Carry Trade Activities: A Multivariate Threshold Model Analysis*

Matthias Gubler†
University of Basel

This Version: March 25, 2011

Abstract

In this empirical study we analyze the relationship between carry trade positions and some key financial as well as macroeconomic variables using a multivariate threshold model. It is often stated that the Swiss franc serves as a funding currency. Therefore, we focus on carry trades based on the currency pairs US dollar/Swiss franc and euro/Swiss franc. Generalized impulse responses differ in magnitude and significance between periods with a large and small interest-rate differential. Furthermore, in periods with a small interest-rate differential, carry trade activities "Granger-cause" the nominal exchange rate. The Granger causality test results further indicate feedback trading. Overall, carry trade positions are driven to a large extend by the expected risk in financial markets and the nominal exchange rate. Liquidity constraints can also be important, whereas the carry itself plays only a minor role.

JEL Codes: C15 C32 E44
Keywords: Carry Trades, Multivariate Threshold Model, Tsay Test, Generalized Impulse Response Functions, Bootstrap Method, Granger Causality

*I would like to thank Daniel Geissmann, Sylvia Kaufmann, Tobias Körner and Iris Oberauner as well as conference audiences at YSEM Bern and RGS Dortmund for valuable and helpful comments and suggestions. Hermione Miller-Moser provided excellent editorial assistance. I am indebted to Peter Kugler for his support and supervision.

†Contact address: University of Basel, Faculty of Business and Economics, Peter-Merian-Weg 6, Postfach, 4002 Basel, Switzerland. Email: matthias.gubler@unibas.ch

1 Introduction

In our paper, we empirically investigate the relationship between speculators' currency carry trade positions and key financial variables which are of macroeconomic interest. The basic idea of a "currency carry trade" (hereinafter "carry trade") involves selling low-interest-rate currencies (e.g., by borrowing money) and investing simultaneously in high-interest-rate currencies. Low-interest-rate currencies, such as the Swiss franc or the Japanese yen, are called funding currencies, whereas high yielding currencies are called target currencies.

Recently, investment strategies to exploit the failure of the uncovered interest rate parity (UIP) have become a major focus of interest not only for financial market participants. Carry trades also appeared on policymakers' agendas, specifically on those of central bankers. For instance, Jean-Pierre Roth, former president of the governing board of the Swiss National Bank, pointed out the crucial role of carry trades in determining the nominal exchange rate in the medium-run (Roth 2007). In our analysis, we focus on two target currencies for which the Swiss franc (CHF) serves as the funding currency: the US dollar (USD) and the euro (EUR).

The UIP states that the gains due to the interest-rate differentials (IRDs) are offset by the loss arising in the depreciation of the target currency. However, several empirical studies emphasize the violation of the UIP ("forward premium puzzle").[1] Meese & Rogoff (1983) compare the out-of-sample forecast accuracy of different structural exchange rate models and conclude that exchange rates follow a "near random walk". Furthermore, Fama (1984) shows that on average the target currency appreciates slightly. This empirical anomaly of the foreign exchange market makes carry trades on average profitable.

While an extensive body of the literature on carry trades examines their profitability, the main contribution of this study is the empirical investigation of the interaction between carry trade activities and financial as well as macroeconomic variables. We choose a multivariate time series model to assess the effects of an unexpected movement of one variable on the others. Carry traders react on shocks to variables which determine the profitability of their investment strategy, such as the interest rate differential (the so-called "carry"), the nominal exchange rate, the risk sentiment, the investment return, and possible liquidity constraints. In addition, these variables can move due to unexpected carry trade activities. Thus, we include these variables, or reasonable proxies, in our model. A similar set of variables is widely chosen in the literature (see e.g., Ranaldo & Söderlind (2010), Brunnermeier et al. (2009) or Nishigaki (2007)). Additionally, we investigate the predictive power of carry trade activities for the other variables and vice-versa.

Therefore, our empirical study is closest to Brunnermeier et al. (2009) and Nishigaki (2007). Brunnermeier et al. (2009) show that in times of reduced funding liquidity and de-

[1] For a literature survey, see for example Engel (1996).

clining risk appetite carry traders are subject to crash risk due to the sudden unwinding of carry trades. These findings are based on univariate and multivariate panel analysis. Nishigaki (2007) examines the yen carry trade. His recursive VAR model analysis implies that the carry has no significant impact on carry trade movements, in contrast to US stock prices. The results also indicate a USD depreciation against the Japanese yen once carry trades unwind. Both of these studies incorporate futures positions to proxy carry trade activities, as we do for the CHF/USD exchange rate. As argued by McGuire & Upper (2007), carry trade positions are not only difficult to detect but also to distinguish from other investment strategies. Furthermore, futures position data with respect to the CHF/EUR exchange rate are not available. Hence, we employ for the Euro market the carry-to-risk ratio (CTR ratio) to proxy carry trade activities. The CTR ratio is an important indicator of the potential profitability of carry trades.

Recent studies highlight the importance of regime-dependent results (see Section 2), and indeed, preliminary analyses of the IRD indicate a nonlinear relationship among the variables in our model. Plots of the IRD reveal periods with larger and smaller carries. Moreover, autoregressive conditional heteroscedasticity (ARCH) tests confirm the descriptive results, as the null hypothesis of no ARCH has to be rejected for almost all error term variances. In addition, professional currency market analysts argue that a threshold level exists for the IRD above which speculators' behavior changes. Finally, the results of a Tsay (1998) test, modified to account for conditional heteroscedasticity, confirm the assumption of nonlinear relationships. Therefore, we apply a multivariate threshold model to account for the possible changes in the dynamic behavior of carry trade activities dependent on the size of the IRD.

By analyzing the generalized impulse response functions (GIRFs), we find the following main results. First, carry trade positions are driven to a large extend by the expected risk in financial markets and the exchange rate. Since the other variables responses to a shock depend on the size of the carry, these differences are carried over to the speculators' carry trade positions. The results indicate that in times with a large carry a positive one-standard deviation shock to the carry itself is not enough to compensate investors for the increased risk. Moreover, in line with the prediction of the UIP, the CHF appreciates instantaneously against the USD in times with high IRDs, but not in the regime with low IRDs. Second, liquidity constraints can be important too, whereas the carry itself plays only a minor role. Third, a sudden unwinding of carry trades has a significant impact on the nominal exchange rate, independent of the size of the IRD. Finally, we conclude that a majority of the responses show similar patterns for the USD/CHF and EUR/CHF exchange rates, although the proxy for carry trade positions differs.

In order to account for small-sample biases we adopt the correction proposed by Kilian (1998b) to compute the confidence intervals. Additionally, we extend the recursive-design wild bootstrap method for univariate models suggested by Goncalves & Kilian (2004) to multivariate models to break up time interdependence between innovations.

As in our study, Klitgaard & Weir (2004) analyze futures position data and find a strong contemporaneous relationship between weekly changes in futures positions and exchange rate movements. However, these net positions do not seem to "Granger-cause" the exchange rate movements of the following week.[2] We follow their approach and apply the Granger causality test to our regime-dependent model and find that past position data help to predict exchange rate movements in periods with small IRDs. Additionally, in samples with the USD as target currency, the exchange rate has very high predictive power for carry trade activities, pointing to feedback trading.[3]

The remainder of this paper is organized as follows. We start in Section 2 with an overview of the related literature. Data sources and variable definitions are presented in Section 3. In Section 4, we outline the methodology used for our empirical study. We provide a detailed discussion on our results for the GIRFs in Section 5 and their robustness analysis in Section 5.3. In Section 5.4, we show the Granger causality test results. Section 6 concludes.

2 Related Literature

A large body of the literature on carry trades examines the profitability of potential carry trade strategies. A few studies conclude that these investment strategies lead to excess returns. These excess returns can be attributed neither to standard risk factors (Burnside et al. 2006), to the exposure to currency crashes (Jurek 2007), nor to disaster risks (Farhi et al. 2009). Instead, market frictions such as the bid-ask spread and price pressure greatly reduce the return on these portfolios (Burnside et al. 2006), or they are not economically significant (Wagner 2008). In contrast, Lustig et al. (2008) argue that carry trade profits are a compensation for systematic risk. Moreover, Darvas (2009) shows that the degree of leverage is crucial for excess returns. Profitability declines with increasing leverage. Furthermore, Kohler (2007) examines the correlation dynamics between returns on global equity portfolios and simple carry trade investment strategies. Based on his results, carry trades are exposed to a severe diversification meltdown in times of global stock markets crisis.

Another stream of the carry trade literature examines other channels to detect carry trade positions, that focus mainly on yen carry trades. For example, Gagnon & Chaboud (2007) emphasize the "canonical yen carry trade" in contrast to the "derivatives carry trade" studied by Brunnermeier et al. (2009) and Nishigaki (2007).[4] Galati et al. (2007)

[2] See also Mogford & Pain (2006) for a similar study.

[3] In contrast, no prediction power is found in samples with the EUR as target currency. This might be due to the definition of the CTR ratio. This issue is discussed in greater detail in Section 5.4.

[4] Gagnon & Chaboud (2007) define canonical carry trades as borrowing low yielding currencies and investing the proceeds in high-interest-rate currencies. In contrast, derivatives carry trades are defined

compare low frequency data from the BIS international banking statistics with higher frequency futures data and find similar insights for carry trade positions. Cai et al. (2001) examine the effects of order flows and macroeconomic news on the dramatic yen/dollar volatility of 1998 with weekly data from the US Treasury on purchases and sales of spot, forward, and futures contracts. Finally, Hattori & Shin (2007) conclude that the waxing and waning of the balance sheets of foreign banks in Japan is related to the state of overall risk appetite. By using descriptive statistics and a simple econometrics analysis, they reveal a positive relationship between the IRD[5] and carry trades.

The importance of regime-dependent results is highlighted, among few others, by Clarida et al. (2009). These authors examine carry trade strategies and identify a robust empirical relationship between their excess returns and exchange rate volatility. Furthermore, they show that the failure of the UIP is only present in low-volatility environments. Jordà & Taylor (2009) argue that more sophisticated conditional carry trade strategies exhibit more favorable payoffs. They adopt a nonlinear regime-dependent model approach and add the fundamental equilibrium exchange rate (FEER) to their model. In distinction to our study, they choose the threshold value exogenously. Christiansen et al. (2010) provide a factor model with regression coefficients dependent on market volatility and liquidity to assess carry trade strategies. In volatile periods the excess returns have much higher exposure to the stock market and also more mean reversion.

To the best of our knowledge, there is only one theoretical contribution in the literature that focuses specifically on carry trades. Plantin & Shin (2010) incorporate funding externalities and carry costs into their model to predict the classic price pattern "going up the stairs, and coming down in the elevator". The increase of carry trade positions is followed by abrupt stochastic reversals.

3 Data

3.1 Variables

We collected data to examine the Swiss franc (CHF) carry trade with the US dollar (USD) or the euro (EUR) as respective target currency. The variables of interest are the interest-rate differential (IRD_{USD}, IRD_{EUR}), the nominal exchange rate (FX_{USD}, FX_{EUR}), the VIX index (VIX), 10-year bond yields (Y_{USD}, Y_{EUR}), stock market prices (P_{USD}, P_{EUR}) and carry trade positions (CTF_{USD} and $CTFO_{USD}$, CT_{EUR}). The majority of the data stems from Datastream.

For the calculation of the IRD_{USD} and IRD_{EUR} we obtain 3-month interbank interest rates. The carries are defined as the difference between the respective target currency

as taking on leveraged positions in derivatives markets. More on this issue is provided in Section 3.1.

[5] The IRD is the difference between the Japanese overnight rate and the average of the US, Euro-zone and Australia policy rates.

interest rate (United States or Euro area) and the Swiss interest rate. Accordingly, we employ the nominal exchange rates CHF/USD (FX_{USD}) as well as CHF/EUR (FX_{EUR}). Furthermore, the VIX volatility index (VIX) from the Chicago Board Options Exchange (CBOE) serves as a proxy for the expected stock market risk.[6]

For an analysis on carry trade positions based on the Swiss and US markets, prices on the US stock exchange market index S&P 500 (P_{USD}) and 10-year constant to maturity Treasury bond yields (Y_{USD}) were collected. If the EUR serves as target currency, prices of the euro stock exchange market index Euro Stoxx 50 (P_{EUR}) and the synthetic euro benchmark bond yield series[7] (Y_{EUR}) are taken.

Trades in the currency markets are usually over-the-counter, making it difficult to find appropriate proxies for carry trade positions. Hence, we rely on data from the U.S. Commodity Futures Trading Commission (CFTC) for carry trade positions with regard to the USD. These contracts are traded on the Chicago Mercantile Exchange (CME). Since October 1992, long and short currency futures positions of non-commercial traders are published periodically. All investors are classified as non-commercial or commercial. Commercial investors have currency risk hedging purposes defined by the CFTC. We are only interested in positions held by those traders who basically trade for speculative purposes.

Burnside et al. (2006) show that a strategy of borrowing the low-interest-rate currency and lending the high-interest-rate currency yields a positive payoff if, and only if, a forward contract has a positive payoff. According to Brunnermeier et al. (2009), only few investors actually implement the carry trade using the spot currency market as futures contracts are economically equivalent.[8]

Our proxy for carry trade positions has several shortcomings. First, these data reflect only a very small fraction of the currency trades.[9] Second, they are not necessarily results from carry trades, and the classification of commercial and non-commercial traders might be inaccurate in some cases (Galati et al. 2007). Finally, Gagnon & Chaboud (2007) show that the timing of changes in these positions might not be perfectly accurate in all cases. For example, the unwinding of the yen carry trades in October 1998 is not displayed in the data.[10] Despite these shortcomings, these futures positions are the best publicly available

[6] The index is based on the stock market index S&P 500 and estimates expected volatility by averaging the weighted prices of options over a wide range of strike prices. Brunnermeier et al. (2009) argue that the index is a useful proxy for investor sentiment or "global risk appetite".

[7] The US benchmark bond yield series from Datastream is almost identical to the 10-year constant to maturity Treasury yields for the US market. Hence, the Euro benchmark bond yield series is a reliable proxy for our purposes.

[8] Futures and forward contracts are similar, yet the former is traded on the stock exchange and the latter over-the-counter. Additionally, they differ in settlement conditions. These differences, however, are not decisive for our purposes.

[9] Following Klitgaard & Weir (2004) a substantial part of the high foreign exchange transaction volume reflects traders' risk management. Hence, the global volume by itself does not preclude the possibility that participations in futures markets might cause currency movements.

[10] The sharp movement to a net long yen position occured one month before the actual carry trade

data (Brunnermeier et al. 2009).

Furthermore, we calculate the so-called "success rate". For the samples considered in our study, we count the observations for which the investors increase the net long futures positions (decrease the net long futures positions) and the CHF appreciates (depreciates) against the USD. The success rate is in the range of 69% and 87%, and above 75% three-quarters of the time. In line with the results of Klitgaard & Weir (2004), we find a strong contemporaneous correlation between changes in net futures positions and exchange rate fluctuations. Thus, knowing the traders actions gives a reasonable chance of correctly estimating the direction of the exchange rate movement during the same week.

A new data set including futures and options was launched from the CME at the end of March 1995. Keeping in mind that an option contract differs in several respects from a futures contract, we use these data for the robustness analysis. From Mogford & Pain (2006) we know that speculative future positions from CME and risk reversals, reflecting the views of options purchasers, move a significant number of times in the same direction.

Carry trade positions are defined as the difference between short and long futures positions (CTF_{USD}) or as the difference between short and long futures and options positions ($CTFO_{USD}$). If the net position is positive (negative), investors are involved in carry trades with the CHF as a funding (target) currency. These currency futures position data are not available for the EUR.[11] Thus, we use the carry-to-risk ratio (CTR ratio) as a proxy for carry trade activities (CT_{EUR}). The CTR ratio is defined as the 3-month interest-rate differential divided by the implied volatility derived from 3-month at-the-money exchange rate options.[12] Data on implied exchange rate volatility are taken from Bloomberg.

The choice of the CTR ratio as proxy for carry trade positions has several caveats as the CTR ratio does not represent (carry trade) positions directly. Nevertheless, professional currency market watchers take it as an important indicator for carry trade activities. Furthermore, Galati et al. (2007) find significant correlations between the CTR ratio and futures positions traded at the CME.[13]

The nominal exchange rates, the stock market prices, the VIX index and the futures (and options) positions are logarithmized.

---

unwinding (Gagnon & Chaboud 2007).

[11]Unfortunately, due to data limitations, we are not able to examine further target currencies such as the Australian dollar or the New Zealand dollar.

[12]We limited our analysis to the currency pair CHF/EUR as data on implied exchange rate volatility are not continuously available for other potential target currencies.

[13]These correlations always involve the USD. Moreover, Brunnermeier et al. (2009) argue that the past return of carry trades is perhaps a better measure for carry trade positions than futures data from CME. In a world with rational market participants, the CTR ratio is, owing to its forward-looking nature, also a good proxy.

3.2 Sample Period and Frequency

The sample period with the USD as target currency starts with 03/28/1995 and ends with 06/24/2008. For our robustness analysis, we estimate the model with different sample lengths. We add observations until the end of 2009 to address the recent financial crises or start already with 10/06/1992.

For model specifications in which the EUR serves as the target currency, we use data for the time period from 01/06/1999 to 06/25/2008.

We determined the data frequency according to the variable with the lowest frequency published, as we expect a strong short-run relationship between the variables included in this study.[14] Futures position data from the CFTC are published weekly, thus leading to a weekly frequency. To ensure comparability along the frequency dimension, we also apply weekly data for the model with the CTR ratio as a proxy for carry trade positions.

4 Methodology

We use a multivariate threshold model to analyze the relationship between key financial and macroeconomic variables focusing on carry trade positions. The choice of the method is based on a descriptive analysis, an econometric test and reported information.

First, the descriptive analysis serves to detect sub-periods separated by an endogenous threshold value of the IRD. The results of this analysis are presented in Figures (1) and (2). The former depicts the 3-month interest-rate differential (IRD_{USD}) between the United States and Switzerland. Until 2001, the IRD_{USD} spread was substantial (about 3 to 4.5%). Subsequently, the difference decreased to around zero percent in November 2001. The following upward trend reaches its maximum of almost 4% at the end of June 2006. The financial crises caused the IRD_{USD} to fall again. Thus, we were able to construct one sub-sample containing large carries and another with smaller differences.[15]

Analogously, Figure (2) illustrates the IRD_{EUR}. The starting point of the sample is the euro launch. The amplitudes of the IRD_{EUR} are not as distinct as for the IRD_{USD}. Nevertheless, three time periods with higher IRD_{EUR} could be identified: the beginning of the sample, the period from mid-2002 to almost the end of 2004 and the end of the sample.

Moreover, these findings are also reflected in the residuals of the estimation with the carry as the dependent variable. The residuals follow a very similar pattern to the interest-rate differentials themselves.

Second, the insights of the descriptive analysis are confirmed by the estimation results

[14] Brunnermeier et al. (2009) include quarterly data, whereas Nishigaki (2007) estimates his model with monthly data.

[15] Note that we allow the sub-periods to be discontinued, i.e. one sub-period is interrupted by the other one.

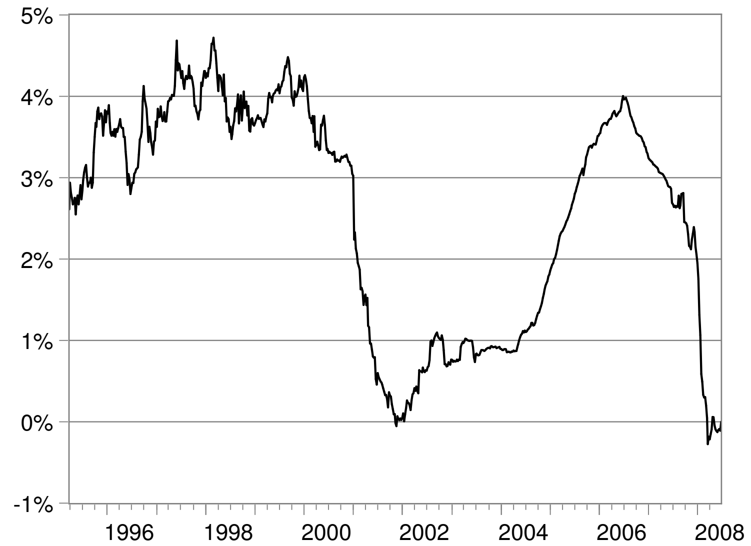

Figure 1: IRD_{USD}

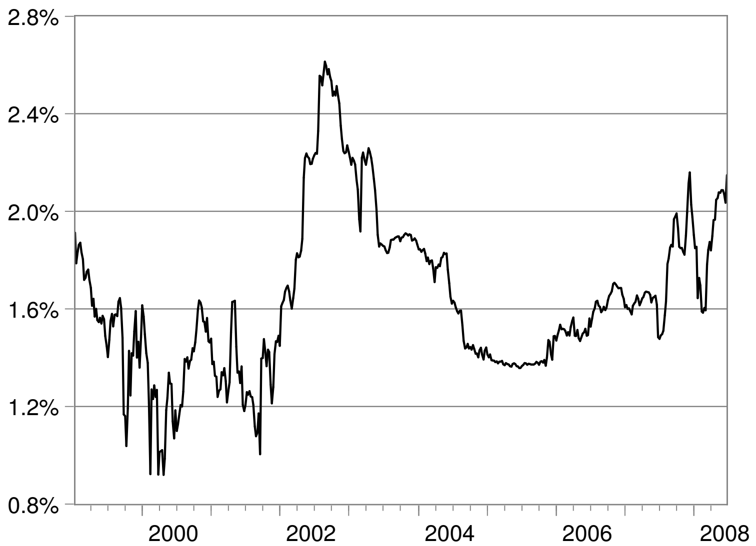

Figure 2: IRD_{EUR}

of a reduced vector autoregressive regression model (VAR) for the whole period. We have to reject the null hypothesis of no autoregressive conditional heteroscedasticity (ARCH) for a majority of the error-term variances.[16] This is not surprising, since we have high frequency financial variables in our model.[17] Nevertheless, this result indicates a nonlinear relationship between the variables considered.

Finally, professional currency market analysts argue that there exists a threshold level for the carry, above which investor behavior changes.[18] We assume that the dynamic behavior of carry trade positions depend on the magnitude of the carry, and therefore apply a multivariate threshold model for our empirical investigation (Tsay 1998). Similar methods to study relationships where nonlinear effects are present are used by Bernholz & Kugler (2010), Canjels et al. (2004) and others.

4.1 Multivariate Threshold Model

Before we turn to the econometric model, we test the appropriateness of a multivariate threshold model by applying a test developed by Tsay (1998). The observations are ordered in descending order of the lagged threshold variable to estimate the recursive residuals. The lag is determined by the threshold delay parameter d. If the dependent variables are linear, then the recursive least squares estimator of the arranged VAR model is consitent, i.e. the coefficients are zero (Tsay 1998). Compared to the standard test, we modify its computation to account for conditional heteroscedasticity (Tsay 1998). The variances of the least squares estimates have to be modified in a way that the correlation between the squared error terms and the elements of $X_t'X_t$ is taken into account. This is done by correcting the weights to standardize the predictive residuals of the recursive least squares estimations. The generalized multivariate threshold model can be written as:

$$\mathbf{y}_t \quad = \quad \mathbf{c}^{(j)} + \Phi_1^{(j)}\mathbf{y}_{t-1} + \cdots + \Phi_p^{(j)}\mathbf{y}_{t-p} + \boldsymbol{\epsilon}_t \quad \text{if } \tau_{j-1} \leq z_{t-d} < \tau_j$$

where $\mathbf{y}_t$ denotes a (6×1) vector containing the values at date t of six variables (IRD, VIX index, carry trade positions, nominal exchange rate, bond yields, stock market index)[19], $\mathbf{c}^{(j)}$ are the constant vectors for the different regimes, and $\Phi^{(j)}$ denotes a (6×6) coefficient matrix of the respective lag and regime. The vector of error terms is denoted as $\boldsymbol{\epsilon}$, and p is the number of lags included. Let $-\infty = \tau_0 < \tau_1 < \cdots < \tau_{s-1} < \tau_s = \infty$.

[16] The ARCH test results are summarized in Table (11) in the appendix.

[17] The variance of the error-term might follow an ARCH/GARCH process when financial variables are included in a model with high frequency data (see e.g., Engle (2001)).

[18] I would like to thank the Head FX Research of a major Swiss bank for this important information.

[19] The variables enter the model either in level or in first differences. More details on the model specifications can be found in Section 5.

Then $j = 1, \ldots, s$ represents the different regimes. We concentrate on models with two regimes, hence, we have only one threshold value and $s = 2$.[20] The multivariate threshold model applied with two regimes has the following form:

$$\mathbf{y}_t = \mathbf{c}^{(1)} + \Phi_1^{(1)}\mathbf{y}_{t-1} + \cdots + \Phi_p^{(1)}\mathbf{y}_{t-p} + \boldsymbol{\epsilon}_t \quad \text{if } z_{t-d} \geq \tau \tag{1}$$

$$\mathbf{y}_t = \mathbf{c}^{(2)} + \Phi_1^{(2)}\mathbf{y}_{t-1} + \cdots + \Phi_p^{(2)}\mathbf{y}_{t-p} + \boldsymbol{\epsilon}_t \quad \text{if } z_{t-d} < \tau \tag{2}$$

The observations of a specific date are included in the first regime (equation 1) if the threshold variable z is above or equal to the threshold value τ. The determination of the delay parameter d is based on the test statistic of the Tsay test. In order to determine the threshold value τ, we use a grid search over a reasonable interval of possible values of the threshold variable. The selection of τ is based on the minimum Akaike Information Criterion (AIC). When τ is known, we can estimate the model by ordinary least squares (OLS). Concretely, we estimate the following model:

$$\mathbf{y}_t = \mathbf{c} + (\Phi_1\mathbf{y}_{t-1} + \cdots + \Phi_p\mathbf{y}_{t-p})D_{t-d} + (\Phi_1\mathbf{y}_{t-1} + \cdots + \Phi_p\mathbf{y}_{t-p})(1 - D_{t-d}) + \boldsymbol{\epsilon}_t$$

where a dummy variable D is defined as being one if $z_{t-d} \geq \tau$, and zero if $z_{t-d} < \tau$.

4.2 Generalized Impulse Response Functions

Since Sims (1980) seminal paper, vector autoregressions (VARs) are routinely carried out to study dynamic systems. In many studies, researchers rely on the Cholesky decomposition to structure the estimation model. Both, Brunnermeier et al. (2009) and Nishigaki (2007) use this approach to examine carry trade positions. The structural shocks are obtained by orthogonalizing the estimated reduced-form error terms. However, the ordering of the variables in the system matters for the results (Pesaran & Shin 1998). In many cases it is very difficult to establish a particular recursive ordering on economic theory or institutional knowledge (Stock & Watson 2001). According to Stock & Watson (2001), researchers are too often tempted to develop economic "theories" that lead to a recursive structure called "Wold causal chain". Therefore, they distinguish between structural and recursive VARs. Without a widely accepted economic theory to help differentiate between correlation and causation ("identification problem"), we prefer the method developed by Koop et al. (1996) and Pesaran & Shin (1998).[21] This alternative approach is invariant

[20]The model was also estimated with two threshold values and with the first difference of the IRD as threshold value. In these cases, the estimation technique does not change, only the notation becomes slightly more complicated.

[21]We follow the approach by Pesaran & Shin (1998) as we correct the estimates for small-sample bias and departures from non-normality of the error terms (Kilian 1998a,b). Furthermore, results from a

to the ordering of the variables, instead, it lacks the possibility to identify a specific shock.

While the recursive structure identifies the shocks through the Cholesky decomposition of the residual variance-covariance matrix, the variance-covariance matrix itself matters for the generalized impulse response functions. The interdependence of the shocks is carried over to the impulse response function. It follows that the method of generalized impulse response analysis is not the preferred approach for policy statements. In our analysis we do not want to indentify specific shocks but rather examine what happens if one of the variables changes unexpectedly. Hence, we let the data speak.

4.3 Confidence Interval: Bootstrap Method

The confidence intervals of impulse responses are routinely computed with bootstrap methods. Kilian (1998b) shows that traditional bootstrap methods such as the frequently applied nonparametric approach developed by Runkle (1987) are inaccurate in the presence of bias and skewness in the small-sample distribution of impulse response estimators. Thus, we adopt his bias-correction (Kilian 1998b), because the construction of sub-periods reduces the number of observations to a great extent.[22] Additionally, Kilian (1998a) demonstrates the outperformance of the bias-corrected confidence intervals if there is evidence of fat tails or skewness in the error distribution, i.e. the innovations' departure from normality. The distribution of a few estimated residuals in our study suffers from non-normality, not only in the full sample but also in the subsamples.

By considering the full samples, we have to reject the null hypothesis of an ARCH-test for a majority of the error term variances (see Section 4). This problem is far less severe in the subsamples, but is still present.[23] Non-normality could be at least partly explained by unknown ARCH/GARCH processes.[24] However, as the bias-correction cannot account for biases introduced by ARCH/GARCH processes (Kilian 1998a), we change the computation of the confidence intervals to deal with unknown ARCH/GARCH processes.

Based on the work by Goncalves & Kilian (2004), we modify the residuals such that we can treat them as *i.i.d.* We extend the recursive-design wild bootstrap method for univariate models proposed by Goncalves & Kilian (2004) to multivariate models. The modification is basically a multiplication of the estimated residuals with an i.i.d. sequence with mean zero and variance one. In our analysis, we draw the sequence from a standard normal distribution. The multiplication implicates that the time interdependence between innovations of an estimation equation breaks up. The application of this method to a multivariate system arises the problem of correctly treating the cross interdependence

recursive VAR consistent with Nishigaki (2007) indicate that the GIRFs are reasonable.

[22]Despite the reduction in the number of observations, they are sufficient for an accurate estimation of the parameters.

[23]Whereas the problem hardly arises in the subsamples with high carries, it is somewhat stronger in the subsamples with low interest-rate differentials.

[24]This is true for the leptokurtosis, but not for the skewness in the residuals (Kilian 1998a).

between innovations of different estimation equations. To overcome this cross interdependence we rely on Pesaran & Shin (1996). In a first step, the residuals are multiplied by the inverse of the Cholesky decomposition:

$$\boldsymbol{\xi} = \mathbf{P}^{-1}\hat{\boldsymbol{\epsilon}}$$

where $\boldsymbol{\xi}$ is a $(m \times T)$ matrix and $\hat{\boldsymbol{\epsilon}}$ are the estimated residuals. T is the number of observations and m the number of variables. The resulting terms in the matrix $\boldsymbol{\xi}$ are independent from each other for every t. After multiplying these terms with the i.i.d. sequence described above, we recover the contemporaneous correlation structure as follows:

$$\hat{\boldsymbol{\epsilon}}^{*} = \mathbf{P}\boldsymbol{\xi}\boldsymbol{\eta}$$

where $\boldsymbol{\eta}$ denotes an $(m \times T)$ matrix with the i.i.d. sequences. Finally, the matrix $\hat{\boldsymbol{\epsilon}}^{*}$ contains modified innovations with the same cross interdependence, but no interdependence over time. We apply this method to the error terms for which we reject the null hypothesis of no ARCH of order one and/or two and/or four at the 5% significance level.[25]

All of these modifications have the property to enlarge the non-centered 95%-confidence intervals of our empirical study. The confidence intervals are based on 11,000 random draws, where the first 1,000 draws are used to compute the bias-correction.[26]

5 Empirical Results

5.1 Preliminary Analysis

In this subsection, we briefly describe the results of the preliminary analysis necessary prior to the estimation of the multivariate threshold model.

5.1.1 Stationarity Tests

In a first step, the time series properties are examined. For this purpose, the test proposed by Phillips & Perron (1988) and the augmented Dickey & Fuller (1979) unit root test are

[25]The computation of the GIRFs requires a constant variance-covariance matrix. The presence of unknown ARCH/GARCH processes might lead to a time-variant variance-covariance matrix. However, we assume that our results are not strongly biased since we draw the innovations from the subsamples in which only a few or even no error term variances follow an unknown ARCH/GARCH process (Koop et al. 1996).

[26]Furthermore, if one of the draws leads to a model with an eigenvalue greater than one (i.e., the model is explosive), the draw is disregarded and repeated.

Table 1: PP and ADF Unit Root Test Results with the USD as Target Currency

	March 1995 - June 2008		March 1995 - Dec 2009	
	PP[2]	ADF[3]	PP[2]	ADF[3]
FX_{USD}[1]	-1.530	-1.539	-1.987	-2.009
P_{USD}[1]	-2.243	-2.178	-2.284	-2.214
VIX	-3.717***	-3.612***	-3.744***	-3.507***
IRD_{USD}	-0.624	-0.720	-0.692	-0.877
Y_{USD}[1]	-3.122	-3.109	-3.420**	-3.385*
Carry Trade Positions				
CTF_{USD}	-6.785***	-6.984***	-7.021***	-7.230***
$CTFO_{USD}$	-6.801***	-6.575***	-7.029***	-7.226***

[1] A deterministic trend is included
[2] Bartlett kernel, Newey-West bandwidth
[3] Lag length selection by modified SIC (Ng & Perron 2001)
*/**/*** denotes significance at 10%, 5% and 1% level, respectively

applied to the variables. Tables (1) and (2) report the results for the models for which the USD serves as the target currency of the carry trades. The results point clearly to stationarity of the carry trade positions and the VIX index, regardless of the sample choice. For the 10-year constant to maturity Treasury bond yields the results are borderline. Even if the null hypothesis cannot be rejected, the test statistic is very close to the critical value of the 10% significance level. The remaining three variables, the CHF/USD exchange rate, the price of the S&P 500 and the interest-rate differential (IRD) are non-stationary.

Table (3) presents the results for the sample with the EUR as target currency. Again, the proxy for carry trade activities is clearly stationary. The results for the VIX index also points to stationarity. All other time series are non-stationary subject to the test results.

All results are confirmed by applying the Kwiatkowski et al. (1992) stationarity test and the two unit root tests from Elliott et al. (1996) and Ng & Perron (2001). Moreover, all of them point to a (weak) stationary IRD between the 3-month interbank interest rates from Switzerland and the Euro area.[27]

The outcomes of tests for non-stationarity of the time series are in line with the findings of other empirical studies (see e.g., Nishigaki (2007)). From a theoretical point of view it is surprising that the null hypothesis cannot be rejected for the difference between the US and Swiss 3-month interbank interest rates. This result implies that the correct model specification includes the first difference of the IRD. However, there is no economical justification for a random walk behavior of the IRD, specifically in the long run. Moreover, as long as the model is stationary and no spurious regression problem

[27] These results are not published but can be obtained from the author upon request.

Table 2: PP and ADF Unit Root Test Results with the USD as Target Currency

	Oct 1992 - June 2008		Oct 1992 - Dec 2009	
	PP[2]	ADF[3]	PP[2]	ADF[3]
FX_{USD}[1]	-1.568	-1.547	-1.946	-1.953
P_{USD}[1]	-1.299	-1.238	-1.292	-1.351
VIX	-3.746***	-3.620***	-3.480***	-3.679***
IRD_{USD}[1]	-2.354	-1.824	-2.200	-2.818
Y_{USD}[1]	-3.197*	-3.043	-3.326*	-3.525**
Carry Trade Positions				
CTF_{USD}	-7.237***	-7.323***	-7.566***	-7.468***

[1] A deterministic trend is included
[2] Bartlett kernel, Newey-West bandwidth
[3] Lag length selection by modified SIC (Ng & Perron 2001)
*/**/*** denotes significance at 10%, 5% and 1% level, respectively

Table 3: PP and ADF Unit Root Test Results with the EUR as Target Currency

	Jan 1999 - June 2008	
	PP[2]	ADF[3]
FX_{EUR}[1]	-2.015	-2.107
P_{EUR}	-1.263	-1.170
VIX	-2.911**	-2.746*
IRD_{EUR}	-2.098	-2.067
Y_{EUR}	-1.709	-1.574
Carry Trade Positions		
CT_{EUR}	-3.461***	-3.603***

[1] A deterministic trend is included
[2] Bartlett kernel, Newey-West bandwidth
[3] Lag length selection by modified SIC (Ng & Perron 2001)
*/**/*** denotes significance at 10%, 5% and 1% level, respectively

arises, the coefficients are estimated consistently, even if the model contains non-stationary variables (Sims et al. 1990). Furthermore, we believe that the divergence of the IRD within the threshold model subsamples is much smaller than in the full sample. Hence, it might be even stationary.[28] Therefore, we assume that the interest-rate differentials are stationary.[29]

Thus, the model contains the nominal exchange rates (ΔFX_{USD}, ΔFX_{EUR}), the prices of the stock market indices (ΔP_{USD}, ΔP_{EUR}) and ΔY_{EUR} in first differences. The IRD (IRD_{USD}, IRD_{EUR}), the VIX index (VIX) and the proxies for carry trade activities enter the model in levels (CTF_{USD} and $CTFO_{USD}$, CT_{EUR}). Furthermore, we assume the 10-year constant to maturity Treasury bond yield series to be trend-stationary and remove the linear trend from the series (Y_{USD}). Following the unit root test results, the series is at least very close to being trend-stationary.[30]. Table (4) displays the definitions of the samples.[31] We do not show all results for the samples constructed to analyze the robustness of the findings.[32]

5.1.2 Threshold Nonlinearity Test and Grid Search

Prior to testing threshold nonlinearity, we determine the number of lags included in the model. According to the Akaike and Schwarz lag length selection test results, the optimal lag length is either one or two. But with very few lags included, the estimated innovations exhibit strong serial correlations, as both multivariate and univariate Lagrange multiplier (LM) test results show. Therefore, we must include more lags to avoid endogeneity problems in our estimates. Thus, the choice of the lag length is based on serial correlation tests for the error terms. We tested for serial correlation in the residuals with the multivariate and univariate LM tests of order one, two and four. The optimal lag length of the samples A_{USD}, C_{USD} and E_{USD} is four. For the sample D_{USD}, we choose five, and for sample B_{EUR}, two lags.[33] Except for sample B_{EUR}, neither including more lags nor reducing the number of lags improves the serial correlation test results noticeably. We estimate sample B_{EUR} with two instead of three lags, because the threshold model cannot be estimated accurately otherwise.[34] Nevertheless, few error terms of the models estimated with the optimal lag length still exhibit serial correlation. The test results for the univariate se-

[28]The sample sizes of the sub-periods are too small to get reasonable results from applying unit root tests. This issue is restated in Section 5.4 where the results of the Granger causality tests are discussed.

[29]We also estimated the model with the first difference of the IRD. In contrast to the model with the IRD in levels, we do not find nonlinear effects for all sample periods. For the periods where we do find nonlinear relationships, the results support our findings.

[30]It is well known that these tests have poor power properties relative to the alternative which follows a persistent stationary stochastic process (see e.g., Christiano et al. (2003))

[31]The subscript to the sample notations indicates the target currency.

[32]These results can be obtained from the author upon request.

[33]The results of sample D_{USD} are robust to the estimation with four lags.

[34] The threshold value determined to detect the two regimes leaves for one regime too few observations for reliable estimations.

Table 4: Sample Definitions

Sample	Period	Variables[1]				
		3-Month LIBOR IRD	Carry Trade Positions	FX	Bond Yields[2]	Stock Market Index
Main Samples						
A_{USD}	March 1995 - June 2008	IRD_{USD}	CTF_{USD}	ΔFX_{USD}	Y_{EUR}	ΔP_{USD}
B_{EUR}	Jan 1999 - June 2008	IRD_{EUR}	CT_{EUR}	ΔFX_{EUR}	ΔY_{EUR}	ΔP_{EUR}
Samples for Robustness Analysis						
C_{USD}	March 1995 - Dec 2009	IRD_{USD}	CTF_{USD}	ΔFX_{USD}	Y_{EUR}	ΔP_{USD}
D_{USD}	Oct 1992 - June 2008	IRD_{USD}	CTF_{USD}	ΔFX_{USD}	Y_{EUR}	ΔP_{USD}
E_{USD}	March 1995 - June 2008	IRD_{USD}	$CTFO_{USD}$	ΔFX_{USD}	Y_{EUR}	ΔP_{USD}

[1] The sources and more details about the variables are described in Section 3.1. All samples additionally include the VIX Index.
[2] The linear trend of the 10-year constant to maturity Treasury bond yields (Y_{USD}) has been removed.

Table 5: Univariate Serial Correlation LM Test Results with Four Lags (Sample A_{USD}) and Two Lags (Sample B_{EUR})

Dependent Variable	Sample A_{USD}[1]			Sample B_{EUR}[1]		
	AR(1)	AR(2)	AR(4)	AR(1)	AR(2)	AR(4)
ΔFX_{USD} / ΔFX_{EUR}	0.040	0.101	3.150	0.081	0.117	5.153
ΔP_{USD} / ΔP_{EUR}	0.026	0.329	5.592	4.828**	6.195*	11.548**
VIX	0.005	0.137	5.006	5.972**	5.966*	7.404
IRD_{USD} / IRD_{EUR}	0.971	2.935	7.269	1.953	6.313**	12.804**
Y_{USD} / ΔY_{EUR}	1.382	2.011	4.522	0.731	2.594	4.238
Carry Trade Positions						
CTF_{USD} / CTF_{EUR}	5.408**	5.335*	5.598	1.737	5.638*	6.131

[1] The samples are described in Table (4).
*/**/*** denotes significance at 10%, 5% and 1% level, respectively

rial correlation LM test are summarized in Table (5). Moreover, the multivariate serial correlation LM test rejects the null hypothesis of no serial correlations of order four for sample A_{USD} at the 5% significance level. For sample B_{EUR}, the null hypothesis of no serial correlations of order one, two and four is rejected at the 10% significance level. The misspecification of a simple linear model might lead to these results.

The Tsay test to detect threshold nonlinearity, corrected for the possibility of conditional heteroscedasticity, is applied with delay parameters (d) equal to one, two and three.[35] For reasons discussed in Section 4, we choose the interest-rate differential as the threshold variable. The findings for all samples are shown in Table (6). Overall, we conclude that for a majority of the model specifications we can reject the null hypothesis of parameter stability. If threshold nonlinearity is present for more than one value of d, we aim to choose d such that it corresponds to the maximum of the Chi-squared test statistic. For different reasons this is not always achievable. The threshold value for sample B_{EUR} with $d = 2$ leaves for one of the two regimes too few observations for an accurate estimation. Hence, we set the delay parameter equal to three. For sample A_{USD} we choose $d = 3$ instead of $d = 2$, because the latter value is preferred for the samples C_{USD} and D_{USD}.[36] Sample E_{USD} is estimated with the delay parameter equal to three for purposes of comparison. As the differences between the test statistics are small, sample A_{USD} is estimated with $d = 1$ and $d = 3$ to check for possible variations in the impulse response functions. Our main model specifications are $A^{d=3}_{USD}$ and $B^{d=3}_{EUR}$. All versions estimated are denoted by extra bold type.

In order to estimate the multivariate threshold model, the threshold values for all model specifications are determined. The selection of the threshold value τ is based

[35] More details on the applied Tsay test is provided in Section 4.1.

[36] The eigenvalue of one regime of the model is greater than one when sample D_{USD} is estimated with $d = 1$.

Table 6: Results of the Tsay Test

Sample[1]	Delay Parameter (d)		
	1	2	3
Main Samples			
A_{USD}	**221.2*****	212.1***	**211.5*****
B_{EUR}	219.6***	274.1***	**273.5*****
Samples for Robustness Analysis			
C_{USD}	225.4***	190.0**	**247.4*****
D_{USD}	238.7***	206.0***	**222.2*****
E_{USD}	214.6***	220.5***	**183.3****

[1] The samples are described in Table (4).
$^{*}/^{**}/^{***}$ denotes significance of the Chi-squared value at 10%, 5% and 1% level, respectively
Note: Estimated models are denoted by extra bold type.

on a grid search to get the minimum Akaike Information Criterion (AIC) value. Table (7) depicts τ for the different models. As shown in Figures (1) and (2), IRD_{USD} and IRD_{EUR} are almost always positive over all sample periods. Therefore, we search for a value which separates two regimes depending on the size of the carry. One regime contains observations with values of the threshold variable greater than or equal to τ, all other observations are collected in the second regime. The threshold values are between 1.84% and 2.94%. Compared to $A^{d=3}_{USD}$, τ falls if additional observations until the end of 2009 are added (sample C_{USD}) or if a smaller delay parameter value is chosen ($d = 1$). The contrary is true for sample D_{USD} starting already with 10/06/1992. The inclusion of options positions does not alter the result.[37]

5.2 Generalized Impulse Response Functions

In this Section, we discuss the generalized impulse response functions (GIRFs) of the main samples $A^{d=3}_{USD}$ and $B^{d=3}_{EUR}$. For sample $A^{d=3}_{USD}$ we calculate the GIRFs for the regime with values of $IRD^{d=3}_{USD}$ greater than or equal to the threshold value of 2.63%. This regime is denoted as H-regime. The GIRFs for the regime with values of $IRD^{d=3}_{USD}$ smaller than 2.63% are shown in the L-regime. The same approach determines the GIRFs of sample $B^{d=3}_{EUR}$ with the threshold variable $IRD^{d=3}_{EUR}$ and the threshold value of 1.84%.

Figures (3) and (4) show the (accumulated) GIRFs of the sample $A^{d=3}_{USD}$ VAR system. The graphs in the first two columns of Figure (3) display the reactions of VIX to a

[37] In addition, for the main samples, we searched for two threshold values instead of one. The mimimum AIC of sample A_{USD} increases in the specification with two threshold values. Therefore, the model with one threshold value is preferred. For sample B_{EUR} the mimimum AIC is smaller. However, as the grid search reveals that one threshold value is almost equal to 1.84% and the other is very close to the minimum value of IRD_{EUR}, we consider only models with one threshold value.

Table 7: Threshold Values (Percentage)

Sample[1]	Delay Parameter (d)	
	1	3
Main Samples		
A_{USD}	2.12	2.63
B_{EUR}		1.84
Samples for Robustness Analysis		
C_{USD}		2.17
D_{USD}		2.94
E_{USD}		2.63

[1] The samples are described in Table (4).

one-standard deviation shock in another variable. The first (second) column includes the GIRFs of the H-regime (L-regime). The impacts on CTF_{USD} are summarized in the next two columns. Figure (4) contains the results for FX_{USD} (column 1 & 2) and P_{USD} (column 3 & 4). In general, we display the reaction of Y_{USD} instead of the impact of an innovation on itself. The GIRFs of IRD_{USD} are not included in these Figures. However, all graphs are summarized in the Figures (7)-(10) in the appendix and have a forecast horizon up to 40 weeks. We refer to these Figures when we discuss long-run effects. The Figures including the GIRFs of sample $\mathrm{B}_{EUR}^{d=3}$ are structured in the same way. All graphs contain the point estimates (solid line), the median of the bootstraps (dashed-dotted line) and the non-centered 95%-confidence interval (dotted lines).[38]

In the H-regime an unexpected increase of IRD_{USD}, through an increase in the US interest rate and/or a decrease in the Swiss interest rate, leads to a statistically significant contemporaneous rise of VIX, a reduction of CTF_{USD} and P_{USD}, as well as an appreciation of the Swiss currency (first row of Figures 3 and 4). The impacts on CTF_{USD} and P_{USD} last slightly longer than one week. While the increased IRD_{USD} improves the environment for a profitable carry trade strategy, other variables such as the risk sentiment and the US stock market indicate a rising risk for a sudden and strong unwinding of carry trades. This result is in line with the finding of Brunnermeier et al. (2009) that the conditional skewness gets more negative after an interest-rate differential shock. The response of FX_{USD} is partially in line with the prediction of the UIP. The immediate appreciation of the low-interest-rate currency could be affected by the fall in CTF_{USD}, among other factors such as the decrease in the investors risk appetite. The so-called save heaven property of the CHF might be an explanation for the lack of the intitial USD appreciation. Clarida et al. (2009) show that in high exchange rate volatility environ-

[38] More information on the bootstrap method used to determine the confidence interval are given in Section 4.3.

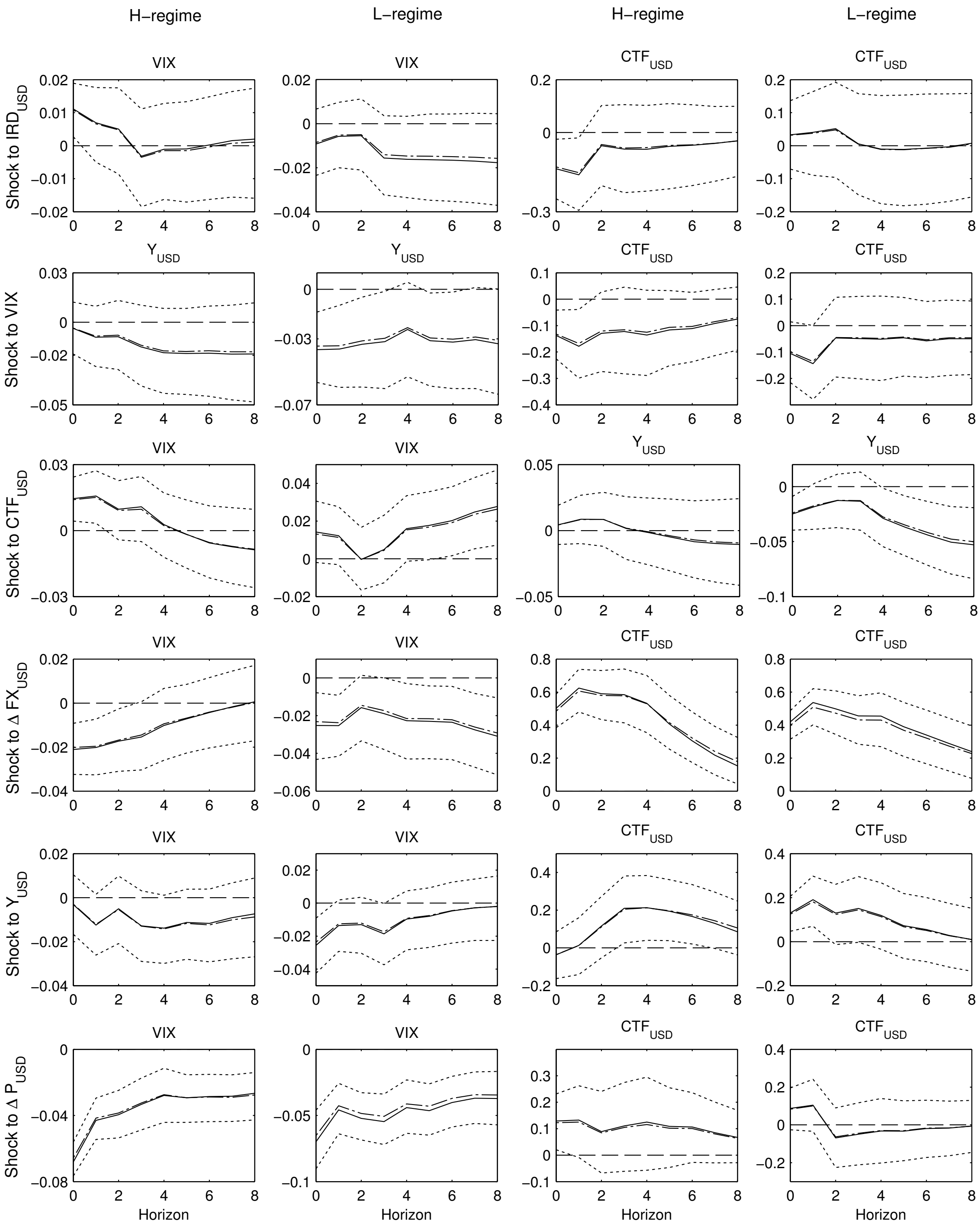

Figure 3: Sample $A_{USD}^{d=3}$: (Accumulated) GIRFs of the variables VIX, CTF_{USD} & Y_{USD}. Solid line: point estimate; dashed-dotted line: bootstrap median; dotted lines: non-centered 95%-confidence interval (small sample bias and GARCH corrected, details are described in Sections 4.2 and 4.3). For more details about sample $A_{USD}^{d=3}$ see Table (4). Number of observations: 418 (H-regime) & 270 (L-regime)

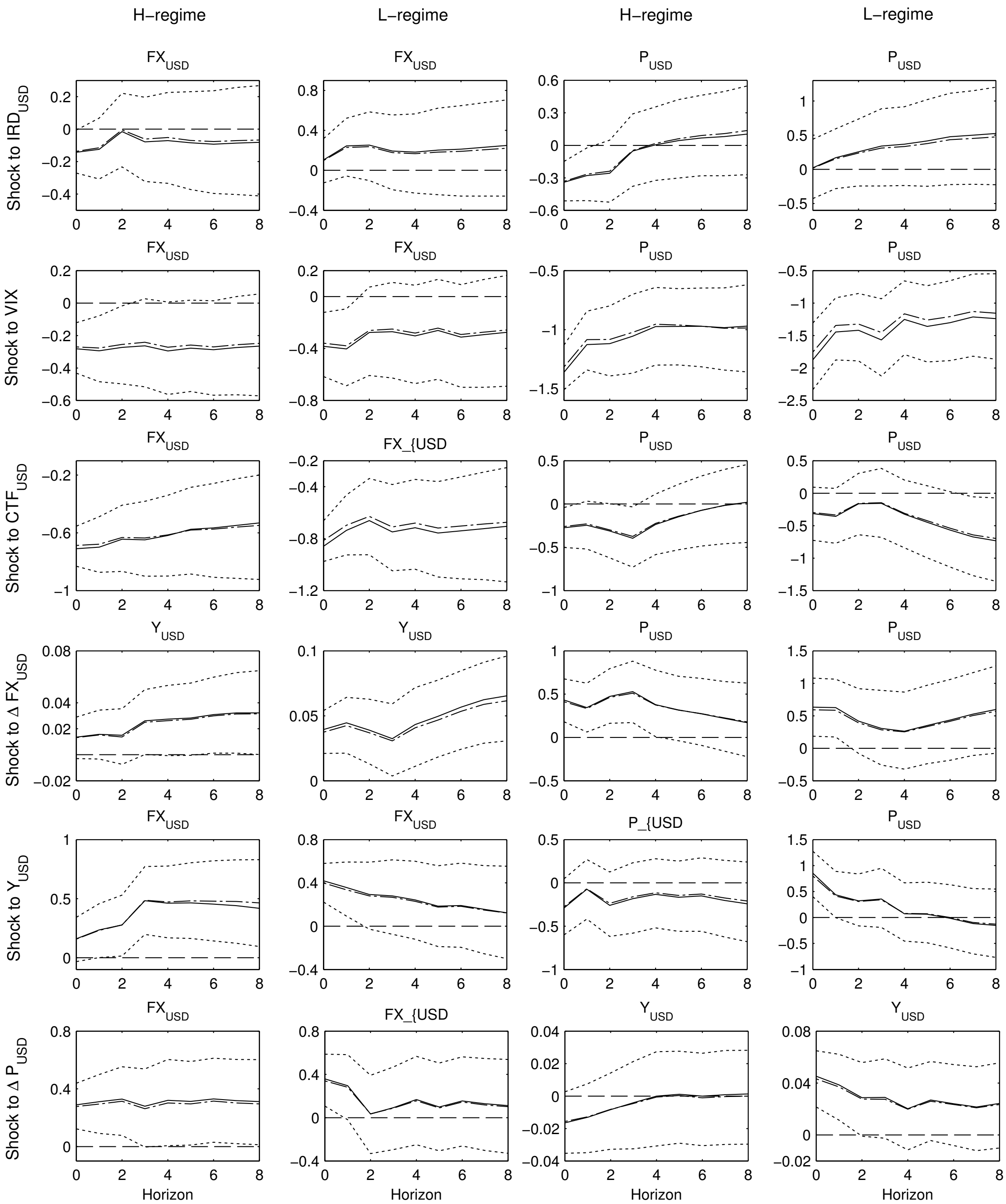

Figure 4: Sample $A^{d=3}_{USD}$: (Accumulated) GIRFs of the variables ΔFX_{USD}, ΔP_{USD} & Y_{USD}. Solid line: point estimate; dashed-dotted line: bootstrap median; dotted lines: non-centered 95%-confidence interval (small sample bias and GARCH corrected, details are described in Sections 4.2 and 4.3). For more details about sample $A^{d=3}_{USD}$ see Table (4). Number of observations: 418 (H-regime) & 270 (L-regime)

ments the low yielding currency tends to appreciate even more than implied by the UIP. Brunnermeier et al. (2009) conclude that carry trade activity in response to the shock is not enough to push up the exchange rate towards the value implied by UIP.

In the L-regime the effects are different. In the short run none of the responses are statistically significant. Nevertheless, some long-run trends are revealed. The shock tends to result in a lower risk sentiment, a continous depreciation of the CHF, pointing to the UIP puzzle, and an increase of P_{USD}. Although in the short run CTF_{USD} hardly moves, in the period between five and ten months after the shock, the buildup of CTF_{USD} is statistically significant. However, the *insignificant* appreciation of the USD on impact and its trend to further appreciate instead of a CHF appreciation as the UIP predicts, could be due to the under reaction of carry trade activity. Similar results for this variable and the nominal exchange rate can be found in Brunnermeier et al. (2009), who do not distinguish between different interest-rate regimes. To summarize, in the H-regime a further increase in the carry leads rather to a fall in CTF_{USD} in the short run and in the L-regime to a rise in the long run. These opposed effects arise due to different risk environments and/or exchange rate fluctuations. While the carry is the key variable determining carry trades in the model of Plantin & Shin (2010), our empirical result suggests that the associated changes in risk and the exchange rate are more important.

A shock to VIX (second row) gives rise to a statistically significant contraction of CTF_{USD} in both regimes, which is in line with the results found by Brunnermeier et al. (2009) and Nishigaki (2007). This result is not surprising, as an increase in VIX represents a higher risk sentiment. The decline is slightly stronger in the H-regime, reflecting an increased risk aversion of the speculators with a larger carry. The effects on the two variables displayed in Figure (4) are similar across both regimes. The initial decrease of FX_{USD} and P_{USD} is somewhat larger in the L-regime.

What happens to the variables in the VAR after a an unexpected unwinding of carry trades? Brunnermeier et al. (2009), for instance, conjecture that sudden exchange rate fluctuations unrelated to fundamental news events can be triggered when investors near funding constraints. We expect a strong appreciation of the CHF as the demand for the Swiss currency rises sharply. The first two graphs in the third row of Figure (4) confirm this assumption. The currency appreciates contemporaneously in both regimes.[39] A shock whose size is twice the standard deviation of CTF_{USD} leads to an immediate appreciation of the CHF of about three percent in the H-regime and four percent in the L-regime. In the L-regime the CHF starts to depreciate after the sudden appreciation. The effect is diminishing over time and ceases to be statistically significant after four months (see Figure 8 in the appendix). In contrast, we find a slight overshooting in the H-regime, and

[39]The variance decomposition based on the Cholesky decomposition ordering in line with Nishigaki (2007)(IRD_{USD}, VIX, CTF_{USD}, FX_{USD}, Y_{USD} and P_{USD}) reveals that the semi-structured carry trade activities shock explains about 25% of FX_{USD} in both regimes. It is the most important shock apart from the own shock.

the Swiss currency remains appreciated against the US currency over the whole period. In the study of Nishigaki (2007), the appreciation of the yen is also statistically significant and lasts for almost two years. Additionally, in both regimes we find an increase of VIX and a fall of P_{USD}. Whereas in the H-regime the effects are statistically significant in the short run, in the other regime they are significant in the medium run too. The latter is also true for the decrease of Y_{USD}.

An unexpected depreciation of the Swiss franc displayed in the fourth row results in a large and statistically significant buildup of CTF_{USD}. The reduction of the positions over time is (marginally) slower in the L-regime. This could be due to the slower mean reversion of VIX, which falls after the shock in both regimes. The increase of Y_{USD} is statistically significant for a longer period in the L-regime, whereas the initial rise of P_{USD} lasts longer in the H-regime.

In the L-regime the investors build up CTF_{USD}, the risk sentiment declines, the USD appreciates and there is a statistically significant rise in P_{USD} over a one to two week period in response to a positive innovation in Y_{USD}. These effects also occur in the H-regime for CTF_{USD} and FX_{USD} with a delay of approximately two weeks.

An unexpected rise of P_{USD}, depicted in the last row, induces a sudden drop of VIX and an appreciation of the US currency in both regimes. Both effects last longer in the H-regime. This might be an explanation for the longer horizon over which CTF_{USD} increases, although not statistically significant for all horizons (see Figure 7 in the appendix). Positive shocks to P_{USD} increases the value of a stock portfolio investors would like to use as collateral for liquidity, to engage in carry trade activities. Nishigaki (2007) finds a persistent fall of yen carry trade positions after a negative US stock market shock.

Now we turn to the results for sample $\mathrm{B}_{EUR}^{d=3}$, shown in the Figures (5) and (6).

The last two graphs of the first row of Figure (6) show IRD_{EUR} in lieu of P_{EUR}. Not surprisingly, a positive innovation to IRD_{EUR} results in a statistically significant rise of CT_{EUR}, which has IRD_{EUR} as its numerator. However, compared to IRD_{EUR} the rise is smaller, hence, the implicit nominal exchange rate volatility increases too. In the long run, depicted in Figure (9) in the appendix, the effect on CT_{EUR} is statistically significant for longer in the L-regime. Apart from the fact that the increase in IRD_{EUR} is statistically significant for a longer period, the negative trend of VIX, and the increase of FX_{EUR}, Y_{EUR} and P_{EUR} might influence this pattern (see Figure (9) in the appendix). This finding is comparable to the results for sample $\mathrm{A}_{USD}^{d=3}$.

Compared to sample $\mathrm{A}_{USD}^{d=3}$, the impulse response functions associated with an innovation to VIX are qualitatively similar, but in the L-regime more pronounced. The fall of CT_{EUR} in the L-regime could be driven by the strong appreciation of the CHF against the EUR. The graph in the third column of Figure (9) in the appendix depicts the rise of CT_{EUR} in the long run. This could be in virtue of the faster mean reversions of VIX and FX_{EUR}, compared to the L-regime.

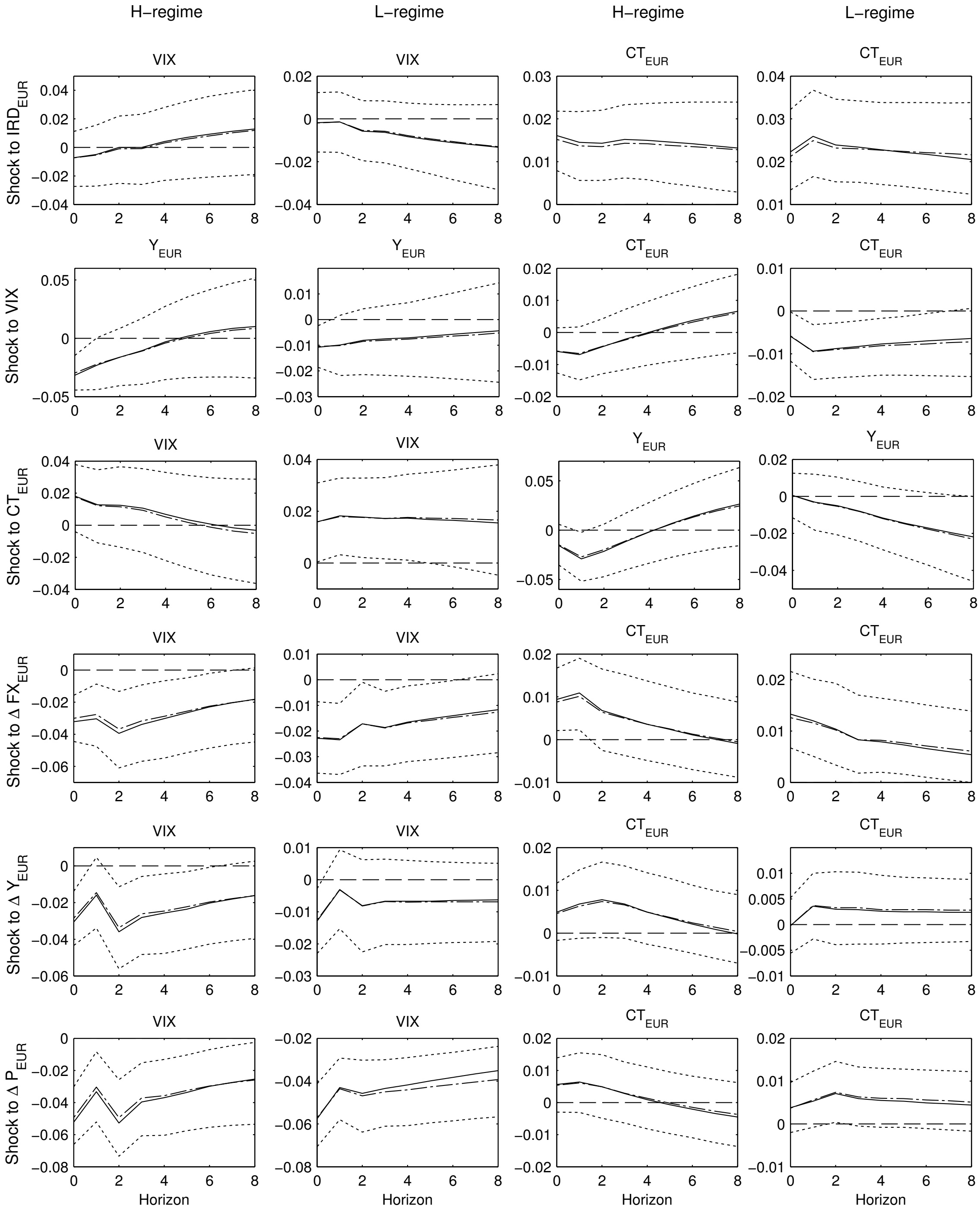

Figure 5: Sample $B_{EUR}^{d=3}$: (Accumulated) GIRFs of the variables VIX, CT_{EUR} & ΔY_{EUR}. Solid line: point estimate; dashed-dotted line: bootstrap median; dotted lines: non-centered 95%-confidence interval (small sample bias and GARCH corrected, details are described in Sections 4.2 and 4.3). For more details about sample $B_{EUR}^{d=3}$ see Table (4). Number of observations: 125 (H-regime) & 367 (L-regime)

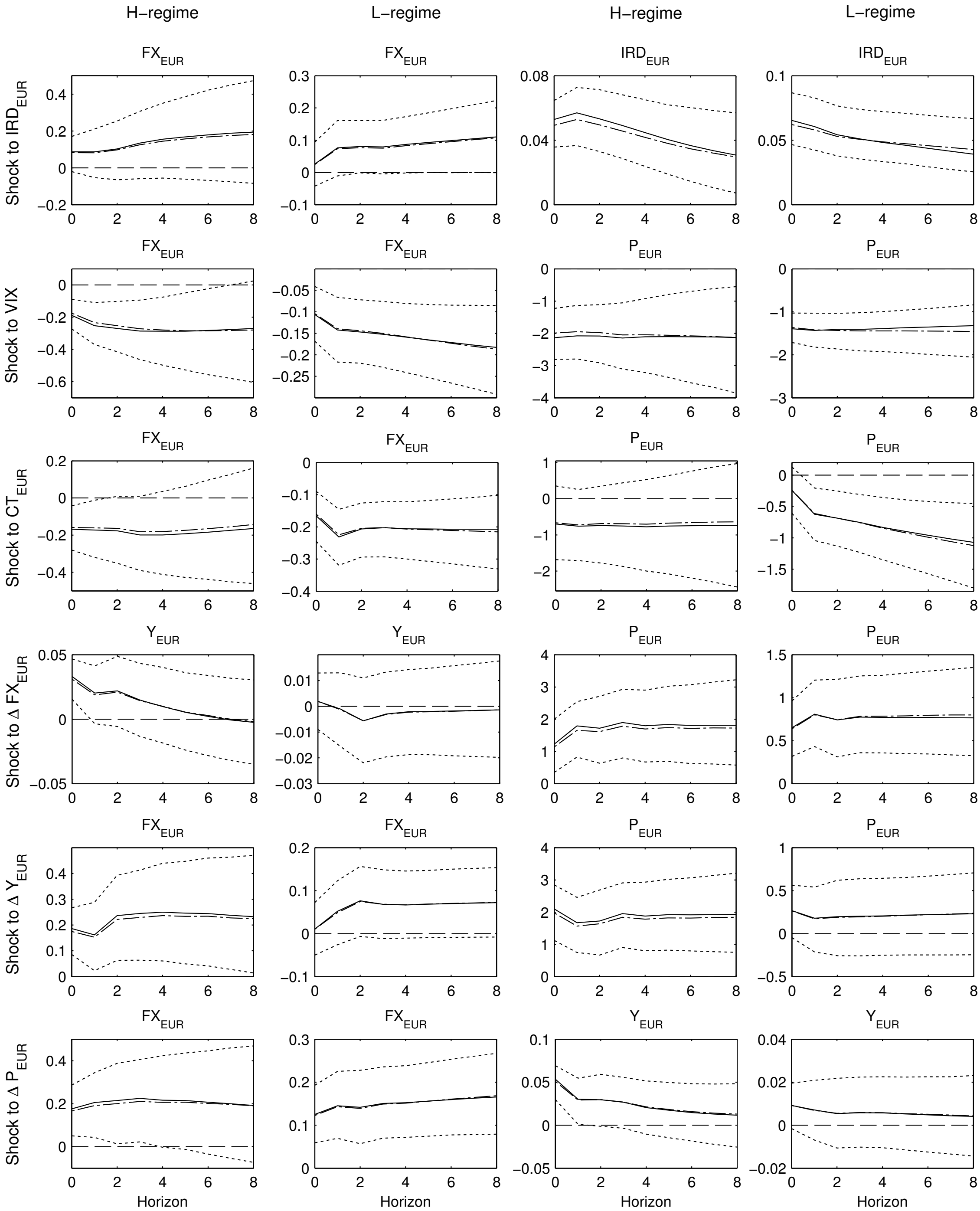

Figure 6: Sample $\mathrm{B}_{EUR}^{d=3}$: (Accumulated) GIRFs of the variables ΔFX_{EUR}, ΔP_{EUR} & ΔY_{EUR}. Solid line: point estimate; dashed-dotted line: bootstrap median; dotted lines: non-centered 95%-confidence interval (small sample bias and GARCH corrected, details are described in Sections 4.2 and 4.3). For more details about sample $\mathrm{B}_{EUR}^{d=3}$ see Table (4). Number of observations: 125 (H-regime) & 367 (L-regime)

A one-standard deviation shock to CT_{EUR} gives rise to an expected appreciation of the CHF.[40] The initial impact is equal for both regimes, but the mean reversion is slower in the L-regime. If the shock equals twice the standard deviation of CT_{EUR} the sudden appreciation of the CHF is slightly more than one percent in both regimes. Moreover, the effect is smaller compared to the sample $\mathrm{A}_{USD}^{d=3}$. Though, as the the proxy for carry trade activities is different, a one-to-one comparison is impossible. Additionally, in both regimes we find an increase of VIX in the short run and a fall of P_{EUR}. However, these impacts are only statistically significant in the L-regime. As in sample $\mathrm{A}_{USD}^{d=3}$, in the L-regime Y_{EUR} decreases in the medium run.

In line with Sample $\mathrm{A}_{USD}^{d=3}$, an unexpected depreciation of the Swiss currency leads to a fall in risk sentiment and an increase in CT_{EUR} and P_{EUR}. However, the effect on P_{EUR} is more pronounced and in the H-regime, CT_{EUR} asymptotes quickly to its steady-state level within two weeks.

The fifth row of the Figures (5) and (6) contains the impulse responses on an innovation to ΔY_{USD}. The rise of CT_{EUR} is hardly statistically significant, contrary to the findings for sample $\mathrm{A}_{USD}^{d=3}$. The CHF tends to depreciate, the VIX to fall and P_{EUR} goes up statistically significant in the H-regime. With the exception of the responses in share prices, the H-regime of both samples show comparable GIRFs.

A shock to ΔP_{EUR} is depicted in the last row. In the short run, the impacts are qualitatively the same as in sample $\mathrm{A}_{USD}^{d=3}$, except for Y_{EUR} in the H-regime. In the L-regime the rise of CT_{EUR} becomes statistically significant two weeks after the shock for one period. The stronger impact here compared to the H-regime may be a consequence of the severe and persistent depreciation of the Swiss currency.

Overall, we note that there are substantial differences across regimes depending on the size of the carry. Furthermore, the comparison of the two samples reveals further that the risk sentiment, the exchange rates, the bond yields and the stock market indices show similar (qualitative) patterns with few exceptions, especially for the exchange rate and the bond yields. Carry traders seem to react also likewise, although the proxies for carry trade activities differ substantially.

5.3 GIRFs: Robustness Analysis

Overall, the robustness analysis demonstrates robust findings across the different samples. In the following, we describe the changes and point out some important qualitative and quantitative divergences from sample $\mathrm{A}_{USD}^{d=3}$.

[40] The variance decomposition based on the Cholesky decomposition ordering in line with Nishigaki (2007)(IRD_{USD}, VIX, CTF_{USD}, FX_{USD}, Y_{USD} and P_{USD}) reveals that the semi-structured carry trade activities shock explains about 5% of FX_{EUR} in the H-regime. Apart from the own shock it is the second most important shock. In the L-regime it is the most important shock apart from the own shock and explains about 16% of FX_{EUR}.

5.3.1 Delay Parameter

Since the Chi-squared test statistic for the delay parameter equal to one is the largest among the different delay parameters (see Table 6), we also estimate sample $A^{d=1}_{USD}$. While the GIRFs of the H-regime reveal no qualitative and quantitative differences, the positive long-run impact of a shock to IRD_{USD} on CTF_{USD} is not statistically significant in the L-regime. This might be due to the somewhat faster mean reversion of the carry, a slightly smaller reduction of VIX and a less pronounced depreciation of the Swiss currency.

5.3.2 Sample Period Selection

We extended the sample period to include observations of the recent financial crises (sample $C^{d=3}_{USD}$). The GIRFs of the H-regime are very robust to this modification. Yet, several GIRFs of the L-regime exhibit distinct differences compared to the results of sample $A^{d=3}_{USD}$. A one-standard deviation shock to IRD_{USD} has no impact on VIX, FX_{USD} and P_{USD} anymore, i.e. the responses show no trend either way. The absence of these trends might explain that investors do not increase CTF_{USD} in the long run. Besides the modification of the sample length, the reduction of the threshold value determines this result (see Table 6). The mean reversion of FX_{USD} after an unexpected unwinding of carry trades takes longer in sample $C^{d=3}_{USD}$. Moreover, the impacts on VIX and P_{USD} are not statistically significant anymore. This also holds when the Swiss currency depreciates unexpectedly. In general, the confidence intervals for the impulse response functions for the L-regime are expanded, pointing to the increased uncertainty during the financial crisis.

As weekly published CME futures positions are available since October 1992, sample $D^{d=3}_{USD}$ contains data from 1992/10/06 until 2008/06/24. The GIRFs of the H-regime are robust to this modification. In contrast to sample $A^{d=3}_{USD}$, the rise of VIX in response to an unexpected decrease of CTF_{USD} only marginally fails to pass the 5% significance level. However, more substantial changes are observed for the L-regime. A shock to IRD_{USD} gives rise to a statistically insignificant increase of CTF_{USD} in the medium run. This lack of significance is somewhat surprising, because the fall of VIX is statistically significant during three weeks. Yet, after an initial tendency to depreciate, the Swiss currency does not continue to follow a depreciation trend in the long run. An unexpected rise of VIX leads to a longer appreciation of the CHF and fall of P_{USD}. Furthermore, the decline of CTF_{USD} is less pronounced and far from being statistically significant. In the medium run, VIX, Y_{USD} and P_{USD} cease to respond statistically significantly to a sudden unwinding of carry trades. Moreover, FX_{USD} exhibits a slower mean reversion.[41] Finally, when P_{USD} goes up unexpectedly, the increase of CTF_{USD} on impact is statistically significant, in contrast to the jump in sample $A^{d=3}_{USD}$. This change arises due to the increase in the

[41] While FX_{USD} is still below its steady-state level after ten months in sample $D^{d=3}_{USD}$, the mean reversion is completed within five months in sample $A^{d=3}_{USD}$.

threshold value, whereas all the other deviations cannot be ascribed to the threshold value change.

5.3.3 Futures and Options Positions

The inclusion of options to the CME futures positions to proxy carry trade activities (sample $E_{USD}^{d=3}$) causes no qualitative change in either regime. However, for the L-regime, the decline of $CTFO_{USD}$ in response to an innovation to VIX is statistical significant for the first week. The same is true for the rise after a shock to P_{USD}. Furthermore, an unexpected increase of Y_{USD} has a longer statistically significant impact of one week on $CTFO_{USD}$. In the H-regime the reaction of $CTFO_{USD}$ to an unexpected increase of IRD_{USD} is slightly less pronounced.

5.3.4 Choice of the Interest Rate

Next, we assess whether the chosen interest rate has any impact on our results. Therefore, we replace the 3-month interbank interest rates with the 1-month interbank interest rates. While this replacement of the interest-rate differentials has no impact on the GIRFs with the USD as target currency, the rise of CTF_{EUR} in response to a sudden increase of IRD_{EUR} is not statistically significant anymore in the H-regime.

5.4 Granger Causality Analysis

In this Section, we shed light on the question of whether one variable in our models moves ahead of the others, i.e. if the variables "Granger-cause" each other. Following Klitgaard & Weir (2004) and Mogford & Pain (2006) position data do not help in anticipating exchange rate movements for the subsequent week. Their insights are based on a Granger (1969) causality test with two variables, the net futures positions and the nominal exchange rate.[42] We extend their analysis in two ways. First, we include additional variables in our model which have the potential to "Granger-cause" another variable. Second and more important, we distinguish the effects between regimes, depending on the size of the interest-rate differential (IRD). If the value of the threshold variable is greater or equal (smaller) than the threshold value, the corresponding observations are assigned to the H-regime (L-regime).[43]

In a first step, the proxy for carry trade positions is excluded from the multivariate threshold model to examine the power of this variable to "Granger-cause" the other variables in the model. Table (8) displays the findings for each regime of all samples.[44]

[42]In their studies, both variables are first differenced prior to the estimation.

[43]Tables (12) and (13) in the appendix show all results for the main samples.

[44]If a VAR model contains one or more random walk series without cointegration relationship, the Granger causality test statistics have a nonstandard limiting distribution (Sims et al. 1990). The unit root tests reveal that the IRD are non-stationary. Nevertheless, we assume this series to be stationary and

Table 8: Granger Causality Test: Carry Trade Positions Excluded

Sample[1]/ Regime[2]	Variable[1]				
	IRD_{USD} / IRD_{EUR}	VIX	ΔFX_{USD} / ΔFX_{EUR}	Y_{USD} / Y_{EUR}	ΔP_{USD} / ΔP_{EUR}
$A_{USD}^{d=3}$					
H	10.560**	9.865**	2.444	3.758	6.493
L	1.390	17.172***	10.206**	14.241***	16.136***
$B_{EUR}^{d=3}$					
H	1.575	0.257	0.373	7.325**	0.617
L	22.717***	1.862	8.281**	7.832**	11.539***
$C_{USD}^{d=3}$					
H	14.376***	5.547	2.563	4.288	4.806
L	2.610	10.271**	2.945	2.015	7.388
$D_{USD}^{d=3}$					
H	12.945**	8.783*	2.365	5.058	6.532
L	2.287	8.363*	6.344	9.874**	4.577
$E_{USD}^{d=3}$					
H	2.894	6.961	4.979	2.072	4.365
L	1.290	15.622***	7.904*	15.800***	14.855***

[1] The samples and variables are described in Table (4).
[2] Observations for which the threshold variable lies above the threshold value are assigned to the H-regime; for values below the threshold values, the observations are included in the L-regime. The threshold values are given in Table (7).
*/**/*** denotes significance of the Chi-squared value at 10%, 5% and 1% level, respectively

In all three models containing futures position data as proxy for carry trade positions, these positions have predictive power for the IRD in the H-regime. The contrary is true for sample $B_{EUR}^{d=3}$, where carry trade activities lead the IRD in the L-regime. This highly statistically significant result, however, has to be interpreted with caution as the IRD is the enumerator of the carry-to-risk ratio (CTR ratio), which is the proxy for carry trade positions.

However, the predictive power of the proxy for carry trade activities is often statistically (more) significant in the L-regime, for example, with respect to nominal exchange rate fluctuations. For all samples, the Chi-squared value for the L-regime is substantially larger and twice statistically significant at the 5% level and once at the 10% level. This result challenges the insights of Klitgaard & Weir (2004) and Mogford & Pain (2006) in the sense that in times with small IRDs there is the possibility that past position data help to predict exchange rate movements. The knowledge about speculative future po-

refer to the standard test statistics since the spread of the IRD is smaller within the regimes compared to the full sample. Further, there is no economic reason for a random walk behavior. The sample sizes of the regimes are too small to get reasonable results from applying unit root tests.

sitions seems to have incremental information about future fluctuations in the exchange rate in line with findings from the literature, pioneered by Evans & Lyons (2002, 2005), that tries to explain and empirically forecast exchange rate movements based on a microstructure approach. The microstructure approach assumes that, apart from common knowledge macroeconomic information (macro approach), heterogeneous beliefs are essential for exchange rate determination as well. In a hybrid view, macroeconomic information influences the exchange rate directly and indirectly through order flow which reveal price-relevant private information such as, for example, heterogeneous interpretations of news or changes in expectations (Rime et al. 2010).[45] Evans & Lyons (2002) provide a theoretical model that integrates both approaches and find empirically that adding order flow as an explanatory variable to a regression of changes in exchange rates on IRDs, serving as a proxy for public macroeconomic information, increases the R-squared from 1-5% to 40-60%. As Evans & Lyons (2005) note, order flow data have not only explanatory but also forecasting power for the exchange rate if the market learns gradually from order flow information. Following the out-of-sample studies by Evans & Lyons (2005) and Rime et al. (2010), order flow is a powerful predictor for exchange rates fluctuations. Like order flow information the CME future position data are not discovered by the market immediately and therefore do not constitute public information. The U.S. Commodity Futures Trading Commission provides the data with a delay of some days (usually three days).

In a second step, the predictive power of all other variables on carry trade positions is determined. The findings are displayed in Table (9). They suggest that exchange rate movements are very important for anticipating carry trade activities, independent of the regime, except for sample $\mathrm{B}_{EUR}^{d=3}$. It is therefore more likely that movements in the exchange rate precede position data than vice versa. This result is in line with the findings reported by Mogford & Pain (2006).[46] The results indicate a basic form of trend-following behavior among the speculative traders at the CME. Movements in the exchange rate FX_{EUR} do not "Granger-cause" the CTR ratio,[47] but the IRD and the CTR ratio seem to “Granger-cause” each other in the L-regime (see also Table 8). This might be due to the calculation of the CTR ratio with the IRD as its enumerator.

Moreover, in all samples, movements in P_{USD} help to predict position data in periods with $IRD_{USD}^{d=3}$ below the threshold value. This might be because the stock market is a proxy for possible liquidity constraints, as the value of investor collateral portfolios descreases.

[45] Order flow is defined as the net of buyer and seller initiated currency transactions. Hence, it is a measure of net buying pressure (Evans & Lyons 2002).

[46] Klitgaard & Weir (2004) also obtain a statistically significant test statistic for the CHF, but not for most other currencies.

[47] We assume that the CTR ratio is an important indicator for carry traders to adjust their positions. However, as long as investors do not follow strictly this indicator we can not rule out potential feedback trading.

Table 9: Granger Causality Test: Which Variables "Granger-cause" Carry Trade Positions?

Sample[1]/ Regime[2]	Variable[1] excluded				
	IRD_{USD} / IRD_{EUR}	VIX	ΔFX_{USD} / ΔFX_{EUR}	Y_{USD} / Y_{EUR}	ΔP_{USD} / ΔP_{EUR}
$A_{USD}^{d=3}$					
H	3.062	2.490	20.562***	7.204	1.602
L	3.455	3.565	24.868***	7.102	16.149***
$B_{EUR}^{d=3}$					
H	2.810	0.942	1.762	2.448	0.139
L	17.838***	4.398	0.144	4.067	2.586
$C_{USD}^{d=3}$					
H	4.422	3.373	22.353***	6.040	3.702
L	3.474	1.234	23.797***	5.237	10.955**
$D_{USD}^{d=3}$					
H	4.564	3.910	21.161***	6.956	1.568
L	6.349	1.841	16.781***	2.855	8.912*
$E_{USD}^{d=3}$					
H	7.098	3.167	33.024***	8.922*	2.298
L	4.998	4.376	29.220***	6.690	18.901***

[1] The samples and variables are described in Table (4).
[2] Observations for which the threshold variable lies above the threshold value are assigned to the H-regime; for values below the threshold values, the observations are included in the L-regime. The threshold values are given in Table (7).
*/**/*** denotes significance of the Chi-squared value at 10%, 5% and 1% level, respectively

Table 10: Granger Causality Test: All but Own Lags Excluded

Sample[1]/ Regime[2]	Variable[1]					
	IRD_{USD} / IRD_{EUR}	VIX	CTF_{USD} / CT_{EUR}	ΔFX_{USD} / ΔFX_{EUR}	Y_{USD} / Y_{EUR}	ΔP_{USD} / ΔP_{EUR}
$A_{USD}^{d=3}$						
H	27.068	28.880*	40.000***	23.204	18.022	28.460*
L	42.165***	51.921***	59.574***	35.185**	35.218**	38.106***
$B_{EUR}^{d=3}$						
H	13.100	13.992	12.948	7.331	26.838***	11.821
L	45.636***	10.682	29.216***	26.503***	12.457	20.118**
$C_{USD}^{d=3}$						
H	30.304*	42.396***	44.700***	21.828	25.673	30.612*
L	35.492**	51.843***	45.201***	33.359**	13.232	29.718*
$D_{USD}^{d=3}$						
H	30.053**	30.581*	42.717***	24.247	16.552	29.637*
L	41.473***	33.136**	46.225***	31.131*	50.045***	32.040**
$E_{USD}^{d=3}$						
H	19.107	25.846	60.500***	25.864	16.279	26.222
L	42.050***	50.183***	65.506***	32.679**	36.892**	36.727**

[1] The samples and variables are described in Table (4).
[2] Observations for which the threshold variable lies above the threshold value are assigned to the H-regime; for values below the threshold values, the observations are included in the L-regime. The threshold values are given in Table (7).
*/**/*** denotes significance of the Chi-squared value at 10%, 5% and 1% level, respectively

Finally, we analyze the Granger causality of all but one series on the remaining variable, i.e. if the value of a specific variable at time t is preceded by the lagged values of all other series. The results are shown in Table (10). With three exceptions, the past values of the others are helpful to anticipate its current value in the L-regime.[48] This insight coincides, except for IRD_{USD}, with the one of Table (8), where solely the carry trade positions are excluded. The Chi-squared test statistics for IRD_{USD} displayed in Table (10) are always larger for the L-regime compared to the H-regime and statistically significant. Though, carry trade position data "Granger-cause" IRD_{USD} in only three samples in the H-regime (see Table 8). In anticipating carry trade activities, the other variables seem to be very helpful, except for the H-regime in sample $\mathrm{B}_{EUR}^{d=3}$. The results of Table (9) indicate that this finding is mainly driven by the nominal exchange rate movements.

6 Summary and Conclusion

We choose a multivariate time series model to assess the effects of an unexpected movement of one variable on the others. Carry traders react to shocks to the variables that

[48] The exceptions are: VIX and Y_{EUR} in sample $\mathrm{B}_{EUR}^{d=3}$ and Y_{USD} in sample $\mathrm{C}_{USD}^{d=3}$

determine the profitability of their investment strategy, such as the interest-rate differential (the so-called "carry"), the nominal exchange rate, the risk sentiment, the investment return and possible liquidity constraints. In addition, these variables can move due to unexpected carry trade activities.

Preliminary analyses of the carry indicate a nonlinear relationship among the variables in our model. Therefore, we apply a multivariate threshold model to account for possible changes in the dynamic behavior of carry trade activities depending on the size of the interest-rate differential.

By analyzing the generalized impulse response functions, we find the following main results. First, carry trade positions are driven to a large extend by the expected risk on financial markets and the exchange rate. Since the other variables responses to a shock depend on the size of the carry, these differences are carried over to the speculators carry trade positions. The results indicate that in times with a large carry a positive one-standard deviation shock to the carry itself is not enough to compensate investors for the increased risk. Moreover, in line with the prediction of the UIP, the CHF appreciates instantaneously against the USD in times with high IRDs, but not in a regime with low IRDs. Second, liquidity constraints can also be important, whereas the carry itself plays only a minor role. Third, a sudden unwinding of carry trades has a significant impact on the nominal exchange rate, independent of the size of the IRD. Finally, we conclude that a majority of the responses show similar patterns for the USD/CHF and EUR/CHF exchange rates, although the proxy for carry trade positions differs.

Furthermore, Granger causality tests reveal that in periods with low-interest-rate differentials past position data help to predict exchange rate movements. On the other hand, in samples with the USD as target currency, the exchange rate has very high predictive power for carry trade activities, pointing to feedback trading.

References

Bernholz, P. & Kugler, P. (2010), 'Financial market integration in the early modern period in Spain: Results from a threshold error correction model', *Economics Letters* . Forthcoming.

Brunnermeier, M. K., Nagel, S. & Pedersen, L. H. (2009), 'Carry trades and currency crashes', *NBER Macroeconomics Annual* **23**(1), 313–348.

Burnside, C., Eichenbaum, M., Kleshchelski, I. & Rebelo, S. (2006), The returns to currency speculation, NBER Working Paper 12489, National Bureau of Economic Research.

Cai, J., Cheung, Y.-L., Lee, R. S. & Melvin, M. (2001), '"Once-in-a-generation" yen volatility in 1998: fundamentals, intervention, and order flow', *Journal of International Money and Finance* **20**, 327–347.

Canjels, E., Prakash-Canjels, G. & Taylor, A. M. (2004), 'Measuring market integration: Foreign exchange arbitrage and the Gold Standard, 1879-1913', *The Review of Economics and Statistics* **86**(4), 868–882.

Christiano, L. J., Eichenbaum, M. & Vigfusson, R. (2003), What happens after a technology shock?, NBER Working Paper 9819, National Bureau of Economic Research.

Christiansen, C., Ranaldo, A. & Söderlind, P. (2010), 'The time-varying systematic risk of carry trade strategies', *Journal of Financial and Quantitative Analysis* . Forthcoming.

Clarida, R., Davis, J. & Pedersen, N. (2009), Currency carry trade regimes: Beyond the Fama regression, NBER Working Paper 15523, National Bureau of Economic Research.

Darvas, Z. (2009), 'Leveraged carry trade portfolios', *Journal of Banking & Finance* **33**, 944–957.

Dickey, D. A. & Fuller, W. A. (1979), 'Distribution of the estimators for autoregressive time series with a unit root', *Journal of the American Statistical Association* **74**, 427–431.

Elliott, G., Rothenberg, T. J. & Stock, J. H. (1996), 'Efficient tests for an autoregressive unit root', *Econometrica* **64**, 813–836.

Engel, C. (1996), 'The forward discount anomaly and the risk preimum: A survey of recent evidence', *Journal of Empirical Finance* **3**, 123–192.

Engle, R. (2001), 'Garch 101: The use of ARCH/GARCH models in applied econometrics', *Journal of Economic Perspectives* **15**(4), 157–168.

Evans, M. D. D. & Lyons, R. K. (2002), 'Order flow and exchange rate dynamics', *The Journal of Political Economy* **110**(1), 170–180.

Evans, M. D. D. & Lyons, R. K. (2005), 'Meese-Rogoff redux: Micro-based exchange-rate forecasting', *American Economic Review* **95**(2), 405–414.

Fama, E. F. (1984), 'Forward and spot exchange rates', *Journal of Monetary Economics* **14**, 319–338.

Farhi, E., Fraiberger, S. P., Gabaix, X., Ranciere, R. & Verdelhan, A. (2009), Crash risk in currency markets, NBER Working Paper 15062, National Bureau of Economic Research.

Gagnon, J. E. & Chaboud, A. P. (2007), What can the data tell us about carry trades in Japanese yen?, International Finance Discussion Papers 899, Board of Governors of the Federal Reserve System.

Galati, G., Heath, A. & McGuire, P. (2007), 'Evidence of carry trade activity', *BIS Quarterly Review* .

Goncalves, S. & Kilian, L. (2004), 'Bootstrapping autoregressions with conditional heteroskedasticity of unknown form', *Journal of Econometrics* **123**, 89–120.

Granger, C. W. J. (1969), 'Investigating causal relations by econometric models and cross-spectral methods', *Econometrica* **37**(3), 424–38.

Hattori, M. & Shin, H. S. (2007), The broad yen carry trade, IMES Discussion Paper Series E-19, Bank of Japan.

Jordà, O. & Taylor, A. M. (2009), The carry trade and fundamentals: Nothing to fear but FEER itself, NBER Working Paper 15518, National Bureau of Economic Research.

Jurek, J. W. (2007), Crash-neutral currency carry trades, AFA 2010 Atlanta meetings paper, The American Finance Association.

Kilian, L. (1998a), 'Confidence intervals for impulse responses under departures from normality', *Econometric Reviews* **17**(1), 1–29.

Kilian, L. (1998b), 'Small-sample confidence intervals for impulse response functions', *The Review of Economics and Statistics* **80**(2), 218–230.

Klitgaard, T. & Weir, L. (2004), Exchange rate changes and net positions of speculators in the futures market, Economic policy review, Federal Reserve Bank of New York.

Kohler, D. (2007), Carry trades: Betting against Safe Haven, Discussion Paper 2007-12, University of St. Gallen Department of Economics.

Koop, G., Pesaran, M. H. & Potter, S. (1996), 'Impulse response analysis in nonlinear multivariate models', *Journal of Econometrics* **74**(1), 119–147.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. & Shin, Y. (1992), 'Testing the null hypothesis of stationary against the alternative of a unit root', *Journal of Econometrics* **54**, 159–178.

Lustig, H., Roussanov, N. & Verdelhan, A. (2008), Common risk factors in currency markets, NBER Working Paper 14082, National Bureau of Economic Research.

McGuire, P. & Upper, C. (2007), 'Detecting FX carry trades', *BIS Quarterly Review* .

Meese, R. A. & Rogoff, K. (1983), 'Empricial exchange rate models of the seventies: Do they fit out of sample?', *Journal of International Economics* **14**, 3–24.

Mogford, C. & Pain, D. (2006), The information content of aggregate data on financial futures positions, Quarterly Bulletin, Bank of England.

Ng, S. & Perron, P. (2001), 'Lag length selection and the construction of unit root tests with good size and power', *Econometrica* **69**(6), 1519–1554.

Nishigaki, H. (2007), 'Relationship between the yen carry trade and the related financial variables', *Economics Bulletin* **13**, 1–7.

Pesaran, H. H. & Shin, Y. (1996), 'Cointegration and speed of convergence to equilibrium', *Journal of Econometrics* **71**, 117–143.

Pesaran, H. H. & Shin, Y. (1998), 'Generalized impulse response analysis in linear multivariate models', *Economics Letters* **58**, 17–29.

Phillips, P. C. & Perron, P. (1988), 'Testing for a unit root in time series regression', *Biometrika* **75**, 335–346.

Plantin, G. & Shin, H. S. (2010), Carry trades and speculative dynamics, Unpublished manuscript, Princeton University and Toulouse School of Economics.

Ranaldo, A. & Söderlind, P. (2010), 'Safe haven currencies', *Review of Finance* **14**(3), 385–407.
URL: *http://rof.oxfordjournals.org/content/14/3/385.abstract*

Rime, D., Sarno, L. & Sojli, E. (2010), 'Exchange rate forecasting, order flow and macroeconomic information', *Journal of International Economics* **80**(1), 72–88.

Roth, J.-P. (2007), Die Rolle des Franken-Wechselkurses im geldpolitischen Konzept der Schweizerischen Nationalbank, Referat im Rahmen des FEW-HSG, Universität St. Gallen.

Runkle, D. E. (1987), 'Vector autoregressions and reality', *Journal of Business and Economic Statistics* **5**, 437–442.

Sims, C. A. (1980), 'Macroeconomics and reality', *Econometrica* **48**(1), 1–48.

Sims, C. A., Stock, J. H. & Watson, M. W. (1990), 'Inference in linear time series models with some unit roots', *Econometrica* **58**(1), 113–144.

Stock, J. H. & Watson, M. W. (2001), 'Vector autoregressions', *Journal of Economic Perspectives* **15**(4), 101–115.

Tsay, R. S. (1998), 'Testing and modeling multivariate threshold models', *Journal of the American Statistical Association* **93**(443), 1188–1202.

Wagner, C. (2008), Risk-premia, carry-trade dynamics, and speculative efficiency of currency markets, Working Paper Series 143, Austrian National Bank.

Appendix

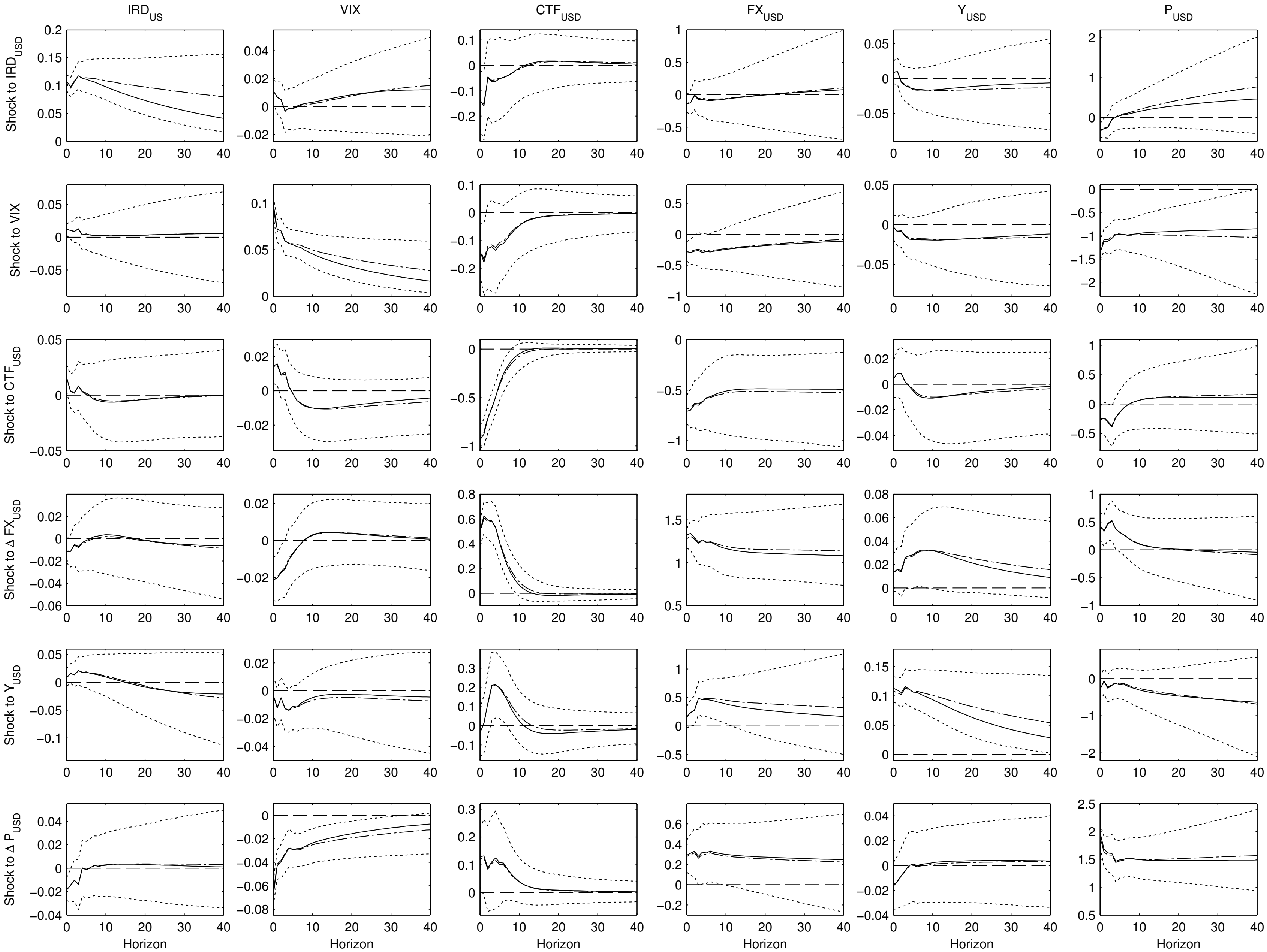

Figure 7: Sample A: (Accumulated) generalized impulse response functions of the H-regime for all variables. Solid line: point estimate; dashed-dotted line: bootstrap median; dotted lines: non-centered 95%-confidence interval (small sample bias and GARCH corrected, details are described in Sections 4.2 and 4.3). For more details about sample A, see Table (4). Number of observations: 418 (H-regime) & 270 (L-regime)

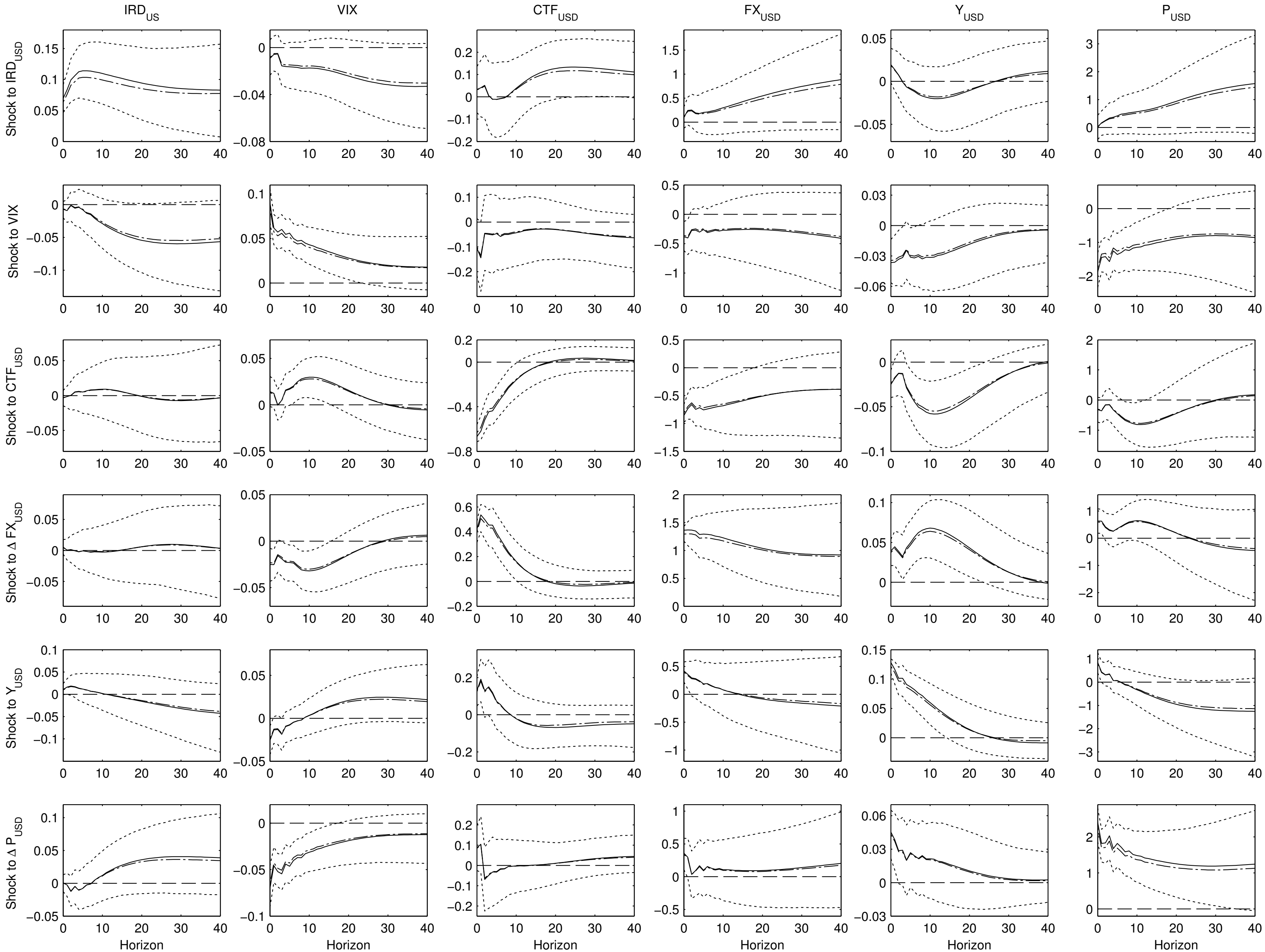

Figure 8: Sample A: (Accumulated) generalized impulse response functions of the L-regime for all variables. Solid line: point estimate; dashed-dotted line: bootstrap median; dotted lines: non-centered 95%-confidence interval (small sample bias and GARCH corrected, details are described in Sections 4.2 and 4.3). For more details about sample A, see Table (4). Number of observations: 418 (H-regime) & 270 (L-regime)

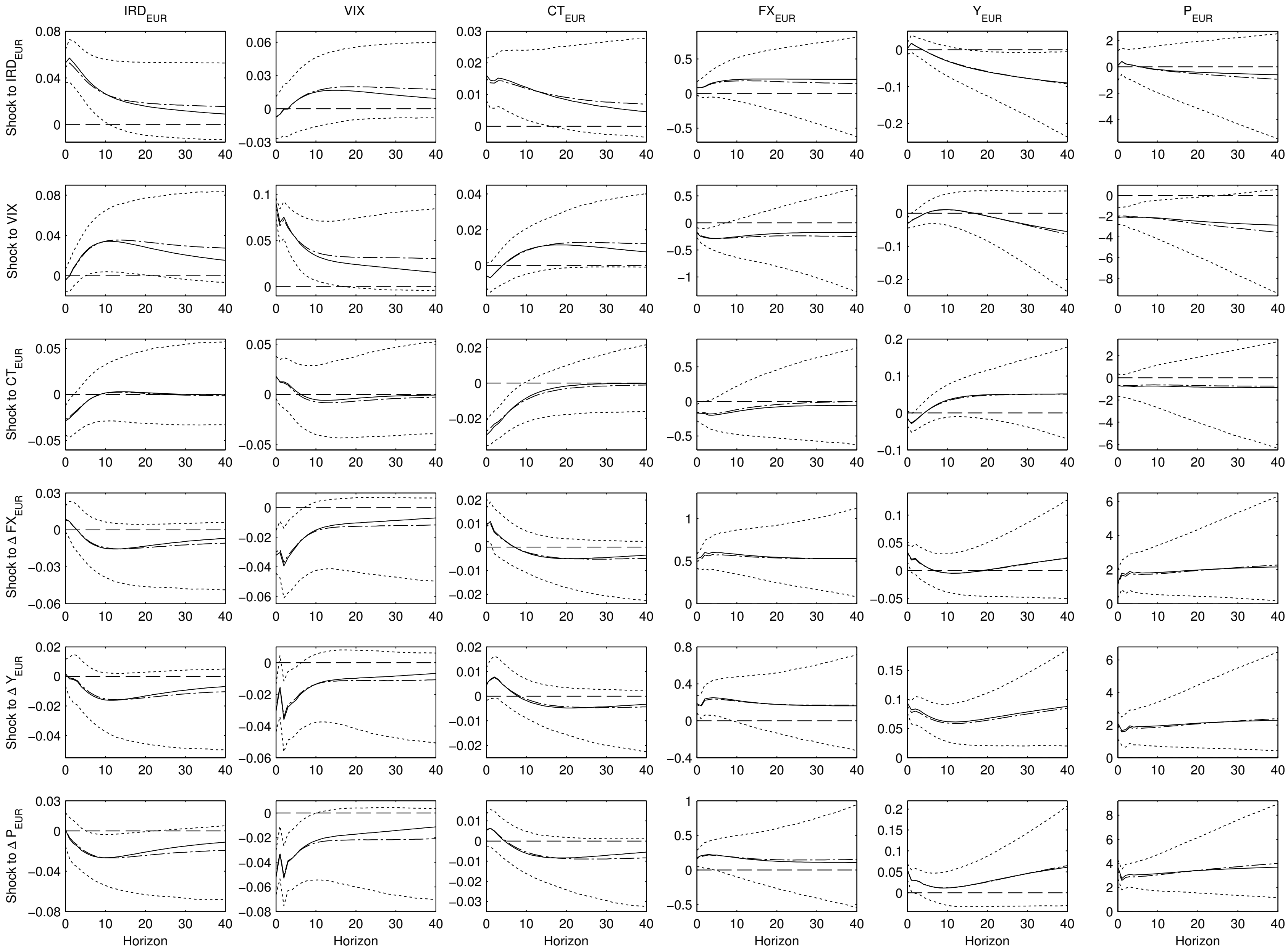

Figure 9: Sample B: (Accumulated) generalized impulse response functions of the H-regime for all variables. Solid line: point estimate; dashed-dotted line: bootstrap median; dotted lines: non-centered 95%-confidence interval (small sample bias and GARCH corrected, details are described in Sections 4.2 and 4.3). For more details about sample B, see Table (4). Number of observations: 125 (H-regime) & 367 (L-regime)

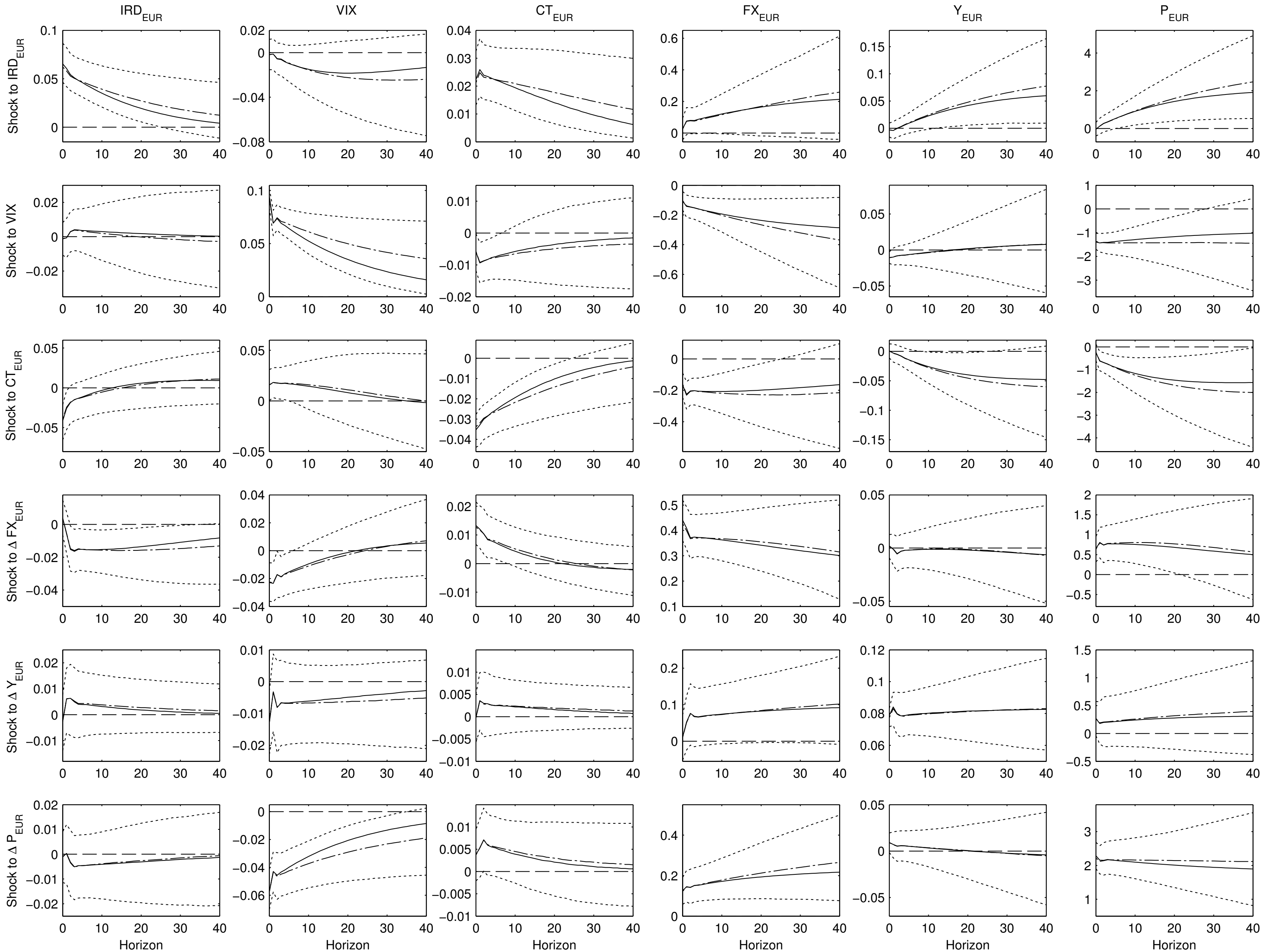

Figure 10: Sample B: (Accumulated) generalized impulse response functions of the L-regime for all variables. Solid line: point estimate; dashed-dotted line: bootstrap median; dotted lines: non-centered 95%-confidence interval (small sample bias and GARCH corrected, details are described in Sections 4.2 and 4.3). For more details about sample B, see Table (4). Number of observations: 125 (H-regime) & 367 (L-regime)

Table 11: ARCH Test Results with USD and EUR as target currencies

Dependent Variable	Target currency: USD[1]			Target currency: EUR[2]		
	ARCH(1)	ARCH(2)	ARCH(4)	ARCH(1)	ARCH(2)	ARCH(4)
ΔFX_{USD} / ΔFX_{EUR}	0.559	7.213**	11.118**	0.112	17.604***	17.728***
ΔP_{USD} / ΔP_{EUR}	42.735***	42.642***	55.007***	41.793***	42.233***	56.998***
VIX	3.015*	5.083*	15.935***	2.997*	4.845*	9.356*
IRD_{USD} / IRD_{EUR}	6.181**	10.394***	10.669**	25.200***	26.300***	32.006***
Y_{USD} / ΔY_{EUR}	4.838**	4.794*	27.997***	1.434	1.644	5.558
Carry Trade Positions						
CTF_{USD} / CT_{EUR}	0.355	1.872	7.328	14.078***	14.185***	15.454***

[1] The model is estimated with four lags from 1995/03/28 until 2008/06/24.
[2] The model is estimated with two lags from 1999/01/06 until 2008/06/25.
*/**/*** denotes significance at 10%, 5% and 1% level, respectively

Table 12: Granger Causality Test for Sample $A_{USD}^{d=3}$

Variable excluded/ Regime[2]	Variable[1]					
	IRD_{USD}	VIX	CTF_{USD}	ΔFX_{USD}	Y_{USD}	ΔP_{USD}
IRD_{USD}						
H		3.062	3.714	4.129	8.496*	6.521
L		3.455	5.418	4.880	6.427	3.539
VIX						
H	1.116		2.490	0.894	1.455	1.843
L	15.669***		3.565	3.862	5.496	0.867
CTF_{USD}						
H	10.560**	9.865**		2.444	3.758	6.493
L	1.390	17.172***		10.206**	14.241***	16.136***
ΔFX_{USD}						
H	4.689	0.244	20.562***		2.136	3.759
L	3.995	3.887	24.868***		6.240	2.798
Y_{USD}						
H	6.440	5.896	7.204	10.900**		9.964**
L	9.368*	10.688**	7.102	5.400		18.378***
ΔP_{USD}						
H	6.396	6.270	1.602	1.994	1.265	
L	11.770**	12.330**	16.149***	13.902***	7.204	
all but own lags						
H	27.068	28.880*	40.000***	23.204	18.022	28.460*
L	42.165***	51.921***	59.574***	35.185**	35.218**	38.106***

[1] The samples and variables are described in Table (4).
[2] The H-regime includes observations where the threshold variable $IRD_{USD}^{d=3}$ is greater or equal to 2.63%. The L-regime includes observations where the threshold variable $IRD_{USD}^{d=3}$ is smaller than 2.63%.
*/**/*** denotes significance of the Chi-squared value at 10%, 5% and 1% level, respectively

Table 13: Granger Causality Test for Sample $B_{EUR}^{d=3}$

Variable excluded/ Regime[2]	Variable[1]					
	IRD_{EUR}	VIX	CT_{EUR}	ΔFX_{EUR}	ΔY_{EUR}	ΔP_{EUR}
IRD_{EUR}						
H		0.793	2.810	1.522	1.030	1.790
L		3.856	17.838***	1.165	2.529	1.532
VIX						
H	5.886*		0.942	1.878	1.543	1.967
L	1.588		4.398	4.780*	1.395	2.461
CT_{EUR}						
H	1.575	0.257		0.373	7.325**	0.617
L	22.717***	1.862		8.281**	7.832**	11.539***
ΔFX_{EUR}						
H	0.174	1.662	1.762	3.745		5.993**
L	5.850*	2.362	0.144	3.694		0.751
ΔY_{EUR}						
H	0.737	2.329	2.448	4.303		0.018
L	6.916**	1.915	4.067	3.975		0.449
ΔP_{EUR}						
H	3.053	1.701	0.139	1.479	6.243**	
L	0.655	0.153	2.586	0.146	0.187	
all but own lags						
H	13.100	13.992	12.948	7.331	26.838***	11.821
L	45.636***	10.682	29.216***	26.503***	12.457	20.118**

[1] The samples and variables are described in Table (4).
[2] The H-regime includes observations where the threshold variable $IRD_{EUR}^{d=3}$ is greater or equal to 1.84%. The L-regime includes observations where the threshold variable $IRD_{EUR}^{d=3}$ is smaller than 1.84%.
*/**/*** denotes significance of the Chi-squared value at 10%, 5% and 1% level, respectively