Grid-Bootstrap Methods vs. Bayesian Analysis. Testing for Structural Breaks in the Conditional Variance of Nominal Interest Rate Spreads - Four Cases in Europe

Pierangelo De Pace*

Johns Hopkins University, Department of Economics

Mergenthaler Hall - 3400, N Charles St. - Baltimore, MD 21218, USA†

January 2006

Abstract

I use numerical methods to test for the presence of one-time structural breaks in the conditional variance of nominal interest rate spreads in four European countries over a period of eleven years (Jan 1988 to Dec 1998). I start with an intuitive approach consisting of a sequence of breakpoint Chow tests performed at subsequent dates over a given subsample of the squared residuals of the autoregressions used to model the yield spreads. Results from this procedure are misleading and spurious to some extent because of the incorrect critical values produced, which make the interpretation of the test statistics basically unreliable. I then switch to large Monte Carlo simulations and to a fixed-regressor grid-bootstrap method to derive the *right* critical values and refine the previous conclusions. Finally, I utilize classical Bayesian econometrics to estimate alternative models for the series of nominal spreads and to detect potential shifts in the innovation variances of the equations describing the data. Outcomes need some interpretation: in the cases of Germany and Spain a break might have occurred in 1990 and 1994 respectively, as derived from the grid-bootstrap approach. Likewise, there is evidence of a shift in the case of France in 1996 according to the Bayesian techniques employed, which also validate the hypothesis of a break for Italian yield spreads in 1995.

JEL Classification: C11, C12, C15; E3, E44.

Keywords: Chow Test; Classical Bayesian Analysis; Conditional Variance; Fixed-Regressor Grid-Bootstrap Method; Structural Breaks.

*I wish to thank A.T. Levin (Board of Governors of the Federal Reserve System), J.M. Piger (Federal Reserve Bank of St. Louis) and T. Woutersen (Johns Hopkins University) for their precious comments.

†*Correspondence:* pierangelo.depace@jhu.edu; Tel. #: +1 410 889 5704.

1 Introduction

The period of time ranging from the last few years of the 1980s to almost the end of the 1990s is plenty of epochal events that had a huge impact on both the political organizations and the economic structures of a vast majority of countries in Europe. The fall of the Berlin Wall in November 1989 only represented the prelude to the decay and disintegration of the communist regimes in the USSR and in all the other Eastern European nations either under the direct or indirect influence of Moscow. The early '90s were characterized by a number of economic happenings and challenges: market economies took over planned economies in former communist areas; Western European countries experienced a period of recession and some of them even survived a currency crisis that first led nations like the United Kingdom and Italy to *temporarily* exit the European Monetary System (EMS) because of the unbearable pressures on their respective currencies in the foreign exchange markets; and then pushed European governments to finally lay the bases for the gradual creation of a unique economic area within which fixed exchange rates would be established. This process culminated with the birth of the European Central Bank (ECB) and the European System of Central Banks (ESCB) on the 1st June 1998 and with the introduction of the Euro as the common currency for eleven (then twelve) countries, as of the beginning of January 1999.

In this paper I study the time evolution and the properties of nominal interest rates in four European countries - namely Germany, Spain, France and Italy - over a period of eleven years, from the beginning of 1988 to the end of 1998. I examine the conditional volatility (variance) of the spread between short-term and long-term interest rates within each of the aforementioned countries in order to determine whether significant structural breaks did occur during the time-span considered, possibly because of some of the circumstances listed above.

The behavior of the yield curve changes along the business cycle. In recessions, premia on long-term bonds are usually high and yields on short-term bonds generally low. As such, recessions are likely to feature upward sloping yield curves. Premia on long bonds are countercyclical because investors do not like to take on risk in bad times. In contrast, yields on short bonds tend to be procyclical because the Central Bank is more willing to lower short yields in recessions in an effort to stimulate economic activity and more willing to raise them when the economy grows faster[1]. On the basis of this simple economic intuition and

[1] This suggests that there should exist positive correlation between real output gaps and nominal spreads defined as the difference between short-term and long-term interest rates.

of the fact that expansions and recessions inevitably alternate, many papers tried to predict GDP growth or GDP recessions through OLS regressions by using the slope of the yield curve (i.e., the spread between short and long nominal interest rates) among the relevant regressors and generally found negative statistically significant correlation between economic growth and spreads. Some works in this field are by Fama (1990), Mishkin (1990); Estrella and Mishkin (1998); Ang, Piazzesi and Wei (2005).

I pay attention to the yield spread, constructed as the difference between short-term and long-term nominal interest rates, since its variability can be considered as a measure of the stability of financial and money markets and since connections might exist between such a variable and the real business cycle. There seems to be some degree of co-movements between nominal interest rate spreads and real output growth rates. Detecting a significant one-time structural break in the conditional variance of the interest rate spread should then translate, under a particular set of assumptions, into the detection of a *potential*[2] change in the regime of conditional volatility of the growth cycle and into a change in the degree of stability of financial markets. In particular, a lower average conditional variance for the time series of the yield spread over a given subsample should imply more stability in financial markets and predict less variability in the economy (growth business cycles) as a whole over that subperiod[3].

The objective of this paper is then to describe a rigorous, though sufficiently simple, statistical framework within which the detection of one-time structural breaks in the innovation variance of the process describing the data is possible. To accomplish my goal, I make use of a broad set of econometric tools. First of all, a fully intuitive (though not completely correct) approach to the problem: this methodology mainly consists of a sequence of break-point Chow tests performed at different dates over a given subsample of the data for the detection of structural shifts in the series of estimated conditional variances. I will argue that this procedure does not produce proper critical values for the relevant statistics, but it serves as the *necessary*[4] starting point for the subsequent kind of analysis. I then utilize classical methods that involve a heavy use of large Monte Carlo simulations for the correct

[2] It is only *potential* since, as I will argue later, the conditional variance of yield spreads is just one component of the total variance of future real output growth rates. In any case, a change of that conditional variability predicts a change in the same direction of the variability of the growth cycle.

[3] I do not directly use any measure for the real output gap to assess switching volatility regimes in the business cycle because of the too many degrees of freedom that one would face while trying to estimate the cyclical and the long-run components of real output. Constructing interest rate spreads is instead much easier and leads unambiguously to well-defined time series with a clear cyclical behavior.

[4] It is necessary since the statistics on which the second approach employed relies on are derived from the sequence of Chow tests performed in the first part of this work.

computation of the test statistics and the corresponding critical values. Finally, I switch to classical Bayesian techniques, which provide a straightforward (at least in principle) way to estimate models and select among alternatives.

At a first stage, the conditional variance of the process describing the interest rate spread of each country is considered as constant over the entire sample of estimation (nothing else than the usual assumption of homoskedasticity of the error term in classical econometrics, a standard feature of any OLS regression). At a second stage, I relax this assumption and allow for potential heteroskedasticity by modelling the variance term in a simple way, that is by assuming the presence of a one-time structural break in the mean of the variance itself. Loosely speaking, the way heteroskedasticity is taken into account in this work requires the hypothesis that the conditional variance under investigation may be consistently modelled by means of a convenient one-step function. This way of modelling the variance is not meant to be exhaustive and completely satisfying under an economic point of view. At the same time, however, it represents an easy and straightforward way to introduce changes in the volatility regime of the series under investigation, to look at them from a different perspective and to reconcile major economic events - which are known to have changed the course of European economies - with a useful econometric setting. I then compare the two alternative specifications for every single country (homoskedasticity vs. heteroskedasticity as modelled) in order to check for the significance of the hypothesized structural break; in some cases, I estimate again the models for the interest rate spreads, namely when evidence of a changed regime in the conditional variability is found. These new estimations are performed assuming heteroskedasticity of the error term (in the form proposed) in the regression equations.

The present paper is organized as follows: in Section 2 I describe the available data by focusing on their time series properties with the aim of providing a rationale for the model I then decide to adopt for the description of their behavior over time. My search for the final results is, at this point, performed in three steps. In Section 3 I test for the presence of a potential structural break in the conditional variance of the nominal interest rate spreads (i.e. I check whether it is possible to abandon the assumption of homoskedasticity of the error term) by utilizing an intuitive approach based on a sequence of breakpoint Chow tests performed at different dates over a subset of the sample period. As I point out later, this procedure is highly unsatisfactory, should be handled with care and be considered only as a useful benchmark to start from for a more sophisticated analysis. Section 4 is entirely devoted to the refinement of the procedure described in Section 3. The so-called grid-bootstrap method developed by Hansen (2000) then becomes a useful tool to compute the correct critical values and probability levels for the statistics already derived (and used) from the previous sequence of Chow tests; and to interpret in a better way the results obtained from the intuitive (naïve) approach. In Section 5 I sketch classical Bayesian techniques, discuss how they should be applied to get an answer to the question of interest and eventually apply them to address the problem. Section 6 describes some of the results obtained using similar procedures on a larger sample (January 1998 to February 2005) and shortly discusses some remaining issues. It is worth noticing that extending the sample over which the analysis is conducted *is likely to raise the probability* that a one-time structural break is actually detected. The period starting in January 1999 (that is in the month immediately after the end of the sample I have decided to concentrate on in this paper) represents a new economic era for European countries, an era characterized by low inflation and low interest rates as never experienced before in the Old Continent. I leave this period out of my main analysis believing that, if included, it would have facilitated the detection of a break in the conditional variance of the yield spreads. Section 6 is used instead for the brief description of such a possibility through Hansen's numerical methods. Section 7 finally comments on the main conclusions of this work and summarizes the most significant findings.

Results regarding estimations and dates of potential structural breaks are reported and interpreted in a historical perspective. Short references to and descriptions of the main theoretical econometric tools used throughout the paper are provided when needed.

2 The Data

I collect time series for short-term and long-term nominal interest rates in four Western European countries, namely Germany (including ex-GDR from 1991), Spain, France and Italy. Such a choice is not random. The four countries represent the four biggest economies in the Euro area[5]. Germany and France have been selected because of their overall economic stability, relatively lower interest rates over the past twenty-five years or so, good performances in terms of inflation rates. Spain and Italy have been picked mainly for the opposite reasons: more pronounced economic problems and difficulties over the last twenty-five years or so (for the most various reasons), relatively more precarious economic stability (at least for some time), higher interest and inflation rates. The idea is to discriminate between two

[5] Germany, France and Italy are even G7-countries.

sets of countries with very different economic patterns over the period of time analyzed that eventually converged to similar parameters.

Monthly data have been assembled from EUROSTAT and are referred to a period that starts in January 1988 (at the eve of the crisis of communist countries) and ends in December 1998, immediately before the European dream and project of a common currency came true. The whole sample is then constituted of a total of 132 time observations for each country. Short-term interest rates are measured as three-month money market rates; the definition of long-term interest rates is given by 10-year government bond yields. The difference between the two variables represents the nominal interest rate spread I make extensive use of in my subsequent analysis.

2.1 What Model for Nominal Interest Rate Spreads?

The two following graphs (Figure 1 and Figure 2) plot the time series of the two definitions of nominal interest rate adopted here for each of the four countries under investigation:

[Figures 1 and 2 about here]

The time pattern is clear. Spanish and Italian interest rates remain well above German and French interest rates over the whole sample, though an evident convergence path, which was completed at the end of 1998, witnesses the efforts of all the four countries to satisfy the requirement of homogeneous interest rates on the date of entrance into the newly born Economic and Monetary Union (EMU). Higher interest rates - in particular, higher interest rates on assets with a long maturity - for Spain and Italy underlie the well-known problems with inflation experienced by these two countries, at least until not much before their accession into the European Monetary Union in 1999. The series of Italian three-month interest rates peaks approximately in the second half of 1992, exactly when the Italian Lira (together with the British Sterling) was pushed out of the *"snake"* of the European Monetary System by the speculative attacks performed in the money and exchange markets by foreign investors. It was in that period that Italy also got to its peak in terms of long-term interest rates, notoriously more sensible to inflation expectations. German short-term interest rates - historically closely tracked by French rates - display a peculiar smooth increase between the last years of the '80s and the first years of the '90s, probably a consequence of the attempts to efficiently deal with the costs of reunification and with the European currency crisis of the 1992-1993 period, mainly aimed at preserving stability and low inflation rates in the country. After that period, a clear convergence path towards permanently lower rates can be fully observed for these two nations.

Both short-term and long-term interest rates rose during the first half of 1995. This is true for 10-year government bond yields in Spain and Italy, which, maybe, felt the effects of soaring expectations about future inflation. As a short historical notice, it was since the early 1980s that long-term government bond yields in the Eurozone in general had declined; a fact that is in line with a similar trend occurred in other industrialized countries around Europe and the world. By the time the Euro currency was formally introduced in 1999, long-term government bond yields across the Eurozone countries (and so across the four nations examined in this paper) had largely converged to that of Germany (the Euro area's largest economy).

It is not redundant to stress the apparently high persistence of nominal interest rates over time. This is a characteristic that possibly reflects the so-called *"interest rate smoothing"*, a known aspect in monetary policy stances of many modern central banks. Very probably, such stances have some degree of influence on the various definitions of interest rates. This is evident from the formal inspection of the correlograms of the series of interest rates, which might suggest the possible presence of unit roots (first partial correlations are close to one in all the cases considered):

[Tables 1 and 2 about here]

Similar rates of persistence can be found in the interest rate spreads of the four countries. Estimated autocorrelations that decay at a very slow rate and estimated partial correlations - which, broadly speaking, are statistically significant only at the first lag or at the first and second lags (depending on the country considered) - strongly indicate that AR models are used to describe in a simple way the time behavior of the yield spreads:

[Figure 3 about here]

[Table 3 about here]

Interest rate spreads seem to be cyclical over time. This intuition is confirmed to some extent by a rough comparison with the quarterly series of real output gaps[6] and with those of future real GDP growth rates expressed at a quarterly frequency[7]. Correlations on common samples indicate that nominal spreads substantially co-move with real output gaps in the two cases of Spain and France (0.416 and 0.443 respectively); this finding is definitely weaker for Germany and Italy, with estimated correlations equal to 0.100 and 0.007 respectively. Correlations between spreads and real output growth rates are generally high and negative as expected: -0.079 for Italy, -0.278 for Germany; -0.348 for Spain and -0.492 for France.

In general, one may claim that greater uncertainty about future rates of inflation, future growth perspectives or future political events are often likely to widen the difference between the short-term and the long-term nominal interest rates. This statement can help understand better why there seems to be evidence (formally not checked in this paper) of co-variance/correlation between real output gaps and nominal spreads; and between real GDP growth rates and nominal spreads. This might also shed some light on my choice of analyzing country-specific yield spreads and their conditional variances as *proxies* for the overall economic performance of each country, primarily in terms of stability, relying on their power in predicting output growth and recessions[8].

A quick glance at cross-correlations gives a further indication about the possibility for the variables under analysis to be modelled in similar ways. Generally high values suggest that this might actually be a feasible and proper option:

[Table 4 about here]

[6] To this purpose, monthly data for interest rate spreads have been converted into quartely data through a simple averaging procedure. Quarterly output gaps have been derived through the application of a Hodrick-Prescott filter to the original series or real output (collected from EUROSTAT).

[7] I use the following definition of real GDP growth rate: $g_{t\rightarrow t+4} = \frac{1}{4}\left(\log GDP_{t+4} - \log GDP_t\right)$.

[8] A popular and effective empirical way to predict output growth rates consists of the specification: $g_{t\rightarrow t+h} = \beta_0 + \beta_1 s_t + u_t$, where s_t is the nominal interest rate spread as defined earlier in this paper and u_t is a normally distributed random error term and $h \geq 0$ (see, for instance, Ang, Piazzesi and Wei (2005)). With quarterly data, a natural choice is $h = 4$. Later in this work, I specify the following process for s_t: $s_t = \mu + \Sigma_{j=1}^{K}\alpha_j s_{t-j} + \varepsilon_t$. It then turns out that, under these assumptions and with $h = 4$, the conditional variance of the yield spreads is related to the variance of future real GDP growth rates conditional on lagged spreads:

$$\begin{aligned} Var\left(s_t|s_{t-j},\ j=1,..,K\right) &= Var\left(\varepsilon_t\right); \\ Var\left(g_{t\rightarrow t+4}|s_{t-j},\ j=1,..,K\right) &= \left(\beta_1\right)^2 Var\left(s_t|s_{t-j},\ j=1,..,K\right) + Var\left(u_t|s_{t-j},\ j=1,..,K\right) \\ &= \left(\beta_1\right)^2 Var\left(\varepsilon_t\right) + Var\left(u_t\right). \end{aligned}$$

This implies that a part of the variability of the future growth cycle can be explained by the conditional variability of the yield curve slope. More specifically, other things being equal, a drop in the conditional variance of yield spreads predicts a fall in the conditional variance of growth rates.

The simplest model for the data, which probably summarizes most of the features detected so far, seems to be a univariate autoregressive (AR) process for each series of nominal spread. Thus, a general specification for an $AR(K)$ process is:

$$s_t = \mu + \Sigma_{j=1}^{K} \alpha_j s_{t-j} + \varepsilon_t \tag{1}$$

where s_t is the yield spread at time t, ε_t is assumed to be a serially uncorrelated, but possibly heteroskedastic, random error term; μ is simply the intercept term of the equation. Andrews and Chen (1994) claim that, given an $AR(K)$ process, the best scalar measure of persistence for the underlying variable is given by $\rho = \Sigma_{j=1}^{K} \alpha_j$, which is likely to provide much better indications than the more usually used largest root of the characteristic equation of the autoregressive process.

Following the approach proposed by Levin and Piger (2004) for the study of inflation rates in industrialized countries, through some algebraic manipulation, equation (1) can be easily rewritten into the following equivalent expression:

$$s_t = \mu + \rho s_{t-1} + \Sigma_{j=1}^{K-1} \phi_j \Delta s_{t-j} + \varepsilon_t \tag{2}$$

In this *ad hoc* formulation, the persistence parameter equals the expression $\rho = \Sigma_{j=1}^{K} \alpha_j$ and each ϕ_j is a simple transformation of the AR coefficients displayed in equation (1). In this way, the process of each spread series can be studied by fitting equation (2) into the data after choosing a suitable lag length (i.e. a suitable value for K) for the autoregressive equation. This can be done in a number of ways. By looking, for instance, at the correlograms presented above, significant partial correlations might indicate the correct lag choice. On this ground, K is equal to 2 in the cases of Germany, Spain and France; possibly equal to 1 (with some reserves) in the case of Italy. Another very way to address this problem is to minimize the Akaike Information Criterion (AIC) or the Schwarz Information Criterion (SIC) of a standard OLS regression. I will look at all the information available in order to choose the correct lag order on a case-by-case basis.

3 Data Mining?

The following OLS regressions are run on the available data by making use of the model described in equation (2). Lag lengths are chosen using the information coming from appropriate criteria and reported in the last column of Table 5. Standard errors are indicated in parenthesis below the corresponding estimated coefficients:

[Table 5 about here]

Persistence parameters, as expected, are found to be very close to unity for at least two countries, namely Germany and Spain. Estimates for France and Italy are somewhat lower (in particular in the case of Italy), but still imply high persistence in the series of the corresponding interest rate spreads. This immediately validates some of the findings already mentioned. In figure 4 I plot the residuals of the regressions above:

[Figure 4 about here]

A visual inspection suggests that the assumption of homoskedasticity for the error terms might not hold for all the four countries and should not be retained for some of them. There seems to be a higher variability of the residuals in the very first part of the sample in the case of Germany as well as in the case of Spain (though the subsample of high volatility looks a little longer in the latter country). The cases of Italy and France are a little more difficult to assess: there are spikes and outliers in both the series, but it is still not easy to figure out whether this actually implies a change in the regimes of volatility for the interest rate spreads in the two nations or just transitory deviations from a rather stable mean[9].

Similar patterns are recognizable from the four graphs of the squared residuals of the OLS regressions. Squared residuals are *proxies* for the estimated conditional variance of the error term and are extensively used later in this section to model potential heteroskedasticity[10]:

[Figure 5 about here]

[9] Modelling outliers with dummy variables might represent a viable solution, perhaps at least in the Italian series (aberrant observations occur in correspondence of the second half of 1992, i.e. when the recalled currency crisis hit Europe). However, this is not the way I chose to proceed.

[10] The logic of this comes from the equivalences that are commonly used to model heteroskedasticity:

$$\begin{aligned} E\left(\varepsilon_t^2|z_t\right) &= \sigma^2 + z_t'\alpha \\ &\therefore \; \varepsilon_t^2 = \sigma^2 + z_t'\alpha + \left[\varepsilon_t^2 - E\left(\varepsilon_t^2|z_t\right)\right] \\ &= \sigma^2 + z_t'\alpha + v_t \end{aligned}$$

with $E\left(v_t|z_t\right) = 0$. If we knew ε_t^2, we could directly use it to model the conditional variance. Given that we do not, a common practice is to replace it with $\widehat{\varepsilon}_t^2$.

German squared residuals generally display higher values in the very first part of the sample and it seems reasonable to assume in this case that, at a certain point in time, German interest rate spreads might have actually switched to a lower regime of conditional variability. A similar finding can be observed in the series of squared residuals from the regression regarding Spanish spreads, even though the point in time at which the conditional volatility should have fallen seems to be in the middle of the sample. It is harder to deduct anything similar from French and Italian data.

An informal way to detect whether there actually exists a structural break in the conditional variability of nominal interest rate spreads consists in a number of steps that are aimed at checking whether some forms of heteroskedasticity of the error term can be allowed by the data and then modelled accordingly. Heteroskedasticity issues can be tackled in various and even rather sophisticated ways. In order to remain consistent with the much more complex kind of analysis that I perform later, in this part of the paper I assume that heteroskedasticity, if any, can be modelled in the following simple way:

$$E\left(\varepsilon_t^2|z_t\right) = \gamma_0 + D_t\gamma_1 \tag{3}$$

where D_t is a dummy variable that controls for the shift in the innovation variance. This means that D_t is a vector of T observations ($T = 132$ in this part of the paper) which contains zeros until a structural break is detected and then contains ones for the remainder of the sample. For simplicity, I allow for a possible one-time structural break only[11], even though the data might actually suggest different assumptions to relax the classical hypothesis of homoskedasticity and to model the variance of the error term. In intuitive words, the unique one-time structural break occurs when the probability of such an event is maximized over

[11] Starting from the series of the squared residuals as generated from the OLS regressions described at the beginning of this section, I regress $\widehat{\varepsilon}_t^2$ on a constant term only and then check for structural shifts in the estimated intercept term using a sequence of breakpoint Chow tests performed at different dates. Moreover, as customary in these kinds of studies, I constrain the potential break to occur in the middle 70% of the whole sample. In my work this means that the sequence of Chow tests is only performed along the subsample 1990:01-1997:04 (boundaries included).

The idea of the breakpoint Chow test is to fit the same equation separately for each subsample and to see whether there are significant differences in the estimated coefficients. A significant difference indicates a structural change in the relationship. To carry out the test, data should be partitioned into two subsamples. Each subsample must contain more observations than the number of coefficients in the equation so that the equation itself can be estimated. The breakpoint Chow test compares the sum of squared residuals obtained by fitting a single equation to the entire sample with the sum of squared residuals obtained when separate equations are fitted to each subsample of the data (F test). Alternatively, a log-likelihood ratio (LR) statistic can be computed by comparing the restricted and unrestricted maximum of the (Gaussian) log-likelihood function.

the sample[12].

Figure 6 plots the results of this recursive procedure. For each of the four countries I report, in separate diagrams, the series of the F statistics and the LR statistics associated with each test performed at a given date. Both statistics deliver qualitatively similar results. Dotted and dashed lines indicate the calculated critical values at the 5% and 1% levels. The null of the test is that no structural break has occurred. Relevant statistics above the critical values show that the null hypothesis can be rejected and that at least a one-time structural break might exist:

[Figure 6 about here]

Some values of the test statistics clearly suggest the presence of structural breaks, with different levels of significance. On a very first and rough approximation, they occur somewhere inside the time intervals in correspondence to which the plotted lines lie above the dotted critical values[13]. Circled observations refer to the date on which the hypothesized unique structural break is *most likely* to be present. In the tables below I briefly report such time intervals together with the highest values for both the test statistics used within the standard Chow test[14]:

[Tables 6, 7 and 8 about here]

However, there exists a major problem in the whole econometric procedure performed until now that can not be ignored. What has been done so far is not distant from pure data mining, this meaning that, even though the approach to this issue looks incredibly appealing and really intuitive when the date of the structural break is unknown and needs to be correctly detected, the sequential investigation implied by repeated breakpoint Chow tests at different points in time is likely to produce inconclusive and even spurious results, simply because it fails to produce the correct critical values[15]. Fortunately, there exists an econometric procedure - which can not be examined without making use of the results obtained in this

[12] This, of course, does not mean that no other breaks have occurred and can be detected. Extensions of the model and of the procedures presented in this paper can be thought in order to account for the possible presence of multiple structural shifts or of a smooth transition from a volatility regime to another one.

[13] **Breakpoint Chow Test (F Statistic):** Germany: 5% critical value=3.915138; 1% critical value=6.837100. Spain: 5% critical value=3.915138; 1% critical value=6.837100. France: 5% critical value=3.915138; 1% critical value=6.837100. Italy: 5% critical value=3.914559; 1% critical value=6.835499.

Breakpoint Chow Test (LR Statistic): Germany: 5% critical value=3.841459; 1% critical value=6.634897. Spain: 5% critical value=3.841459; 1% critical value=6.634897. France: 5% critical value=3.841459; 1% critical value=6.634897. Italy: 5% critical value=3.841459; 1% critical value=6.634897.

[14] The maximum value of the Chow F test statistic (max F or sup F) is usually known as the Quandt test statistic or Andrews test statistic. For further details, see: Quandt, R. (1960) and Andrews, D.W.K. (1993).

[15] The test statistics derived do not have standard distributions. The correct ones should then be simulated through Monte Carlo methods.

section - to address this inconvenience. It is described in details in the next section.

Sequential breakpoint Chow tests are at least capable of selecting possible time intervals within which the actual break may be found. By focusing on the maximal statistics only, it is possible to isolate individual dates in correspondence to which the probability of the presence of a simple one-time break is maximized. In the case of Germany this date is found to have occurred in March 1990, just a few months after the fall of the Berlin Wall, at the very beginning of the process of reunification of the two parts of the country. This could be the event that might have caused the process describing the evolution of the nominal interest rate spreads over time to permanently change its volatility regime. Spain might have changed such a regime in October 1994, once the bases for the next birth of the European Monetary Union had been built and the efforts of compliance with the Maastricht Treaty became substantial for all the nations involved in the process. France and Italy, instead, may have respectively experienced such changes in September 1992 (the beginning of the European currency crisis) and in March 1993 (a few months after the currency crisis, when it was clear that the Italian Lira - which had to exit the EMS in September 1992 - would not re-enter the system so soon as expected, and the new course to recovery had already been instituted by Italian political and economic authorities).

4 A "Fixed-Regressor Bootstrap Method" to Detect Structural Shifts

In Section 3 I presented an intuitive and naïve way to test for and detect the presence of a one-time structural break (at an unknown date) in the innovation variance of the processes describing the nominal interest rate spreads of four different European countries over a given time period. As already argued, the problem with that approach is that usual critical values that come out of standard breakpoint Chow tests can not be used because of their incorrectness. Though the procedure might seem appealing and even give some approximate indication of whether and when the shift has occurred, the inference based on it is likely to be spurious. In any case, it represents the *necessary* starting point for the analysis performed in this section.

Over the last twenty years econometricians developed a large set of new instruments to test for structural changes of unknown timing in regression models. Among them, particular attention should be paid to those which rely on the *sup F statistic* of Andrews (1993), the *exp F* and *ave F statistics* of Andrews-Ploberger (1994). Hansen (2000) proved that the statistics generated by the procedure described earlier have non-standard distributions that should be correctly computed by means of Monte Carlo simulations. He refined Andrews and Ploberger's findings by showing that their statistics may vary to structural changes in the regressors of the test equations. Hansen (1999) developed a *"fixed-regressor grid-bootstrap"*[16] procedure to derive the first-order asymptotic distribution for the statistics of interest. The attractive feature in Hansen's approach is that his grid-bootstrap method allows for arbitrary structural changes in the regressors, including simple structural shifts, as in the case described in this work; and for lagged dependent variables and heteroskedastic error processes.

To the purposes of this paper, I slightly modify the framework presented by Hansen. By making use of a similar methodology and keeping the empirical set-up described in the previous section of this paper, I calculate the relevant statistics for my analysis on the estimated conditional variances[17]. After defining the F (Wald) statistic of the breakpoint Chow test at time t as F_t, Andrews and Ploberger's statistics are then computed as[18]:

$$\begin{aligned}
\sup F &= \sup_{t\in[t_1,t_2]} F_t && \text{(Quandt/Andrews Statistic)}\\
\exp F &= \ln\left[\int_{t_1}^{t_2} e^{\left(\frac{F_t}{2}\right)} dw\left(t\right)\right] && \text{(Exp. Weighted F Statistic)}\\
\text{ave } F &= \int_{t_1}^{t_2} F_t dw\left(t\right) && \text{(Average F Statistic)}
\end{aligned}$$

where w is a measure that puts weight $\frac{1}{t_2-t_1}$ on each integer t in the interval $[t_1;t_2]$, with t_1 and t_2 representing the boundaries of the time interval along which the Chow test is executed. Usually, $t_1 = 0.15 \cdot T$ and $t_2 = 0.75 \cdot T$, with T being the total sample length. The $\exp F$ statistic is proved to be optimal against distant alternatives, whereas the *ave F*

[16] Hansen proposed a "grid-bootstrap" method to construct confidence intervals with an improved performance over conventional bootstrap methods when the sampling distribution depends upon the parameter of interest. The basic idea is to calculate the bootstrap distribution over a grid of values of the parameter of interest and form the confidence interval by the no-rejection principle. This framework perfectly applies to autoregressive models, where it is known that conventional bootstrap methods fail to provide correct first-order asymptotic coverage when an autoregressive root is close to unity. In contrast, the grid bootstrap is first-order correct globally in the parameter space.

The bootstrap method employed here treats all the regressors as exogenous even when they contain lagged values of the dependent variable.

[17] I use the squared residuals $\hat{\varepsilon}_t^2$ - a by-product of the regressions already performed - to model potential heteroskedasticity exactly as described in equation (3). More precisely, I regress the series of the squared residuals $\hat{\varepsilon}_t^2$ on a constant term only and then perform a sequence of breakpoint Chow tests at different dates to derive the corresponding F statistics.

[18] Note that the computation of these statistics would not be possible without the sequence of Chow tests run in Section 3.

statistic is optimal against very local alternatives.

The probability levels for each statistic are computed following Hansen's indications and by making use of large Monte Carlo simulations. This approach, by construction, either confirms or rejects the findings of the naïve procedure without providing different results in terms of detected break-dates. In particular, it returns an estimated breakpoint date that corresponds to the time observation for which the highest F statistic has been derived in the sequence of breakpoint Chow tests[19].

4.1 Hansen's Method - The Results

Assuming homoskedasticity, the regressions already discussed in Section 3 provide estimated conditional variances over the whole sample as reported in Table 9:

[Table 9 about here]

In table 10 I present the results coming from Hansen's grid-bootstrap procedure[20]:

[Table 10 about here]

Given the null, low p-values (possibly below 5% or even below 1%) indicate that at least a structural break in the conditional variance of the process can be assumed. This seems to be true for Germany and Spain: all the test statistics are *sufficiently large*, so that the null can be rejected and heteroskedasticity of the error terms can be allowed for and taken into account in the form proposed. With some exceptions (most of them occurring if the heteroskedasticity-corrected p-values are considered) the same conclusion can not be stated in the two other cases of France and Italy[21].

[19] Formally it is $t^* : F_{t^*} = \sup_{t \in [t_1, t_2]} F_t$.

[20] The three statistics described earlier are computed for each country and for each test regression run. The bootstrap approach also produces potential dates on which a shift in the average conditional variance might have occurred and p-values for the null that no structural breaks exist. Potential one-time structural shifts are detected at the point in time where the probability mass accumulates the most, i.e. where the probability of such an occurrence is maximized

[21] The three types of p-values are computed assuming three different asymptotic distributional approximations for the test statistics. The first p-value is derived by utilizing the asymptotic approximation used by Andrews (1993). As easily noticeable, it is very close to the second one, which, instead, is based on the homoskedastic fixed-regressor bootstrap approximation proposed by Hansen (2000). They both provide the same qualitative results in terms of rejection of the null. The third p-value is calculated on the basis of a heteroskedastic fixed-regressor bootstrap approximation, still introduced in Hansen (2000). However, the performance of all the test statistics under this last assumption deteriorates significantly in the two cases of France and Italy, failing not to reject the null of no structural break in the innovation variance. This is immediately explained by the natural conclusion that can be inferred for these countries after the tests performed

For Germany and Spain I then re-estimate the OLS regressions seen in Section 3 after introducing heteroskedasticity as modelled so far. I also provide numerical values for the means of the estimated conditional variances for both the countries over the two subsamples as determined by the detected breakpoint dates:

[Tables 11 and 12 about here]

A much more appealing result regards the estimated conditional variances of the regressions over the two subsamples considered (before and after the shift). In both the cases the conditional variance declines substantially, indicating a probable improved stability of financial markets for the two countries along their roads towards the accession into the EMU and a positive contribution to the stabilization of output growth over the cycle.

The grid-bootstrap methodology confirms in some way the results sketched in the section devoted to the naïve approach to the problem. It *obviously* returns the same breakpoint dates inferred from the maximum Chow test statistics previously shown. It rejects the null of no structural breaks in the innovation variance only for those countries (Germany and Spain) whose F and LR statistics reach (and/or go beyond) the respective *wrong* 1% critical values. This outcome could not be correctly assessed in Section 3 without first correcting p-values using the appropriate econometric tools introduced in this part. One should also take into account the fact that the rejection of the null does not automatically imply homoskedasticity of the error term, but just a preference for this assumption against the alternative of heteroskedasticity modelled as indicated in equation (3).

5 A Bayesian Approach to Model Heteroskedasticity and Detect Potential Structural Shifts

In this part of the paper I explicitly refer to the methodology used by Piger and Levin (2004) as a useful alternative to classical hypothesis testing. Here, I make use of classical Bayesian methods to model potential conditional heteroskedasticity in the AR model proposed in previous sections to fit the data regarding nominal interest rate spreads for the same four countries. In particular, following the same broad guidelines already analyzed in earlier parts

so far: the best assumption is probably that of homoskedasticity of the error terms in the two regressions describing the time evolution of French and Italian nominal interest rate spreads; as such, changing regimes in the innovation variance should not be introduced and hetero-corrected p-values are inconclusive.

of this work, I assume that, when testing for a possible one-time break in the innovation variance of a selected AR process, the conditional variance under investigation has the form expressed by equation (3)[22].

I estimate the following two models[23]:

$$s_t = \mu + \rho s_{t-1} + \Sigma_{j=1}^{K-1} \phi_j \Delta s_{t-j} + \varepsilon_t \tag{2.a}$$

(with ε_t serially independent and homoskedastic)

$$s_t = \mu + \rho s_{t-1} + \Sigma_{j=1}^{K-1} \phi_j \Delta s_{t-j} + \varepsilon_t \tag{2.b}$$

(with ε_t serially independent and heteroskedastic)

The relevant aspect in the estimation of the proposed AR models in a Bayesian fashion is given by the full specification in terms of probability distributions that must be assigned to each of the parameters involved. I assume that the error terms in equations (2.a) and (2.b) are normally distributed with zero mean and variance σ_i^2. In the first model, σ_i^2 is simply thought to be constant, that is $\sigma_i^2 = \sigma^2$. In equation (2.b) I model conditional heteroskedasticity in ε_t by allowing for a one-time structural shift in its variance, so that $\sigma_i^2 = \sigma_0^2 (1 - D_t) + \sigma_1^2 D_t$. In the last expression, D_t is a dummy variable that, as in a previous section, controls for the shift in the innovation variance.

In order to compute the marginal likelihood of (2.b), further assumptions are needed. I track Chib (1998) and assume that D_t is a discrete latent variable with Markov-transition probabilities $\Pr(D_{t+1} = 0 | D_t = 0) = q$ and $\Pr(D_{t+1} = 1 | D_t = 1) = 1$ with, of course, $q \in (0, 1)$. This means that there is a constant positive probability $(1 - q)$ for a break to occur

[22] I first estimate the model for each time series assuming homoskedasticity in the error term; I then perform the estimation of an alternative model which allows for a one-time shift in the innovation variance. I stick to the usual practice of restricting the break not to occur neither in the initial fifteen percent nor in the final fifteen percent of the full sample period. Finally, I compare the two competing models by computing the Bayes factor, that is the ratio between the marginal likelihood associated with the model with a break in the innovation variance and the marginal likelihood associated with the model without breaks. In general, a Bayes factor that is constructed in this way and that turns out to be bigger than one represents an evidence for preferring the model with a shift. The greater the Bayes factor (which, on the basis of the expression used to derive it and presented later, is always positive), the more likely the model with a break is with respect to the model without breaks. On the other hand, a Bayes factor that is less than one points in favor of the model that does not allow for any breaks.

[23] Note that Model (2.a) coincides with Model (2), but it is estimated in a Bayesian fashion. Model (2.b) differs from Model (2) in how potential heteroskedasticity is modelled. Equation (3) is not sufficient anymore to detect the potential break in the innovation variance. Even though the general form is kept, the Bayesian technique involved here requires the introduction of a latent variable with the definition of proper transition probabilities.

in any period, if it has not occurred yet. Through some simple algebra it is possible to show that the expected duration (n) of the number of periods prior to the break is equal to $E(n) = \frac{1}{1-q}$. All this eventually implies that once the break has occurred at a specific date n, then $D_t = 1$, $\forall t \geq n$ (*absorbing state*)[24]. Bayesian estimates generate a posterior distribution for q: the breakpoint date is promptly derived from the expression stating its expected value by using the posterior mean of q.

I adopt the distributions below to fully specify model (2.a):

$$
\begin{aligned}
\mu|\sigma^2 &\sim N\left(0, 3\sigma^2\right) \\
\rho|\sigma^2 &\sim N\left(1, 3\sigma^2\right) \\
\phi_j|\sigma^2 &\sim N\left(0, 3\sigma^2\right), \forall j \\
\sigma^2 &\sim InvGamma\,(1, 2)
\end{aligned}
$$

and the following to describe model (2.b):

$$
\begin{aligned}
\mu &\sim N\,(0, 3) \\
\rho &\sim N\,(1, 3) \\
\phi_j &\sim N\,(0, 3), \forall j \\
\sigma_0^2 &\sim InvGamma\,(1, 2) \\
\sigma_1^2 &\sim InvGamma\,(1, 2) \\
q &\sim Beta\,(8, 0.05)
\end{aligned}
$$

with μ, ρ and ϕ_j statistically independent of each other. The choice of relatively informative priors for the parameters is intended to provide a compromise between the need of *letting the data speak* and the necessity of incorporating the *a-priori* information coming from the analysis of correlograms[25]. The distributional structure imposed to the model without breaks assigns priors for μ, ρ and ϕ_j that are elicited conditional on σ^2. Final estimates are

[24] For further technical details about how to estimate Markov-Switching models in a Bayesian setting through Gibbs sampling, see Chapter 9 in Kim, C.J. and C.R. Nelson (1999).

[25] The values of first partial autocorrelations are very close to one; higher-order partial autocorrelations are generally close to zero. The standard Beta ensures that the domain of the probability measure q is over the interval $[0, 1]$. The parameters chosen imply that much of the mass of the distribution is spread around values that are very close to one. This specification gives more prior probability to late breakpoint dates in the sample considered. Estimated breakpoint dates suggest, instead, that the likelihood (i.e. the data) outweighs the prior in the computation of the posterior. Different calibrations for the prior of q did not alter the estimated changepoints.

not influenced by this assumption and have been proved to be generally robust to different specifications of the prior distributions. This lets linear model (2.a) fit the so-called *Normal-Gamma* framework[26] and makes the computation of many relevant quantities analytically feasible.

For each model I choose the lag order K that maximizes the corresponding marginal likelihood, with $1 \leq K \leq 4$[27]. The equations are then estimated by making use of the Gibbs

[26]The Normal-Gamma framework is a particular case of a two-level hierarchical Bayesian model, in which a conjugate prior distribution is specified at the first stage and a non-informative prior is generally assumed at the second stage.

[27]Given a set of data Y, the marginal likelihood of model M is analytically computed as (see: Chib, S. (1995) and Bos, C.S (2002):

$$m\left(Y|M\right) = \int_{\theta} \mathcal{L}\left(Y|\theta; M\right) \cdot \pi\left(\theta|M\right) d\theta$$

that is as the integral over the whole parameter space of the likelihood function of model M multiplied by the joint prior distribution of the parameters θ. The Bayes factor (see: Clarke, K.A. (2000)) for comparing model M_1 against model M_2 is then given by the expression:

$$BF_{1,2} = \frac{m\left(Y|M_1\right)}{m\left(Y|M_2\right)}$$

Computing the Bayes factor for the comparison of two alternative models is equivalent to the computation of an odds ratio for which the noninformative choice of selecting equal prior model probabilities has been initially made. The evaluation of the integral above is not always feasible, unless the model takes specific forms. Numerical methods must be usually employed in order to get an estimate of the marginal likelihood of a model. I derive the natural logarithm of what Chib calls the *"basic marginal likelihood"*. Bayes Theorem in its continuous version and in terms of the data and of the vector of parameters θ is:

$$p\left(\theta|Y; M\right) = \frac{\mathcal{L}\left(Y|\theta; M\right) \cdot \pi\left(\theta|M\right)}{m\left(Y|M\right)}$$

where $p\left(\theta|Y; M\right)$ is the joint posterior density of the parameters θ and the marginal likelihood $m\left(Y|M\right)$ is here interpreted as a normalizing factor. Solving for the normalizing constant:

$$m\left(Y|M\right) = \frac{\mathcal{L}\left(Y|\theta; M\right) \cdot \pi\left(\theta|M\right)}{p\left(\theta|Y; M\right)}$$

which is the *"basic marginal likelihood identity"*. By taking the logs of both the sides of the equation above:

$$\ln m\left(Y|M\right) = \ln \mathcal{L}\left(Y|\theta; M\right) + \ln \pi\left(\theta|M\right) - \ln p\left(\theta|Y; M\right)$$

which is the marginal log-likelihood of model M. Estimating the latter is really all the MCMC method used here is about. This requires the correct choice of θ^* from the posterior density of the vector of parameters. Chib argues that the selection of θ^* is not really a critical aspect of the procedure. However, it should be taken from a point of *"high density"*. Usually, the posterior mean is used to this purpose. Thus, the estimated marginal log-likelihood is equal to:

$$\ln \hat{m}\left(Y|M\right) = \ln \mathcal{L}\left(Y|\theta^*; M\right) + \ln \pi\left(\theta^*|M\right) - \ln \hat{p}\left(\theta^*|Y; M\right)$$

Since a Normal-Gamma specification is used for the model without breaks, its marginal log-likelihood has a convenient analytical expression, can be calculated exactly and does not need to be simulated. The Bayes factor that is finally computed is given by:

$$\widehat{BF}_{1,2} = e^{\left[\ln \hat{m}(Y|M_1) - \ln \hat{m}(Y|M_2)\right]}.$$

The marginal likelihood of model M_2 (hereafter the model without breaks as described by (2.a)) can be derived analitically, as briefly argued earlier. As such, the corresponding Bayes factors that are actually used have the following simpler form:

$$\widehat{BF}_{1,2} = e^{\left[\ln \hat{m}(Y|M_1) - \ln m(Y|M_2)\right]}.$$

sampler, a Markov Chain Monte Carlo (MCMC) technique that computes marginal posterior distributions for the parameters through the likelihood function of the model under analysis and by means of complex numerical methods that simulate draws from the joint posterior[28].

5.1 Bayesian Analysis - The Results

The following few tables briefly summarize the main outcomes obtained from the Bayesian analysis shortly described above. Values corresponding to the 5^{th}, 50^{th} and 95^{th} percentiles of the estimated marginal posterior distribution for the persistence parameter ρ are reported here, together with its estimated mean and standard deviation. Marginal log-likelihoods are then computed and Bayes factors calculated. In this section, M_1 denotes the model with a one-time break in the innovation variance and M_2 the model without breaks, so that values of the Bayes factor that are greater than 1 indicate a preference for the model with a break[29]:

[Tables 13 and 14 about here]

Classical Bayesian methods estimate the same breakpoint dates already detected using Hansen's procedure for Germany and Spain. Such dates are slightly different for France and Italy. There seems to be evidence in favor of a model with a break in the innovation variance (July 1995) for Italy, a finding that reverses the results highlighted through the grid-bootstrap procedure, which could not find significant proof of this.

Similar claims can be stated for the Bayesian model describing French yield spreads, which clearly favors the presence of a break in the corresponding conditional variance in May 1996. On the other hand classical Bayesian econometrics does not detect *statistically*

[28] In order to assess the likelihood of a model with respect to the other and then interpret the results as suggested by the computed Bayes factors, I follow the *"rule of thumb"* proposed by Jeffreys (1961) and Raftery (1996):

$$0 \leq 2 \cdot \log \widehat{BF}_{1,2} \leq 2.2 \rightarrow \text{very weak evidence for } M_1;$$
$$2.2 \leq 2 \cdot \log \widehat{BF}_{1,2} \leq 5 \rightarrow \text{weak to moderate evidence for } M_1;$$
$$5 \leq 2 \cdot \log \widehat{BF}_{1,2} \leq 10 \rightarrow \text{moderate to strong evidence for } M_1;$$
$$2 \cdot \log \widehat{BF}_{1,2} \geq 10 \rightarrow \text{decisive evidence for } M_1.$$

One should bear in mind that the interpretation of the Bayes factor is never so obvious. Bayesian econometrics returns probability distributions as *estimates* of the parameters of a model and something similar is true for alternative models, for which posterior model probabilities can be derived. Moreover, odds ratios give the researcher the possibility of *believing* more in a model with respect to another by simply specifying different prior model probabilities for each of them and, as a result, of drastically influencing his inference.

[29] Some of the results from Bayesian analysis in this section might be slightly different from their theoretical values. This is because the MCMC method is not an exact way of sampling the target distributions, but rather one for which realizations can be made as close as one likes to exact samples by observing the converging Markov chain for sufficiently long.

significant shifts in the conditional variance of German and Spanish nominal interest rate spreads. Evidence for these conclusions is moderate for Italy and quite strong for France. As for Germany and Spain, inferences derived from previous sections do not seem to find validation in Bayesian results[30].

The Bayesian model with a break detects exactly the same dates for Germany and Spain (but not the same for Italy and France) as those produced by Hansen's procedure. Combining the results from these two sources - this will be even clearer in the next section - suggests that, even though somewhat controversial and conditional on the methodology employed for their detection, structural shifts might have happened in all the four countries at the aforementioned dates.

Tables 15 and 16 summarize Bayesian estimations by proposing a comparison of the estimated average conditional variances for the four countries over the whole sample and for France and Italy only over the two subsamples, as determined by their respective break-dates. Since the output of any Bayesian procedure is always a probability distribution, relevant percentiles are reported for the estimated variable under investigation. In both the cases for which a shift has been detected and validated through Bayesian econometrics, estimates of the conditional variances substantially drop from the first to the second subsample. This empirical evidence is then compatible with the hypothesis (and the common belief) that the latest years of the '90s were characterized by less volatile and stabler financial markets in Europe.

[Tables 15 and 16 about here]

6 Extending the Sample - Hansen's Approach (again)

In the introduction to this paper I explained the reasons for which the focus of my analysis is on the sample period starting in January 1988 and ending in December 1998. I intuitively argued that considering longer samples for addressing the main issue of this work would be likely to raise the probability of actually detecting a one-time structural shift in the conditional variance of yield spreads. In this section I consider the following extension: I pick a longer sample, from January 1988 to February 2005, for the same four countries and I sketch the main results of Hansen's fixed-regressor grid-bootstrap method.

[30] Indeed, Bayesian techniques detect shifts in the cases of Germany and Spain, too. As pointed out earlier, these shifts are exactly the same as those previously found in Section 4, but they are not *significant* here since higher posterior probabilities are calculated for the corresponding models without breaks.

Beginning from January 1999, all the countries in the Euro area started sharing the same money market, directly influenced by the policy of the European Central Bank. Money-market interest rates (three-month rates in this study) have been common to all the countries in the EMU since January 1999. Long-term rates (referred to assets which are still country-specific, not included in the money market and not subject to the direct control of the ECB) have, instead, kept *some degree of independence*, even after the birth of the monetary union. Government bonds are, in fact, still issued by individual countries. However, more and more harmonized economic policies and integrated financial markets at the European level have implied that 10-year interest rates are very close to each other over the extension of the sample. Despite some difficulties under the point of view of the overall economic performance, the years after the settlement of the ECB in Europe have been notoriously characterized by low interest rates, low inflation rates and higher stability, particularly in financial and money markets. If one considers that the opposite is generally true for the period before that date (i.e., the period already analyzed in depth in the previous parts of this paper), it is legitimate to expect the detection of a one-time structural break in the conditional variance of yield spreads with a bigger probability.

After fitting equation (2) into the extended set of data and deriving the new series of the squared residuals, a graphical inspection of the recursive breakpoint Chow tests performed over the subsample 1990:07-2002:07 indicates this possibility. Only F statistics are reported below[31]:

[Figure 7 about here]

F-statistics above the *incorrect* critical values delivered by this procedure are more numerous than in the case considered in Section 3. This causes the intervals within which the potential structural break might have occurred to be larger and almost directly translates into a higher probability of estimating a one-time shift in correspondence of the highest values of the test statistics. Such values are found to be at the following dates: 1990:08 for Germany, 1994:10 for Spain, 1996:02 for France and 1993:03 for Italy.

Hansen's methodology adjusts the critical values just derived and provides additional statistics to test for the usual null of no structural breaks in the conditional variance of the

[31] **Breakpoint Chow Test (F Statistic):** Germany: 5% critical value=3.888375; 1% critical value=6.763299. Spain: 5% critical value=3.887906; 1% critical value=6.762011. France: 5% critical value=3.888139; 1% critical value=6.762652. Italy: 5% critical value=3.888139; 1% critical value=6.762652.

interest rate spreads, as indicated in Table 17:

[Table 17 about here]

Adjusted p-values reveal that a one-time shift is a better assumption for all the countries at their respective estimated dates. The dates of potential break detected for Spain and Italy are exactly the same as those estimated when the smaller sample had been considered. In that specific case, the null hypothesis for Italy could not be accepted. The date for Germany is close to that derived earlier[32]. As for France, the relevant date slightly switches from September 1992 (the non-significant break detected in Section 4) to February 1996.

Tables 18 and 19 present the estimated variances of the OLS regressions performed for each of the four countries over the whole sample and over the two subsamples as determined by the breakpoint dates detected:

[Tables 18 and 19 about here]

Estimated variances over the full sample are lower than the corresponding estimates over the smaller sample examined in Section 4 and this is true for all the four countries. This is reasonable, since the end of the old century and the beginning of the new one have clearly been stabler than the years from 1989 to 1998. As expected, estimated variances turn out to be higher over the subsamples before the occurrence of the structural breaks than over the subsamples after those breaks. Interestingly, lower variances (with some exceptions) are generally found in the cases of Germany and France. This might be evidence of overall better stability properties for German and French financial markets and economies with respect to Italian and Spanish ones over the period considered. A finding that confirms again the common belief about the characteristics of the four countries investigated.

7 Conclusions

I have presented several econometric alternatives to test for one-time structural shifts in the conditional variance of nominal interest rate spreads for four European countries, currently

[32] In that framework, such a date is found in March 1990; now, the break is estimated to occur in August 1990. The two results are still consistent with the same event (the fall of the Berlin Wall and the beginning of the reunification process of the two parts of the country). The reason for which they differ is easily explained: in this section, the sequence of Chow tests from which each test statistic is derived starts at a later date than in Section 3. It is clear from the results just shown that much of the *probability distribution* of the structural break under investigation in the case of Germany accumulates towards the beginning of 1990.

members of the European Monetary Union, over an eleven-year period of time starting in January 1988 and ending in December 1998. Autoregressive equations are used to model the time behavior of interest rate spreads and assumptions on the error terms are made in order to relax, when needed, the classical hypothesis of homoskedasticity and introduce a convenient form of conditional heteroskedasticity with the aim of taking the possibility of a shift in the innovation variance into account.

Results are homogeneous, to some degree, across the three methodologies used. A sequence of breakpoint Chow tests performed over the mid-70% of the entire sample period returns approximate intervals within which the shift might have occurred. The fixed-regressor grid-bootstrap method proposed by Hansen - which uses all the F statistics delivered by the previous procedure - refines those findings, provides the correct critical values and probability levels for the test statistics used and isolates exact change-point dates. Such a methodology highlights the presence of a break in Germany in March 1990 and of a break in Spain in October 1994. Classical Bayesian analysis validates these results to some extent by estimating only *non-significant* breaks in the innovation variance on the same dates in Germany and Spain. Some evidence of a decreased conditional variability of the interest rate spreads is found for these two countries after the possible break has arisen.

The cases of Italy and France are somewhat more ambiguous. The naïve approach points in favor of a very weak evidence of a structural break in Italy either at the beginning of 1993 or in mid-1995, but the first finding is confirmed by Hansen's grid-bootstrap only when an extended sample is used. Classical Bayesian econometrics estimates, instead, such a break in July 1995, but evidence of this is only moderate. French spreads seem to experience a shift in May 1996 according to Bayesian methods, but this break fails to be detected otherwise.

A simple extension presented in the last part of the paper considers a larger sample ending in February 2005. Extending the sample increases the probability of detecting a structural break in the innovation variance of the process describing yield spreads. Hansen's procedure applied to this broader case isolates one-time breaks for all the four countries considered. Breakpoint dates found in this new framework are generally consistent with those previously listed. The only significant difference is represented by the series of French spreads, for which a break in the conditional variance is estimated in February 1996. This is not in line with the results of the bootstrap method executed on the smaller sample but finds some confirmation in what obtained through Bayesian techniques. Table 20 briefly and informally summarizes the outcomes of the econometric analysis performed so far:

[Table 20 about here]

The methodologies adopted may suggest interesting extensions for this work. One of them (not examined here) might, for instance, consist in modelling potential heteroskedasticity in alternative ways, by specifying ARCH or GARCH models for the conditional variance of the yield spread. It would also be possible to study the stability properties of the parameters estimated, possibly allowing for multiple breaks over the whole sample or for smooth and slower adjustments from a volatility regime to another.

I derived a set of few dates, singled out by distinct approaches, in correspondence to which the researched structural shifts might have happened. Results are not always straightforward; their interpretation may depend on the technique used, but, for each of the nations considered, they are linked to specific events. Evidence of a decreased volatility in European financial markets (as measured by the conditional variability of the yield curve slope) has also been highlighted several times through the estimation of lower average conditional variances for those countries displaying structural breaks (all of them) over the subsamples starting from the dates of switching volatility regime. On the basis of the breakpoint dates detected, I find that the fall of the Berlin Wall, the currency crisis at the beginning of the '90s and the convergence towards the economic criteria required to join the Economic and Monetary Union in Europe in mid-1998 might have played a major role in determining the occurrence of those shifts and have started a period of higher stability in European financial markets and growth cycles at the end of the last century.

8 References

1. Andrews, D.W.K. (1993). *"Tests for Parameter Instability and Structural Change with Unknown Change Point"*; Econometrica 61 (4), 821-856;

2. Andrews, D.W.K. and H.Y. Chen (1994). *"Approximately Median-Unbiased Estimation of Autoregressive Models"*; Journal of Business and Economic Statistics 12 (2), 187-204;

3. Andrews, D.W.K. and W. Ploberger (1994). *"Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative"*; Econometrica 62 (6), 1383-1414;

4. Ang, A.; M. Piazzesi and M. Wei (2005). *"What Does the Yield Curve Tell Us about GDP Growth?"*; Journal of Econometrics (forthcoming);

5. Bos, C.S. (2002). *"A comparison of Marginal Likelihood Computation Methods"*; N. 02-084/4, Tinbergen Institute Discussion Papers from Tinbergen Institute;

6. Chib, S. (1995). *"Marginal Likelihood from the Gibbs Output"*; Journal of the American Statistical Association 90, 1313-1321;

7. Chib, S. (1998). *"Estimation and Comparison of Multiple Change-Point Models"*; Journal of Econometrics 86 (2), 221-241;

8. Clarke, K.A. (2000). *"The Effect of Priors on Approximate Bayes Factors from MCMC Output"*; Unpublished Manuscript, University of Michigan;

9. Côté, D. and C. Graham (2004). *"Convergence of Government Bond Yields in the Euro Area: The Role of Policy Harmonization"*; Bank of Canada - Working Paper Series N. 04-23;

10. Estrella, A. and F.S. Mishkin (1998). *"Predicting U.S. Recessions: Financial Variables as Leading Indicators"*; Review of Economics and Statistics 1, 45–61;

11. Fama, E.F. (1990). *"Term Structure Forecasts of Interest Rates, Inflation and Real Returns"*; Journal of Monetary Economics 25, 56–76;

12. Hansen, B.E. (1997). *"Approximate Asymptotic P Values for Structural-Change Tests"*; Journal of Business and Economic Statistics 15 (1); ABI/INFORM Global pg. 60;

13. Hansen, B.E. (1999). *"The Grid Bootstrap and the Autoregressive Model"*; The Review of Economics and Statistics 81 (4), 594–607;

14. Hansen, B.E. (2000). *"Testing for Structural Change in Conditional Models"*; Journal of Econometrics 97, 93-115;

15. Jeffreys, H. (1961). *"Theory of Probability (3 ed.)"*; Oxford: Oxford University Press;

16. Kim, C.J. and C.R. Nelson (1999). *"State-Space Models with Regime Switching - Classical and Gibbs-Sampling Approaches with Applications"*; MIT Press;

17. Levin, A.T. and J.M. Piger (2004). *"Is Inflation Persistence Intrinsic in Industrial Economies?"*; ECB - Working Paper Series N. 334;

18. Mishkin, F.S. (1990). *"The Information in the Longer Maturity Term Structure about Future Inflation"*; Quarterly Journal of Economics 55, 815–828;

19. Quandt, R. (1960). *"Tests for the Hypothesis that a Linear Regression Obeys Two Separate Regimes"*; Journal of the American Statistical Association 55, 324–330;

20. Raftery, A.E. (1996). *"Approximate Bayes Factors and Accounting for Model Uncertainty in Generalized Linear Models"*; Biometrika 83, 251-266.

Figures

Figure 1: Time Evolution of Three-Month Money Market Interest Rates

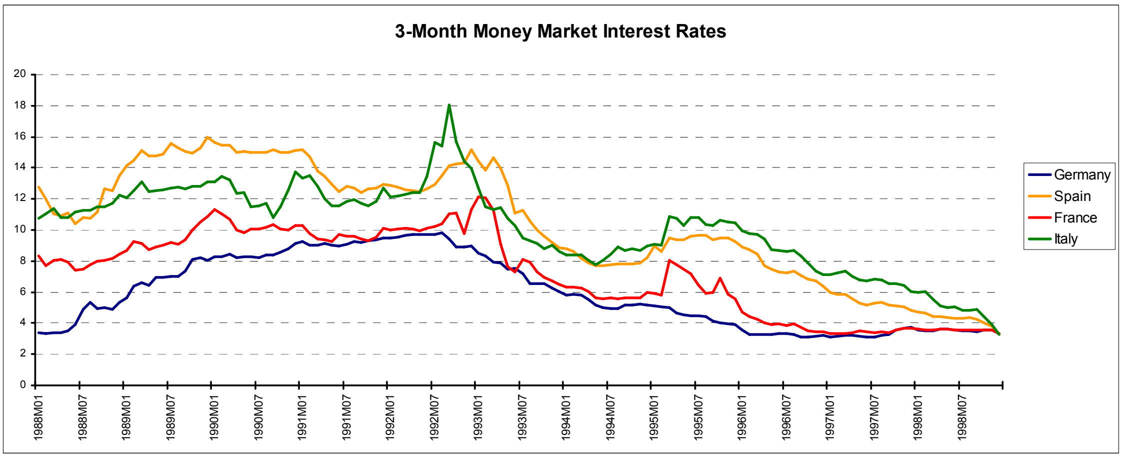


Figure 2: Time Evolution of 10-Year Government Bond Yields

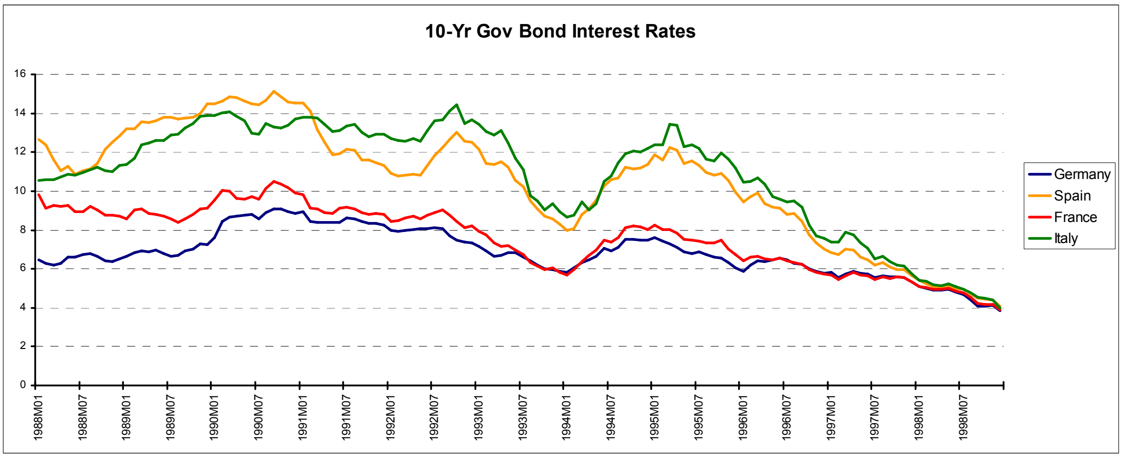


Figure 3: Time Evolution of Nominal Yield Spreads

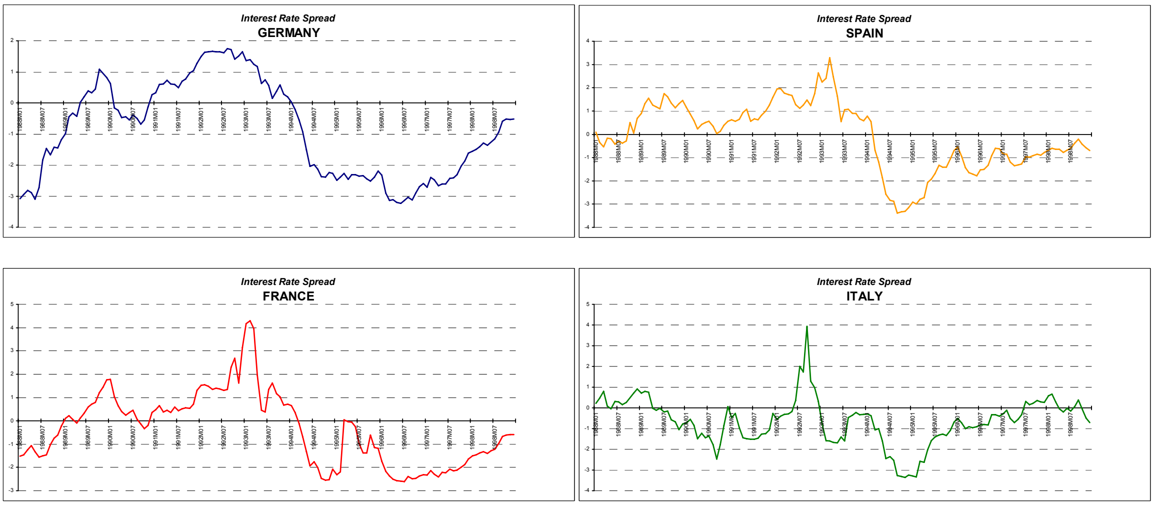


Figure 4: Equation $s_t = \mu + \rho s_{t-1} + \Sigma_{j=1}^{K-1} \phi_j \Delta s_{t-j} + \varepsilon_t$ - Estimated Residuals

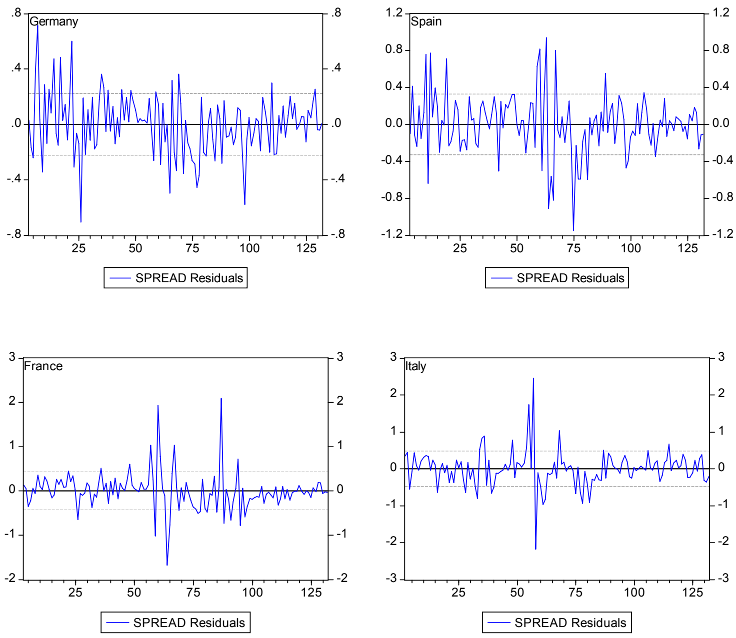


Figure 5: Equation $s_t = \mu + \rho s_{t-1} + \Sigma_{j=1}^{K-1} \phi_j \Delta s_{t-j} + \varepsilon_t$ - Estimated Squared Residuals

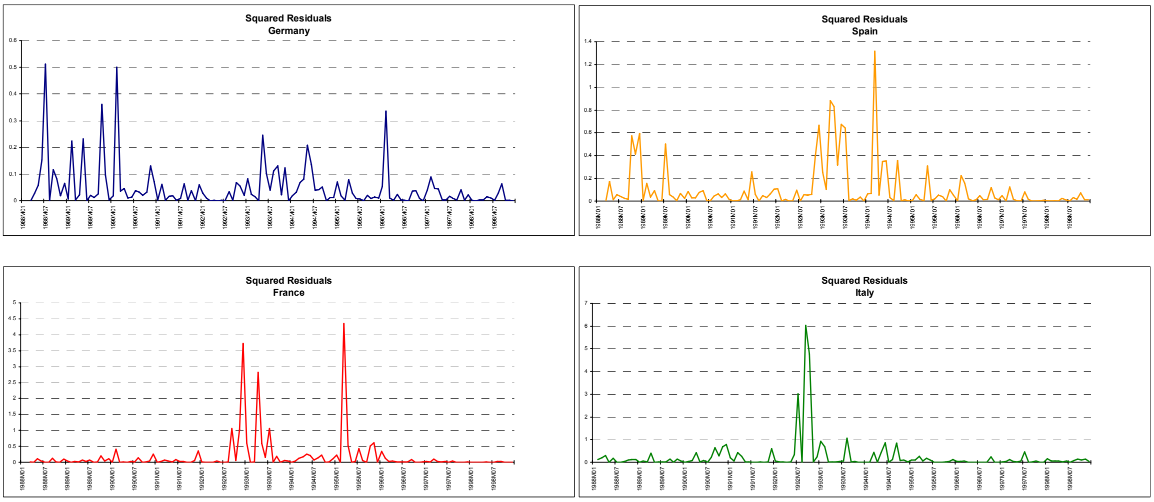


Figure 6: Naïve Approach - Sequence of Chow Tests and Sup Statistics

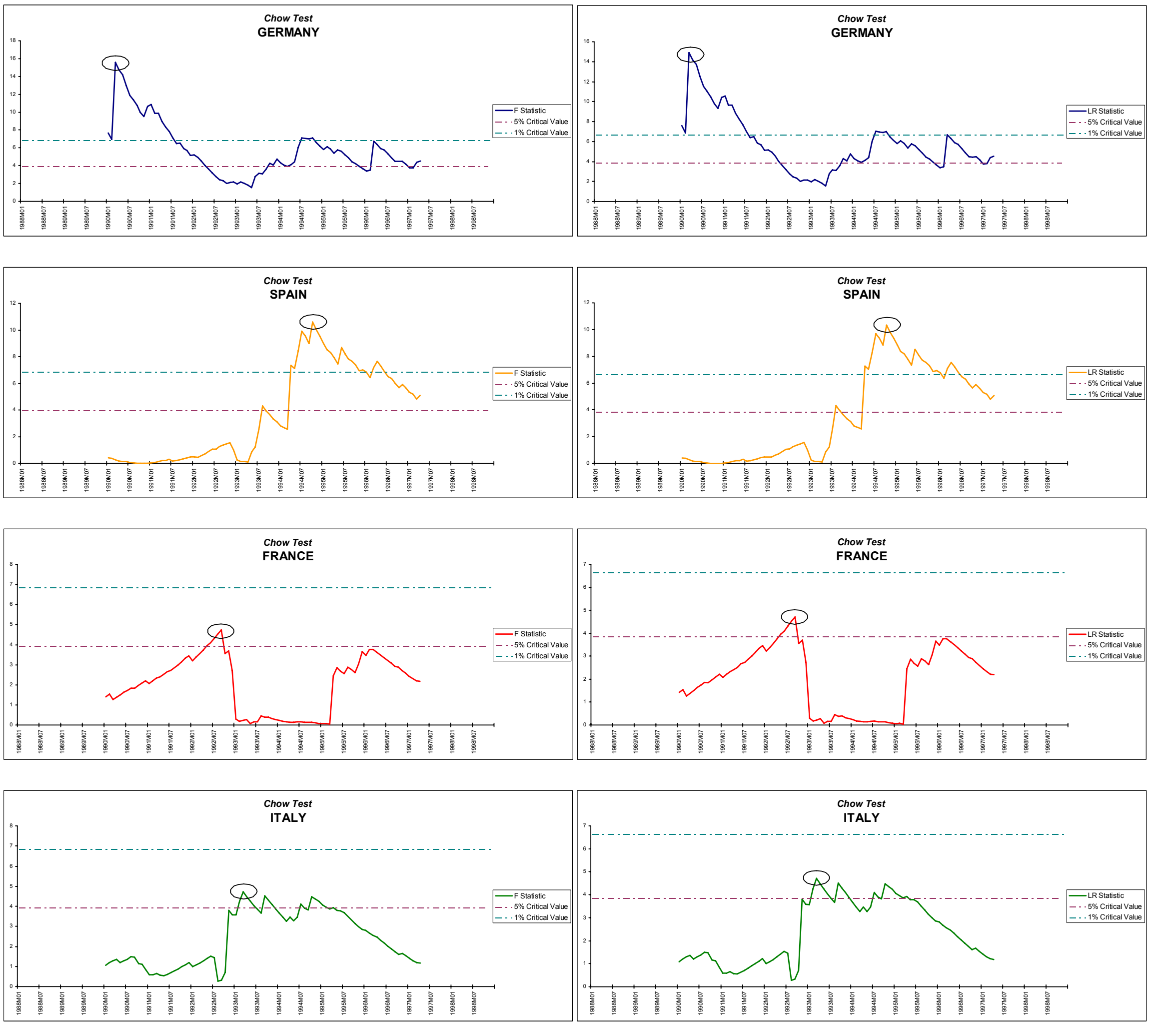
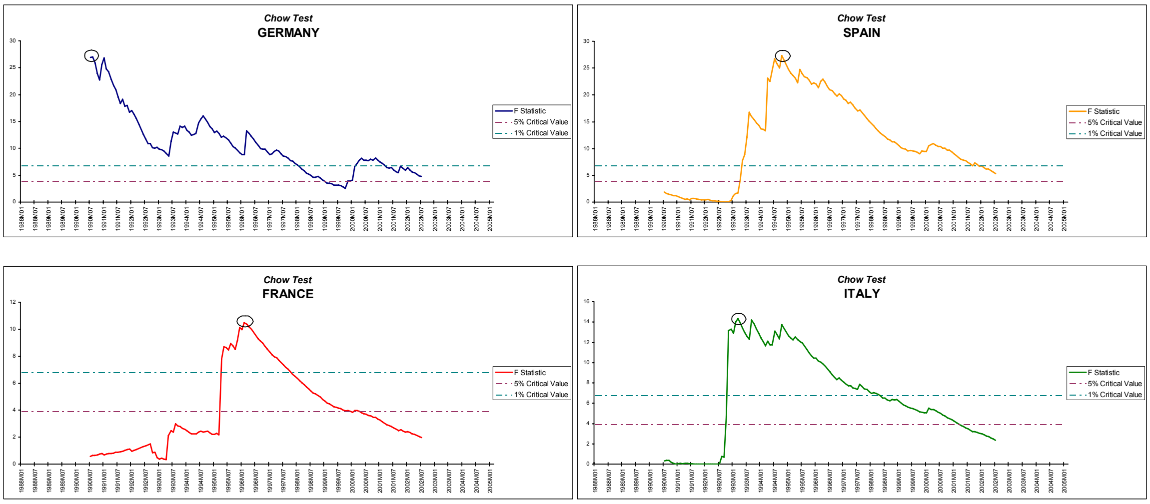


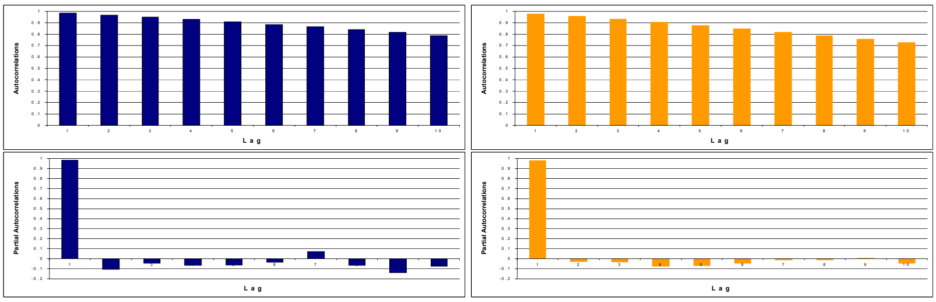
Figure 7: Naïve Approach - Sequence of Chow Tests and Sup Statistics (extended sample)



Tables

Table 1: 3-Month Money Market Interest Rates (Correlogram)

Germany *Spain*



France *Italy*

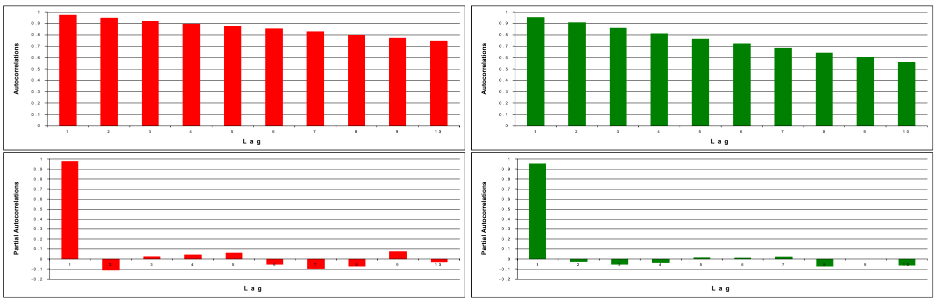
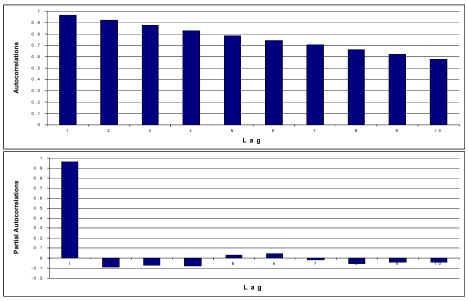
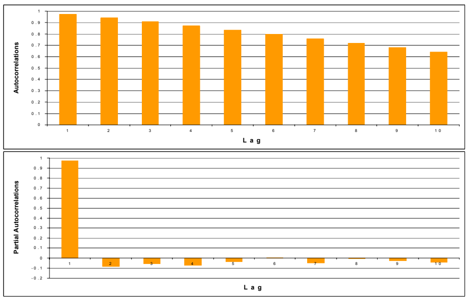


Table 2: 10-Yr Gov. Bond Interest Rates (Correlogram)

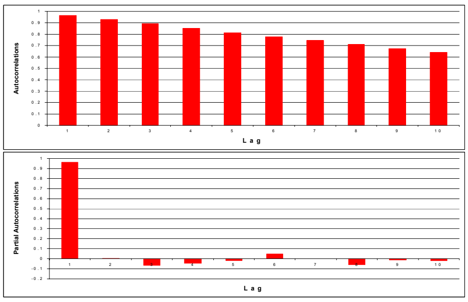
Germany



Spain



France



Italy

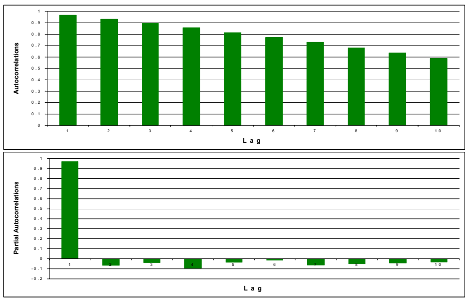
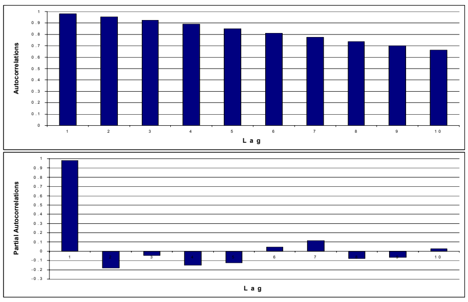
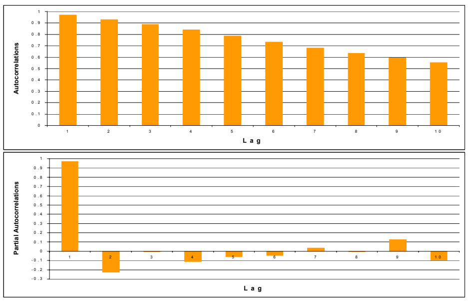


Table 3: Interest Rate Spreads (Correlogram)

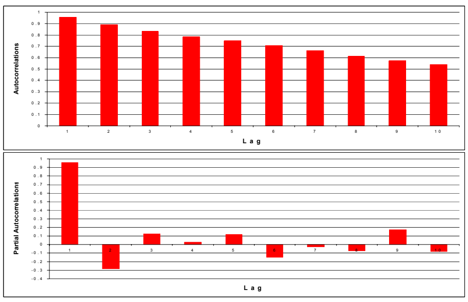
Germany



Spain



France



Italy

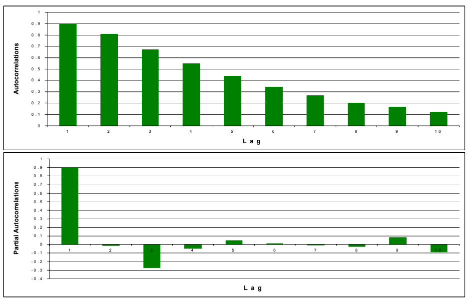


Table 4: Correlation Matrices

Correlation Matrix (3-month M.M. Int. Rates)				
	Germany	Spain	France	Italy
Germany	1	0.812	0.887	0.766
Spain	0.812	1	0.950	0.923
France	0.887	0.950	1	0.896
Italy	0.766	0.923	0.896	1

Correlation Matrix (10-year Gov. Bond Yields)				
	Germany	Spain	France	Italy
Germany	1	0.848	0.877	0.888
Spain	0.848	1	0.946	0.945
France	0.877	0.946	1	0.889
Italy	0.888	0.945	0.889	1

Correlation Matrix (Spreads)				
	Germany	Spain	France	Italy
Germany	1	0.834	0.898	0.205
Spain	0.834	1	0.835	0.461
France	0.898	0.835	1	0.169
Italy	0.205	0.461	0.169	1

Table 5: OLS Regressions ($s_t = \mu + \rho s_{t-1} + \Sigma_{j=1}^{K-1} \phi_j \Delta s_{t-j} + \varepsilon_t$)

Country	$\boldsymbol{\mu}$	$\boldsymbol{\rho}$	$\boldsymbol{\phi_1}$	**K**
Germany	-0.004798	0.979946***	0.327503***	2
	(0.022474)	(0.012733)	(0.083086)	
Spain	-0.004623	0.965689***	0.245094***	2
	(0.028942)	(0.020323)	(0.085550)	
France	-0.014550	0.944633***	0.297665***	2
	(0.038654)	(0.024414)	(0.084524)	
Italy	-0.070745	0.899499***	–	1
	(0.048719)	(0.038015)	–	

***: the estimated coefficient is significant at the 1% level.

Table 6: Chow Tests, F-Statistics and Dates of Potential Break

Country	Chow Test - F Statistic	Dates of potential break
Germany	above 5% Critical Value	1990:01-1992:05; 1993:10-1995:10; 1996:03-1996:12; 1997:03-1997:04
	above 1% Critical Value	1990:01-1991:07; 1994:07-1994:10
Spain	above 5% Critical Value	1993:08-1993:09; 1994:04-1997:04
	above 1% Critical Value	1994:04-1996:01; 1996:03-1996:06
France	above 5% Critical Value	1992:05-1992:09
	above 1% Critical Value	-
Italy	above 5% Critical Value	1993:02-1993:06; 1993:09-1993:11; 1994:07; 1994:10-1995:02; 1995:04
	above 1% Critical Value	-

Table 7: Chow Tests, LR-Statistics and Dates of Potential Break

Country	Chow Test - LR Statistic	Dates of potential break
Germany	above 5% Critical Value	1990:01-1992:04; 1993:10-1995:11; 1996:03-1996:12; 1997:03-1997:04
	above 1% Critical Value	1990:01-1991:07; 1994:07-1994:10; 1996:03
Spain	above 5% Critical Value	1993:08-1993:09; 1994:04-1997:04
	above 1% Critical Value	1994:04-1996:01; 1996:03-1996:06
France	above 5% Critical Value	1992:05-1992:09
	above 1% Critical Value	-
Italy	above 5% Critical Value	1993:02-1993:06; 1993:09-1993:12; 1994:07; 1994:08-1994:11; 1995:04
	above 1% Critical Value	-

Table 8: Chow Tests, Sup-Statistics and Most Likely Dates of Potential Break

Country		**Test Statistic**	**Date**	**Significance**	**Event**
Germany	sup F	15.572	1990:03	99%	German Reunification
	sup LR	14.925	1990:03	99%	Process
Spain	sup F	10.587	1994:10	99%	Convergence to
	sup LR	10.331	1994:10	99%	EMU
France	sup F	4.727	1992:09	95%	European Currency
	sup LR	4.714	1992:09	95%	Crisis
Italy	sup F	4.726	1993:03	95%	Italian Currency
	sup LR	4.714	1993:03	95%	Crisis

Table 9: OLS Regressions - Estimated Conditional Variances

	Estimated Average Conditional Variance (full sample)
Germany	0.049
Spain	0.106
France	0.181
Italy	0.231

Table 10: Hansen's Fixed-Regressor Grid-Bootstrap Method - Results

Germany (Date of potential break: 1990:03)

	Test Statistic	Andrews P-Value	Bootstrap P-Value	Hetero-Corrected P-Value
sup F	15.5720	0.00190	0.00100	0.03600
exp F	4.5867	0.00113	0.00100	0.03200
ave F	5.7995	0.00324	0.00300	0.01400

Spain (Date of potential break: 1994:10)

	Test Statistic	Andrews P-Value	Bootstrap P-Value	Hetero-Corrected P-Value
sup F	10.5870	0.02007	0.01600	0.00600
exp F	3.1345	0.01273	0.00900	0.00200
ave F	3.5859	0.02700	0.02100	0.00400

France (Date of potential break: 1992:09)

	Test Statistic	Andrews P-Value	Bootstrap P-Value	Hetero-Corrected P-Value
sup F	4.7271	0.27747	0.26300	0.08000
exp F	1.2124	0.14516	0.14500	0.01700
ave F	1.9913	0.11348	0.11500	0.01000

Italy (Date of potential break: 1993:03)

	Test Statistic	Andrews P-Value	Bootstrap P-Value	Hetero-Corrected P-Value
sup F	4.7262	0.27759	0.25700	0.08800
exp F	1.3843	0.11489	0.11700	0.01800
ave F	2.3087	0.08324	0.08500	0.00400

Table 11: OLS Regressions ($s_t = \mu + \rho s_{t-1} + \Sigma_{j=1}^{K-1} \phi_j \Delta s_{t-j} + \varepsilon_t$)

Heteroskedastic Error Terms

Country	$\boldsymbol{\mu}$	$\boldsymbol{\rho}$	$\boldsymbol{\phi_1}$	**K**
Germany	-0.006936	0.986140***	0.318812***	2
	(0.020307)	(0.011404)	(0.080196)	
Spain	-0.015547	0.953406***	0.281473***	2
	(0.026946)	(0.017550)	(0.081185)	

***: the estimated coefficient is significant at the 1% level.

Table 12: OLS Regressions

(Estimated Conditional Variances before and after the break)

	Estimated Average Conditional Variance (1^{st} subsample - before the break)	Estimated Average Conditional Variance (2^{nd} subsample - after the break)
Germany	0.107	0.036
Spain	0.151	0.036

Tables 13: Bayesian Estimations ($s_t = \mu + \rho s_{t-1} + \Sigma_{j=1}^{K-1} \phi_j \Delta s_{t-j} + \varepsilon_t$)

Country		**Model with a break** (M_1) $\boldsymbol{\rho}$	**Model without breaks** (M_2) $\boldsymbol{\rho}$
Germany	5^{th} percentile	0.9651	0.9574
	50^{th} percentile	0.9851	0.9799
	95^{th} percentile	1.0062	1.0020
	Mean	0.9856	0.9799
		(0.0124)	(0.0137)
	K	2	2
Spain	5^{th} percentile	0.9248	0.9310
	50^{th} percentile	0.9601	0.9658
	95^{th} percentile	0.9961	1.0010
	Mean	0.9605	0.9658
		(0.0213)	(0.0211)
	K	2	2
France	5^{th} percentile	0.9317	0.9039
	50^{th} percentile	0.9719	0.9446
	95^{th} percentile	1.0039	0.9854
	Mean	0.9706	0.9446
		(0.0224)	(0.0248)
	K	1	2
Italy	5^{th} percentile	0.8372	0.8369
	50^{th} percentile	0.9068	0.8996
	95^{th} percentile	0.9679	0.9626
	Mean	0.9044	0.8996
		(0.0382)	(0.0383)
	K	1	1

Table 14: Bayesian Estimations - Model Selection

Country		Model with a break (M_1)	Model without breaks (M_2)
Germany	Marginal Log-Likelihood	-17.967	5.403
	Estimated Breakpoint Date	1990:03	-
	$\widehat{BF}_{1,2}$	7.0904E-11	
	$2\cdot \log \widehat{BF}_{1,2}$	-20.2987	
Spain	Marginal Log-Likelihood	-57.645	-45.803
	Estimated Breakpoint Date	1994:10	-
	$\widehat{BF}_{1,2}$	7.1948E-06	
	$2\cdot \log \widehat{BF}_{1,2}$	-10.2860	
France	Marginal Log-Likelihood	45.414	-81.339
	Estimated Breakpoint Date	1996:05	-
	$\widehat{BF}_{1,2}$	1.1175E+55	
	$2\cdot \log \widehat{BF}_{1,2}$	110.0965	
Italy	Marginal Log-Likelihood	-89.760	-95.825
	Estimated Breakpoint Date	1995:07	-
	$\widehat{BF}_{1,2}$	430.3864	
	$2\cdot \log \widehat{BF}_{1,2}$	5.2677	

Table 15: Bayesian Estimations - Estimated Conditional Variances

	Estimated Average Conditional Variance (full sample)			
	5^{th} percentile	50^{th} percentile	95^{th} percentile	Mean
Germany	0.047	0.057	0.071	0.057
Spain	0.094	0.115	0.142	0.114
France	0.156	0.191	0.236	0.190
Italy	0.194	0.237	0.291	0.235

Table 16: Bayesian Estimations

(Estimated Conditional Variances before and after the break)

Estimated Average Conditional Variance (1^{st} subsample - before the break)				
	5^{th} percentile	50^{th} percentile	95^{th} percentile	Mean
France	0.212	0.266	0.339	0.264
Italy	0.258	0.327	0.422	0.324

Estimated Average Conditional Variance (2^{nd} subsample - after the break)				
	5^{th} percentile	50^{th} percentile	95^{th} percentile	Mean
France	0.030	0.044	0.069	0.044
Italy	0.055	0.076	0.111	0.075

Table 17: Hansen's Fixed-Regressor Grid-Bootstrap Method - Results - Extended Sample

Germany (Date of potential break: 1990:08)

	Test Statistic	Andrews P-Value	Bootstrap P-Value	Hetero-Corrected P-Value
sup F	27.0420	0.00001	0.00000	0.00400
exp F	10.1730	0.00000	0.00000	0.00300
ave F	10.5800	0.00000	0.00000	0.00000

Spain (Date of potential break: 1994:10)

	Test Statistic	Andrews P-Value	Bootstrap P-Value	Hetero-Corrected P-Value
sup F	27.3050	0.00001	0.00000	0.00000
exp F	10.5940	0.00000	0.00000	0.00000
ave F	11.5970	0.00000	0.00000	0.00000

France (Date of potential break: 1996:02)

	Test Statistic	Andrews P-Value	Bootstrap P-Value	Hetero-Corrected P-Value
sup F	10.4720	0.02117	0.01100	0.00500
exp F	3.1551	0.01238	0.00700	0.00300
ave F	4.0425	0.01830	0.01100	0.00100

Italy (Date of potential break: 1993:03)

	Test Statistic	Andrews P-Value	Bootstrap P-Value	Hetero-Corrected P-Value
sup F	14.3380	0.00342	0.00400	0.00100
exp F	5.1803	0.00028	0.00200	0.00400
ave F	6.6860	0.00100	0.00100	0.00100

Table 18: OLS Regressions - Estimated Conditional Variances - Extended Sample

	Estimated Average Conditional Variance (full sample)
Germany	0.040
Spain	0.078
France	0.125
Italy	0.151

Table 19: OLS Regressions

(Estimated Conditional Variances before and after the break)

Extended Sample

	Estimated Average Conditional Variance (1^{st} subsample - before the break)	Estimated Average Conditional Variance (2^{nd} subsample - after the break)
Germany	0.101	0.031
Spain	0.151	0.032
France	0.239	0.026
Italy	0.357	0.067

Table 20: Summary of the Detected Structural Breaks

GERMANY	Original Sample (1988:01 - 1998:12)		Extended Sample (1988:01 - 2005:02)	
Approach	Date of Potential Break	Significant?	Date of Potential Break	Significant?
Naïve	1990:03	Yes	1990:08	Yes
Fixed-Regressor Grid-Bootstrap	1990:03	Yes	1990:08	Yes
Bayesian Analysis	1990:03	No	n.a.	-

SPAIN	Original Sample (1988:01 - 1998:12)		Extended Sample (1988:01 - 2005:02)	
Approach	Date of Potential Break	Significant?	Date of Potential Break	Significant?
Naïve	1994:10	Yes	1994:10	Yes
Fixed-Regressor Grid-Bootstrap	1994:10	Yes	1994:10	Yes
Bayesian Analysis	1994:10	No	n.a.	-

FRANCE	Original Sample (1988:01 - 1998:12)		Extended Sample (1988:01 - 2005:02)	
Approach	Date of Potential Break	Significant?	Date of Potential Break	Significant?
Naïve	1992:09	Yes	1996:02	Yes
Fixed-Regressor Grid-Bootstrap	1992:09	No	1996:02	Yes
Bayesian Analysis	1996:05	Yes	n.a.	-

ITALY	Original Sample (1988:01 - 1998:12)		Extended Sample (1988:01 - 2005:02)	
Approach	Date of Potential Break	Significant?	Date of Potential Break	Significant?
Naïve	1993:03	Yes	1993:03	Yes
Fixed-Regressor Grid-Bootstrap	1993:03	No	1993:03	Yes
Bayesian Analysis	1995:07	Yes	n.a.	-