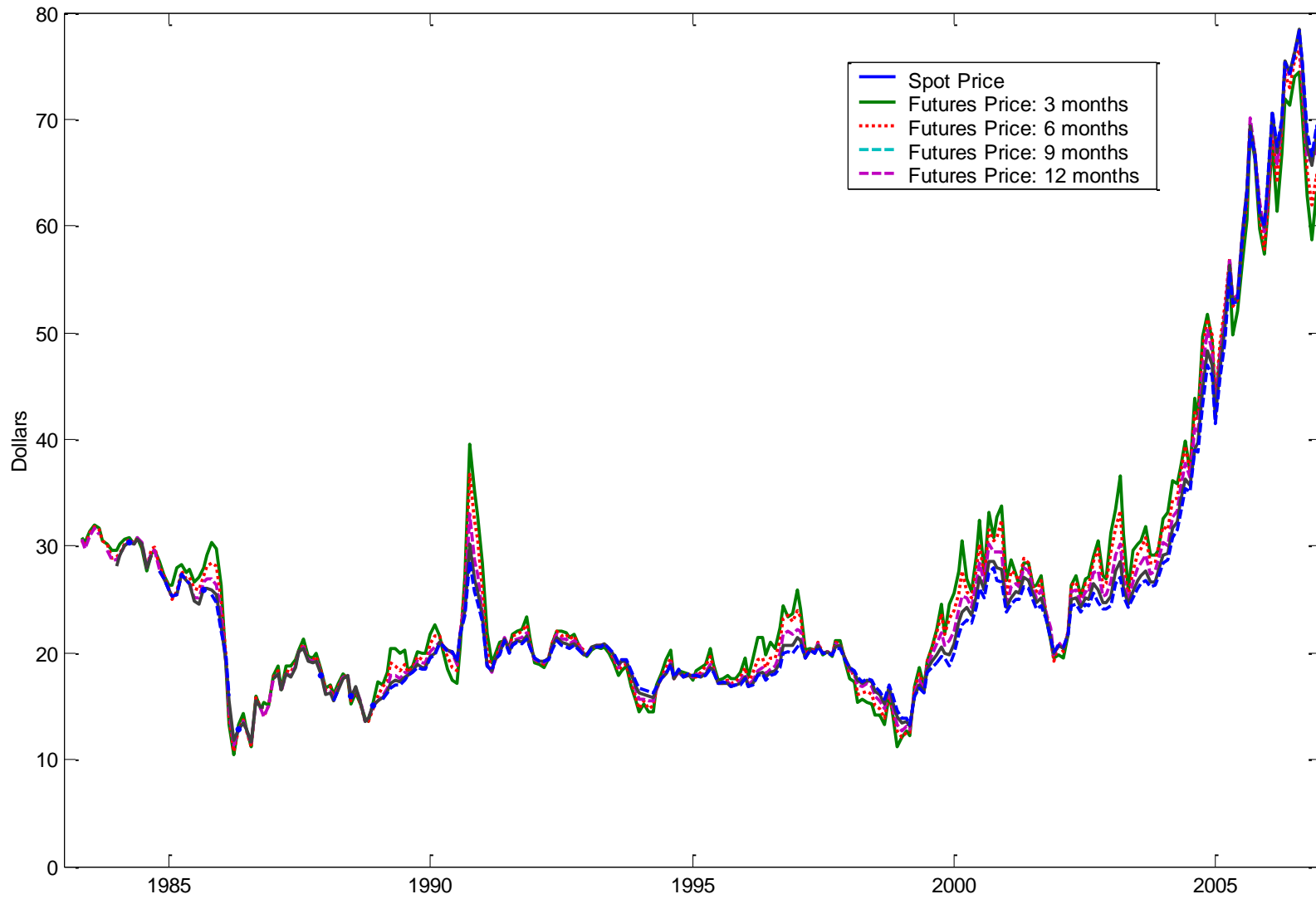
Notes: The sample for the 3-month forecasts is 1983.4-2007.2; and that for the 12-month forecast is 1990.1-2007.2, reflecting the data constraints. The *p*-values of the test for a zero mean are based on HAC standard errors. The measure of persistence is the sum of the autoregressive coefficients proposed by Andrews and Chen (1994). The autoregressive lag order is determined using the AIC with an upper bound of 24 lags.

Table 2-8: Contemporaneous Correlation of the Negative Spread and the VAR Estimate of the Precautionary Demand Component of Real Spot Price of Crude Oil
(Percent)

Horizon	1989.1-2006.12	1989.1-2003.12
3	39.1	57.9
6	49.7	69.9
9	56.4	75.8
12	61.4	79.4

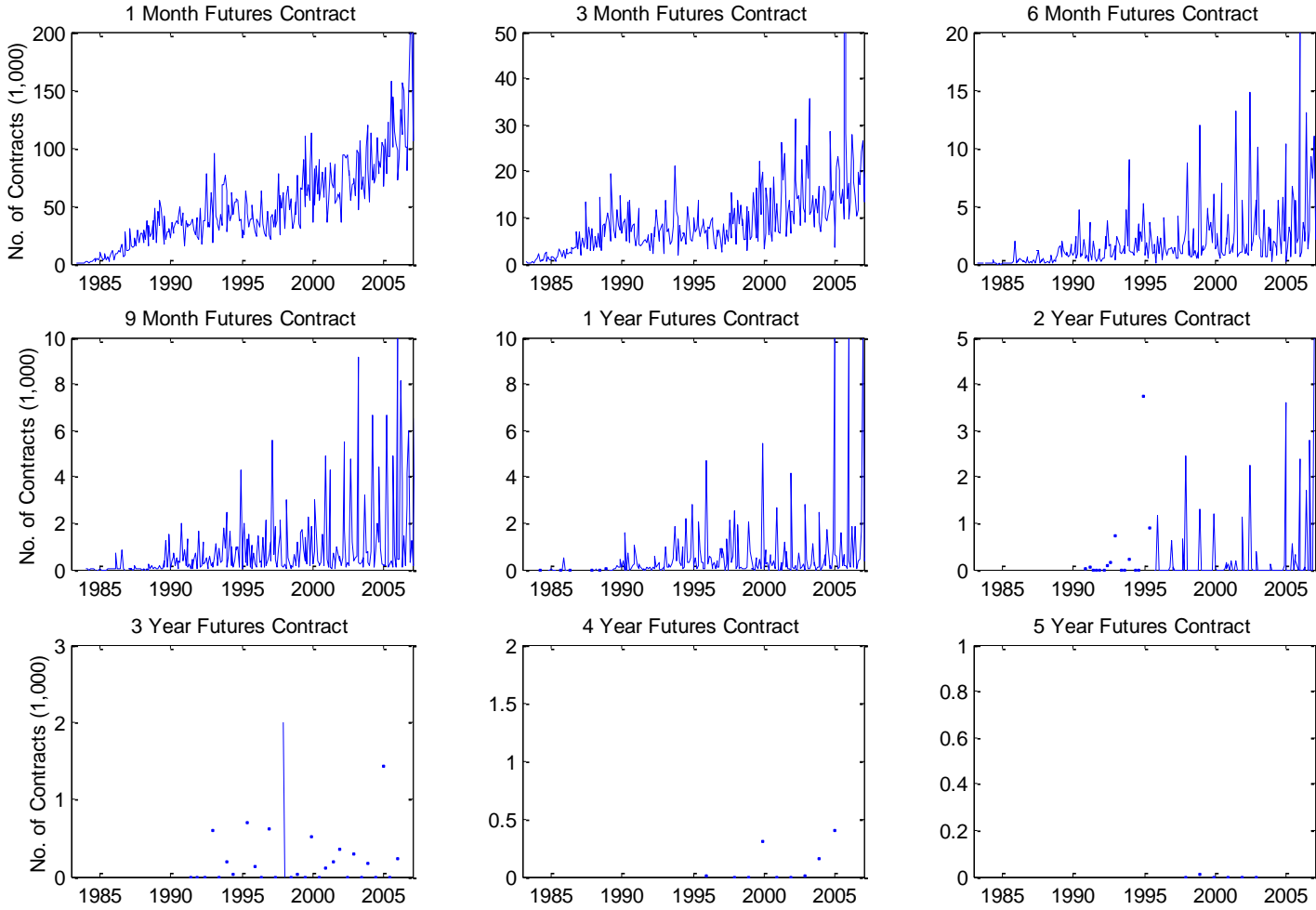
Notes: Computed based on Figure 5 and the VAR estimates of the precautionary demand component of the spot price of crude oil in Kilian (2007).

Figure 2-1: Prices of Oil Futures Contracts and Spot Price of Oil, 1983.3-2007.2



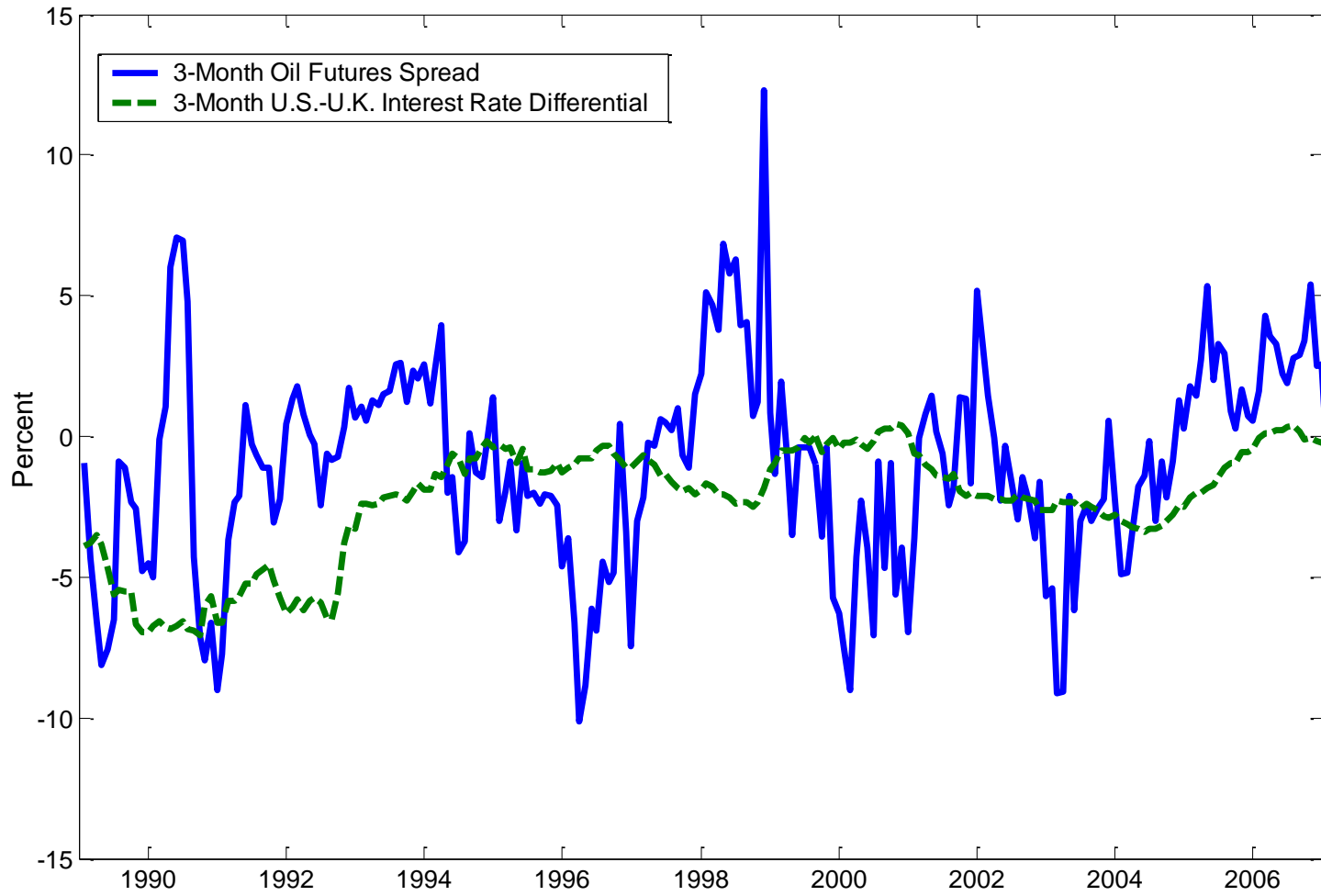
Source: Computed as described in the text based on daily NYMEX oil futures prices and the daily WTI spot price

Figure 2-2: Volume of NYMEX Oil Futures Contracts



Source: *Price-data.com*

Figure 2-3: Oil Futures Spread and Foreign Exchange Futures Spread
3-Month Horizon



Notes: The interest rates are end-of-month Treasury bill rates from the Bank of England and the Federal Reserve Board

Figure 2-4a: The Effect of an Increase in Uncertainty on the Marginal Convenience Yield

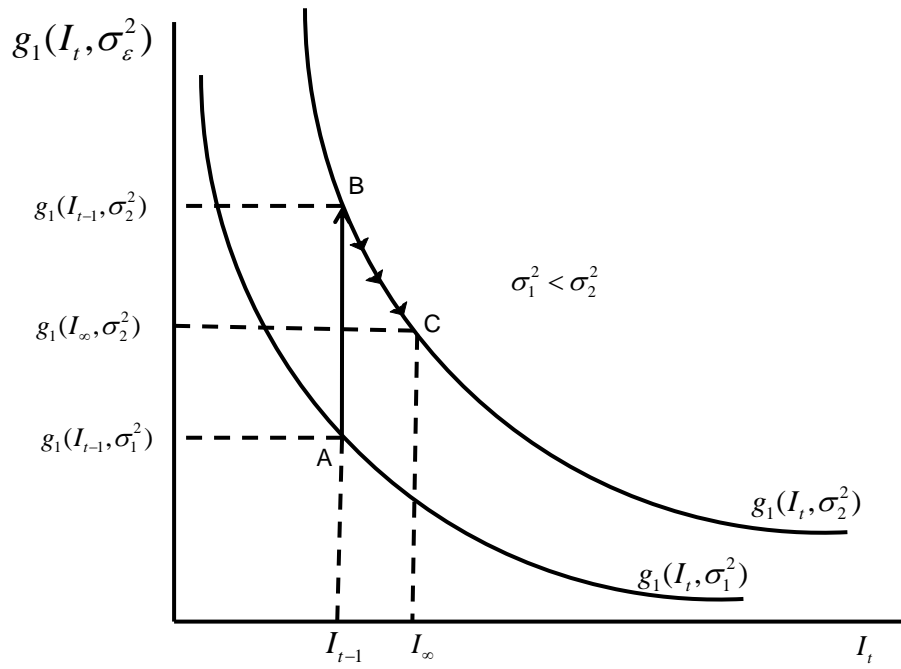


Figure 2-4b: The Effect of an Increase in Uncertainty on the Demand for Oil

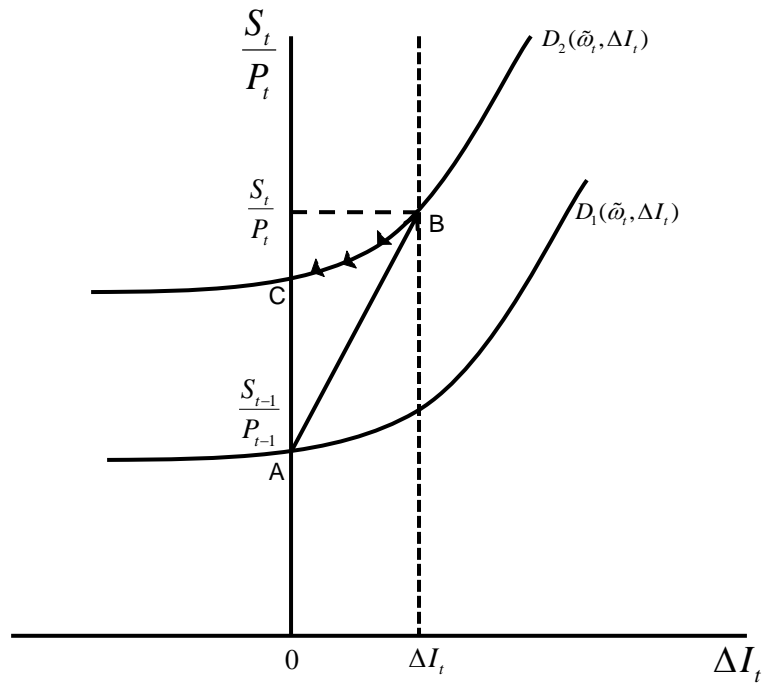


Figure 2-5: Negative Spread by Horizon, 1989.1-2007.2

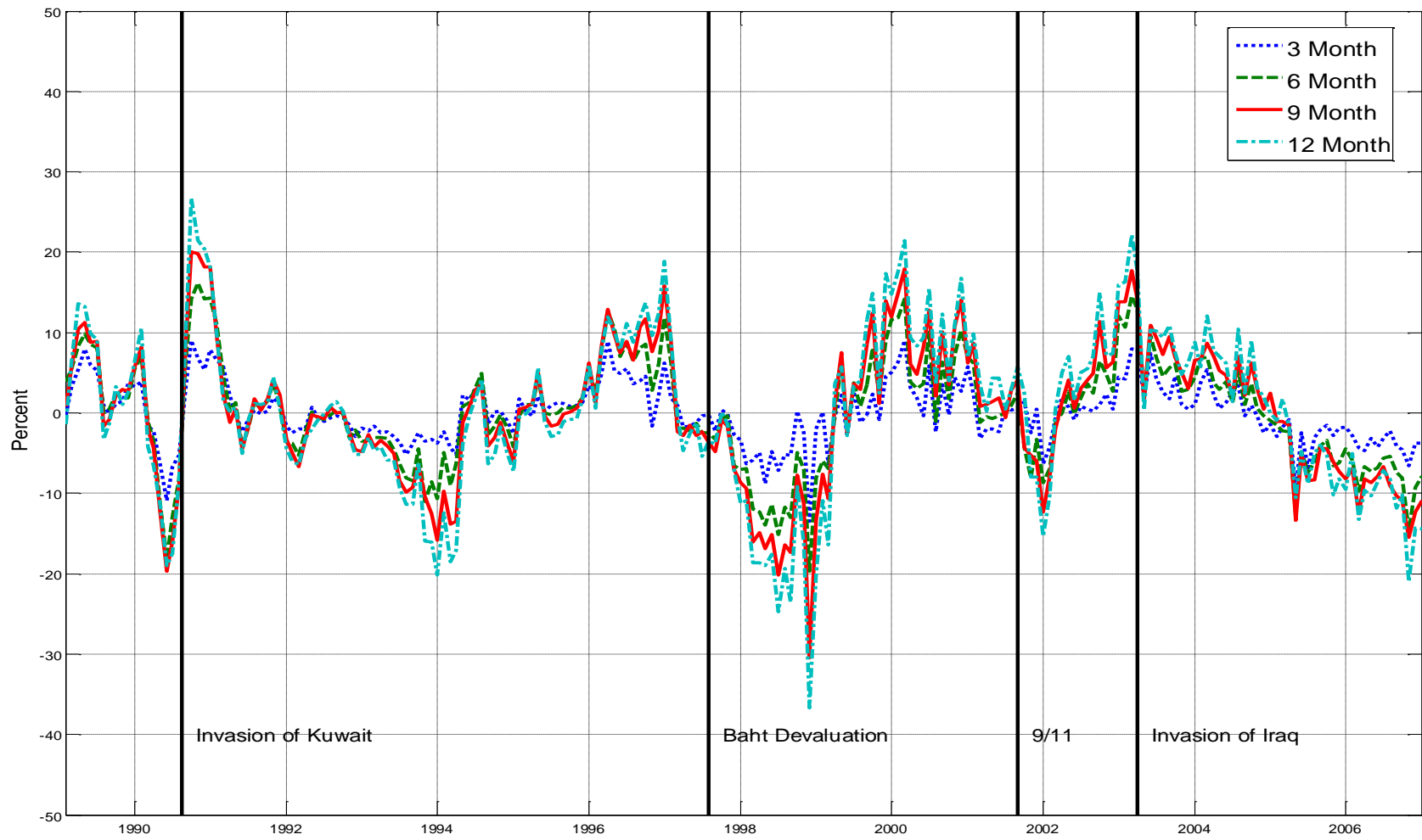
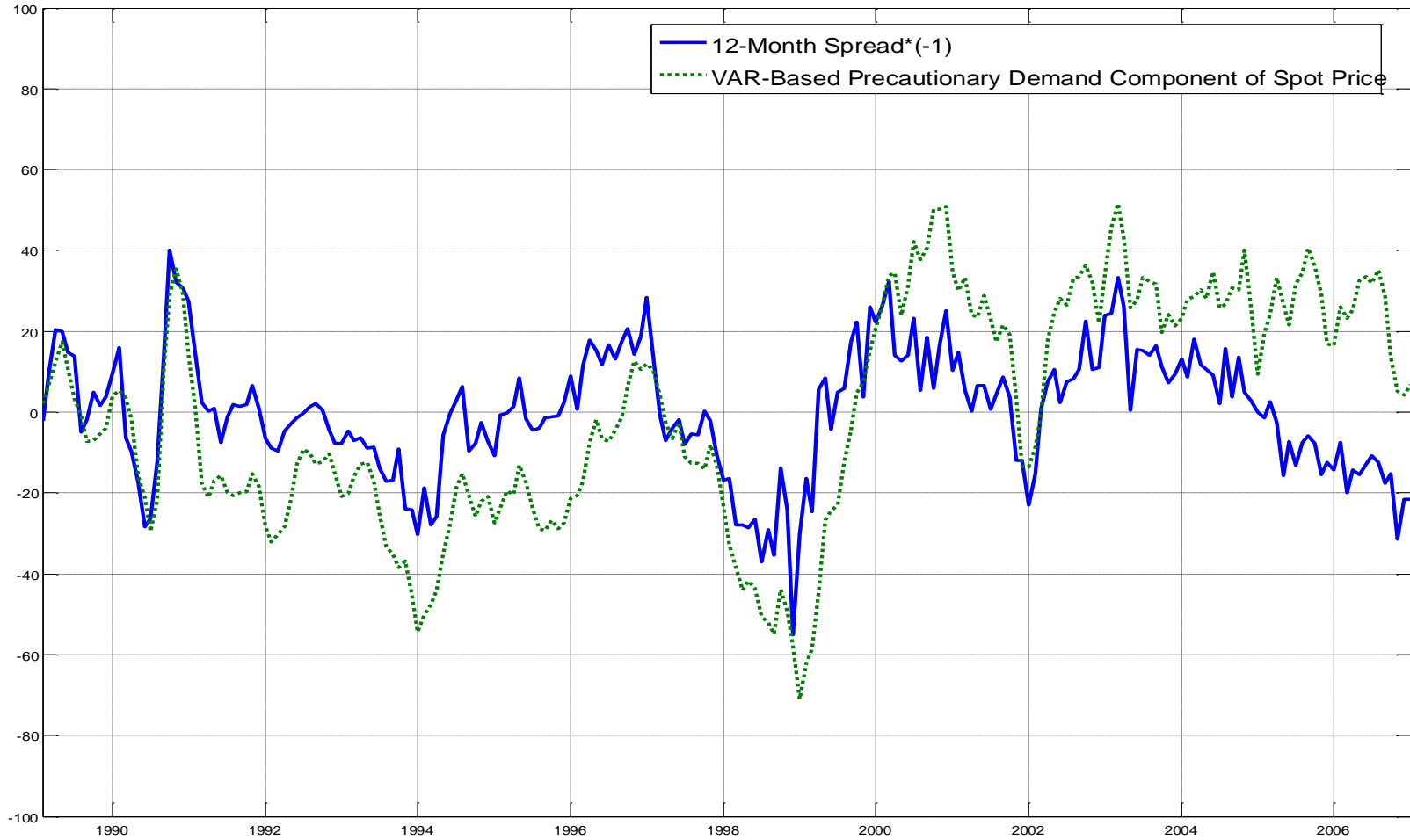


Figure 2-6: Negative Spread and VAR-based Estimate of Precautionary Demand Component of Spot Price, 1989.1-2006.12

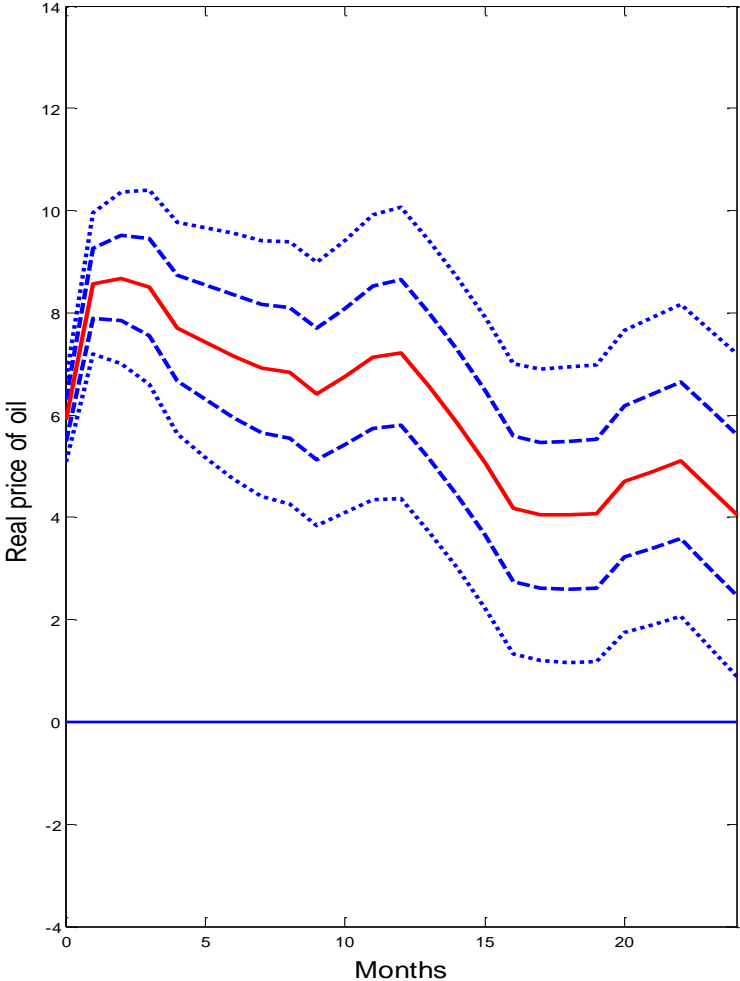


49

Notes: The spread has been scaled by -1.5 to improve the readability of the graph. Since the spread is essentially an index that transformation does not involve any loss of generality.

Figure 2-7: Response of the Real Price of Oil to a Positive Precautionary Demand Shock

Point Estimate with One- and Two-Standard Error Bands



Source: Kilian (2007a). The sample period is 1973.2-2006.12. The model is described in the text.

References

- Andrews, D.W.K. and H.-Y. Chen (1994), "Approximately Median-Unbiased Estimation of Autoregressive Models," *Journal of Business and Economic Statistics*, 12(2), 187-204.
- Bailey, W. and K.C. Chan (1993), "Macroeconomic Influences and the Variability of the Commodity Futures Basis," *Journal of Finance*, 48(2), 555-573.
- Barsky, R.B., and L. Kilian (2002), "Do We Really Know that Oil Caused the Great Stagflation? A Monetary Alternative," in: *NBER Macroeconomics Annual 2001*, B.S. Bernanke and K. Rogoff (eds.), MIT Press: Cambridge, MA, 137-183.
- Barsky, R.B., and L. Kilian (2004), "Oil and the Macroeconomy Since the 1970s," *Journal of Economic Perspectives*, 18(4), 115-134.
- Bernanke, B.S (2004), "Oil and the Economy," speech presented at Darton College, Albany, GA,
<http://www.federalreserve.gov/boarddocs/speeches/2004/20041021/default.htm>
- Bernanke, B.S. (2006). "Energy and the Economy," remarks before the Economic Club of Chicago, Chicago, IL,
<http://www.federalreserve.gov/boarddocs/speeches/2006/200606152/default.htm>
- Brennan, M.J. (1991), "The Cost of Convenience and the Pricing of Commodity Contingent Claims," in D. Lund and B. Øksendal (eds.), *Stochastic Models and Option Values*, North Holland: Amsterdam.
- Chernenko, S.V., K.B. Schwarz, and J.H. Wright (2004), "The Information Content of Forward and Futures Prices," *International Finance Discussion Paper No. 808*.
- Chinn, M.D., M. LeBlanc, and O. Coibion (2005), "The Predictive Content of Energy Futures: An Update on Petroleum, Natural Gas, Heating Oil and Gasoline," NBER Working Paper No. 11033.
- Chinn, M.D. and J.A. Frankel (1995), "Patterns on Exchange Rates for 25 Currencies," *Journal of Money, Credit, and Banking*, 26, 759-770.
- Clark, T.E. and K.D. West (2006), "Using Out-of-Sample Mean Squared Prediction Errors to Test the Martingale Difference Hypothesis," *Journal of Econometrics*, 135(1-2), 155-186.
- Considine, T. J. (1997), "Inventories Under Joint Production: An Empirical Analysis of Petroleum Refining," *Review of Economics and Statistics*, 79(3), 493-502.

Davies, P. (2007), "What's the Value of an Energy Economist?" Speech presented at the International Association for Energy Economics, Wellington, New Zealand.

Deaton, A., and G. Laroque (1992), "On the Behavior of Commodity Prices," *Review of Economic Studies*, 59, 1-23.

Deaton, A., and G. Laroque (1996), "Competitive Storage and Commodity Price Dynamics," *Journal of Political Economy*, 104, 896-923.

Diebold, F.X., and L. Kilian (2000), "Unit Root Tests are Useful for Selecting Forecasting Models," *Journal of Business and Economic Statistics*, 18(3), 265-273.

Diebold, F.X. and R.S. Mariano (1995), "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13(3), 253-263.

Dornbusch, R. (1976), "Expectations and Exchange Rate Dynamics," *Journal of Political Economy*, 84, 1161-1176.

Elliott, G., and A. Timmermann (2007), "Economic Forecasting," forthcoming: *Journal of Economic Literature*.

Elliott, G., I. Komunjer, and A. Timmermann, 2005, "Estimation and Testing of Forecast Rationality under Flexible Loss," *Review of Economic Studies*, 72, 1107-1125.

Fama, E.F. and K.R. French (1987), "Commodity Futures Prices: Some Evidence on Forecast Power, Premia, and the Theory of Storage," *Journal of Business*, 60(1), 55-73.

Fama, E.F. and K.R. French (1988), "Business Cycles and the Behavior of Metals Prices," *Journal of Finance*, 43(5), 1075-1093.

Froot, K.A., and R.H. Thaler (1990), "Anomalies: Foreign Exchange," *Journal of Economic Perspectives*, 1(3), 179-192.

Gibson, R., and E.S. Schwartz (1990), "Stochastic Convenience Yield and the Pricing of Oil Contingent Claims," *Journal of Finance*, 45(3), 959-976.

Gramlich, E.M. (2004), Oil Shocks and Monetary Policy, Annual Economic Luncheon, Federal Reserve Bank of Kansas City, Kansas City, Missouri.
<http://www.federalreserve.gov/boarddocs/speeches/2004/20040916/default.htm>

Greenspan, A. (2004a), "Energy," Remarks by Chairman Alan Greenspan Before the Center for Strategic & International Studies, Washington, D.C.
<http://www.federalreserve.gov/boarddocs/speeches/2004/20040427/default.htm>

Greenspan, A. (2004b), "Oil," Speech presented at the National Italian American Foundation, Washington, DC.

<http://www.federalreserve.gov/boarddocs/speeches/2004/200410152/default.htm>

Greenspan, A. (2005), "Energy," remarks before the Japan Business Federation, Tokyo, Japan <http://www.federalreserve.gov/boarddocs/speeches/2005/20051017/default.htm>

Hotelling, H. (1931), "The Economics of Exhaustible Resources," *Journal of Political Economy*, 39(2), 137-175.

Inoue, A., and L. Kilian (2004), "In-Sample or Out-of-Sample Tests of Predictability: Which One Should We Use?" *Econometric Reviews*, 23(4), 371-402.

International Monetary Fund (2005), *World Economic Outlook*, Washington, DC.

International Monetary Fund (2007), *World Economic Outlook*, Washington, DC.

Kilian, L. (2007a), "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market," mimeo, Department of Economics, University of Michigan.

Kilian, L. (2007b), "The Economic Effects of Energy Price Shocks," mimeo, Department of Economics, University of Michigan (invited by the *Journal of Economic Literature*).

Kilian, L. (2008), "Exogenous Oil Supply Shocks: How Big Are They and How Much Do They Matter for the U.S. Economy?" forthcoming: *Review of Economics and Statistics*.

Kohn, D.L. (2007), "The Economic Outlook," remarks at the Atlanta Rotary Club, Atlanta, GA,

<http://www.federalreserve.gov/boarddocs/speeches/2007/20070108/default.htm>

Litzenberger, R. H. and N. Rabinowitz (1995), "Backwardation in Oil Futures Markets: Theory and Empirical Evidence," *Journal of Finance*, 50(5), 1517-1545.

Neuberger, A. (1999), "Hedging Long-Term Exposures with Multiple Short-Term Futures Contracts," *Review of Financial Studies*, 12(3), 429-459.

New York Mercantile Exchange (2007a), "Light Sweet Crude Oil: Market Data," http://www.nymex.com/lSCO_fut_descri.aspx.

New York Mercantile Exchange (2007b), "Light Sweet Crude Oil: Contract Detail," http://www.nymex.com/WS_spec.aspx.

Penrose, E. (1976), "The Development of Crisis," in R. Vernon (ed.), *The Oil Crisis*, New York: Norton.

Pesaran, M.H. and A. Timmermann (1992), "A Simple Nonparametric Test of Predictive Performance," *Journal of Business and Economic Statistics*, 10(4), 461-465.

Pindyck, R.S. (1994), "Inventories and the Short-Run Dynamics of Commodity Prices," *RAND Journal of Economics*, 25(1), 141-159.

Pindyck, R.S. (2001), "The Dynamics of Commodity and Futures Markets: A Primer," *Energy Journal*, 22(3), 1-29.

Ramey, V. (1994), "Inventories as Factors of Production and Economic Fluctuations," *American Economic Review*, 79(3), 338-354.

Svensson, L.E.O. (2005), "Oil Prices and ECB Monetary Policy," mimeo, Department of Economics, Princeton University, see <http://www.princeton.edu/svensson/>

Taylor, M.P. (1989), "Covered Interest Arbitrage and Market Turbulence," *Economic Journal*, 99, 376-391.

Terzian, P. (1985), *OPEC. The Inside Story*. London: Zed Books.

Wu, T., and A. McCallum (2005), "Do Oil Futures Prices Help Predict Future Oil Prices?" *Federal Reserve Bank of San Francisco Economic Letter*, 2005-38.

Chapter 3

Did Adhering to the Gold Standard Reduce the Cost of Capital?

3.1. Introduction

Governments cannot simultaneously maintain an open capital market, retain monetary policy autonomy, and fix the exchange rate. This trilemma forces governments to make trade-offs among these goals by choosing two of the three objectives (Obstfeld et al. 2005). In the 19th century, governments opted for the perceived benefits of open capital markets and fixed exchange rates at the cost of an independent monetary policy.

The costs of the gold standard are well known. For example, Temin (1991), Eichengreen (1992), and Bernanke (1995) lay much of the blame for the spread and severity of the Great Depression on the gold standard. On the other hand, the benefits of the gold standard remain the subject of debate. Until recently the most widespread theory viewed a country's adherence to the gold standard as a credible signal of a country's commitment to prudent fiscal and monetary policy (Bordo and Rockoff 1996). This "good-housekeeping" interpretation of the gold standard generates a testable implication: Countries that tied their currencies to gold should have been rewarded with a lower cost of capital. Available 19th century asset data were too sparse for researchers to evaluate this hypothesis directly by comparing the returns of assets issued by countries on and off gold. Instead, past tests of the good housekeeping hypothesis compared observable coupon yields or capital flows between countries. These data have proven insufficient to resolve the question. Researchers have reached different conclusions on the effect of the gold standard on borrowing costs depending upon the choice of sample countries, empirical specification, and fiscal and monetary control variables.

For example, Bordo and Rockoff (1996) examined the coupon yields of sovereign bonds and found that pre-World War I yields "differed substantially from country to country" and that "these differences were correlated with a country's long-term

commitment to the gold standard.” Their findings are consistent with the country studies of Martin-Acena (1993) and Sussman and Yafeh (2000), and have been confirmed in a larger sample of countries by Obstfeld and Taylor (2003).

On the other hand, Ferguson and Schularick (2006) have argued that membership in the British Empire, rather than adhering to the gold standard, was the key to reducing borrowing costs. Clemens and Williamson (2004) examined capital flows rather than bond yields and concluded that “gold was nowhere near the most important determinant of [capital] flows” and gold standard adherence paled in importance compared to the fundamental determinants of capital productivity. Flandreau and Zumer (2004) viewed the good-housekeeping hypothesis most pessimistically. They argued that international lenders focused almost exclusively on a country's ability to repay its foreign obligations and paid little attention to a country's monetary regime. Like previous authors, Flandreau and Zumer examined sovereign bond coupon yields but found that adhering to the gold standard had a negligible influence once they control for measures of sound public finances.

Past tests of the good housekeeping hypothesis have reached mixed conclusions about the effect of the gold standard on borrowing costs due to differences in the sample of countries, model specification, and controls for non-gold risk factors. It is unlikely that a consensus will be reached with the existing data. Adjusting for risk is difficult when working with coupon yields rather than expected returns and no a priori criteria exist to disentangle competing specifications.

I introduce a more direct test of the good-housekeeping hypothesis based on a new dataset. My data consist of the holding period returns of almost every stock and sovereign bond traded in London between 1870 and 1907. Using realized returns is novel for tests of the good housekeeping hypothesis, but my test methodology is routine in the finance literature when evaluating the hypothesis that the market demands a risk premium for holding assets with an observable trait.¹² Data limitations have prevented past authors

¹² For example, Cornell and Green (1991) and Elton, Gruber, and Blake (1995) use realized return to measure the effect of bond ratings on the cost of capital; Fama and French (1995) use realized return to measure the effect of firm size on the cost of capital; and Lewellen (1999) uses realized return to measure the effect of book-to-market ratio on the cost of capital.

from employing return-based tests of the good housekeeping hypothesis. My dataset addresses this shortcoming.

The dataset contains over 450,000 realized 28-day stock and sovereign bond returns collected from late 19th and early 20th century financial publications. These data allow us to apply to this question the same empirical methods that researchers use to identify characteristics associated with risk and return in modern financial markets. My principal finding is that adherence to the gold standard did not reduce the cost of capital. Across a variety of specifications and samples, there is no link between a country being on the gold standard and the realized risk-adjusted return of its stock and sovereign debt. Conditional on business-cycle risk, the returns of assets issued by countries on and off gold are statistically indistinguishable from each another. An investor who selected assets based on gold standard adherence would not outperform an investor who selected assets at random.

The next section (Section 2) discusses the empirical implications of the theory of the gold standard as a repeated game. Section 3 discusses the differences between tests based on coupon yield and expected return. Section 4 describes my data and empirical specifications. Section 5 reports my results. Section 6 summarizes my findings and concludes.

3.2. The Gold Standard as a Repeated Game

Bordo and Kydland's (1995) model postulates that the gold standard functioned as a credible commitment mechanism to overcome the time-inconsistency problem associated with international borrowing and lending. Countries and international lenders were engaged in a repeated game, and each period the government could raise funds with a mix of taxes and borrowing. A future government that chose the optimal mix of borrowing, taxation, and inflation to minimize the dead-weight loss would choose to tax assets that were inelastically supplied, and a welfare-maximizing government had the incentive to tax bondholders by inflating away the value of its outstanding debt. Investors recognized that future government policy was time-inconsistent and were unwilling to loan funds today without a credible guarantee that the future government would repay its debt in real terms.

To overcome this time-inconsistency problem, the bond market needed a mechanism to monitor the government's behavior. In this repeated game, a government could signal the international bond market that it was following sound fiscal and monetary policies. This amounted to a signal that the government was pursuing a policy mix that made it unlikely that it would be forced to devalue the currency. Adherence to the gold standard could serve as such a commitment device. Thus the empirical implication of the good housekeeping hypothesis is that the capital market assigned a lower price, and demanded commensurately higher returns, to assets issued by countries that had abandoned gold.

A problem with the good housekeeping explanation is that it required investors to act collectively. The gold standard's repeated game equilibrium relies upon the capital market collectively forgoing current profits to punish governments that deviated from sound financial policies. With many participants, this equilibrium required a collective action mechanism to prevent arbitrage-seeking investors from pushing the prices of otherwise identical on- and off-gold assets together. Large institutional investors, who were both sufficiently patient to punish the country and large enough to influence equilibrium prices, were good candidates to punish countries that abandoned gold. Historical candidates include the Council of Foreign Bondholders and large investment banks. The available archival evidence suggests that the Council was effective at organizing lenders when borrowers defaulted and renegotiating debt with the larger borrowers, such as Argentina, Brazil, and Turkey (Mauro and Yafeh 2003). Whether such organizations were sufficiently large to affect equilibrium prices and punish cheaters is the empirical question that my paper addresses.

3.3. The Good-Housekeeping Hypothesis, Coupon Yields, and Expected Returns

Given a bond with future expected cash flow $E_0 \{X_t\}_{t=1}^T$, the current price of this bond is equal to the discounted stream of its cash flow

$$(1) \quad P_0 = E_0 \left\{ \sum_{t=1}^T \frac{X_t}{(1+r_t)^t} \right\}.$$

The expected rate of return, r_t , is the variable of interest. I would like to determine if the market demanded a higher expected rate of return for assets issued by countries that

violated the gold standard's rules, but the expected rate of return is unobservable. Its value depends on the market's expectation of future default embedded in the expected cash flow sequence $E_0 \{X_t\}_{t=1}^T$. Even if the $\{X_t\}_{t=1}^T$ sequence were known with certainty, the r_t sequence would be identified by the bond price under the assumption that r_t is constant for all t .

The expected one-period gross return is

$$(2) \quad E_0\{R_1\} = P_0^{-1} E_0\left\{X_1 + \sum_{t=2}^T \frac{X_t}{(1+r_t)^t}\right\} = E_0\left\{\frac{X_1 + P_1}{P_0}\right\}.$$

I can rewrite (2) to express the expected gross return of a bond as the sum of the expected capital gain $E_0\left\{\frac{P_1}{P_0}\right\}$ and the coupon yield $E_0\left\{\frac{X_1}{P_0}\right\}$. Past coupon yield-based tests of the good housekeeping hypothesis have ignored the capital gain and evaluated the effect of adhering to the gold standard based on coupon yield alone. Coupon yield is correlated with expected return but it ignores capital gains and understates the true cost of capital when bonds are issued below par or contain sinking fund provisions, two features shared by the sovereign bonds of many 19th century countries.

The expected return is unobservable but it can be estimated from ex post realized returns. If financial markets are informationally efficient, the realized ex post return of an asset is the sum of its expected return and an independent mean zero random error

$$(3) \quad r_{t+1} = E_t[r_{t+1}] + \varepsilon_{t+1}$$

Sample average returns converge to their expected value, and the average ex post holding period return is a good proxy for the actual unobservable expected return.

3.4. Controlling for Risk with a Relative Pricing Model

The good-housekeeping hypothesis predicts that, holding all else equal, assets of countries adhering to the gold standard trade at higher prices (lower expected return) than assets issued by non-gold standard countries. As I discussed above, this hypothesis is not easy to evaluate because expected returns are unobservable. Consequently, economists must use proxies such as coupon yield or realized holding period returns. Both expected returns and coupon yields of assets vary for reasons other than the credibility of the issuing government's monetary regime. Past tests of the good-housekeeping hypothesis

have relied on available bond yields and attempted to control for other factors by estimating regressions of the form:

$$(4) \quad Yield_{it} = \alpha_i + \delta GS_{it} + \lambda_t Marketyield_t + X'_{it} \beta_i + \varepsilon_{i,t}$$

where $Yield_{it}$ is a proxy for expected return such as coupon yield; GS_{it} is a dummy variable equal to one if the country is on gold; $Marketyield_t$ captures common changes in all yields and is the yield on British consols or an average of the coupon yields on all bonds; and X_{it} is a vector of country-specific fiscal and monetary variables intended to capture differences in economic fundamentals. The test of the good-housekeeping hypothesis amounts to testing that $\delta < 0$.

Authors have estimated different versions of equation (4) and reached contradictory conclusions depending upon their choice of sample countries, the use of pooled regressions or fixed effects, and fiscal and monetary variables. The disagreement hinges on what variables 19th century investors used to evaluate the riskiness of a foreign loan. For example, the difference between the findings of Flandreau and Zumer (2004) and Obstfeld and Taylor (2004) amounts to a disagreement about the variables that 19th century investors relied upon to measure risk. With no a priori criteria to sort competing models, I am pessimistic that this debate will be resolved with existing methods and data.

Instead I suggest a different test based on realized holding period returns. Both the coupon yield and the realized return are correlated with unobservable expected return, but realized return is a better proxy for expected return. The realized return also allows us to control for confounding risk factors with the same methods that economists use to investigate the performance of contemporary bond portfolios.

My method sidesteps the need to determine the variables that 19th century investors used to determine an asset's risk. I use a relative pricing model that compares the returns of assets issued by countries on or off gold to a control group of British assets. In the contemporary asset-pricing literature, economists use these methods to compare the expected returns of assets selected by observable criteria such as fund manager skill (Jensen 1967), accounting fundamentals (Fama and French 1992), and past returns (Jegadeesh 1990). In a paper similar in spirit to mine, Cornell and Green (1991) use this methodology to control for confounding risk factors when comparing the performance of

bonds sorted by investment grade. In addition, Elton, Gruber, and Blake (1995) employ similar methods to examine the risk factors in modern bond index portfolios.

3.5. Data and Empirical Specification

According to the good-housekeeping hypothesis, international lenders punished countries that did not adhere to the gold standard by demanding a higher risk-adjusted return. I test this hypothesis by forming portfolios that mimic the returns an investor would have earned had he purchased a portfolio of assets issued by countries off gold and sold short a second portfolio comprised of assets issued by countries on gold. Put differently, I postulate that an investor wakes up every 28 days, sells a portfolio consisting of assets issued by countries adhering to the gold standard and buys a portfolio of assets issued by countries off gold. If the investor demanded a higher return for the assets of countries off gold, this strategy would generate risk-adjusted excess returns. My empirical test amounts to testing the hypothesis that an investor could use his knowledge that a given country adhered to the gold standard to beat the market on a risk-adjusted basis.

The key words in the previous paragraph are “risk-adjusted”. I do not take a stand on the underlying determinants of risk and return. Instead I take returns as given and compare the return of a foreign asset to the return of a similarly risky portfolio of British assets. The main advantage of this approach is that it addresses the problem that countries did not leave gold randomly. If a country wanted to remain on gold but was forced off due to a negative business-cycle shock, I need to be careful not to conflate the business-cycle risks with the repeated game punishment. Business-cycle shocks may be correlated across countries and British investors may legitimately demand higher expected returns as compensation for holding greater business-cycle risk. In this case, the good-housekeeping hypothesis could be false but the return on the assets of countries that remained on the gold standard would be smaller due to smaller exposure to business-cycle risk. By comparing foreign returns to a control group of British assets, I disentangle the two effects and assess if investors demanded a premium due to gold standard adherence or business-cycle risk.

I control for risk by estimating the regression:

$$(5) \quad R_{it} - R_{ft} = \alpha_i + \beta_1(R_{con,t} - R_{ft}) + \beta_2(R_{SM,t} - R_{ft}) + \beta_3(R_{BM,t} - R_{ft}) + \varepsilon_i,$$

where $R_{con,t}$ is the time t return on the British consol; $R_{SM,t}$ is the return on the value-weighted portfolio of all British stocks at time t ; and $R_{BM,t}$ is the value-weighted portfolio of British corporate bonds. I use the London banker's bill rate as a proxy for the risk-free rate R_{ft} . α_i is referred to as Jensen's alpha, after Jensen (1967) who suggested this framework for controlling for risk while evaluating the stock-picking abilities of mutual fund managers. α_i measures the difference between portfolio i 's return and the return of the portfolio of British assets with percentage weights β_1 invested in the British Consol, β_2 invested in the value-weighted British stock market portfolio, β_3 invested in the value-weighted British bond market portfolio and $(1 - \beta_1 - \beta_2 - \beta_3)$ invested in the London banker's bill. α_i therefore measures the difference between portfolio i 's return and the return of the portfolio of British assets with the same business-cycle risk.

If an investor actively managed the portfolio in equation (5), α_i is a measure of the manager's ability to beat a portfolio of British assets with the same risk. The good-housekeeping hypothesis implies that an investor could outperform British assets by selecting assets based on gold standard adherence. I test this hypothesis by forming an excess return portfolio that is long assets issued by countries off of gold and short assets issued by countries on gold. The resulting regression equation is given by:

$$(6) \quad R_{off,t} - R_{on,t} = \alpha_i + \tilde{\beta}_1(R_{con,t} - R_{ft}) + \tilde{\beta}_2(R_{SM,t} - R_{ft}) + \tilde{\beta}_3(R_{BM,t} - R_{ft}) + \varepsilon_t,$$

where $\tilde{\beta}_k = \tilde{\beta}_{k,off} - \tilde{\beta}_{k,on}$. Thus a test of the good housekeeping hypothesis amounts to a test that α_i is greater than zero.

3.5.1. Data

My tests require a monthly database of asset returns, and data limitations have prevented past researchers from applying such tests to the good-housekeeping hypothesis. I address this shortcoming by collecting a large sample of stocks, sovereign bonds, and colonial bonds trading on the London Stock Exchange between 1870 and 1907. These data consist of the bid and ask prices and dividend (coupon) payments for almost every foreign and British stock and foreign government bond regularly quoted on the exchange. The data

set includes 213 sovereign bonds issued by 36 countries, 110 colonial bonds issued by 12 British colonies, and 1,808 stocks issued by firms in 44 countries. The prices were sampled every 28-days from the official quotation list published in the *Money Market Review* and the *Economist*. I use the price and dividend (coupon) data to compute a time series of realized holding period returns corrected for dividends, stock splits, and defaults.¹³

3.5.2. Managed Portfolios

I form portfolios that mimic the return an investor would have realized had he actively managed his portfolio by purchasing all assets issued by countries off gold and sold short all assets issued by countries on gold. When a country adopts the gold standard, I remove the country's assets from the off-gold portfolio and add it to the on-gold portfolio.

I rely upon both contemporary and modern sources to date each country's gold-standard adherence. A detailed list of sources and the dates that I used are in Appendix 2. In cases where it is difficult to determine de jure versus de facto adherence to the gold standard, I code the country as adhering to gold from the de jure convertibility date. In many cases, I am able to date gold standard adherence quite precisely, identifying the month and sometimes even the day, on which the currency became convertible into gold. In the cases where I can only identify the year in which a country adopted the gold standard, I date gold standard adherence from January 1 of that year.¹⁴

3.5.3. Sample of Countries

One potential explanation for the lack of consensus among previous tests of the good housekeeping hypothesis is that each study uses a different sample of countries. I have assets from a much larger sample of countries than previous authors. To evaluate the extent to which my results are driven by the countries selected, I evaluate the hypothesis with my full sample of countries and with subsamples corresponding as closely as possible to the countries used in Bordo and Rockoff (1996), Bordo and Schwartz (1996),

¹³ I use the Council of Foreign Bondholders and default data provided by Obstfeld and Taylor (2003) to date defaults on foreign debt.

¹⁴ In a previous draft of this paper, I assessed the robustness of our conclusions to alternative dating conventions. The results are very similar regardless of how we date gold standard adherence. These regression results are available upon request.

Obstfeld and Taylor (2003), and Flandreau and Zumer (2004). In all cases, I am able to mimic closely each of these samples. The countries contained in each of these subsamples are documented in Appendix 1.

3.5.4. Country Weights

When forming the portfolios, I must determine how much weight to give to the assets of each country. I formed value-weighted portfolios, which weight each asset by its market value. For example, if the market value of all Russian bonds is twice the value of Italian bonds, Russia receives twice the weight of Italy.¹⁵

3.5.5. Comparing Stocks and Bonds

Past tests of the good housekeeping hypothesis have focused exclusively on sovereign bonds. This is not surprising in light of the theoretical framework that postulates the gold standard as a repeated game between governments and the international capital market. Furthermore, coupon yield based tests cannot be applied to stocks that pay variable and unknown dividends. In contrast, my return-based tests can be applied to portfolios of stocks and bonds. I compute both managed sovereign bond portfolios and managed stock portfolios. Exchange rate volatility affects the private cost of capital in modern day developing markets (Solnik 1983). Adding stock portfolios allow us to assess the effects of adhering to the gold standard on the private cost of capital.

3.6. Empirical Results

Table 1 contains average returns and regression results for the value-weighted portfolios. Panel A reports the results for portfolios comprised of sovereign bonds. The strategy of buying assets issued by countries off gold and shorting assets issued by countries on gold did generate positive returns, but virtually all of these excess returns were due to differences in business-cycle risk. Controlling for differences in business-cycle risk, the return of off- and on-gold portfolios are indistinguishable from one another.

Depending on the sample of countries, an investor who selected assets based on gold standard adherence would have earned an average 28-day excess return between 12

¹⁵ To evaluate the robustness of the results, I also formed equally weighted portfolios. The results are very similar regardless of how we formed the portfolios. These regression results are available upon request.

and 19 basis points, amounting to an excess return of 1.6%-2.5% per year. This excess return is attributable to differences in business-cycle risk and does not reflect a market punishment for violating the rules of the gold standard. Once I control for risk by comparing the managed portfolios to similarly risky British assets, the excess returns vanish. Had an investor purchased a portfolio of British assets with the same business-cycle risk, he would have earned the same return as an investor who selected assets based upon adherence to the gold standard. Regardless of sample countries, the alpha is economically small and statistically insignificant.

With one exception, the betas in Panel A are always positive. The off-gold portfolios were more sensitive to movements in the return on the British consol, stock, and bond indices than the on-gold portfolios. Movements in these well-diversified British indexes largely reflect business-cycle shocks, suggesting that the higher unconditional returns associated with holding the off-gold country assets were compensation for bearing this business-cycle risk. During business-cycle contractions the assets of countries off gold tended to lose more value than the assets of countries on gold. British investors demanded a premium to hold this additional business-cycle risk, but it did not reflect a punishment for abandoning gold. A portfolio of British stocks and bonds with the same exposure to the British indexes generated statistically identical returns. Because these British assets were not being punished for violating the rules of the gold standard, I conclude that this excess return was compensation for business-cycle risk and the amount of compensation per unit of beta exposure was identical across assets regardless of whether the country adhered to gold or not.

The results are similar in Panel B. Holding equity in a firm located in a country that did not adhere to gold did not yield a higher return than holding equity in a firm in a country on gold. Again, this result is robust to the sample of countries.

The low R-squared statistics in Table 1 warrant a closer look. The R-squared statistics are small because the model does a poor job of explaining the time series variation of my managed portfolios returns. This outcome is exactly what one would expect if adherence to the gold standard had no effect on asset prices. The R-squared statistics may be small because variations in the British portfolios do a poor job of explaining variation in my managed portfolios or because gold standard adherence did

not matter and selecting assets based on gold is equivalent to selecting assets at random. If adhering to the gold standard had no effect on returns, selecting assets based on gold would be equivalent to assigning assets to portfolios at random. As the number of assets randomly assigned to each portfolio grows, the portfolio β 's would converge to population betas and the β 's of the on gold portfolio minus off portfolio would converge to zero as well.

I expect the R-squared statistics to be small when the British portfolios do a poor job of explaining non-diversifiable risk or when gold standard adherence does a poor job of explaining returns. I can disentangle these competing causes by estimating regression (5) with the on- and off-gold returns and comparing the results to the excess return regression in equation (6). Table 2 reports the individual regression results for the on- and off-gold value- and equally weighted portfolios. The model does a good job of pricing both portfolios. The adjusted R-squared statistics are respectable for monthly data – between 15%-20%. In addition, the estimated coefficients in the separate on- and off-gold portfolios tend to be quite close to one another. Thus, when one computes excess returns by taking the difference between the off- and on-gold portfolios, the estimated coefficients and R-squared statistics are close to zero – exactly what one would expect if the gold standard did not matter.

In sum, a 19th century investor did not demand a risk premium for holding countries assets issued by countries off of gold. Had an investor tried to beat the market by betting that off-gold countries would outperform on-gold countries, he would have been unable to do so.

3.6.1. Robustness Check: Random Portfolios

Portfolios selected with knowledge of adherence to the gold standard did not outperform portfolios of British assets with the same exposure to business-cycle risk. I base this conclusion on the small α_i values in Table 1. A test that α_i is equal to zero is a joint test of the good housekeeping hypothesis and the risk and return model implied by the excess return regression. To be certain that the small alphas are not due to model misspecification, I compare the alphas of the managed portfolios to the alpha an investor would have obtained if he had selected assets purely at random.

For each type of asset and country sample in Table 1, I compute excess return portfolios by randomly assigning the same sample of assets to one of two portfolios. Assets are assigned in the same proportion as the proportion of gold standard adherence. I compute 1000 random portfolios and report the proportion of times the portfolio selected using the gold standard criterion earns a higher excess return than a portfolio selected at random.

For example, in the value-weighted Bordo-Rockoff sample of sovereign bonds, 56% of the observed returns were associated with bonds on gold. I compute a random excess return portfolio with the same assets by randomly buying 44% of the bonds each period and shorting the other 56%. I compute the random portfolio's α_i and repeat the random selection 1000 times. The result is 1000 random alphas.

If gold standard adherence didn't matter, I would expect portfolios selected based on gold standard adherence to do no better than portfolios selected at random. The simulation, which sorts the assets randomly to one or the other portfolio, directly addresses the question of whether sorting on the gold standard does better than sorting randomly. Table 1 reports the success rate – the proportion of times the managed portfolios beat portfolios formed at random. The results are consistent with the conclusion that gold standard adherence did not matter. On average, portfolios selected by gold standard adherence do no better than portfolios selected at random.

3.6.2. Further Robustness Checks: Perfect-Foresight and Cheater Portfolios

I also assess the robustness of the results to both anticipation effects and allowing investors to punish countries that defaulted. The managed portfolios may reflect the market's anticipation that countries within each portfolio would resume or suspend the gold standard. Thus the return of the off-gold portfolio minus the on-gold portfolio will be biased upwards if my dating procedure lags market expectations and the null hypothesis is true. In addition, I coded a country as adhering to the gold standard if they were on gold in the preceding tests. However, some countries were on the gold standard, suspended convertibility for a brief period, and returned to gold at a lower par (Bordo and Kydland 1995). If the capital market considered countries on gold at a lower par to be cheating, my managed portfolios based on gold convertibility miscodes some countries.

To check the robustness of the results to anticipation affects, I formed both value- and equally weighted portfolios derived in the same manner as above, with the exception that changes in gold standard status are reflected one year before the actual change in status occurs. I thus assume that the Victorian investor perfectly anticipates whether a country joins or suspends the gold standard one year in advance. I omit these regression results for brevity, but they are very similar to those presented in Table 1 and are available upon request. Managed portfolios that select assets based on gold standard adherence do not generate excess returns even when the portfolio manager perfectly anticipates a change in gold standard status.

To assess the effect of temporarily suspending convertibility, I recode the countries as adhering to gold if it is currently on gold at its original par. If a country is on gold, suspends convertibility, and returns to gold at a lower par, I recode this country as a “cheater”. I consider this country as off gold for the period that it is off gold or on gold at a lower par. Countries that suspend convertibility are thus considered off gold until they return at the original par. With this new coding, I recomputed my results, which are omitted for brevity but are available upon request. The results are robust to this new coding scheme. The off-gold minus the on-gold managed portfolios are statistically indistinguishable from British assets even when I use memory of past cheating to sort countries.

3.7. Conclusion

The good-housekeeping hypothesis predicts that the international capital market rewarded countries that adhered to the classical gold standard with a favorable cost of capital. Due to data limitations, past tests of this hypothesis did not measure actual returns. I address this shortcoming by introducing a new data set consisting of the 28-day holding period returns of sovereign bonds, colonial bonds, and stocks trading in London between 1870 and 1907. The new sample allows us to measure actual holding period returns and apply modern asset-pricing models to evaluate the testable implications of the good housekeeping hypothesis.

I find no evidence in favor of the good-housekeeping hypothesis. The returns on assets issued by countries on and off gold are statistically identical to each other.

Portfolios formed by shorting assets issued by countries on gold and purchasing assets issued by countries off of gold do not outperform similarly risky British assets. Furthermore, portfolios selected with knowledge about gold standard adherence do not outperform portfolios selected completely at random.

These results are very robust. I find no evidence of a gold-standard effect across asset classes or country samples. These results persist even when I form portfolios with perfect foresight of gold standard regime change or assume the international capital market punished past departures from the rules of the game. Overall, these new data and tests suggest that countries that adopted the gold standard were not rewarded with a low cost of capital. Thus the gold standard did not necessarily serve as a transparent signal of prudent financial policies.

More broadly, these results shed new light on the perceived benefits of fixed exchange rate regimes. A widely cited benefit of the gold standard is that it reduced borrowing costs. I find that the choice of exchange rate regime had no impact on the cost of capital.

Notes: The regression is equation (6) in the text. The success rate measures the proportion of times that the gold portfolio earns a higher excess return than a randomly selected portfolio. Robust t -statistics are in parentheses. *** (**) (*) indicates significance at the 1% (5%) (10%) level.

Table 3-1: 28-Day Excess Return Regressions: Off-Gold Portfolio – On Gold Portfolio (cont.)

Dependent Variable: $R_{Off,t} - R_{On,t}$

Panel A: Value-Weighted Sovereign Bond Portfolios					
	Alquist-Chabot	Bordo-Rockoff	Bordo-Schwartz	Obstfeld-Taylor	Flandreau-Zumer
Average Return	0.0012	0.0018	0.0012	0.0017	0.0019
$\hat{\alpha}_i$	-0.0004	0.0004	0.0001	-0.0003	0.0006
	(0.45)	(0.36)	(0.05)	(0.23)	(0.50)
Consol	0.226**	0.138	0.054	0.215***	0.312***
	(2.55)	(1.40)	(0.59)	(2.12)	(3.07)
Market	0.103	0.000	0.170**	0.136	-0.012
	(1.30)	(0.00)	(2.06)	(1.50)	(0.13)
Corp. Bond	0.265***	0.386***	0.148	0.365***	0.271**
	(2.76)	(3.61)	(1.50)	(3.33)	(2.46)
\bar{R}^2	0.062	0.045	0.028	0.070	0.047
Success Rate	24.1%	68.9%	48.4%	39.1%	78.7%

Table 3-1: 28-Day Excess Return Regressions: Off-Gold Portfolio – On Gold Portfolio (concl.)

Dependent Variable: $R_{Off,t} - R_{On,t}$

Panel B: Value-Weighted Stock Portfolios					
	Alquist-Chabot	Bordo-Rockoff	Bordo-Schwartz	Obstfeld-Taylor	Flandreau-Zumer
Average Return	0.0005	0.0005	0.0003	0.0005	0.0013
$\hat{\alpha}_i$	0.0005	0.0009	0.0005	0.0006	0.0007
	(0.44)	(0.59)	(0.35)	(0.46)	(0.48)
Consol	-0.124	-0.031	0.004	-0.080	0.053
	(1.11)	(0.22)	(0.03)	(0.69)	(0.40)
Market	-0.016	-0.155	-0.083	-0.040	0.122
	(0.16)	(1.27)	(0.72)	(0.38)	(1.03)
Corp. Bond	0.079	0.085	0.042	0.068	0.020
	(0.66)	(0.58)	(0.30)	(0.54)	(0.14)
\bar{R}^2	0.003	0.004	0.001	0.002	0.005
Success Rate	66.3%	75.8%	62.2%	65.4%	80.9%

Table 3-2: 28-Day Excess Return Regressions: On- and Off-Gold Portfolios

Dependent Variable: $R_{i,t} - R_{ft}$

Panel A: Sovereign Bonds				
	<u>Value-Weighted</u>		<u>Equally Weighted</u>	
	On-Gold	Off-Gold	On-Gold	Off-Gold
Average Return	0.0041	0.0053	0.0043	0.0054
$\hat{\alpha}_i$	0.0025*** (3.64)	0.0020** (2.12)	0.0028*** (5.72)	0.0014 (1.10)
Consol	0.126** (2.04)	0.352*** (4.03)	0.147*** (3.33)	0.314*** (2.66)
Market	0.295*** (5.31)	0.399*** (5.07)	0.235*** (5.88)	0.723*** (6.78)
Corp. Bond	0.094 (1.41)	0.359*** (3.80)	0.095** (1.98)	0.206 (1.61)
\bar{R}^2	0.118	0.203	0.171	0.177

Panel B: Foreign Stocks				
	<u>Value-Weighted</u>		<u>Equally Weighted</u>	
	On-Gold	Off-Gold	On-Gold	Off-Gold
Average Return	0.0042	0.0047	0.0055	0.0059
$\hat{\alpha}_i$	0.0007 (0.65)	0.0012 (1.22)	0.0032*** (3.37)	0.0028** (2.51)
Consol	-0.309*** (3.19)	0.185** (2.02)	0.078 (0.92)	0.034 (0.33)
Market	0.759*** (8.69)	0.743*** (8.99)	0.581*** (7.57)	0.744*** (8.13)
Corp. Bond	0.017 (0.17)	0.097 (0.97)	0.016 (0.17)	0.081 (0.74)
\bar{R}^2	0.222	0.224	0.148	0.166

Notes: Robust t -statistics are in parentheses. *** (**) (*) indicates significance at the 1% (5%) (10%) level.

Appendix 3-1: Sample of Countries

Sovereign Bond Data

The countries included in my dataset on sovereign bonds are: Argentina; Australia; Austria-Hungary; Belgium; Brazil; British Guiana; Bulgaria; Canada; Ceylon; Chile; China; Colombia; Costa Rica; Denmark; Ecuador; Egypt; France; Germany; Greece; Guatemala; Hawaii; Honduras; Hong Kong; Italy; Jamaica; Japan; Liberia; Mauritius; Mexico; Netherlands; New Zealand; Nicaragua; Norway; Orange Free State; Paraguay; Peru; Portugal; Russia; Saint Lucia; Santo Domingo; South Africa; Spain; Straits Settlements; Sweden; Trinidad; Turkey; United States; Uruguay; and Venezuela.

The British colonies comprise a subset of the countries in my sample. They are: Australia; British Guiana; Canada; Ceylon; Hong Kong; Jamaica; Mauritius; New Zealand; Saint Lucia; South Africa; and Trinidad.

Stock Data

In addition to the countries listed above, the countries included in my dataset on stocks are: Cuba; El Salvador; Dutch East Indies; India; Malta; Morocco; Persia; Philippines; Romania; and Uganda.

Subsamples

I formed subsamples of countries based upon previous work on the gold standard. I am able to mimic closely each of the samples of countries listed below.

Bordo and Rockoff (1996) : Argentina; Australia; Brazil; Canada; Chile; Italy; Portugal; Spain; and United States.

Bordo and Schwartz (1996) : Argentina; Australia; Belgium; Brazil; Canada; Chile; Denmark; Finland; Greece; Italy; Japan; Netherlands; Norway; Portugal; Sweden; and Switzerland. The four core countries are France, Germany, the UK, and the United States. I exclude Finland and Switzerland due to lack of data.

Obstfeld and Taylor (2003): Argentina; Australia; Austria-Hungary; Brazil; Canada; Chile; Egypt; Greece; India; Italy; Japan; Mexico; New Zealand; Norway; Portugal; South Africa; Spain; Sweden; Turkey; the United States; and Uruguay.

Flandreau and Zumer (2004): Argentina; Austria-Hungary; Belgium; Brazil; Denmark; France; Germany; Greece; Italy; Netherlands; Norway; Portugal; Spain; Sweden; Switzerland; and Russia. I exclude Switzerland due to lack of data.

I adopt the definition of cheater from Bordo and Schwartz's Table 1 (Bordo and Schwartz 1996). I code a country as a cheater if it suspended gold convertibility because of war, lax fiscal policy, a financial crisis, failed convertibility, or some combination of the four. The countries that fall into this category are Argentina, Brazil, Chile, Greece, Italy, and Portugal.

Appendix 3-2: Gold Standard Adherence

Argentina: January 3, 1867-May 1876: Dated from establishment of Office of Exchange to control amount of paper pesos outstanding. This event did not result in a change of parity. July 1883-December 1884: de facto suspension of convertibility by banks. This event did not result in a change of parity. October 31, 1899: Law establishing external convertibility. Source: della Paolera and Taylor (2001), pp. 25, 41, 47.

Australia: Adopted the free convertibility of currency into gold before 1870. Source: Meissner (2005).

Austria-Hungary: August 2, 1892. Source: Helfferich (1927), pp. 200.

Belgium: November 5, 1878. Source: Helfferich (1927), pp. 180.

Brazil: October 1888-October 1890. This event did not result in a change of parity. December 31, 1906: Conversion Office opened. Sources: Martin-Acena (2000), 2000, pp 155; and Subercaseaux (1931).

British Guiana: Adopted the free convertibility of currency into gold before 1870. Source: Officer (2001).

Bulgaria: Did not adopt the free convertibility of currency into gold before 1914. Source: Meissner (2005).

Canada: Adopted the free convertibility of currency into gold before 1870. Newfoundland did not adopt the free convertibility of currency into gold until 1895, according to Officer (2001). I date the bonds from Newfoundland from January 1, 1895. Sources: Meissner (2005).

Ceylon: 1898. I date Ceylon as adhering to the gold standard from January 1, 1898. Source: Officer (2001).

Chile: Law passed February 11, 1895 providing for conversion from June 1, 1895. Suspended convertibility July 31, 1898. This event resulted in a change of parity. Source: Kemmerer (1926).

China: Did not adopt the free convertibility of currency into gold before 1914. Source: Bloomfield (1959).

Colombia: From 1903, Colombia has fixed gold parities, but it did not adopt a complete gold standard until 1923, when the first Kemmerer Mission intervened. I code Colombia

as not adopting the free convertibility of currency into gold before 1914. Source: Martin-Acena (2000), pp. 237.

Costa Rica: Law passed in October 1896, but currency was not convertible into gold until July 15, 1900. I date it from November 1896. Source: Young (1925), pp. 193-96.

Cuba: 1898. I date Cuba as adhering to the gold standard from January 1, 1898. Source: Officer (2001).

Denmark: Agreement reached December 18, 1872. Formed monetary union with Sweden in May 1873. I date it from June 1873. Sources: Helfferich (1927), pp. 175; and Jonung (1984), pp. 367.

Dutch East Indies: 1877. I date the Dutch East Indies (Indonesia) as adhering to the gold standard from January 1, 1877. Meissner (2005).

Ecuador: 1900. I date Ecuador as adhering to the gold standard from January 1, 1900. Source: Meissner (2005).

Egypt: 1885. I date Egypt as adhering to the gold standard from January 1, 1885. Source: Officer (2001).

El Salvador: 1892. I date El Salvador as adhering to the gold standard from January 1, 1892. Meissner (2005).

France: November 5, 1878. Source: Helfferich (1927), pp. 180

Germany: December 4, 1871. Source: Helfferich (1927), pp. 156-7.

Greece: January 1885-September 1885. This event did not result in a change of parity. Greece did not join the gold standard again until March 1910. Sources: Bordo and Schwartz (1999), pp. 251; and Lazaretou (2005).

Guatemala: Did not adopt the free convertibility of currency into gold before 1914. Source: Bulmer-Thomas (2003), pp 112.

Hawaii: Adopted the Gold Law of 1884, which made the gold coins of the United States legal tender. I date Hawaii as adhering to the gold standard from January 1, 1884. Source: Tate (1965), pp. 69.

Honduras: Did not adopt the free convertibility of currency into gold before 1914. Source: Bulmer-Thomas (2003), pp 112.

Hong Kong: Did not adopt the free convertibility of currency into gold before 1914. Source: Tom (1989).

India: 1898. I date India as adhering to the gold standard from January 1, 1898. Source: Officer (2001).

Italy: Law establishing convertibility April 12, 1884 passed March 1, 1883. Affidavit introduced on second semester coupon payments in July 1893. Required Italians to swear rendita coupon payments received abroad did not belong to Italian citizens. Introduced incentives for lenders to redeem in Milan. This event resulted in a change in parity. Sources: Helfferich (1927), pp. 175; and Fratianni and Spinelli (1984), pp. 415.

Jamaica: Adopted the free convertibility of currency into gold before 1870. Source: Officer (2001).

Japan: Law passed March 29, 1897 providing for conversion between October 1, 1897- July 31, 1898. Sources: Helfferich (1927), pp. 201; and Laughlin (1900).

Liberia: Pegged to US dollar at the exchange rate L\$1 = US\$1. I code it the same as the US. Source: <http://users.erols.com/kurrency/africa.htm>

Malta : Adopted the free convertibility of currency into gold before 1870. Source: Officer (2001).

Mauritius : 1898. I date Mauritius as adhering to the gold standard from January 1, 1898. Source: Officer (2001).

Mexico: Law passed December 9, 1904. Decree enforced March 25, 1905. Source: Helfferich (1927), pp. 204.

Morocco: I find no evidence that Morocco adhered to the gold standard during the sample period. I code it as not adhering to the gold standard.

Nicaragua: Did not adopt the free convertibility of currency into gold before 1914. Sources: Bulmer-Thomas (2003), pp 115; and Young (1925), pp. 119-30; Appendix E.

Netherlands: June 6, 1875. Source: Helfferich (1927), pp. 176.

New Zealand: Adopted the free convertibility of currency into gold before 1870. Source: Officer (2001).

Norway: Agreement reached December 18, 1872. Norway formed a monetary union with Denmark and Sweden in October 1875. Sources: Helfferich (1927), pp. 175; and Jonung (1984), pp. 367.

Orange Free State: 1870-December 14, 1899. December 15, 1899-March 14, 1900: Gold standard suspended. March 15, 1900-1910: Gold standard re-established. Source: <http://users.erols.com/kurrency/africa.htm>

Paraguay: Did not adopt the free convertibility of currency into gold before 1914. Source: Bulmer-Thomas (2003), pp 114.

Persia: Did not adopt the free convertibility of currency into gold before 1914. Source: Meissner (2005).

Peru: Did not adopt the free convertibility of currency into gold before 1914. Source: Meissner (2005).

Philippines: 1903. I code the Philippines as adhering to gold from January 1, 1903. Source: Meissner (2005).

Portugal: 1854-May 1891. This event resulted in a change in parity. Source: Reis (2000), pp 94.

Romania: 1890. I code Romania as adhering to gold from January 1, 1890. Source: Meissner (2005).

Russia: February 1897. Source: Anonymous (1897).

St. Lucia: Adopted the free convertibility of currency into gold before 1870. Source: Officer (2001).

Santo Domingo: Did not adopt the free convertibility of currency into gold before 1914. Sources: Laughlin (1894); and Meissner (2005).

South Africa: Adopted the free convertibility of currency into gold before 1870. Source: Officer (2001).

Spain: Suspended summer 1883 because of decline in foreign investment after January 1882 stock market crash. The precise date is vague, and I date it from the end of June 1883. Source: Martin-Acena (2000), pp 128.

Straits Settlements: 1903. I date the Straits Settlements as adhering to the gold standard from January 1, 1903. Source: Meissner (2005).

Sweden: Agreement reached December 18, 1872. Formed monetary union with Denmark in May 1873. Sources: Helfferich (1927), pp. 175; and Jonung (1984), pp. 367.

Trinidad: Adopted the free convertibility of currency into gold before 1870. Source: Officer (2001).

Turkey: 1881. I date Turkey as adhering to the gold standard from January 1, 1881. Sources: Pamuk (2000), pp. 218.

Uganda: 1898. I date Uganda as adhering to the gold standard from January 1, 1898. Source: Officer (2001).

Uruguay: 1885. I date Uruguay as adhering to the gold standard from January 1, 1885. Source: Meissner (2005).

United States: January 1, 1879. Source: Officer (2001).

Venezuela: Did not adopt the free convertibility of currency into gold before 1914. Source: Meissner (2005).

References

- Anonymous. "Specie Resumption in Russia." *Journal of Political Economy* 5, no. 2 (1897): 241-244.
- Bernanke, Ben. "The Macroeconomics of the Great Depression: A Comparative Approach." *Journal of Money, Credit, and Banking* 27, no. 1 (1995): 1-28.
- Bloomfield, Arthur I. *Monetary Policy Under the Gold Standard*. New York: Federal Reserve Bank of New York, 1959.
- Bordo, Michael D., and Anna J. Schwartz. "The Operation of the Specie Standard: Evidence for Core and Peripheral Countries: 1880-1990." In *The Gold Standard and Related Regimes: Collected Essays*, edited by Michael D. Bordo, 238-317. New York: Cambridge University Press, 1999.
- Bordo, Michael D., and Finn Kydland. "The Gold Standard as a Rule." *Explorations in Economic History* 32, no. 4 (1995): 423-464.
- Bordo, Michael D., and Hugh Rockoff. "The Gold Standard as a Good Housekeeping Seal of Approval." *Journal of Economic History* 56, no. 2 (1996): 389-428.
- Bulmer-Thomas, Victor. *The Economic History of Latin America Since Independence*. New York: Cambridge University Press, 2003.
- Clemens, Michael A. and Jeffrey G. Williamson. "Wealth Bias in the First Global Capital Market Boom, 1870-1913." *Economic Journal* 114, no. 495 (2004): 304-337.
- Cornell, Bradford, and Kevin Green. "The Investment Performance of Low-Grade Bond Funds." *Journal of Finance* 46, no. 1 (1991): 29-48.
- della Paolera, Gerardo, and Alan Taylor. *Straining at the Anchor: The Argentine Currency Board and the Search for Macroeconomic Stability: 1880-1935*. Chicago: University of Chicago Press, 2001.
- Eichengreen, Barry. *Golden Fetters: The Gold Standard and the Great Depression, 1919-1939*. New York: Oxford University Press, 1992.
- Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake. "Fundamental Economic Variables, Expected Returns, and Bond Fund Performance." *Journal of Finance* 50, no. 4 (1995): 1229-1256.

- Fama, Eugene F., and Kenneth R. French. "The Cross-Section of Expected Stock Returns." *Journal of Finance* 47, no. 2 (1992): 427-465.
- Fama, Eugene F., and Kenneth R. French. "Size and Book-to-Market Factors in Earnings and Returns." *Journal of Financial Economics* 50, no. 1 (1995): 131-155.
- Ferguson, Niall, and Moritz Schularick. "The Empire Effect: The Determinants of Country Risk in the First Age of Globalization, 1880-1913." *Journal of Economic History* 66, no. 2 (2006): 283-312.
- Flandreau, Marc, and Frederic Zumer. *The Making of Global Finance, 1880-1913*. Paris: Organization for Economic Cooperation and Development, 2004.
- Helfferich, Karl. *Money*, 2 vols. Translated by Louis Infield. New York: Adelphi Company Publishers, 1927.
- Jegadeesh, Narasimhan. "Evidence of Predictable Behavior of Security Returns." *Journal of Finance* 45, no. 3 (1990), 881-898.
- Jensen, Michael C. "The Performance of Mutual Funds in the Period 1945-1964." *Journal of Finance* 23, no. 2 (1967), 389-416.
- Jonung, Lars. "Swedish Experience under the Classical Gold Standard: 1873-1914." In *A Retrospective on the Classical Gold Standard: 1821-1931*, edited by Michael D. Bordo and Anna J. Schwartz, 361-399. Chicago: University of Chicago Press, 1984.
- Kemmerer, Edwin Walter. "Chile Returns to Gold Standard." *Journal of Political Economy* 34, no. 3 (1926): 265-273.
- Laughlin, J. Laurence. "Gold and Silver in Santo Domingo." *Journal of Political Economy* 2, no. 4 (1894): 536-560.
- _____. "Report on the Adoption of the Gold Standard in Japan." *Journal of Political Economy* 8, no. 3 (1900): 424-427.
- Lazaretou, Sophia. "The Drachma, Foreign Creditors, and the International Monetary System: Tales of a Currency during the 19th and 20th Centuries." *Explorations in Economic History* 42, no. 2 (2005): 202-236.
- Lewellen, Jonathan W. "The Time-Series Relations Among Expected Return, Risk, and Book-to-Market." *Journal of Financial Economics* 54, no. 1 (1999): 5-43.

Martin-Acena, Pablo. "Spain During the Classical Gold Standard Years, 1880-1914." In *Monetary Regimes in Transition*, edited by Michael D. Bordo and Forrest Capie, 135-172. New York: Cambridge University Press, 1993.

Martin-Acena, Pablo. "The Spanish Monetary Experience: 1848-1914." In *Monetary Standards in the Periphery*, edited by Pablo Martin Acena and Jaime Reis, 112-151. New York: St. Martin's Press, 2000.

Mauro, Paolo, and Yishay Yafeh. "The Corporation of Foreign Bondholders." IMF Working Paper No. 03/107, Washington, D.C., May 2003.

Meissner, Chris. "A New World Order: Explaining the International Diffusion of the Gold Standard, 1870-1913." *Journal of International Economics* 66, no. 2 (2005): 385-406.

Obstfeld, Maurice, and Alan M. Taylor. "Sovereign Risk, Credibility, and the Gold Standard: 1870-1913 versus 1925-31." *Economic Journal* 113, no. 487 (2003): 241-275.

Obstfeld, Maurice, Jay C. Shambaugh, and Alan M. Taylor. "The Trilemma in History: Tradeoffs Among Exchange Rates, Monetary Policies, and Capital Mobility." *The Review of Economics and Statistics* 87, no. 3 (2005): 423-438.

Officer, Lawrence. "Gold Standard." EH.Net Encyclopedia, edited by Robert Whaples. October 1, 2001. URL <http://eh.net/encyclopedia/article/officer.gold.standard>

Pamuk, Sevket. *A Monetary History of the Ottoman Empire*. New York: Cambridge University Press, 2000.

Reis, Jamie. "The Gold Standard in Portugal: 1854-1891." In *Monetary Standards in the Periphery*, edited by Pablo Martin-Acena and Jaime Reis, 69-111. New York: St. Martin's Press, 2000.

Solnik, Bruno. "International Arbitrage Pricing Theory." *Journal of Finance* 38, no. 2 (1983): 449-457.

Subercaseaux, Guillermo. "The Modern Gold Standard with Illustrations from South America." *American Economic Review* 21, no. 2 (1931): 249-259.

Sussman, Nathan, and Yishay Yafeh. "Institutions, Reforms, and Country Risk: Lessons from Japanese Government Debt in the Meiji Period." *This Journal* 60, no. 2 (2000): 442-467.

Tate, Merze. *The United States and the Hawaiian Kingdom*. New Haven and London: Yale University Press, 1965.

Temin, Peter. *Lessons from the Great Depression*. New York: MIT Press, 1991.

Tom, C. F. Joseph. *Monetary Problems of an Entrepot: The Hong Kong Experience*. New York: Peter Lang Publishing, 1989.

Young, John Parke. *Central American Currency and Finance*. Princeton: Princeton University Press, 1925.

Chapter 4

How Important Is Liquidity Risk for Sovereign Bond Risk Premium?

4.1. Introduction

Liquidity risk – the risk that the price of a security will decline when the market as a whole becomes illiquid – is important for understanding the behavior of sovereign bond prices. “Liquidity” describes the ease with which one can trade large quantities of a security quickly, at low cost, and without altering the security’s price. Liquidity risk and the associated premium that a risk-averse investor demands for bearing such risk thus reflect the sensitivity of the sovereign bond’s price to unexpected changes in an aggregate measure of the ease with which one can trade securities. For example, during periods of financial turbulence, there is often a “flight to liquidity,” a phenomenon that has become increasingly important in international financial markets. In a flight to liquidity, market participants shift their asset holdings to more liquid sovereign bonds rather than less liquid ones. A recent example of a flight to quality was the sharp increase in the price of US Treasuries relative to less-liquid debt instruments in response to Russian default in August 1998.

To the extent that sovereign bonds are differentially exposed to the risk associated with credit crunches and the attendant decline in market liquidity, asset-pricing theory predicts that risk-averse investors will demand a liquidity premium for bearing such risk. Sovereign bonds that are more exposed to liquidity risk will trade at a discount relative to bonds that are less exposed to such risk. The available evidence for the US suggests that there is a liquidity premium embedded in the price of US government bonds, suggesting that liquidity risk is an important determinant of sovereign bond risk premia (Amihud and Mendelson 1991; Krishnamurthy 2002; and Longstaff 2004). Liquidity risk is also a good candidate for explaining the cost of sovereign borrowing from abroad.

The rapid growth of the international market for sovereign debt since 1980 has made studying the influence of liquidity risk on sovereign bond risk premia salient. By “sovereign bonds,” I mean bonds issued internationally by governments or government agencies, whose payments are guaranteed by these governments. Most sovereign bonds are issued in foreign currency and only a small number of developed countries in North America and Western Europe are able to issue debt in their own currency. As developed countries have run large and persistent deficits since the 1970s, these countries have adopted more liberal policies concerning the issuance and trading of their sovereign debt, which in turn facilitated the dramatic increase in the size of the international debt market. Similarly, the composition of capital flows to developing countries has shifted from syndicated bank loans – the primary form of sovereign lending during the 1970s (Folkerts-Landau 1985) – to debt instruments. Between 1990 and 2007, capital flows to developing countries in the form bonds have increased at an average annual rate of 20.0% per year whereas equities and syndicated bank loans have increased at average annual rates of 31.9% and 13.6% per year (Table 1). The data in Table 1 do not discriminate between capital flows to private borrowers and sovereign borrowers, but independent evidence shows that most of these flows of bond finance to developing countries consist of sovereign bonds as these borrowers have switched away from syndicated bank loans (Eichengreen and Mody 2000).

This evidence underscores that the volume of outstanding sovereign debt has increased and highlights the pressing need for understanding how this debt is valued. As a benchmark for the private cost of capital, the price of sovereign debt is a key price for countries that use the international bond market to meet their external financing needs. Understanding the determinants of this price is thus a first-order concern.

It may seem surprising that modern data are inadequate to understand the role of liquidity risk as a determinant of sovereign bond prices, but there were no centralized markets for trading such debt until recently. Until the late 1970s, interest rate and foreign exchange controls in the major financial centers prevented the development of an international sovereign debt market. Since then, active secondary markets in New York, London, and Tokyo for foreign debt have emerged, but observable bid-ask spreads, a key variable for studying the influence of liquidity risk on the price of sovereign debt, are not

widely available. Sovereign bonds are primarily sold “over the counter,” which means that prices are quoted privately to the customer by the salesman rather than openly quoted on the stock market. Indeed, a 1999 Bank for International Settlements study of government bond market liquidity in the G10 countries found it difficult to compare bid-ask spreads among these countries because market data on bid-ask spreads were unavailable and had to be estimated by the countries’ central banks (Bank for International Settlements 1999). This lack of data from the contemporary period makes studying the influence of liquidity risk on sovereign borrowing costs problematic.

This paper overcomes this difficulty by studying the sovereign debt market in London prior to World War One. This market is ideally suited to shedding light on the relationship between liquidity risk and the price of sovereign. In fact, there are strong similarities between that era of extensive financial integration and the current one. Capital flows during that era were as large as, or larger than, they are today, suggesting that capital mobility in the current period is comparable to the levels experienced during the late 19th century (Mauro, Sussman, and Yafeh 2002; and Taylor 2002). Moreover, sovereign borrowers during that period relied exclusively on bonds as their principal means for obtaining external finance. To the extent that today’s international bond market continues to be important for sovereign borrowers to obtain external finance, characterizing the relationship between sovereign bond risk premia and liquidity risk prior to World War One provides us with instructive lessons for understanding this relationship today.

The London sovereign debt market possesses unique features that make it an excellent laboratory for studying effect of liquidity risk on sovereign bond risk premia. Observable bid-ask spreads denominated in a common currency for a heterogeneous set of sovereign borrowers are available from 1870 until 1914. There is substantial time-series variation in the behavior of sovereign bond risk premia and market liquidity during this period, which encompasses the peaks and troughs of several international business cycles, including the depression that began in 1873; financial crises that affected the major sovereign borrowers; and at least three lending booms in the London sovereign debt market – one in the mid-1870s, one in the mid- to late 1880s, and one in the decade before World War One (Suter 1990). This time-series variation encompasses not only

several business cycles but also a number of complete lending cycles and market liquidity, making this environment a natural setting in which to study the influence of liquidity risk on sovereign bond risk premia.

The results are highly informative. There is strong evidence that sovereign bond risk premia reflect exposure to liquidity risk in addition to business-cycle risk and interest-rate risk. The data reject the null hypothesis that excluding unanticipated changes in liquidity improves the statistical fit of the model, suggesting an independent role for liquidity risk as a driver of the price of risk. The contribution of liquidity risk-to-risk premium associated with sovereign borrowing is about the same value in absolute magnitude as the contribution of business-cycle risk and it is larger than the contribution of interest-rate risk. These results therefore indicate that the liquidity premium is one of the key drivers of the price of sovereign debt, in addition to the business-cycle risk and interest-rate risk.

The principal lesson for sovereign borrowers is that lenders value liquidity in sovereign bond issues and that the liquidity premium is quantitatively important. Sovereign borrowers should find methods of promoting the liquidity of their bond issues by, for example, bundling together a series of small issues into a single large issue. Although such a recommendation seems obvious, the evidence presented in this paper establishes the quantitative importance of liquidity risk for sovereign borrowing costs.

Section 2 reviews the theoretical relationship between unexpected changes in liquidity risk and equity returns and makes the case for postulating a similar relationship in the sovereign bond market. Section 3 discusses the features of the 19th century sovereign debt that make it a natural laboratory for studying the relationship between sovereign bond risk premia and liquidity risk. Section 4 describes the data, the construction of the market liquidity measure, and some of the characteristics of this liquidity measure. Section 5 presents the asset-pricing methods for studying sovereign bond risk premia. Section 6 reports the evidence documenting the liquidity premium embedded in sovereign bond prices. Section 7 concludes.

4.2. Sovereign Bond Risk Premia and Liquidity Risk

Although economists have studied the relationship between risk premia and changes in market liquidity in the equity market, they have paid little attention to this relationship in the sovereign debt market. The evidence from the equity market suggests that stocks are differentially exposed to liquidity risk and that liquidity risk is a priced state variable (see, for example, Pástor and Stambaugh 2003; Acharya and Pedersen 2005; and Liu 2006).

The theoretical logic of this relationship is well-established. A security's risk premium is a linear function of its sensitivity to unexpected changes in pervasive risks (Merton 1973; Ross 1976; and Breeden 1979). Models that explicitly incorporate liquidity as a pervasive risk are consistent with such reasoning (Acharya and Pedersen 2005). In these models, unanticipated changes in pervasive liquidity covary positively with the business cycle. Risk-averse investors therefore demand a premium for holding securities whose payoffs also covary positively with these changes in pervasive liquidity. Thus a security's return is linearly related to its sensitivity to changes in unexpected changes in pervasive liquidity in the same way that the consumption-based capital asset pricing model predicts a security's return is linearly related to its sensitivity to unexpected changes in macroeconomic state variables.

The available evidence from the US Treasury market suggests a role for liquidity in the pricing of this debt. More liquid US Treasury issues trade at a premium compared to less liquid issues (Amihud and Mendelson 1991). On-the-run, or more recently issued, 30-year US Treasury bonds trade at a discount relative to off-the-run, or less recently issued, 30-year US Treasury bonds (Krishnamurthy 2002). In addition, there is indirect evidence of the influence of liquidity on sovereign bond prices in international financial markets. During periods of financial turbulence, US Treasuries act as a safe haven for international investors. Financial capital exhibits a flight to liquidity in conjunction with financial crises and the attendant decreases in liquidity (Longstaff 2004). A recent example of such a flight to liquidity with broad repercussions for international financial markets was the increase in demand for US Treasuries in August 1998 because of the Russian default.

This evidence is important because it suggests a role for liquidity in pricing sovereign debt that is traded internationally, although little attention has been paid to this relationship. Several papers study sovereign bond risk premia in the debt market for developed countries (Ilmanen 1995; and Barr and Priestley 2004), developing countries (Edwards 1986; and Durbin and Ng 2005), or the risk premia associated with syndicated bank loans to developing countries (Boehmer and Megginson 1990; and Eichengreen and Mody 2000). The papers that focus on the sovereign borrowing costs during earlier periods of extensive financial integration either study the sovereign lending boom that began in the 1920s and ended with the onset of the Great Depression (Eichengreen and Portes 1986; and Stone 1991), or they examine different questions (Mauro, Sussman, and Yafeh 2002). None of these papers studies liquidity risk as a determinant of sovereign borrowing costs.

This paper studies the influence of liquidity risk on sovereign bond risk premia. It shows that liquidity risk has an economically important and statistically significant effect on the magnitude of the sovereign bond risk premium and suggests an independent role for unexpected changes in liquidity as a pervasive risk in determining sovereign bond risk premia, above and beyond business-cycle risk and interest-rate risk.

These findings are important for two reasons. First, since the price of sovereign debt serves as a benchmark for the cost of private-sector debt in the economy, modeling the influence of liquidity risk on the cost of sovereign debt helps us to understand better the determinants of the price of corporate debt in these borrower countries, particularly those that rely on the international capital market to meet their financing needs (Dittmar and Yuan forthcoming). Second, to the extent that decreases in market liquidity are associated with financial distress and sovereign defaults, such evidence sheds light on the channels through which these episodes affect borrowing costs. In particular, the evidence presented in this paper isolates the effect of unexpected changes in liquidity on sovereign borrowing costs, in addition to other pervasive risks.

4.3. The 19th Century London Sovereign Debt Market as a Natural Experiment

The quality of the data set that this paper uses to study sovereign bond risk premia is hard to match even by contemporary standards. It contains the entire population of sovereign

bond prices available on the London Stock Exchange during the late 19th century. In light of London's importance as a financial center during this period, these data represent a substantial fraction of the universe of bonds traded in the world during that time. This time series is long enough to conduct asset-pricing tests and the cross-section of bonds diverse enough to guarantee dispersion in the factor sensitivities across sovereign portfolios. Moreover, since the bonds were traded in a single, centralized market, these data enable me to identify, in the econometric sense, the relationship between sovereign bond risk premia and liquidity risk without having to control for different levels of market liquidity in geographically separated markets. These features of the data make the London Stock Exchange during that period a natural environment for characterizing the effect liquidity risk on sovereign borrowing costs.

This paper uses these data not out of antiquarian interest but because conducting such a study with contemporary data is problematic. Although countries today increasingly rely on the international bond market for external finance, this development is recent. Prior to the 1970s, the last time that sovereigns relied on the international bond market was the 1920s and, prior to that, the period before World War One. Between 1930 and 1970, interest rate and foreign exchange controls sharply curtailed private international capital flows in general and cross-border bond investment in particular. Starting in the 1970s, the major financial centers relaxed these restrictions on international capital flows, facilitating the gradual expansion of the international bond market. Reliable sovereign prices, and bid-ask spreads in particular, are at best available from the late 1970s and, in practice, may be only available from the 1980s or even the 1990s: Although the secondary market for the sovereign debt of developing countries has existed since the mid-1980s, it was an over the counter market in which each transaction was negotiated in a complex and time-consuming process. This market structure ensured that trading was dominated by a small set of firms in New York and London that matched buyers and sellers as well as trading for their own accounts. Because of this market's lack of transparency, the data prior to the 1990s are considered unreliable (Stone 1991). There are thus at most 18 years of price data available for developing countries and, in practice, there are typically fewer data points available. Bid-ask spreads or other measures of market liquidity are often unavailable (see, for example, Durbin and Ng 2005). Thus,

even with monthly data and under the most optimistic assumptions about data availability, these time series are too short to conduct asset-pricing tests with statistical power. We are in the remarkable position that data from the past are more complete in this particular dimension than contemporary data and enable us to address a question of current interest.

This paper's contribution is to use the unique economic environment of the London Stock Exchange as a laboratory to study the relationship between sovereign bond risk premia and liquidity risk. These price data provide us with an opportunity to learn more about the pervasive risks that affect the price of sovereign debt. As this paper documents, liquidity risk has an economically important and statistically significant effect on the price of sovereign debt.

4.4. Data

The data consist of the closing bid and ask prices as well as the coupon payments for all of the sovereign bonds traded on the London Stock Exchange between 1866 and 1907. This data set contains 213 bonds issued by 38 countries and 110 bonds issued by 11 British colonies. The prices were sampled every 28 days from the official quotation list published in the *Money Market Review* between 1866 and 1907, when such detailed price quotations ceased to be available, and were supplemented with quotations from the *Times of London* and the *Economist*. These price series comprise the entire universe of sovereign bonds traded on the London Stock Exchange during that time. The number of sovereign bond prices quoted in the *Money Market Review* does not exceed four until late 1870 and there are fewer than 10 sovereign bond price quotations until December 1870. Appendix 1 contains further details on the construction of the data set and a complete list of all of the countries represented in it.

4.4.1. Holding-Period Returns

Because yields from the London bond market are unavailable during this period, I estimate borrowing costs using holding-period returns. The net holding-period return r_{t+1} is

$$1 + r_{t+1} = E_t \left[\frac{P_{t+1} + C_{t+1}}{P_t} \right].$$

where P_{t+1} denotes the price of the sovereign bond at $t + 1$; and C_{t+1} denotes the coupon payment, if any, at $t + 1$. Expected returns are not observable, but they can be estimated using ex post realized returns. If expectations are rational, then the realized ex post return of the bond is its expected return plus an independent mean zero random error term $r_{t+1} = E_t[r_{t+1}] + \varepsilon_{t+1}$. The sample average of holding-period returns thus converges to its expected value and average ex post holding-period returns are a good proxy for unobservable expected returns.

When the sovereign bond is actively traded in a secondary market, holding-period returns take account of the changing opportunity cost of borrowing and lending over time, thereby capturing the time-series variation in sovereign borrowing costs. These price changes can reflect two types of risk. The first type of risk is devaluation risk, which captures debtor countries' incentive to inflate away the value of their debts that is denominated in domestic currency. This risk exists in economic environments with floating exchange rates. The second type of risk is the risk associated with outright default or repudiation of the foreign debt. Most of the sovereign bonds traded in London were denominated in pounds, although there were important exceptions to this rule. Some countries in Western Europe and North America such as France, Germany, and the US issued debt denominated in their domestic currencies. To the extent that there was devaluation risk embedded in these bond prices, the gold standard, the fixed exchange rate regime that prevailed during the late 19th century, ensured that these two risks were, for practical purposes, perfectly correlated: The countries that defaulted on their debts did so in conjunction with suspending the gold standard, which entailed devaluing the domestic currency (Bordo and Rockoff 1996). Empirically, every sovereign borrower who defaulted also left gold. In this economic environment, these two types of risk are not separately identifiable, but the external constraints that the gold standard imposed on sovereign borrowers and the empirical features of default in the 19th century suggest that the risk premium on sovereign debt reflected default risk rather than currency risk.

Thus changes in the bond's price, adjusted for coupon payments and the appropriate risk-free interest rate, are good indicators of changing expectations about the price of a sovereign bond. Alternative measures of borrowing costs such as the coupon yield do not capture the time-series variation in capital gains and losses associated with price changes and ignore information about changes in market expectations, as reflected in such price changes. Moreover, empirical asset-pricing methods can decompose holding-period returns into the components due to pervasive risks and assess each component's importance for pricing sovereign debt.

4.4.2. Market Liquidity

Studies have adopted alternative measures of liquidity depending on data availability (see, for example, Pástor and Shambaugh 2003; and Bekaert, Harvey, and Lundblad 2007), and there is no strong consensus which empirical measure of liquidity is best (Korajczyk and Sadka 2008). This study constructs a measure of liquidity based on the bid-ask spread – that is, the difference between the prices at which a trader is able to buy and sell a security relative to the midpoint of these two prices.

The economic logic of using the bid-ask spread as a measure of the liquidity relies upon models in which market makers set these spreads to protect themselves against trading with informed traders. The bid-ask spread for an individual security emerges as the market maker's compensation for his losses from trading with these informed traders (Copeland and Galai 1983; Glosten and Milgrom 1985; and Kyle 1985). It is indicative of the amount of asymmetric information associated with each security and the market-wide liquidity index based on each individual security's spread is intended to measure the aggregate amount of asymmetric information in the market. During periods of market turbulence, this index should widen, indicating greater asymmetric information in the market (see, for example, Chordia, Sarkar, and Subrahmanyam 2005).

To construct a measure of market liquidity in the London sovereign bond market, I use the bid-ask spreads from the bonds trading at a given point in time. The bid and ask prices from which I compute the bid-ask spread and the measure of market liquidity span

January 1871 to December 1907, yielding 482 time-series observations.¹⁶ The bid-ask spread for bond i is:

$$(1) \quad S_{it} = \frac{P_{it}^a - P_{it}^b}{0.5(P_{it}^a + P_{it}^b)}$$

where P_{it}^a is the ask price of the bond; and P_{it}^b is the bid price of the bond. The equally weighted bid-ask spread index for the entire market is an index of market liquidity for the entire market:

$$(2) \quad L_t = N_t^{-1} \sum_{i=1}^{N_t} S_{it}$$

where N_t is the number of bonds trading in period t . With the market values of each bond in a particular period, one could construct a value-weighted index of bid-ask spreads. A value-weighted portfolio would assign weights to each bid-ask spread based upon the market value of the shares outstanding. For example, if the market value of Argentine bonds were ten times the value of Danish bonds, Argentina would receive ten times the weight of Denmark.

There is an advantage to using the equally weighted bid-ask spread rather than the value-weighted spread as a measure of market liquidity, however. Doing so ensures the index is not skewed toward bonds with large market capitalizations, which tend to be immune to market-wide decreases in liquidity. The equally weighted index is a broader measure of market liquidity than one that weights larger, more liquid bonds more heavily.

To capture the time-series variation in liquidity, I define the unexpected change in market liquidity between $t-1$ and t as:

$$\Delta L_t \equiv L_t - E_t[L_{t-1}]$$

That is, I take the change between the current market liquidity and expected market liquidity. To avoid introducing a negative correlation between this unexpected change in market liquidity and sovereign returns, I set $E_t[L_{t-1}] = L_{t-2}$ to obtain

$$(3) \quad \Delta L_t = L_t - L_{t-2}$$

¹⁶There are 482 rather than 481 observations because there were 14 rather than the usual 13 issues of the Money Market Review published in 1897.

This measure of market liquidity is an ex ante indicator of changes in market liquidity because it postulates that expectations are formed prior to period $t - 1$ and does not use information that would help to predict excess returns on sovereign bonds in period t , which are a function of the price at $t - 1$. Using innovations to liquidity stems from the rationale that only unanticipated changes in pervasive risks are priced states variables and are thus the variables of interest in asset-pricing tests (Chen, Roll, and Ross 1986).

4.4.3. Time-Series Features of Bond Market Liquidity

Table 2 summarizes the time-series features of market liquidity in the London sovereign bond market between 1871 and 1907. The first column shows the properties of the level series. The average market liquidity is about 2.4%, and it is statistically different from zero. The range between the minimum and maximum values for market-wide bond liquidity indicates the bid-ask spread lies between about 1.5 and 4.5%. To put these values into perspective, these bid-ask spreads are much wider than the spreads in the secondary market for Japanese government bonds during the first half of the 1980s (Takagi 1987) or the bid-ask spreads in the market for US Treasury bills during the second half of the 1990s (Fleming 2002). In both markets the bid-ask spreads are smaller than 0.5% and most were smaller than 0.10%. These differences make sense insofar as the Japanese and US government bond markets serve as international benchmarks and are among the largest and most liquid in the world. The level series is highly persistent. The sum of the autoregressive coefficients for this measure of market liquidity is 0.93, which is consistent with the magnitude documented by studies of the equity market in the current period (Korajczyk and Sadka 2008). Figure 1 provides visual evidence that this measure of market liquidity is persistent.

The second column in Table 2 shows the properties of the change in the level of the bid-ask spread in percentage points. One cannot reject the null hypothesis that the mean of the series is zero. The differenced series is not persistent. The differenced series appears stationary with a zero mean.

Figure 2 and Figure 3 provide visual evidence that the level and the change in market-wide sovereign bond liquidity identify periods in which investors' perceptions of the risks associated with foreign lending shifted. These two figures plot the series and the

dates associated with selected major defaults during this period. As the figures illustrate, bid-ask spreads tend to be wide during periods when sovereign borrowers default on their external obligations; these defaults are also associated with sharp increases in the spread. For example, the largest spike in the plot during 1875-76 is associated with the Ottoman Empire's default in October 1875 and Egypt's default in April 1876, two of the largest borrowers in the sovereign debt market during that period. These defaults involved lengthy negotiations: The Ottoman Empire reached a final settlement with its creditors in 1881 and Egypt in 1880 (Suter 1990). Market-wide bid-ask spreads also widened in anticipation of the Argentine default in April 1890, which precipitated the Baring crisis in November of that year, and the Brazilian default in June 1898. Although the market-wide bid-ask spread did not widen in anticipation of Greece's default in December 1893, they did widen in the month of the actual default. Overall, this evidence suggests that the equally weighted bid-ask spread broadly captures market liquidity conditions.

4.5. Empirical Methods

Arbitrage-pricing theory predicts that the expected returns on securities are a linear function of their exposure to pervasive risks and that these betas describe the cross-section of security returns. Asset pricing models have the generic form

$$(4) \quad E[r_i] = \lambda_0 + b_{i1}\lambda_1 + \dots + b_{ik}\lambda_k$$

where $E[r_i]$ is the expected return of security i ; λ_0 is expected zero-beta or risk-free rate of return; b_{ki} is the beta of security i with respect to the k th pervasive risk; and λ_k is the risk premium associated with the k th pervasive risk (Ross 1976).

To estimate the models of the form in equation (4), I first characterize the exposure of sovereign bond returns to unexpected changes in market liquidity over time. This exercise establishes that unexpected changes affect sovereign bond returns and the magnitude of the effect is a function of the sovereign bond's observable characteristics. To do so, I run time-series regressions on test portfolios sorted by the ex ante bid-ask spread and the size of the bond, as measured by the bond's market capitalization. Sorting the individual sovereign bonds by this observable characteristic generates cross-sectional

dispersion in the portfolios' sensitivities to pervasive risks and econometrically identifies this variability across portfolios.

Starting in January 1871, I assign the returns of each sovereign bond into one of ten portfolios based on the size of the each bond's bid-ask spread. I link the returns from the constituent bonds in a set of ten value- and equally weighted test portfolios. I repeat the procedure each month, thereby generating ten separate time series of returns. I estimate the time-series regression

$$(5) \quad r_{it} - r_{ft} = \alpha_i + \beta_B(r_{Bt} - r_{ft}) + \beta_G(r_{Gt} - r_{ft}) + \beta_L \Delta L_t + \epsilon_i$$

for each of the $i=1, \dots, 10$ value- or equally weighted portfolios of sovereign bond portfolios sorted by the ex ante bid-ask spread. In equation (5), r_{it} is the return on portfolio i ; r_{ft} is the proxy for the risk-free rate, the London banker bill rate; α_i is the portfolio-specific risk premium and reflects the risk associated with holding portfolio i ; r_{Bt} is the return of the value-weighted foreign sovereign bond index, which is the market portfolio of all the sovereign bonds traded on the London Stock Exchange; and r_{Gt} is the return of the British consol; and ΔL_t is the unexpected change in liquidity in the sovereign bond market, constructed as in equation (3).

I include the value-weighted foreign bond index as a pervasive risk based on evidence that the excess returns of a bond market index drive most of the time-series variation in sovereign bond risk premia (Ilmanen 1995; and Barr and Priestley 2004); and I include the excess returns of the British consol to capture changes in British interest rates as a pervasive risk because models of sovereign borrowing postulate the price of sovereign debt reflects the return on alternative investments, as proxied by the world interest rate (Fernández-Arias 1995). The available evidence supports this hypothesis: Much of the variation in the developing-country spreads can be accounted for by changes in US interest rates (Uribe and Yue 2006).

Including the unexpected change in market liquidity as an additional factor in (5) enables me to decompose the sovereign bond risk premia into the components attributable to business-cycle risk, as measured by the excess returns of the foreign bond index, interest-rate risk, as measured by the excess returns of the British consol index,

and the risk associated with unexpected changes in liquidity. In particular, it identifies the component of the risk premium associated with such unanticipated changes in liquidity and the variation in the liquidity beta β_{L_i} across the bid-ask spread-sorted test portfolios. Economic theory predicts the liquidity betas will be negative – an increase in the market-wide bid-ask spread is a decrease in market liquidity, which depresses contemporaneous security prices and reduces security returns (Acharya and Pedersen 2005). As I discuss below, the evidence from these regressions supports this hypothesis.

In light of this evidence that innovations in liquidity explain the time-series variation in bond risk premia, I form a liquidity-mimicking portfolio IML_t that reflects the time-series behavior of pervasive liquidity risk to assess if this liquidity factor is a priced state variable. This liquidity-mimicking portfolio is the difference in returns between the least liquid and the most liquid portfolio from the ten test portfolios sorted by the bid-ask spread

$$(6) \quad IML_t = r_{It} - r_{Lt}$$

where r_{It} denotes the returns of the most illiquid sovereign bond portfolio; and r_{Lt} denotes the returns of the most liquid sovereign bond portfolio. This liquidity-mimicking portfolio is identical to postulating that an investor holds a leveraged portfolio that invests £1 in the least liquid sovereign bond portfolio and sells £1 of the most liquid sovereign bond portfolio.

I use the liquidity-mimicking portfolio IML_t rather than ΔL_t to test if market liquidity is a priced state variable. The rationale for constructing such a liquidity-mimicking portfolio is that the returns associated with this portfolio mimic the returns associated with hedging liquidity risk and allows me to compute the risk premium $\hat{\lambda}_L$ associated with liquidity risk. Such a procedure captures the returns associated with a hedging portfolio that is exposed to particular pervasive risks (see, for example, Chen, Roll, and Ross 1986; Breeden, Gibbons, and Litzenberger 1989; and Fama and French 1996).

I postulate a three-factor model for sovereign bond returns based on the excess returns of the foreign bond index; the excess returns of the British consol; and the return

differential on the most illiquid and the most liquid portfolios. I estimate the risk premia $\hat{\lambda}_B$, $\hat{\lambda}_C$, $\hat{\lambda}_L$ implied by (4) using efficient Generalized Method of Moments (GMM). Appendix 2 discusses this procedure in more detail. Using GMM to estimate the risk premia enables me to test formally the fit of the set of models and the ability of the liquidity premium to explain cross-sectional differences in sovereign bond risk premia using the D -test (Newey and West 1987b). The evidence provided by the results of these tests suggests that liquidity risk is a priced state variable that is important for pricing sovereign debt.

4.6. Empirical Evidence

4.6.1. Spread-Sorted Test Portfolios

Table 3 and Table 4 contain average excess returns and regression results for the ten value-weighted portfolios sorted by the bid-ask spread. Table 3 does not include the unexpected change in market liquidity as an independent variable whereas Table 4 does include it. The results in the first row of Table 4 show that, unconditionally, the portfolio with the least liquid sovereign bonds earns, on average, an annualized excess return of 5.10% whereas the portfolio with the most liquid securities earns, on average, an annualized excess return of 2.48%. Less liquid sovereign bonds thus command a higher return. Controlling for differences in business-cycle risk and interest-rate risk, the returns from the portfolio of illiquid sovereign bonds are larger than the returns from the portfolio of liquid bonds. The last column in Table 3 shows that a strategy of investing £1 in the least liquid sovereign bond portfolio and selling £1 of the most liquid sovereign bond portfolio would have generated an average annual excess return of 4%. This evidence suggests that the risk premium is larger for illiquid securities. The most liquid bonds are more sensitive to business-cycle risk, as captured by the excess returns of the foreign bond index, than the least liquid bonds. At almost 50 basis points per month, this difference is substantial and statistically significant, suggesting that illiquid sovereign bonds are more sensitive to business-cycle risk than liquid sovereign bonds. The least liquid bonds are more exposed to the risk associated with changes in British interest rates, but this difference is not statistically different from zero. These findings indicate that

there are systematic differences between liquid and illiquid sovereign bonds and that they are attributable to their differential exposure to the excess returns of the foreign bond index. As Table 4 reports, an unexpected decrease in liquidity reduces contemporaneous returns, which is consistent with economic theory. When the bid-ask spread unexpectedly widens and market liquidity declines, sovereign bond prices fall and generate negative returns. The least liquid sovereign bonds are more sensitive to decreases in market-wide bond liquidity than the most liquid bonds: On average, a one percentage point increase in the market-wide bid-ask spread reduces the returns on the least liquid bonds by 1.68% in a given month whereas a similar change in market-wide bond liquidity reduces the returns on the most liquid bonds by 0.73% in a given month. Both betas are statistically different from zero. The last column shows, however, that the difference between the two betas is not statistically significant at conventional levels.

Table 5 and Table 6 show the results for equally weighted sovereign bond portfolios as test portfolios. They are similar to those reported in Table 3 and Table 4, but at about 3.1% per year the premium for holding illiquid sovereign bonds is somewhat smaller than that implied by the value-weighted portfolios. Again, the most liquid sovereign bonds are more exposed to the risk associated with changes in the business-cycle risk, as proxied by the excess returns of the foreign bond index, and this difference is statistically significant. The difference between the exposure to changes in British interest rates of the least liquid and the most liquid portfolios is smaller in magnitude than reported in Table 3 and Table 4, and it remains statistically insignificant.

Table 6 indicates that including the unexpected change in market liquidity as a nontraded factor does not materially alter the results compared to those reported in Table 4, but the sensitivity of the least liquid portfolio to innovations in market liquidity is statistically different from the sensitivity of the most liquid portfolio to innovations in market liquidity. A one percentage point increase in the market-wide bid-ask spread reduces the returns on the most liquid sovereign bond portfolio by 1.72% more than such a change reduces the return on the most liquid sovereign bond portfolio. This finding is intuitive: Less liquid sovereign bonds are harder to trade during a credit crunch and, consequently, such bonds tend to carry larger betas on unexpected changes in market liquidity.

The excess returns of the illiquid portfolio do not reflect the differential exposure of these bonds to the excess returns of the foreign bond index. That the least liquid bonds are less sensitive to business-cycle risk than the most liquid bonds suggests that the larger excess returns these bonds earn is not compensation for bearing more aggregate risk. Instead these excess returns reflect compensation for bearing the risk associated with the least liquid portfolio's sensitivity to market liquidity and changes in British interest rates.

4.6.2. Size-Sorted Test Portfolios

Table 7 and Table 8 report the regression results where the test portfolios are value-weighted portfolios sorted by size. Table 7 shows that the small sovereign bond portfolio earns a 3.47% premium relative to the large sovereign bond portfolio, after adjusting for business-cycle risk and interest-rate risk. Large sovereign bonds tend to be more exposed than small sovereign bonds to the risk associated with the excess returns on the foreign bond portfolio but less exposed to risk associated with changes in British interest rates. Both differences are statistically significant, but the difference in the betas associated with the excess returns of the foreign bond index is larger in magnitude. On average, a 1% increase in the excess return of the foreign bond portfolio reduces the returns of the leveraged portfolio by 1.2% per month. By way of contrast, on average a 1% increase in the returns on the British consol index increases the returns of the leveraged portfolio by 0.23% in a given month.

Table 8 repeats the regressions from Table 7 but adds the portfolio IML_t as a traded factor. In every case except one, the estimated beta on the liquidity-mimicking portfolio is statistically different from zero. The largest and smallest sovereign bond portfolios carry different sensitivities to the liquidity-mimicking portfolio. The portfolio of the smallest sovereign bonds carries a loading that is 15 basis points larger than the loading associated with the portfolio of the largest sovereign bonds, suggesting that smaller sovereign bonds are more exposed to liquidity risk. Smaller sovereign bonds are also less sensitive to excess returns of the foreign bond index and more sensitive to changes in British interest rates. These finding makes sense insofar as smaller sovereign bond issues are less liquid than larger sovereign bond issues and are consistent with the results in Tables 3-6.

Table 9 and Table 10 repeat the regressions in Table 7 and Table 8 but use equally weighted portfolios of sovereign bond returns sorted by size. The results are very similar. Again, the portfolio of the smallest sovereign bonds is more sensitive to unexpected changes in market liquidity in the sovereign bond market, as captured by the liquidity-mimicking portfolio, although the difference between the sensitivities of the smallest and the largest sovereign bond portfolios to the liquidity-mimicking portfolio is not statistically different from zero. This finding indicates that the betas in this specification exhibit less dispersion than the betas in the specification with value-weighted portfolios and suggests that there is larger estimation error associated with the smaller sovereign bond issues, which receive the same weight as the larger bond issues, and whose returns tend to be more volatile.

Taken together, these results suggest that the difference in returns between the least liquid portfolio and the most liquid portfolio can explain the differences in the variation of sovereign bond risk premia that are sorted by size. Smaller sovereign bonds are more exposed to liquidity risk than larger sovereign bonds, and this difference is both economically and statistically significant.

4.6.3. Sovereign Bond Risk Premia and the Liquidity Premium

Table 11 reports the efficient GMM estimates of the risk premia. The first column in the table (Unrestricted) reports the estimates for the risk premia from the model with all three risk factors. All of the estimated risk premia are statistically different zero. A portfolio that had a beta of one with respect to the excess returns of the foreign bond portfolio and betas of zero with respect to the other two pervasive risks would have earned about 3.8% per year; a portfolio that had a beta of one with respect to the excess returns of the British consol and betas of zero with respect to the other two pervasive risks would have earned about 4.1% per year; and a portfolio that had a beta of one with respect to the excess returns of the liquidity-mimicking portfolio IML_t and betas of zero with respect to the other two pervasive risks would have earned about 22.6% per year. The component of a portfolio i 's expected return that is attributable to liquidity risk is given by the product of the liquidity beta with the risk premium $\beta_{L_i} \lambda_L$. The contribution of liquidity risk to the expected return of the SML portfolio is 3.38% per year (0.15×0.226). To put this

value into perspective, the contribution of business-cycle risk is -4.48% per year (-1.14×0.0375) and that for interest-rate risk is 0.79% per year (0.19×0.0413). The negative contribution of business-cycle risk arises because the *SML* portfolio has a negative exposure to business-cycle risk: The beta with respect to the foreign bond index on the small bond portfolio is smaller in size than the beta with respect to the foreign bond index associated with the large bond portfolio. The contribution of liquidity risk to expected returns is thus comparable in magnitude to the contribution of business-cycle risk and larger than the contribution of interest-rate risk.

The second column (Restricted 1) shows the estimated risk premia for the model excluding the excess returns of the foreign bond portfolio. Both of the risk premia are economically and statistically significant. The estimated risk premium for the British consol is about 9.8% per year; and the estimated risk premium for the liquidity-mimicking factor is about 14.2% per year. Although smaller in magnitude than the estimated liquidity premium in the unrestricted specification, the liquidity premium in this specification remains sizable. Again, this evidence indicates the importance of the liquidity premium in explaining sovereign bond risk premia. The results of the *D*-test suggest that one can reject the null hypothesis that the excess return of the British consol and the liquidity-mimicking factor by themselves provide a statistically adequate characterization of the sovereign bond risk premia.

The third column (Restricted 2) shows the estimated risk premia for the specification that excludes the excess returns of the British consol. The estimated risk premium for the foreign bond index is about 3.9% per year and that for the liquidity premium is about 22.5% per year. The *D*-test rejects the null hypothesis that including the excess returns of the British consol in the unrestricted model does not statistically improve the performance of this model.

The fourth column (Restricted 3) reports the estimates from the specification excluding the liquidity-mimicking factor. The estimated risk premia for the foreign bond index and the British consol are both statistically significant. Importantly, the *D*-test rejects this specification in favor of the unrestricted model that includes the excess returns on the liquidity-mimicking factor. This finding therefore suggests that one should

include liquidity-mimicking factor in the model to provide an adequate statistical characterization of sovereign bond risk premia.

All of the evidence points in the same direction. The liquidity premium is statistically significant and, based on the estimates from the unrestricted model, the contribution of liquidity risk to the expected returns of the *SML* portfolio is economically meaningful. In each of these specifications, the liquidity premium is statistically significant. Moreover, one cannot reject the null hypothesis that all three of the pervasive risk factors are needed to characterize adequately the data. This finding is important because it shows that innovations in liquidity are a priced state variable, above and beyond business-cycle risk, as captured by the excess returns of the foreign bond index, and interest rate risk, as captured by the excess returns of the British consol. Lenders in the London sovereign debt market valued liquidity and demanded a premium for bearing the risk associated with holding sovereign debt exposed to such risk.

4.7. Conclusion

Events such as the flight to liquidity during the 1998 Russian financial crisis suggest a strong role for the liquidity risk on the price of sovereign debt traded internationally, particularly during periods of financial upheaval. Notwithstanding the importance of uncovering the quantitative importance of liquidity risk in the price of sovereign debt, contemporary data are ill-suited to addressing studying the relationship between sovereign bond risk premia and liquidity risk. In contrast, the data used in this paper from the London Stock Exchange can shed light on this relationship. The late 19th century was the last time that a large and diverse sample of sovereign bonds with observable bid-ask spreads was traded on a single, centralized market, denominated in a single currency, and for a period sufficiently long to conduct asset-pricing tests.

Empirical analysis of these data establish that sovereign bonds with wide bid-ask spreads earn about 3-4% more per year than sovereign bonds with narrow bid-ask spreads. This premium for holding less liquid sovereign bonds reflects the greater sensitivity of such bonds to variations in unexpected changes in liquidity than more liquid sovereign bonds. Moreover, the data show that the excess returns associated with a leveraged portfolio long the least liquid bonds and short the most liquid bonds is a state

variable important for pricing sovereign bond risk premia. That is, cross-sectional differences in sovereign bond risk premia are linearly related to the sensitivities of returns to fluctuations in market liquidity, as captured by this liquidity-mimicking portfolio. The liquidity premium's contribution to the risk premium associated with sovereign debt is economically large and statistically significant. It is comparable in absolute magnitude to the contribution of business-cycle risk to expected returns and larger than the contribution of interest-rate risk. The significance of liquidity risk is robust to controlling for business-cycle risk and interest rate risk, suggesting an independent role for liquidity risk, above and beyond these other two pervasive risks. Overall, this evidence suggests a strong role for liquidity risk as a driver of sovereign bond risk premia.

From a policy perspective, the quantitative importance of the liquidity premium's contribution to expected returns suggests that sovereign borrowers can reduce their borrowing costs by floating more liquid issues. As this paper shows, larger issues tend to be less exposed to liquidity risk, so one way of improving the liquidity of a bond issue to increase its size. Thus, to the extent that sovereign borrowers can roll smaller bond issues into a single large issue, doing so would reduce the issue's exposure to liquidity risk and hence sovereign borrowing costs. Although this recommendation is consistent with economic intuition, the evidence in this paper indicates that the liquidity premium can add substantially to borrowing costs and identifies it as a key determinant of the sovereign risk premium.

Table 4-1. Developing Country Bond Issues, Equity Issues, and Loan Commitments, 1990-2007 (cont.)

(USD billions)

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Bonds	7.8	13.9	24.4	62.7	56.5	59.2	103.0	126.2	79.5	82.4	80.5
Equities	1.2	5.6	7.2	11.9	18.0	10.0	17.8	26.2	9.4	23.2	41.8
Loans	28.4	50.7	42.5	43.0	55.2	82.0	89.0	122.5	60.0	58.1	94.2

Source: International Monetary Fund.

Table 4-1. Developing Country Bond Issues, Equity Issues, and Loan Commitments, 1990-2007 (concl.)

(USD billions)

	2001	2002	2003	2004	2005	2006	2007
Bonds	89.4	65.0	100.5	135.6	186.6	179.9	207.9
Equities	11.2	16.5	27.7	45.8	85.5	121.4	170.7
Loans	64.2	82.5	97.6	148.3	189.7	252.1	282.9

Source: International Monetary Fund.

Table 4-2. Time-Series Features of Bond Market Liquidity

	Level	Change
	(percent)	(percentage points)
Mean	2.38	0.00
(<i>p</i> -value)	(0.00)	(0.48)
Mean Abs. Dev.	2.38	0.14
Min.	1.48	-1.16
Max.	4.46	1.56
Persistence	0.93	0.15

Notes: The sample is 1871.1-1907.12 (482 observations). The *p*-values of the test for a zero mean are based on heteroskedasticity and autocorrelation robust standard errors (Newey and West 1987a). The measure of persistence is the sum of the autoregressive coefficients (Andrews and Chen 1994). The autoregressive lag order is determined using the AIC with an upper bound of 26 lags (2 calendar years).

Table 4-3. Value-Weighted Sovereign Bond Portfolios Sorted by Spread

Dependent Variable: $r_{it} - r_{ft}$

	Liquid	2	3	4	5	6	7	8	9	Illiquid	<i>IML</i>
Annualized $\bar{r}_i - \bar{r}_f$	2.48	3.13	2.88	2.50	2.60	1.90	2.49	4.58	3.50	5.10	2.62
Annualized $\hat{\alpha}_i$	0.79	2.74	2.22	2.16	2.39	1.51	1.93	4.08	3.64	4.79	4.00
	(0.23)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)	(0.03)
$\hat{\beta}_B$	0.56	0.12	0.21	0.10	0.06	0.12	0.17	0.16	-0.06	0.09	-0.47
	(0.00)	(0.00)	(0.09)	(0.24)	(0.28)	(0.20)	(0.14)	(0.20)	(0.41)	(0.40)	(0.05)
$\hat{\beta}_C$	0.14	0.41	0.42	0.33	0.28	0.31	0.45	0.33	0.53	0.40	0.26
	(0.05)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)	(0.16)
\bar{R}^2	0.34	0.11	0.16	0.10	0.08	0.11	0.18	0.08	0.08	0.03	0.09

Notes: The estimation period is 1871.1-1907.12 (482 observations). *IML* denotes the leveraged portfolio that is long £1 of the most illiquid portfolio and short £1 of the most liquid portfolio. All of the returns are taken in excess of the risk-free rate r_{ft} , the London banker bill rate. The p -values in parentheses are based on the heteroskedasticity and autocorrelation robust standard errors (Newey and West 1987a) and the asymptotic critical values.

Table 4-4. Value-Weighted Sovereign Bond Portfolios Sorted by Spread

Dependent Variable: $r_{it} - r_{ft}$

	Liquid	2	3	4	5	6	7	8	9	Illiquid	<i>IML</i>
Annualized $\bar{r}_i - \bar{r}_f$	2.48	3.13	2.88	2.50	2.60	1.90	2.49	4.58	3.50	5.10	2.62
Annualized $\hat{\alpha}_i$	0.77	2.74	2.22	2.14	2.38	1.51	1.92	4.07	3.63	4.76	3.98
	(0.23)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)
$\hat{\beta}_B$	0.57	0.12	0.21	0.11	0.07	0.12	0.18	0.16	-0.06	0.10	-0.47
	(0.00)	(0.28)	(0.08)	(0.22)	(0.26)	(0.19)	(0.13)	(0.19)	(0.41)	(0.38)	(0.05)
$\hat{\beta}_C$	0.14	0.41	0.42	0.33	0.28	0.31	0.45	0.33	0.53	0.40	0.26
	(0.05)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)	(0.16)
$\hat{\beta}_L$	-0.73	-0.10	-0.09	-0.84	-0.45	-0.26	-0.49	-0.42	-0.57	-1.68	-0.95
	(0.08)	(0.38)	(0.42)	(0.07)	(0.07)	(0.23)	(0.21)	(0.28)	(0.16)	(0.07)	(0.14)
\bar{R}^2	0.34	0.11	0.16	0.11	0.08	0.11	0.18	0.09	0.08	0.05	0.10

Notes: The estimation period is 1871.1-1907.12 (482 observations). *IML* denotes the leveraged portfolio that is long £1 of the most illiquid portfolio and short £1 of the most liquid portfolio. All of the returns are taken in excess of the risk-free rate r_{ft} , the London banker bill rate. The p -values in parentheses are based on the heteroskedasticity and autocorrelation robust standard errors (Newey and West 1987a) and the asymptotic critical values.

Table 4-5. Equally Weighted Sovereign Bond Portfolios Sorted by Spread

Dependent Variable: $r_{it} - r_{ft}$

	Liquid	2	3	4	5	6	7	8	9	Illiquid	<i>IML</i>
Annualized $\bar{r}_i - \bar{r}_f$	2.31	2.78	2.76	2.01	2.38	1.83	2.18	3.42	2.99	4.50	2.19
Annualized $\hat{\alpha}_i$	1.48	2.50	2.29	1.74	2.28	1.57	1.75	3.25	3.22	4.58	3.10
	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.02)
$\hat{\beta}_B$	0.27	0.08	0.15	0.08	0.03	0.08	0.13	0.05	-0.10	-0.04	-0.31
	(0.04)	(0.49)	(0.11)	(0.20)	(0.37)	(0.23)	(0.16)	(0.41)	(0.36)	(0.45)	(0.03)
$\hat{\beta}_C$	0.31	0.35	0.24	0.22	0.24	0.25	0.38	0.31	0.51	0.37	0.06
	(0.00)	(0.00)	(0.03)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)	(0.36)
\bar{R}^2	0.24	0.09	0.07	0.10	0.08	0.12	0.13	0.05	0.08	0.02	0.06

Notes: The estimation period is 1871.1-1907.12 (482 observations). *IML* denotes the leveraged portfolio that is long £1 of the most illiquid portfolio and short £1 of the most liquid portfolio. All of the returns are taken in excess of the risk-free rate r_{ft} , the London banker bill rate. The p -values in parentheses are based on the heteroskedasticity and autocorrelation robust standard errors (Newey and West 1987a) and the asymptotic critical values.

Table 4-6. Equally Weighted Sovereign Bond Portfolios Sorted by Spread

Dependent Variable: $r_{it} - r_{ft}$

	Liquid	2	3	4	5	6	7	8	9	Illiquid	<i>IML</i>
Annualized $\bar{r}_i - \bar{r}_f$	2.31	2.78	2.76	2.01	2.38	1.83	2.18	3.42	2.99	4.50	2.19
Annualized $\hat{\alpha}_i$	1.48	2.49	2.28	1.73	2.28	1.56	1.73	3.23	3.20	4.54	3.07
	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.02)
$\hat{\beta}_B$	0.27	0.09	0.16	0.09	0.03	0.08	0.14	0.06	-0.08	-0.02	-0.29
	(0.04)	(0.31)	(0.09)	(0.18)	(0.35)	(0.22)	(0.14)	(0.39)	(0.37)	(0.47)	(0.03)
$\hat{\beta}_C$	0.31	0.35	0.23	0.22	0.24	0.325	0.38	0.31	0.51	0.37	0.06
	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)	(0.37)
$\hat{\beta}_L$	-0.21	-0.13	-0.78	-0.59	-0.34	-0.31	-0.65	-0.75	-1.14	-1.94	-1.72
	(0.25)	(0.32)	(0.04)	(0.14)	(0.06)	(0.16)	(0.20)	(0.11)	(0.01)	(0.05)	(0.05)
\bar{R}^2	0.24	0.09	0.08	0.07	0.08	0.11	0.14	0.09	0.09	0.05	0.08

Notes: The estimation period is 1871.1-1907.12 (482 observations). *IML* denotes the leveraged portfolio that is long £1 of the most illiquid portfolio and short £1 of the most liquid portfolio. All of the returns are taken in excess of the risk-free rate r_{ft} , the London banker bill rate. The p -values in parentheses are based on the heteroskedasticity and autocorrelation robust standard errors (Newey and West 1987a) and the asymptotic critical values.

Table 4-7. Value-Weighted Sovereign Bond Portfolios Sorted by Size

Dependent Variable: $r_{it} - r_{ft}$

	Small	2	3	4	5	6	7	8	9	Large	<i>SML</i>
Annualized $\bar{r}_i - \bar{r}_f$	4.05	2.98	3.84	3.94	4.58	3.26	3.97	2.52	3.44	4.16	-0.11
Annualized $\hat{\alpha}_i$	3.83	3.45	4.03	3.89	4.31	3.11	3.42	2.15	2.77	0.37	3.47
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.15)	(0.00)
$\hat{\beta}_B$	0.07	-0.17	-0.07	0.01	0.08	0.04	0.18	0.11	0.21	1.28	-1.21
	(0.06)	(0.15)	(0.34)	(0.48)	(0.29)	(0.28)	(0.20)	(0.22)	(0.17)	(0.00)	(0.00)
$\hat{\beta}_C$	0.15	0.37	0.22	0.30	0.27	0.38	0.34	0.38	0.43	-0.08	0.23
	(0.00)	(0.00)	(0.03)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
\bar{R}^2	0.02	0.16	0.02	0.06	0.03	0.09	0.08	0.18	0.19	0.96	0.74

Notes: The estimation period is 1871.1-1907.12 (482 observations). *SML* denotes the leveraged portfolio that is long £1 of the smallest portfolio and short £1 of the largest portfolio. All of the returns are taken in excess of the risk-free rate r_{ft} , the London banker bill rate. The p -values in parentheses are based on the heteroskedasticity and autocorrelation robust standard errors (Newey and West 1987a) and the asymptotic critical values.

Table 4-8. Value-Weighted Sovereign Bond Portfolios Sorted by Size

Dependent Variable: $r_{it} - r_{ft}$

	Small	2	3	4	5	6	7	8	9	Large	<i>SML</i>
Annualized $\bar{r}_i - \bar{r}_f$	4.05	2.98	3.84	3.94	4.58	3.26	3.97	2.52	3.44	4.16	-0.11
Annualized $\hat{\alpha}_i$	3.31	2.85	2.97	3.89	3.59	2.45	2.93	1.80	2.27	0.43	2.88
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.08)	(0.00)
$\hat{\beta}_B$	0.13	-0.10	0.05	0.04	0.17	0.12	0.23	0.15	0.27	1.27	-1.14
	(0.02)	(0.25)	(0.31)	(0.48)	(0.08)	(0.11)	(0.09)	(0.11)	(0.08)	(0.00)	(0.00)
$\hat{\beta}_C$	0.11	0.33	0.15	0.28	0.22	0.33	0.31	0.36	0.40	-0.07	0.19
	(0.01)	(0.00)	(0.04)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
$\hat{\beta}_L$	0.13	0.15	0.26	0.06	0.18	0.16	0.12	0.09	0.13	-0.02	0.15
	(0.05)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.31)	(0.05)
\bar{R}^2	0.11	0.19	0.23	0.08	0.12	0.23	0.11	0.18	0.25	0.96	0.76

Notes: The estimation period is 1871.1-1907.12 (482 observations). *SML* denotes the leveraged portfolio that is long £1 of the smallest portfolio and short £1 of the largest portfolio. All of the returns are taken in excess of the risk-free rate r_{ft} , the London banker bill rate. The p -values in parentheses are based on the heteroskedasticity and autocorrelation robust standard errors (Newey and West 1987a) and the asymptotic critical values.

Table 4-9. Equally Weighted Sovereign Bond Portfolios Sorted by Size

Dependent Variable: $r_{it} - r_{ft}$

	Small	2	3	4	5	6	7	8	9	Large	<i>SML</i>
Annualized $\bar{r}_i - \bar{r}_f$	3.72	2.41	3.00	3.29	3.60	2.39	3.10	2.11	3.07	3.04	0.68
Annualized $\hat{\alpha}_i$	3.56 (0.00)	3.01 (0.00)	3.20 (0.00)	3.24 (0.00)	3.33 (0.00)	2.23 (0.00)	2.53 (0.01)	1.73 (0.00)	2.43 (0.00)	1.45 (0.00)	2.11 (0.01)
$\hat{\beta}_B$	0.05 (0.09)	-0.22 (0.11)	-0.08 (0.32)	0.01 (0.48)	0.08 (0.29)	0.04 (0.38)	0.18 (0.19)	0.11 (0.23)	0.20 (0.17)	0.53 (0.00)	-0.47 (0.00)
$\hat{\beta}_C$	0.13 (0.00)	0.37 (0.00)	0.23 (0.02)	0.29 (0.00)	0.24 (0.00)	0.39 (0.00)	0.34 (0.01)	0.37 (0.00)	0.43 (0.00)	0.28 (0.00)	-0.15 (0.01)
\bar{R}^2	0.03	0.12	0.02	0.06	0.03	0.09	0.08	0.16	0.19	0.67	0.42

Notes: The estimation period is 1871.1-1907.12 (482 observations). *SML* denotes the leveraged portfolio that is long £1 of the smallest portfolio and short £1 of the largest portfolio. All of the returns are taken in excess of the risk-free rate r_{ft} , the London banker bill rate. The p -values in parentheses are based on the heteroskedasticity and autocorrelation robust standard errors (Newey and West 1987a) and the asymptotic critical values.

Table 4-10. Equally Weighted Sovereign Bond Portfolios Sorted by Size

Dependent Variable: $r_{it} - r_{ft}$

	Small	2	3	4	5	6	7	8	9	Large	<i>SML</i>
Annualized $\bar{r}_i - \bar{r}_f$	3.72	2.41	3.00	3.29	3.60	2.39	3.10	2.11	3.07	3.04	0.68
Annualized $\hat{\alpha}_i$	3.19 (0.00)	2.39 (0.00)	2.24 (0.00)	2.98 (0.00)	2.56 (0.00)	1.60 (0.03)	2.03 (0.02)	1.38 (0.01)	1.92 (0.00)	1.35 (0.00)	1.84 (0.01)
$\hat{\beta}_B$	0.09 (0.01)	-0.14 (0.18)	0.04 (0.37)	0.04 (0.37)	0.17 (0.08)	0.11 (0.10)	0.24 (0.09)	0.16 (0.11)	0.26 (0.08)	0.54 (0.00)	-0.44 (0.00)
$\hat{\beta}_C$	0.10 (0.00)	0.33 (0.00)	0.17 (0.02)	0.27 (0.00)	0.19 (0.01)	0.35 (0.00)	0.31 (0.00)	0.34 (0.00)	0.40 (0.00)	0.28 (0.00)	-0.17 (0.01)
$\hat{\beta}_L$	0.09 (0.05)	0.16 (0.00)	0.24 (0.00)	0.07 (0.00)	0.19 (0.00)	0.16 (0.00)	0.13 (0.00)	0.09 (0.00)	0.13 (0.00)	0.03 (0.27)	0.07 (0.16)
\bar{R}^2	0.11	0.22	0.23	0.08	0.14	0.20	0.12	0.21	0.25	0.67	0.43

Notes: The estimation period is 1871.1-1907.12 (482 observations). *SML* denotes the leveraged portfolio that is long £1 of the smallest portfolio and short £1 of the largest portfolio. All of the returns are taken in excess of the risk-free rate r_{ft} , the London banker bill rate. The p -values in parentheses are based on the heteroskedasticity and autocorrelation robust standard errors (Newey and West 1987a) and the asymptotic critical values.

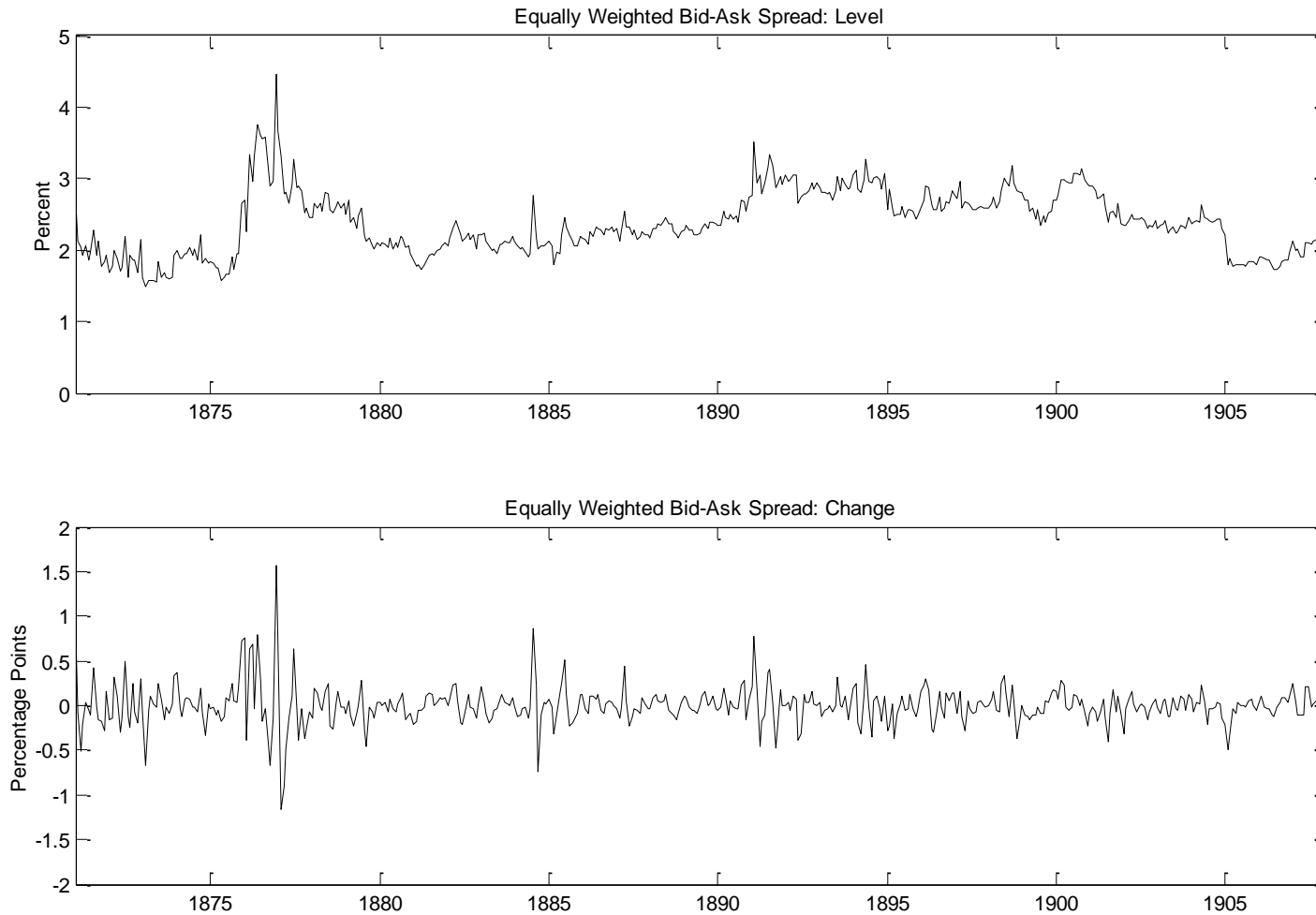
Table 4-11. Risk Premia for Value-Weighted Sovereign Bond Portfolios

Dependent Variable: $r_i - r_f$

	Unrestricted	Restricted 1	Restricted 2	Restricted 3
$\hat{\lambda}_B$	3.75 (0.01)		3.88 (0.01)	3.89 (0.00)
$\hat{\lambda}_C$	4.13 (0.01)	9.79 (0.00)		9.82 (0.01)
$\hat{\lambda}_L$	22.56 (0.00)	14.23 (0.02)	22.45 (0.00)	
J_T -test	12.74 (0.08)	13.90 (0.08)	12.36 (0.14)	16.92 (0.03)
D -test		22.83 (0.00)	4.87 (0.03)	24.26 (0.00)
Reject H_0 ?		Yes	Yes	Yes

Notes: The estimated risk premia are annualized. $\hat{\lambda}_L$ refers to the estimated risk premium associated with the liquidity-mimicking portfolio IML_t . J_T -test refers to the Hansen (1982) test of the overidentifying restrictions. The null hypothesis is that the moment conditions are all satisfied. D -test refers to the Newey-West (1987b) model specification test. The null hypothesis is that including the additional variables in the unrestricted model does not improve the performance of the restricted model. The estimates in the restricted model are based on optimal weighting matrix from the unrestricted model. The p -values in parentheses are based on the heteroskedasticity and autocorrelation robust standard errors.

Figure 4-1. Market Liquidity in the London Sovereign Bond Market, 1871.1-1907.12



Notes: The change in the market-wide bid-ask spread is defined in equation (3) of the text.

Figure 4-2. Market Liquidity and Major Defaults, 1871.1-1907.12

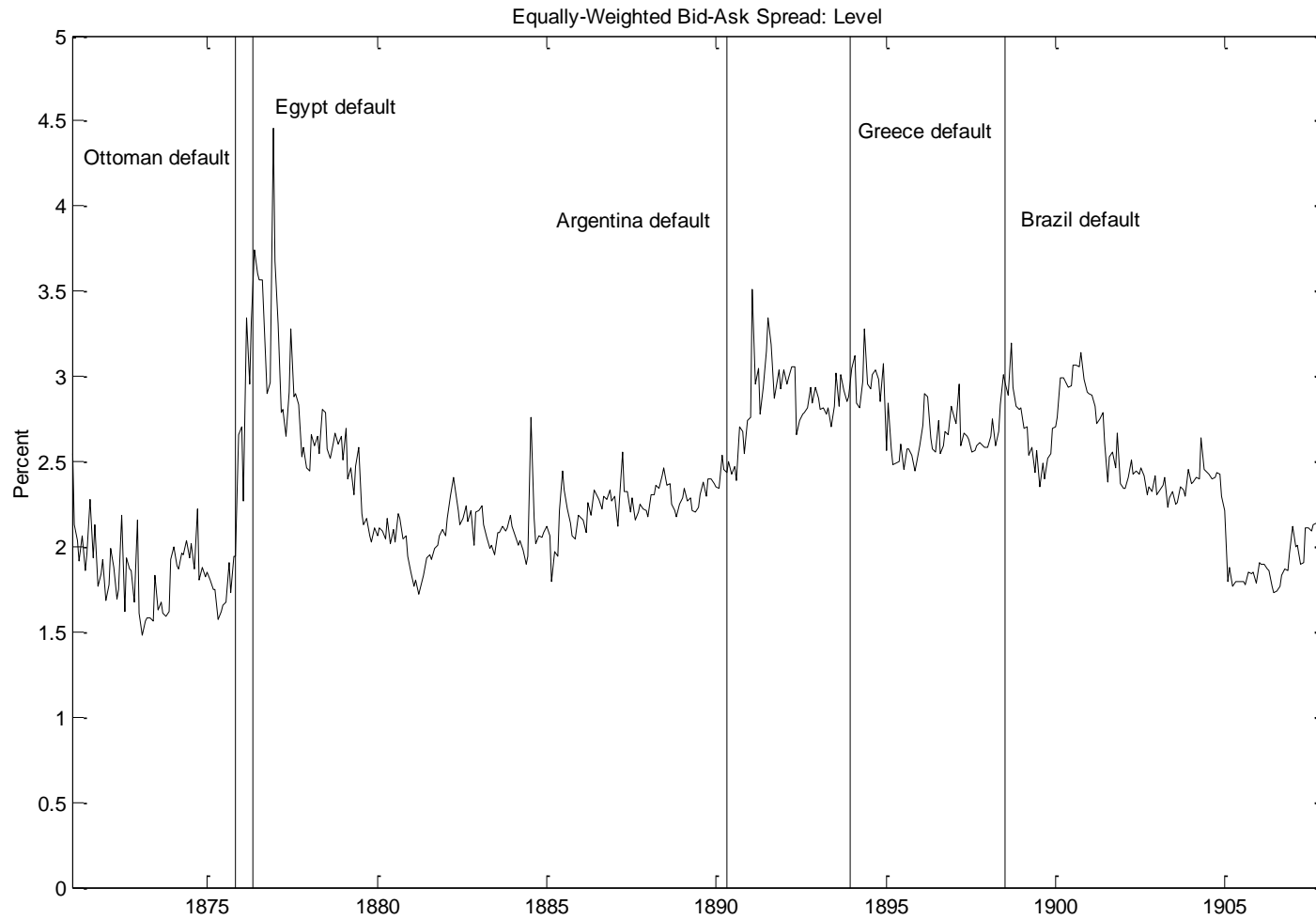
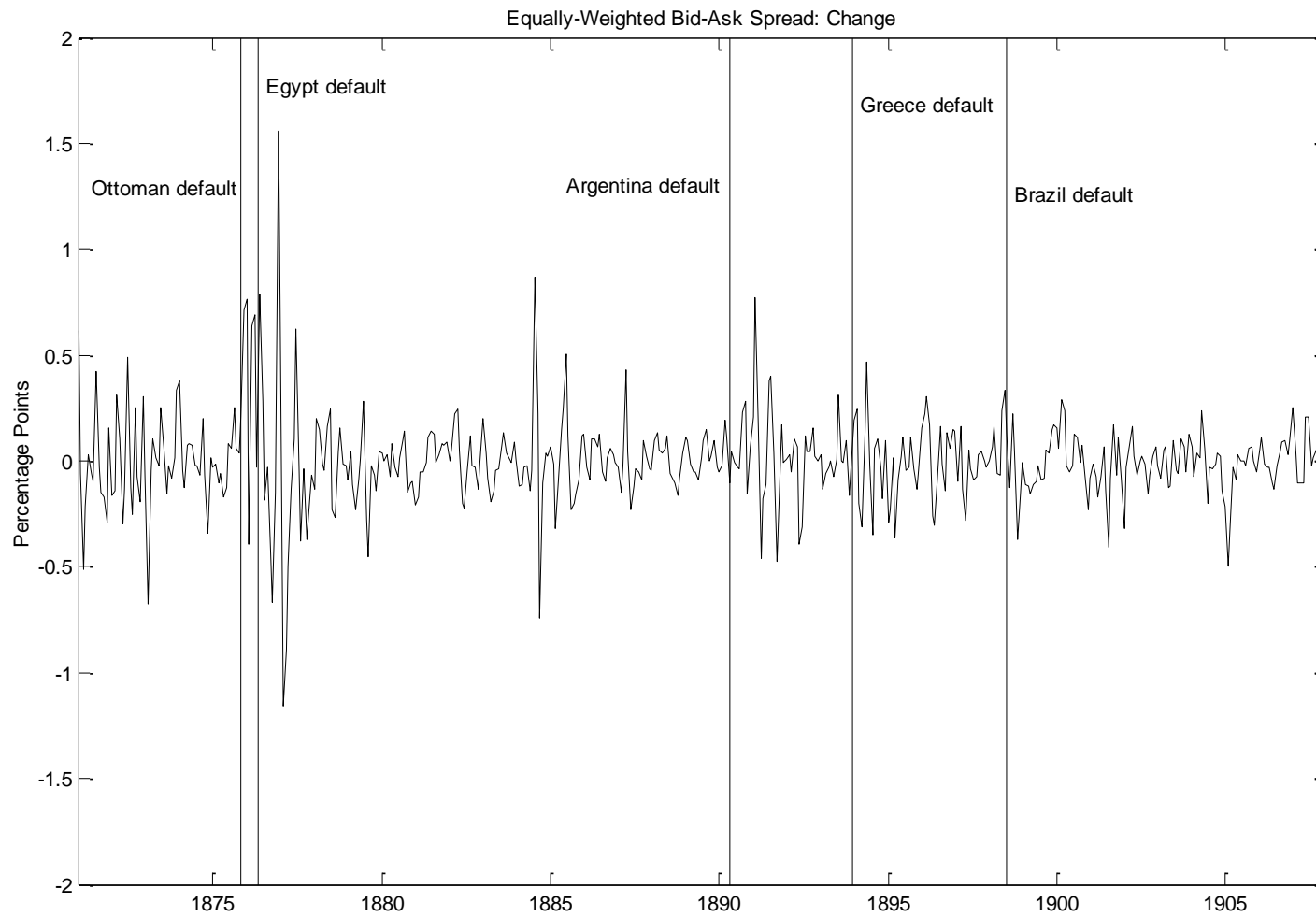


Figure 4-3. Changes in Market Liquidity and Major Defaults, 1871.1-1907.12



Appendix 4-1

Sovereign Bond Price Data

For each of the sovereign bonds, I use the price and coupon data to compute the time series of realized holding-period returns corrected for sovereign defaults. To date these defaults, I rely on the dates provided by Suter (1990), but I also consulted the annual reports issued by the Council of Foreign Bondholders. Although the sovereign bonds that were in default were not actively traded, it was common the underwriting firms to keep these prices on the books. The prices of these sovereign bonds were quoted in the *Money Market Review* and the *Economist*, even though they ceased paying coupons and were not actively traded on the exchange. Depending on how negotiations between the debtor country and its creditors proceeded, these bonds could remain on the underwriters' books for years. These sovereign bonds pose a challenge because they tend to have very wide bid-ask spreads and very low returns, biasing the returns of the most illiquid portfolio downwards and the returns of the most liquid portfolio upwards. To address this problem, I removed all of the sovereign bonds trading below £40 or 40% of par.

Sovereign Bond Data

The countries included in the sovereign-bond dataset are Argentina; Australia; Austria-Hungary; Belgium; Brazil; British Guiana; Bulgaria; Canada; Ceylon; Chile; China; Colombia; Costa Rica; Denmark; Ecuador; Egypt; France; Germany; Greece; Guatemala; Hawaii; Honduras; Hong Kong; Italy; Jamaica; Japan; Liberia; Mauritius; Mexico; Netherlands; New Zealand; Nicaragua; Norway; Orange Free State; Paraguay; Peru; Portugal; Russia; Saint Lucia; Santo Domingo; South Africa; Spain; Straits Settlements; Sweden; Trinidad; Turkey; United States; Uruguay; and Venezuela.

The British colonies comprise a subset of the countries in this sample. They are Australia; British Guiana; Canada; Ceylon; Hong Kong; Jamaica; Mauritius; New Zealand; Saint Lucia; South Africa; and Trinidad.

Appendix 4-2

Generalized Method of Moments

In the Generalized Method of Moments procedure, the model error vector is given by

$$(A2.1) \quad g(b) = E[r_t - r_{ft}] - E[(r_t - r_{ft}) f_t'] \psi$$

where $E[r_t - r_{ft}] = E[r_{1t} - r_{ft}, \dots, r_{10t} - r_{ft}]'$ is the vector of excess returns of each of the ten test portfolios; $f_t = [r_{Bt} - r_{ft}, r_{Gt} - r_{ft}, r_{Lt} - r_{ft}]'$ is the vector excess returns of each the three factors; and $b = [b^B, b^C, b^L]'$ of factor sensitivities. The sample analogue is

$$(A2.2) \quad g_T(b) = T^{-1} \sum_{t=1}^T [r_t - r_f - (r_t - r_f) f_t' b]$$

In the first step, b_1 is the solution to minimization problem

$$(A2.3) \quad b_1 = \arg \min g_T'(b) W_T g_T(b)$$

where $W_T = I_T$. If we define the sample gradient as

$$(A2.4) \quad D_T = T^{-1} \sum_{t=1}^T (r_t - r_f) f_t',$$

the analytical solution to (A2.3) is

$$(A2.5) \quad \hat{b}_1 = (D_T' D_T)^{-1} D_T' T^{-1} \sum_{t=1}^T (r_t - r_{ft}).$$

In the second step, b_2 is the solution to the minimization problem

$$(A2.6) \quad b_2 = \arg \min g_T'(b_2) W_T^* g_T(b_2)$$

where the optimal weighting matrix W_T^* is a consistent estimator of the long-run variance-covariance matrix of the moments from the first-stage $[T \text{var}(g_T(\hat{b}_1))]^{-1} = S_T^{-1}$, such as the one proposed by Newey and West (1987a). The analytical solution to (A2.6) is

$$\hat{b}_2 = (D_T' S_T^{-1} D_T)^{-1} D_T' S_T^{-1} T^{-1} \sum_{t=1}^T (r_t - r_{ft}).$$

References

Acharya, Viral V. and Lasse Heje Pedersen. "Asset Pricing with Liquidity Risk." *Journal of Financial Economics* 77, no. 2 (2005): 375-410.

Amihud, Yakov and Haim Mendelson. "Liquidity, Maturity, and the Yields on US Treasury Securities." *Journal of Finance* 46, no. 4 (1991): 1411-1425.

Andrews, Donald W. K. and Hong-Yuan Chen. "Approximately Median-Unbiased Estimation of

Autoregressive Models." *Journal of Business and Economic Statistics* 12, no. 2 (1994): 187-204.

Bank for International Settlements. *Market Liquidity: Research Findings and Selected Policy Implications*. Basel: Bank for International Settlements, 1999.

Barr, David G. and Richard Priestley. "Expected Returns, Risk, and the Integration of International Bond Markets." *Journal of International Money and Finance* 23, no. 1 (2004): 71-97.

Bekaert, Geert, Campbell R. Harvey, and Christian Lundblad. "Liquidity and Expected Returns: Lessons from Emerging Markets." *Review of Financial Studies* 20, no. 6 (2007): 1783-1831.

Boehmer, Ekkehart and William L. Megginson. "Determinants of Secondary Prices for Developing Country Syndicated Loans." *Journal of Finance* 45, no. 5 (1990): 1517-1540.

Bordo, Michael D. and Hugh Rockoff. "The Gold Standard as a 'Good Housekeeping Seal of Approval'." *Journal of Economic History* 56, no. 2 (1996): 389-428.

Breeden, Douglas T. "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities." *Journal of Financial Economics* 7, no. 3 (1979): 265-296.

Breeden, Douglas T., Michael R. Gibbons, and Robert H. Litzenberger. "Empirical Tests of the Consumption-Oriented CAPM." *Journal of Finance* 44, no. 2 (1989): 231-262.

Calvo, Guillermo A., Leonardo Leiderman, and Carmen M. Reinhart. "Inflows of Capital to Developing Countries in the 1990s." *Journal of Economic Perspectives* 10, no. 2 (1996): 123-139.

- Chen, Nai-Fu, Richard Roll, and Stephen A. Ross. "Economic Forces and the Stock Market." *Journal of Business* 59, no. 3 (1986): 383-403.
- Chordia, Tarun, Asani Sarkar, and Avanidhar Subrahmanyam. "An Empirical Analysis of Stock and Bond Market Liquidity." *Review of Financial Studies* 18, no. 1 (2005): 85-129.
- Copeland, Thomas E. and Dan Galai. "Information Effects on the Bid-Ask Spread." *Journal of Finance* 38, no. 5 (1983): 1457-1469.
- Dittmar, Robert F. and Kathy Yuan. "Do Sovereign Bonds Benefit Corporate Bonds in Emerging Markets?," *Review of Financial Studies*, forthcoming.
- Durbin, Erik and David Ng. "The Sovereign Ceiling and Emerging Market Corporate Spreads." *Journal of International Money and Finance* 24, no. 4 (2005): 631-649.
- Edwards, Sebastian. "The Pricing of Bonds and Bank Loans in International Markets: An Empirical Analysis of Developing Countries' Foreign Borrowing." *European Economic Review* 30, no. 3 (1986): 565-589.
- Eichengreen, Barry J. and Richard Portes. "Debt and Default in the 1930s." *European Economic Review* 30, no. 3 (1986): 599-640.
- Eichengreen, Barry J. and Ashoka Mody. "Lending Booms, Reserves, and the Sustainability of Short-Term Debt: Inferences from the Pricing of Syndicated Bank Loans." *Journal of Development Economics* 63, no. 1 (2000): 5-44.
- Fama, Eugene F. and Kenneth R. French. "Multifactor Explanations of Asset-Pricing Anomalies." *Journal of Finance* 51, no. 1 (1996): 55-84.
- Fernández-Arias, Eduardo. "The New Wave of Private Capital Inflows: Push or Pull?" *Journal of Development Economics* 48, no. 2 (1996): 389-418.
- Fleming, Michael J. "Are Larger Treasury Issues More Liquid? Evidence from Bill Reopenings." *Journal of Money, Credit, and Banking* 34, no. 3 (2002): 707-735.
- Folkerts-Landau, David. "The Changing Role of Bank Lending in Development Finance." *International Monetary Fund Staff Papers* 32, no. 2 (1985): 317-363.
- Glosten, Lawrence R. and Paul R. Milgrom. "Bid, Ask, and Transaction Prices in Specialist Market with Heterogeneously Informed Trader." *Journal of Financial Economics* 21, no. 1 (1985): 71-100.

- Hansen, Lars Peter. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica* 50, no. 4 (1982): 1029-1054.
- Ilmanen, Antti. "Time-Varying Expected Returns in International Bonds." *Journal of Finance* 50, no. 2 (1995): 481-506.
- International Monetary Fund. *International Capital Markets: Developments, Prospects, and Key Policy Issues*. Washington, DC: IMF, various years.
- International Monetary Fund. *Global Financial Stability Report*. Washington, DC: IMF, various years.
- Korajczyk, Robert A. and Ronnie Sadka. "Pricing the Commonality Across Alternative Measures of Liquidity." *Journal of Financial Economics* 87, no. 1 (2008): 45-72.
- Krishnamurthy, Arvind. "The Bond/Old-Bond Spread." *Journal of Financial Economics* 66, no. 2-3 (2002): 463-506.
- Kyle, Albert S. "Continuous Auctions and Insider Trading." *Econometrica* 53, no 6 (1985): 1315-1335.
- Liu, Weimin. "A Liquidity-Augmented Capital Asset Pricing Model." *Journal of Financial Economics* 82, no. 3 (2006): 621-671.
- Longstaff, Francis. "The Flight-to-Liquidity Premium in US Treasury Bond Prices." *Journal of Business* 77, no. 3 (2004): 511-526.
- Mauro, Paolo, Nathan Sussman, and Yishay Yafeh, "Emerging Market Spreads: Then Versus Now." *Quarterly Journal of Economics* 117, no. 2 (2002): 695-733.
- Merton, Robert C "An Intertemporal Capital Asset Pricing Model." *Econometrica* 41, no. 5 (1973): 867-887.
- Newey, Whitney and Kenneth D. West. "A Simple, Positive Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55, no. 3 (1987a): 703-708.
- Newey, Whitney and Kenneth D. West. "Hypothesis Testing with Efficient Method of Moments Estimation." *International Economic Review* 28, no. 3 (1987b): 777-787.
- Pástor, Lubos and Robert F. Stambaugh. "Liquidity and Expected Stock Returns." *Journal of Political Economy* 111, no. 3 (2003): 642-685.

Ross, Stephen. "The Arbitrage Pricing Theory of Capital Asset Pricing." *Journal of Economic Theory* 13, no. 3 (1976): 341-360.

Stone, Mark R. "Are Sovereign Debt Secondary Market Returns Sensitive to Macroeconomic Fundamentals? Evidence from the Contemporary and Interwar Markets." *Journal of International Money and Finance* 10, no. 1 (1991): 100-122.

Suter, Christian. *Schuldenzyklen in der dritten Welt: Kreditaufnahme, Zahlungskrisen und Schuldenregelungen peripherer Länder im Weltsystem von 1820 bis 1986*. Frankfurt, a. M.: Anton Hain, 1990.

Takagi, Shinji. "Transaction Costs and the Term Structure of Interest Rates in the OTC Bond Market in Japan." *Journal of Money, Credit, and Banking* 19, no. 4 (1987): 515-527.

Taylor, Alan M. "A Century of Current Account Dynamics." *Journal of International Money and Finance* 21, no. 6 (2002): 725-748.

Uribe, Martín and Vivian Z. Yue. "Country Spreads and Emerging Countries: Who Drives Whom?" *Journal of International Economics* 69, no. 1 (2006): 6-36.

Chapter 5

Conclusion

The three papers that comprise my dissertation study international asset markets. The first paper establishes that crude oil future contracts are poor predictors of the future spot price of oil but that the spread between the current futures price and the current spot price can be viewed as an indicator of changes in the precautionary demand for crude oil. The second paper shows that the exchange rate regime matters less for both the sovereign and the private cost of capital than previously thought. The third paper finds that liquidity risk is a quantitatively important determinant of sovereign borrowing costs in the market for international debt.

By extending our knowledge of how particular international asset markets function, these papers contribute to the ongoing discussion about the implications of international financial integration. As long as international financial integration remains an important feature of the current economic landscape, such an understanding will remain critical not only for those who wish to make sense of this environment but also for those who wish to influence the direction that it will take.