Banco Central de Chile
Documentos de Trabajo

Central Bank of Chile
Working Papers

N° 613

Febrero 2011

MODELING COPPER PRICE: A REGIME – SWITCHING APPROACH

Javier García – Cicco   Roque Montero

BANCO CENTRAL DE CHILE

CENTRAL BANK OF CHILE

La serie Documentos de Trabajo es una publicación del Banco Central de Chile que divulga los trabajos de investigación económica realizados por profesionales de esta institución o encargados por ella a terceros. El objetivo de la serie es aportar al debate temas relevantes y presentar nuevos enfoques en el análisis de los mismos. La difusión de los Documentos de Trabajo sólo intenta facilitar el intercambio de ideas y dar a conocer investigaciones, con carácter preliminar, para su discusión y comentarios.

La publicación de los Documentos de Trabajo no está sujeta a la aprobación previa de los miembros del Consejo del Banco Central de Chile. Tanto el contenido de los Documentos de Trabajo como también los análisis y conclusiones que de ellos se deriven, son de exclusiva responsabilidad de su o sus autores y no reflejan necesariamente la opinión del Banco Central de Chile o de sus Consejeros.

The Working Papers series of the Central Bank of Chile disseminates economic research conducted by Central Bank staff or third parties under the sponsorship of the Bank. The purpose of the series is to contribute to the discussion of relevant issues and develop new analytical or empirical approaches in their analyses. The only aim of the Working Papers is to disseminate preliminary research for its discussion and comments.

Publication of Working Papers is not subject to previous approval by the members of the Board of the Central Bank. The views and conclusions presented in the papers are exclusively those of the author(s) and do not necessarily reflect the position of the Central Bank of Chile or of the Board members.

MODELING COPPER PRICE: A REGIME – SWITCHING APPROACH

Javier García – Cicco
Banco Central de Chile
y Universidad Católica de Argentina

Roque Montero
Banco Central de Chile

Abstract

This paper explores the virtues of Markov-Switching models to characterize the behavior of copper price. In particular, we study the performance of several univariate specifications of this type of models, both in and out of sample, comparing them also with constant parameter models such as ARMA and GARCH. The main finding is that allowing for a regime-switching variance in the error term is most relevant in explaining the behavior of this price.

Resumen

Este artículo explora las virtudes de modelos de cambios de régimen (Markov Switching) para caracterizar el comportamiento del precio del cobre. En particular, se analiza el desempeño de especificaciones univariadas de este tipo de modelos, tanto dentro como fuera de muestra, comparándolos también con modelos de parámetros constantes, como ARMA y GARCH. El resultado principal es que, a la hora de modelar el precio del cobre, es importante considerar el cambio de régimen en la varianza.

1   Introduction

Commodity prices generally exhibit large and persistent swings, displaying periods of relatively stable prices and times of high volatility. The price of copper not an exception to this general characterization, as shown in figure 1 where the log of the monthly spot price of copper is displayed. Given this behavior, the goal of this paper is to analyze if the evolution of this price can be characterized by a Markov Switching (MS) model, i.e. a model in which parameters change according to an unobserved Markov process. Moreover, the evolution of this price in the last 5 years, particularly the large swings experienced during the recent global financial crisis, further motives the evaluation of this kind of model as a contender to explain the behavior of this variable.

The analysis is divided in two parts. On one hand, we analyze the virtues of an autoregressive MS model in sample, using several specifications for this model and comparing them with both ARMA and GARCH models. In particular, we compare specifications according to information criteria, evaluate the estimated parameters of the model and describe the inference about the unobserved Markov states. On the other hand, we evaluate the forecasting ability of these alternatives, both in terms of point and density forecast. A distinctive feature of the analysis, motivated by the non-linearities and non-normality intrinsic in MS models, is the use of Markov Chain Monte Carlo methods to both characterize the distribution of the parameters and to evaluate the forecast density.

The main result of the paper is that, in modeling copper price, to consider a regime-switching variance is most relevant. Every MS specification evaluated that includes a time-changing variance outperforms others MS alternatives that do not allow for this feature. Moreover, MS model are superior to ARMA and GARCH models in sample. Out of sample, MS appears to improve, in terms of root-mean-squared forecast error, over ARMA models but is similar to GARCH specifications. In terms of coverage, the forecast confidence bands of the MS are slightly better than those of both alternatives, but not statistically different. Finally, the forecast variance decomposition reveals the importance of accounting for the uncertainty coming from the unobserved state in characterizing the uncertainty about the forecast.

This paper is related with some recent studies that propose to use MS models for copper price. Heaney (2006) uses an MS model to characterize the ratio of future to spot price of copper, presenting both univariate and structural models, analyzing only the in-sample performance of these models. Hong Chan and Young (2009) use a GARCH-Jump models with MS variance to explain the daily return of future prices, both in sample and out of sample, but focusing only on point forecast and not on its density. Choi and Hammoudeh (2010) specify a model considering only a MS variance, evaluating its in-sample performance. Relative to these studies, our analysis provides a more thorough model comparison exercise and, as emphasized before, use simulation-based method for the estimation and the characterization of the density forecast. Finally, our paper is also related with a growing literature in macroeconomics that emphasizes the importance of allowing for time-varying variance in explaining the behavior of several macro aggregates (see, for instance, the recent survey by Fernández-Villaverde and Rubio-Ramírez,

2010).

The rest of the paper is organized as follows. Section 2 discusses some basics of MS models and describes our approach for the empirical implementation of these models, while a more detailed discussion of the methodology is presented in the Appendix. Section 3 presents the in-sample analysis, both in terms of the specification of the MS model and the comparison with other constant-parameter models . The forecasting ability of these kind of models is analyzed in Section 4. Section 5 concludes.

2   Methodology

We begin with a brief description of kind of MS models that we are going to consider. For simplicity, we just describe here the case of a first-order autoregressive model with two states, while in the appendix we show a more general fomrulation and also describe some details associated with the estimation.[1] Given an observed variable of interest, y_t, this model can be written as

$$y_t = c_{St} + \phi_{St} y_{t-1} + \sigma_{St} \varepsilon_t,$$

where ε_t is an $i.i.d.$ process with standard normal distribution. The discrete variable $S_t = 1, 2$ denotes the unobserved state of the economy, which is determined by a Markov process. The characteristic element of the transition probability matrix of this process is given by

$$p_{ij} \equiv \Pr(s_t = j | s_{t-1}, s_{t-2}, ..., y_{t-1}, y_{t-2}, ...) = \Pr(s_t = j | s_{t-1} = i),$$

satisfying $\sum_i p_{ij} = 1$. Notice that these assumptions imply that the transition probability is independent from the values observable variable, and also that the previous value taken by the state is a sufficient statistic to characterize the transition probability. Finally, the notation for the parameters $(c_{St}, \phi_{St}, \sigma_{St})$ denotes the values of, respectively, the constant, the lag coefficient and the standard error of the shock in each possible state S_t.

In addition to the parameters, another statistic of interest is the probability of being in a given state at date t, implied by the available observations up to date t and for given values for the parameters. Therefore, we will be interested in characterizing the probability

$$\xi_{jt|t} \equiv \Pr(s_t = j | \Omega_t; \theta)$$

where Ω_t denotes the set of observations up to time t and θ collects all the parameters of the model $(c_{St}, \phi_{St}, \sigma_{St}, p_{ij})$. Given that the states are unobserved, we can use filtering techniques to infer these probabilities.

The parameters to be estimated are, therefore, $(c_{St}, \phi_{St}, \sigma_{St})$ and the transition probabilities p_{ij} (e.g. in this simple model, eight parameters have to be estimated). While the likelihood function can be easily evaluated numerically, as shown in the appendix, it will be a highly non-

---

[1]For a more detailed treatment see, for instance, Hamilton (1994).

linear function of the parameters in the model, limiting the ability of gradient-based numerical methods to find the global maximum of the objective function. For this reason, we choose to work with a Markov Chain Monte Carlo (MCMC) approach, using the Metropolis-Hastings algorithm to characterize the likelihood function. Therefore, we can give either a frequentist interpretation to our estimated parameters, according to the Laplace-type estimators proposed by Chernozhukov and Hong (2003), or a Bayesian interpretation under the assumption of flat priors.

The MCMC approach is also useful to perform inference about the parameters. Although asymptotic tests are available for these models (see, for instance, Garcia, 1998), given the non-linearity and non-normality of the model, it is likely for asymptotic inference to be inappropriate in short samples. The MCMC procedure, instead, will allow us to compute confidence sets that are more appealing in short samples while still being asymptotically valid. Moreover, the model comparison tests that are available also rely on asymptotic distributions (see, for instance, the survey in Hamilton, 2008), and for the same reasons we consider them unreliable for short samples. We will use instead three different information criteria for model comparison and selection: Bayesian (BIC), Akaike (AIC) and Hannan-Quinn (HQC).[2]

3   In-Sample Analysis

We divide the in-sample analysis in two parts. First, we start by selecting and characterizing the preferred specification among different MS models, including also a model comparison exercise with ARMA and GARCH models. We then explore the robustness of the results to different subsamples and repeat the in-sample analysis using the growth rate of the copper price instead of the level.

3.1   Specification of the MS Model

The data used corresponds to the log of the monthly spot price of copper (in dollars) at the London Market (the source is IMF-IFS), from January 1975 to January 2010. Our Benchmark for the comparison will be a linear AR(2) model. This specification was chosen based on both Box-Jenkins specification analysis and on information criteria, comparing ARMA models with up to 12 lags in both AR and MA components. We also performed a battery of unit root tests, using information criteria for lag selection that are robust to the local-to-unity problem. Although these tests cannot reject the null of a unit root, we choose not to work with a random walk specification as a benchmark for at least two reasons. First, if the appropriate model is one with regime-switching parameters (as our later analysis suggests), the typical Augmented Dickey-Fuller test for unit roots will be biased. In addition, given that throughout the paper

---

[2]Smith et al. (2006) also explore the use of information criteria in MS models. From a Bayesian perspective, the approach of using information criteria as a test for model selection can be justified based on the results of, for instance, Hong and Preston (2008), who show that comparing models based on BIC is asymptotically equivalent to perform hypothesis tests based on posterior-odds ratios.

our approach for model selection is based on information criteria, none of the random walk specifications that we tried were able improve over the AR(2) according to the three criteria that we consider.

In terms of the MS model, we will carry our analysis with an autoregressive model of order two. On one hand, this choice ensures a cleaner comparison with the linear specification. On the other hand, although not reported, we have also estimated all the variants of the MS models also for up to four lags, but the analysis based on all the information criteria suggested two lags as the appropriate choice for the MS models as well. We will start by evaluating seven different cases, each of them differing in the type of coefficients that change according to a two-regime Markov process. In particular,

- Case 1: All parameters change.

- Case 2: Only the constant changes.

- Case 3: Only the lags change.

- Case 4: Only the variance changes.

- Case 5: Only the constant and the lags change.

- Case 6: Only the constant and the variance change.

- Case 7: Only the lags and the variance change.

Table 1 displays the values for the three information criteria obtained for each of these seven cases, as well as other alternatives described below. Among these seven, all the criteria point to the specification where only the variance changes as the preferred one (case 4). However, the difference with the alternatives where the variance is also changing (cases 1, 6 and 7) is much smaller than with the cases in which the variance remains fixed across regimes. This is the first piece of evidence emphasizing the role of a changing variance to model copper price.[3]

In all the cases we have considered so far, when more than one parameter is changing, it was assumed that all the parameters change according to the same Markov process. Alternatively, we can consider that different parameters change according to different processes. Given the emphasis in the variance, we estimate two additional cases. In Case 8, the variance changes according to a Markov process that can take two values, while both the constant and the lags coefficients move according to a different (and independent from that of the variance) two-states Markov process. Additionally, Case 9 is similar to this last one, with the difference that we

---

[3]We have also considered, although not reported, these same seven cases in a specification in which parameters can change according to a three-state Markov process. Two conclusions can be drawn from that exercise. First, among these alternatives with three states, the relative ranking in terms of information criteria of the seven cases is the same as in the case with two states. Second, the best specification with three states (also case 4) is not better in this dimension than any of the models with two states that allow for the change in the variance. Given these results, we discarded this alternative three-state alternative.

fixed the lags parameters to be constant across states.[4] The information criteria for these cases are also presented in Table 1. The difference between these two alternatives and cases 4 and 6 is much smaller (in terms of the AIC these two are even slightly better).

The final exercise in terms of information criteria is to compare the MS models with other alternatives that do not consider regime switching. A first natural comparison is against the AR(2) model. Additionally, given the importance of the variance, we also consider a GARCH model. In particular, after comparing information criteria for different specifications of this type of model, we choose an AR(2)-GARCH(1,1) as the preferred model. The information criteria for these two alternatives are reported in the final lines of Table 1. As we can see, while the GARCH improves over the AR specification both models are clearly outperformed by the MS alternatives that include a changing variance. We interpret this as evidence in favor of the MS specifications.

We next turn to the analysis of the estimated parameters. Tables 2 and 3 report the estimated coefficients for cases 1, 4, 6, 8 and 9, as well as those for the AR(2) specification for comparison. We focus on these cases based on the results obtained using information criteria. Starting with case 1, where all parameters change according to the same Markov process, we can see that the second state is identified as the one associated with the higher variance. Additionally, while all the coefficients have small standard errors and the point estimates seem to differ across regimes, we can also see that the confidence set for the difference of the coefficients includes zero for all the parameters except for the variance. In terms of transition probabilities, both regimes appear to be quite persistent, particularly the one associated with the low variance (regime 1): the point estimate for $p_{1,1}$ is 0.953 and for $p_{2,2} = 1 - p_{2,1}$ is 0.78. However, it should be noticed that, while the confidence set for $p_{1,1}$ is quite tight, $p_{2,1}$ is estimated with smaller precision.[5]

The results for cases 4 and 6 display a similar results. The estimated coefficients for the variance in both regimes are almost indistinguishable from those in case 1, and the confidence sets seem to indicate that the variance is indeed different across states. On the contrary, in case 6 it seems that the estimated values for the constant are not significantly different across regimes. In terms of the transition probabilities, the estimated values for these two cases are quite similar to those obtained in case 1, and we can also see how the confidence set for $p_{2,1}$ is thinner than in case 1; reflecting the increase in power that we obtained by constraining some parameters to remain constant across states.

---

[4]These two cases can be considered in the general framework as a four-state Markov process. To see this, let $S_{1t} = 1, 2$ denote the two-state process governing the change in the variance, while $S_{2t} = 1, 2$ denotes the two-state process governing the change in other parameters. Then $S_t = \{S_{1t}, S_{2t}\}$ can take four values: $\{1, 1\}, \{1, 2\}, \{2, 1\}, \{2, 2\}$.

[5]This is a quite common pattern observed in MS applications, and it is due to the fact that, because of the small sample, one generally tends to observe fewer periods with changing states than periods when a given state last on time. Therefore, a general drawback of MS models is the somehow limited power of inference for the probability of moving from one state to the other. Moreover, this feature further emphasizes the use of methods (like MCMC) that allow to characterize short sample distributions, instead of using asymptotically-based inference.

In terms of models that are driven by two independent Markov processes (cases 8 and 9), it is also the case that the estimated variances in both cases are similar to those in the other cases. Additionally, while the point estimates of the other parameters tend to differ across states, the confidence set for the difference across regimes includes the zero in both cases. The transition probabilities of the Markov process that govern the evolution of parameters other than the variance are less precisely estimated than those probabilities for the variance state, particularly for case 9 where the confidence set for these probabilities include almost all the parameter space.

As we mentioned before, one of the interesting features of the MS model is that we can use filtering techniques to characterize the probability of being in a given state at given period, conditional on the whole sample (i.e. $\xi_{jt|T}$). The different plots in Figure 2 show the smoothed probability of being in the low-variance state, for all the cases considered. The upper-left graph displays this probability for case 4, along with the price of copper. As we can see, for most of the sample the price remained in the low-variance state, with some exceptions: for three months early in 1980, for two months in mid 1982, for almost a year staring on November 1987, for a month in June 1996, and after 2006 (although with some interruptions). We will further analyze this last period latter. Additionally, the upper-right graph reproduces the smoothed probability of being in the low-variance state, now including also a 95% confidence band for this estimated probability. As can be see, this probability is estimated quite precisely, with some exceptions at the end of the of the sample.[6]

Given the similarities in terms of goodness of fit between cases 1, 4, 6 and 7, it is interesting to see whether the inference about the unobserved state is similar in these cases. This is reported in the lower-left panel of figure 2, where we can see that the smoothed probability is virtually identical in these four cases. On the contrary, if we compute this probability for the other cases that do not include a changing variance (cases 2, 3 and 5), we can see in the lower-right panel of the figure that the inference is quite different relative to the other cases, and that these estimated probabilities are more erratic.

A similar pattern can be found if we analyze cases 8 and 9, presented in figure 3: while the inference for the low variance probability is comparable with that obtained for case 4, the filtered probability for the Markov process for the constant and lag coefficients cannot clearly identify the presence of different states for these parameters.[7]

The conclusion of this part of the analysis is that, in explaining the in-sample evolution of copper price, considering a regime-switching variance is crucial. Moreover, it is less clear that allowing with regime switches in the other parameters of the model can significantly improve over a model that just accounts for switching in the variance of the error term.

---

[6]The fact that the inference is less precise at the end of the sample was in part expected, for we are conditioning in a smaller information set latter in the sample given the two-sided filtered probability.

[7]Moreover, if we were to include the confidence bands for these filtered probabilities, they will show significant uncertainty associated with the smoothed probabilities presented in the left graph of the figure.

3.2 Robustness

Here we present several robustness exercises, guided by the following observation. At first glance, the plot of the copper price in figure 1 would suggest a change in the unconditional mean of the series, starting somewhere near 2005. The results so far, on the other hand, had emphasized the role of the variance but not of the mean or lags coefficients. One might then wonder whether the model is somehow confusing changes in the variances for those in the unconditional mean.

A starting point is to test whether the unconditional mean is significantly different across regimes in the models we have estimated so far. The confidence set of the differences across regimes is presented in table 4 for cases 6, 8 and 9, where one can appreciate that these models do not imply a significant difference in this statistic across regimes.

A second exercise consists in re-estimating the models in two subsamples: one finishing in December 2004, where the price reached values closed to its then historical maximum period attained during 1989, and another one finishing in December 2007, so that we eliminate the observations associated with the 2008 global financial crisis and its aftermath. Table 5 reports the information criteria for cases 1 to 7, where we can see that the ranking of models found for the full sample remains in these two cases. In particular, case 4 (only the variance changing) highlights as the preferred model, and the difference with the close competitors is somehow wider.

In terms of parameters, table 6 displays the point estimates for case 4. The coefficients are really similar in the different samples, with only a minor reduction in the variance in the high-variance state, which was expected given that we are eliminating a highly volatile period in both samples. Finally, figure 4 plots the filtered probability of the low-variance state in the different samples. As can be seen, the inference in terms of this probability is comparable with that obtained with the full sample.

As a final exercise we estimated a model for the log-difference of copper price. In such a setting, a one-time regime change in the unconditional mean will just represent an outlier in the sample and, therefore, if the model is really confusing variance with unconditional mean, the time-varying variance should not be identified if we use the log-difference for estimation. Table 7 presents the information criteria analysis for cases 1 to 7 using this alternative series, where results indicate that case 4 is also the preferred specification in this case, and that models that include a regime-switching variance outperform those who do not. Moreover, in terms of the filtered probabilities, figure 5 replicates the analysis presented in figure 2 for the log of copper price. The two patterns detected before can be found here as well: all the models that include a regime switch in the variance provide the same inference for this probability, while the other models estimate really erratic probabilities that do not allow to identify periods where the model clearly indicates the presence of one of the regimes.

The results of these exercises reinforce the findings previously presented and allow us to say, with some degree of confidence, that the model is truly detecting a change in the variance that is not being confused by a change in the unconditional mean.

4 Out-of-Sample Analysis

In this section, we evaluate the forecasting ability of the MS model, in terms of both point forecast as well as in terms confidence sets. The inclusion of confidence-set analysis is motivated by the importance previously highlighted of the regime-switching variance. The exercises we perform consist in comparing the MS model of case 4 with both AR and GARCH models used also in the previous section. We focus in case 4 only just to make the analysis more clear and shorter. Nevertheless, we have compared also case 4 with the other cases and found no significant improvements with these other alternatives.

In terms of point forecast, we run a progressive estimation starting in January 2004 (i.e. dropping the last 15% of the sample), and for each new sample we estimated the model by maximizing the likelihood and computed the forecast (up to twelve month ahead) using the likelihood mode. Notice that, at least *ex-ante*, it is not clear which of the three models should dominate; for in all cases the estimators of the parameters should be consistent, and the correct specification of the variance should only matters for efficiency. Table 8 displays the root-mean-squared forecast errors (RMSFE) for each of the models. In addition, we present the statistics of the tests by Diebold and Mariano (1995), and the refinement proposed by Harvey et al. (1997), of the null that both models have the same RMSFE. On one hand, we can see that the MS model seems to outperforms the AR model at all horizons. On the other hand, the advantage over the GARCH model is less clear, with the MS model providing a significantly smaller RMSFE only at two and three month ahead horizons, while for the other both models seem to be equally good.

In terms of the confidence sets and coverage analysis, we choose an approach that allows us to characterize all the sources of uncertainty that are present in forecasting with MS models. Given the model, a forecast starting on period T of future values of the observed variables up to $T + J$, $\{y_{T+h}\}_{h=1}^{J}$, can be formed as a function of the parameters, θ, current values of the unobserved states, S_T, a sequence of shock, $\{\varepsilon_{T+h}\}_{h=1}^{J}$, and the history of observables up to date T, y^T. The object of interest is then the distribution of the forecasted variables conditional on observed information up to period T, i.e. $p(\{y_{T+h}\}|y^T)$. A convenient way to write this distribution is

$$
\begin{aligned}
p(\{y_{T+h}\}|y^T) &= \int p(\{y_{T+h}\}, \theta|y^T)d\theta = \int p(\{y_{T+h}\}|y^T, \theta)p(\theta|y^T)d\theta = \\
&= \int p(\{y_{T+h}\}|y^T, \theta, S_T)p(S_T|y^T, \theta)p(\theta|y^T)d\theta = \\
&= \int p(\{y_{T+h}\}|y^T, \theta, \{\varepsilon_{T+h}\}, S_T)p(\{\varepsilon_{T+h}\}|\theta)p(S_T|y^T, \theta)p(\theta|y^T)d\theta, \quad (1)
\end{aligned}
$$

where $p(\theta|y^T)$ is the distribution of the parameters inferred in the estimation process and the other probabilities are obtained from the structure of the model. This decomposition highlights that the uncertainty about the forecast can be divided in three sources: parameter uncertainty $p(\theta|y^T)$, initial unobserved-state uncertainty $p(S_T|y^T, \theta)$, and shock uncertainty $p(\{u_{T+h}\}|\theta)$.

We are interested in confidence sets (percentiles) associated with $p(\{y_{T+h}\}|y^T)$. However, as the integral in (1) cannot be computed algebraically, simulation based methods can be used. We propose the following pseudo-algorithm:

1. Draw a parameter vector θ^i from $p(\theta|y^T)$.[8]

2. Draw an initial state vector $S_T^{i,j}$ from $p(S_T|y^T, \theta^i)$.

3. Draw a sequence $\varepsilon_{T+1}^{i,j,n}, ..., \varepsilon_{T+J}^{i,j,n}$ from the distribution $p(\{\varepsilon_{T+h}\}|\theta^i)$.

4. Use the model, θ^i, $S_T^{i,j}$, and $\varepsilon_{T+1}^{i,j,n}, ..., \varepsilon_{T+J}^{i,j,n}$ to generate the forecast $y_{T+1}^{i,j,n}, ..., y_{T+J}^{i,j,n}$.

5. Repeat 3 and 4 for $n = 1, ..., N_\varepsilon$ for the same θ^i and $s_T^{i,j}$.

6. Repeat 2 to 5 for $j = 1, ..., N_S$ for the same θ^i.

7. Repeat 1 to 6 for $i = 1, ..., N_\theta$.

This algorithm will give us $N = N_\varepsilon \cdot N_S \cdot N_\theta$ draws from the distribution of interest $p(\{y_{T+h}\}|y^T)$ that can be used to construct confidence sets, where all sources of uncertainty are taken into account. In the exercises we preset below, we used $N_\varepsilon = N_S = N_\theta = 50$, giving us a total of 125K simulations to construct the confidence sets.[9]

The starting date for this forecast exercise starts in December 2004, and we repeat the algorithm presented before for each new sample consisting on adding one observation at a time.[10] To have a first look at how inference about confidence sets can be different with different models, figure 6 presents two examples of a 90% confidence set: one when the forecast starts at January 2005 (a period identified to be a low-variance state) and the other starting in January 2008 (a period identified to be a high-variance state). We can see that in both cases the MS provides a tighter confidence set than the other two alternatives. The difference between the two pictures is the behavior of the AR and GARCH models: in the low-variance state, the GARCH model estimates a smaller confidence set than the AR model, while the opposite happens in the high-variance state.

Regardless of the width of the confidence set, a more formal analysis is to compare the coverage of these sets (i.e. the percentage of times that the actual observation turned up to be inside the set). This is presented for the three models in table 9, were we also included the p-values of the Giacomini and White (2006) test of predictive ability, using a quadratic

---

[8]This is obtained as a by-product of the MCMC procedure used for estimation.

[9]As a robustness check, we also tried, for some starting dates for the forecast, increasing the number of draws up to $N_\varepsilon = N_S = N_\theta = 500$, which gives a total of 125M simulations. This alternative, which is considerably more time consuming, yields similar results to those presented here. For the AR and GARCH models, $N_S = 0$ by definition, and we choose $N_\theta = 1000$ and $N_\varepsilon = 100$, for a total of 100K simulations. The draws from $p(\theta|y^T)$ for these two models were obtained by the same MCMC procedure used for the MS alternative.

[10]A difference with the exercise used for the RMSFE analysis is that we do not re-estimate the model for each new date we add. We choose to do this to save computational time, because to re-estimate will entail running the MCMC procedure for each sample, which is quite time consuming, thinking also that after doing the MCMC we would need to run the 125K simulations to construct the confidence set.

coverage accuracy loss function, of the null that a given model provides the same coverage that the MS model. We can see that the MS model seems to have better coverage properties than both of the alternatives, particularly compared to the GARCH specification, although the test appears to indicate that these difference are not significant.

A final exercise that can be obtained as a by product of the algorithm used to compute the confidence set is a forecast-error variance decomposition, allowing us to identify how important are the alternative sources of uncertainty in characterizing the density of the forecast.[11] The results are presented in table 10. For all models, the mayor source of forecast uncertainty is the variance of the error term. The difference is that, while in both AR and GARCH models parameter uncertainty plays a non-negligible role, particularly at long horizons, in the MS model the uncertainty about parameters is relatively unimportant, while not knowing the initial state can have a relevant impact on the variance of the forecast.

5   Conclusions

We have presented a thorough analysis of the virtues of Markov switching models in explaining the time series of copper price. The main message of the paper is that including a regime-switching variance is most relevant in modeling this price. Moreover, MS models that include this feature seem to outperform both ARMA and GARCH specifications in sample, and are slightly better out of sample relative to these alternatives.

We conclude by suggesting two alternative routs for future research related to our analysis. On one hand, a natural extension is to consider a Markov switching vector auto-regression model that includes copper price along with other relevant determinants such as inventories, exchange rates, world interest rates, among others, and to evaluate the role of MS models in explaining and forecasting copper price in such a framework. On the other hand, in terms of forecasting, it would be interesting to study whether the forecast and its density can be improved by pooling MS with other models such as ARMA or GARCH.

---

[11]The appendix shows how to construct this decomposition from outcome of the simulation exercise.

References

Chernozhukov, V. and H. Hong (2003), "An MCMC approach to classical estimation." *Journal of Econometrics*, 115, 293–346.

Choi, K. and S. Hammoudeh (2010), "Volatility behavior of oil, industrial commodity and stock markets in a regime-switching environment." *Energy Policy*, 38, 43884399.

Diebold, F.X. and R.S. Mariano (1995), "Comparing predictive accuracy." *Journal of Business and Economic Statistics*, 13, 253–263.

Fernández-Villaverde, J. and J. Rubio-Ramírez (2010), "Macroeconomics and volatility: Data, models, and estimation." *NBER Working Paper*, No. 14284.

Garcia, R. (1998), "Asymptotic null distribution of the likelihood ratio test in markov switching models." *International Economic Review*, 39, 763–788.

Giacomini, R. and H. White (2006), "Tests of conditional predictive ability." *Econometrica*, 74, 1545–1578.

Hamilton, J. (1994), *Time Series Analysis*. Princeton University Press.

Hamilton, J. (2008), "Regime switching models." *New Palgrave Dictionary of Economics*, 2nd Edition.

Harvey, D., S. Leybourne, and P. Newbold (1997), "Testing the equality of prediction mean squared errors." *International Journal of Forecasting*, 13, 281–291.

Heaney, R. (2006), "An empirical analysis of commodity pricing." *The Journal of Futures Market*, 26 (4), 391–415.

Hong, H. and B. Preston (2008), "Bayesian average, prediction and nonnested model selection." *NBER Working Paper*, No. 14284.

Hong Chan, W. and D. Young (2009), "A new look at copper markets: A regime-switching jump model." *University of Alberta, Working Papers*, WP 2009-13.

Smith, A., P. Naik, and C. Tsai (2006), "Markov-switching model selection using kullback-leibler divergence." *Journal of Econometrics*, 134, 553–577.

A   Technical Appendix

Here we present several details regarding the specification and estimation of MS models. In general, an autoregressive model of order p, where parameters follow a N-states Markov process, $MS(N)AR(p)$, can be written as

$$y_t = c_{St} + \phi_{1,St}y_{t-1} + ... + \phi_{p,St}y_{t-p} + \sigma_{St}\varepsilon_t$$

where $\varepsilon_t \sim \mathcal{N}(0,1)$ and $S_t = \{1,2,...,N\}$ is the state variable with transition probability matrix

$$P = \begin{bmatrix} p_{11} & p_{21} & \cdots & p_{N1} \\ p_{12} & p_{22} & \cdots & p_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1N} & p_{2N} & \cdots & p_{NN} \end{bmatrix}$$

where the characteristic element is defined as $p_{ij} = \Pr(s_t = j|s_{t-1} = i)$.

Let $\theta = (c_1, ..., c_N, \phi_{1,1}, ..., \phi_{p,N}, \sigma_1, ..., \sigma_N, p_{11}, ..., p_{NN})$ be the vector that groups all the parameter of the model. We can also define the probability of being in state j at time t, conditional on the parameter vector and on the data up to data t, as $\xi_{jt} = \Pr(s_t = j|\Omega_t; \theta)$. The procedure to evaluate the likelihood of the model can be summarized as follows:

1. Given θ and an initial value for ξ_{j0}, the density for period t in the state j,

$$\eta_{jt} = f(y_t/s_t = j, \Omega_{t-1}; \theta) = \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left[-\frac{(y_t - c_j - \phi_{1,j}y_{t-1} - ... - \phi_{p,j}y_{t-p})^2}{2\sigma_j^2}\right]$$

2. The conditional likelihood for observation t is defined as:

$$f(y_t|\Omega_{t-1}; \theta) = \sum_{i=1}^{N}\sum_{j=1}^{N} p_{ij}\xi_{i,t-1}\eta_{jt}$$

3. Using these, the inference over the states of the model in the period t is

$$\xi_{jt} = \frac{\sum_{i=1}^{2} p_{ij}\xi_{i,t-1}\eta_{jt}}{f(y_t|\Omega_{t-1}; \theta)}$$

4. Repeat these steps for $t = 1, 2, ..., T$.

Through these iterations it is possible to calculate the whole sequence of ξ_{jt} and the conditional density $f(y_t|\Omega_{t-1}; \theta)$ for each observation. Therefore, we are able to write and evaluate the conditional log likelihood of the data given θ as

$$\log f(y_1, ..., y_T|y_0; \theta) = \sum_{t=1}^{T} \log f(y_t|\Omega_{t-1}; \theta).$$

Two alternatives can be used as the initial value of ξ_{j0}. The first one is to take an arbitrary value for it, and the second uses the unconditional probability of each state implied by the parameter vector. The latter is preferred because the initial value of ξ_{j0} changes endogenously in each iterations of the numerical algorithm. The unconditional probability vector π can be obtained by solving the system

$$\pi = (A'A)^{-1} A'e_{N+1},$$

where $A = \begin{bmatrix} I_N - P \\ \mathbf{1} \end{bmatrix}$, $\mathbf{1}$ denotes a $N \times 1$ vector of ones and e_{N+1} is the $N + 1$th column of I_{N+1}.

Another relevant issue for the initialization of the maximization algorithm is the starting values for the parameters in each of the regimes. In particular, we try four different alternatives. First, we run a QLR test of structural break in the parameters and compute the different values of the coefficients before and after the break. The disadvantage of this alternative is that it detects only on break that is deterministic. As an alternative, we also try two threshold models (TAR), where the regime depends on whether the variable is above or below of, in one alternative, its mean or, in the other, its median. Finally, we also considered a SETAR model in which the threshold value is estimated.[12] To choose among these alternatives, we ran a Monte Carlo experiment were data was generated by an artificial MS model, and check which of the alternatives used to initialize the algorithm generate maximum likelihood estimates closer to the true parameters. According to the results of this experiment, we choose to use the SETAR alternative to find the initial values for the estimation of the MS model

As we mentioned, our estimation approach consist in characterizing the likelihood function using MCMC methods. In particular, we proceed in two steps. First, the likelihood is maximized, using the optimization algorithm `csminwel` developed by Chris Sims.[13] The resulting mode is used as the starting value of a Random Walk Metropolis-Hastings algorithm, using a $\mathcal{N}(0, c\Sigma)$ as the proposal distribution.[14] The parameter c is calibrated to obtain an acceptance ratio close to 30% and the convergence of the chain is analyzed by checking recursive means. For each estimated alternative we generate 300K draws from the posterior, eliminating the first half of the chain to reduce the dependence from initial values.

A.1 Forecast Error Variance Decomposition

A relevant question in these models is what fraction of the total forecast uncertainty can be attributed to each of the sources described in the text. In what follows we show a theoretical decomposition of the forecast variance and suggest an implementation based on the outcome from the algorithm used to construct the confidence sets. For this, the following result would

---

[12]The estimation of these last three alternatives was implemented using concentration methods to minimize the sum of squared residuals.

[13]Available at http://sims.princeton.edu/yftp/optimize/.

[14]Σ is the inverse of the posterior's Hessian evaluated at the mode computed in the first step.

prove useful. Let x_t and z_t be two random vectors, then

$$V_x(x_t) = E_z\left[V_{x|z}(x_t|z_t)\right] + V_z\left[E_{x|z}(x_t|z_t)\right], \tag{2}$$

where $E_w(\cdot)$ and $V_w(\cdot)$ denote, respectively, the expectation and variance-covariance operator computed over the distribution of a generic random vector w_t.[15]

We are interested in $V_{y_{T+h}|y_T}(y_{T+h}|y_T)$. Using the previous results, we can write

$$V_{y_{T+h}|y_T}(y_{T+h}|y_T) = E_{\theta|y^T}\left[V_{y_{T+h}|\theta,y_T}(y_{T+h}|\theta,y_T)\right] + V_{\theta|y^T}\left[E_{y_{T+h}|\theta,y_T}(y_{T+h}|\theta,y_T)\right].$$

The first term $E_{\theta|y^T}\left[V_{y_{T+h}|\theta,y_T}(y_{T+h}|\theta,y_T)\right]$ represents the average uncertainty in the forecast when parameters are assumed to be known. The second term $V_{\theta|y^T}\left[E_{y_{T+h}|\theta,y_T}(y_{T+h}|\theta,y_T)\right]$ therefore represents the additional volatility that comes from parameter uncertainty. Applying again (2) to the first term we get,

$$\begin{aligned}
V_{y_{T+h}|y_T}(y_{T+h}|y_T) &= E_{\theta|y^T}\left\{E_{S_T|\theta,y^T}\left[V_{y_{T+h}|S_T,\theta,y_T}(y_{T+h}|S_T,\theta,y_T)\right]\right\} \\
&+ E_{\theta|y^T}\left\{V_{S_T|\theta,y^T}\left[E_{y_{T+h}|S_T,\theta,y_T}(y_{T+h}|S_T,\theta,y_T)\right]\right\} \\
&+ V_{\theta|y^T}\left[E_{y_{T+h}|\theta,y_T}(y_{T+h}|\theta,y_T)\right].
\end{aligned}$$

Now, the second term $E_{\theta|y^T}\left\{V_{S_T|\theta,y^T}\left[E_{y_{T+h}|S_T,\theta,y_T}(y_{T+h}|S_T,\theta,y_T)\right]\right\}$ represents the average uncertainty brought about by not knowing the current state S_T, while the first term $E_{\theta|y^T}\left\{E_{S_T|\theta,y^T}\left[V_{y_{T+h}|S_T,\theta,y_T}(y_{T+h}|S_T,\theta,y_T)\right]\right\}$ can then be interpreted as the uncertainty related to the exogenous shocks.

The outcome from the algorithm proposed in section 4 can be used to compute each of these terms as follows:[16]

- Computing $E_{\theta|y^T}\left\{E_{S_T|\theta,y^T}\left[V_{y_{T+h}|S_T,\theta,y_T}(y_{T+h}|S_T,\theta,y_T)\right]\right\}$ (shock uncertainty):

    - $E_{y_{T+h}|S_T,\theta,y_T}(y_{T+h}|S_T,\theta,y_T) \approx N_\varepsilon^{-1}\sum_n y_{T+h}^{i,j,n} \equiv \bar{y}_{T+h}^{i,j}, \quad [N_\theta \cdot N_s]$
    - $V_{y_{T+h}|S_T,\theta,y_T}(y_{T+h}|S_T,\theta,y_T) \approx N_\varepsilon^{-1}\sum_n \left(y_{T+h}^{i,j,n} - \bar{y}_{T+h}^{i,j}\right)^2 \equiv V_{T+h}^{i,j}, \quad [N_\theta \cdot N_s]$
    - $E_{S_T|\theta,y^T}\left[V_{y_{T+h}|S_T,\theta,y_T}(y_{T+h}|S_T,\theta,y_T)\right] \approx N_S^{-1}\sum_j V_{T+h}^{i,j} \equiv \bar{V}_{T+h}^i, \quad [N_\theta]$
    - $E_{\theta|y^T}\left\{E_{S_T|\theta,y^T}\left[V_{y_{T+h}|S_T,\theta,y_T}(y_{T+h}|S_T,\theta,y_T)\right]\right\} \approx N_\theta^{-1}\sum_i \bar{V}_{T+h}^i. \quad [1]$

- Computing $E_{\theta|y^T}\left\{V_{S_T|\theta,y^T}\left[E_{y_{T+h}|S_T,\theta,y_T}(y_{T+h}|S_T,\theta,y_T)\right]\right\}$ (initial-state uncertainty):

    - $E_{S_T|\theta,y^T}\left[E_{y_{T+h}|S_T,\theta,y_T}(y_{T+h}|S_T,\theta,y_T)\right] \approx N_S^{-1}\sum_j \bar{y}_{T+h}^{i,j} \equiv \bar{y}_{T+h}^i, \quad [N_\theta]$

---

[15]A proof of this claim for the univariate case is as follows (the extension for vectors is straightforward):

$$\begin{aligned}
V_x(x_t) &= E_x\left\{[x_t - E_x(x_t)]^2\right\} = E_x(x_t^2) - [E_x(x_t)]^2 = E_z\left[E_{x|z}(x_t^2|z_t)\right] - \left\{E_z\left[E_{x|z}(x_t|z_t)\right]\right\}^2, \\
&= E_z\left[E_{x|z}(x_t^2|z_t)\right] - E_z\left\{[E_{x|z}(x_t|z_t)]^2\right\} + V_z\left[E_{x|z}(x_t|z_t)\right] = E_z\left[V_{x|z}(x_t|z_t)\right] + V_z\left[E_{x|z}(x_t|z_t)\right],
\end{aligned}$$

where the third equality follows from the law of iterated expectations, and the fourth and fifth use the formula for the variance $V_w(w_t) = E_w(w_t^2) - [E_w(w_t)]^2$.

[16]The number of simulations left in each case is shown in brackets at the end of each line.

$$- V_{S_T|\theta,y^T}\left[E_{y_{T+h}|S_T,\theta,y_T}(y_{T+h}|S_T,\theta,y_T)\right] \approx N_S^{-1}\sum_j \left(\bar{y}_{T+h}^{i,j} - \bar{y}_{T+h}^i\right)^2 \equiv V_{T+h}^i, \quad [N_\theta]$$

$$- E_{\theta|y^T}\left\{V_{S_T|\theta,y^T}\left[E_{y_{T+h}|S_T,\theta,y_T}(y_{T+h}|S_T,\theta,y_T)\right]\right\} \approx N_\theta^{-1}\sum_i V_{T+h}^i. \quad [1]$$

– Computing $V_{\theta|y^T}\left[E_{y_{T+h}|\theta,y_T}(y_{T+h}|\theta,y_T)\right]$ (parameter uncertainty):

$$- E_{y_{T+h}|\theta,y_T}(y_{T+h}|\theta,y_T) \approx (N_\varepsilon N_S)^{-1}\sum_{n,j} y_{T+h}^{i,j,n} \equiv \bar{y}_{T+h}^i, \quad [N_\theta]$$

$$- V_{\theta|y^T}\left[E_{y_{T+h}|\theta,y_T}(y_{T+h}|\theta,y_T)\right] \approx N_\theta^{-1}\sum_i \left(\bar{y}_{T+h}^i - N_S^{-1}\sum_\theta \bar{y}_{T+h}^i\right)^2. \quad [1]$$

$$- V_{y_{T+h}|y_T}(y_{T+h}|y_T) \approx N^{-1}\sum_{n,j}\left(y_{T+h}^{i,j,n} - N^{-1}\sum_{n,j} y_{T+h}^{i,j,n}\right)^2, \quad [1]$$

B   Tables and Figures

Table 1: Information Criteria

Case	Parameters changing	Number of states	AIC	HQC	BIC
1	c, ϕ, σ	2	-2.9262	-2.8881	-2.8299
2	c	2	-2.7621	-2.7354	-2.6946
3	ϕ	2	-2.8376	-2.8072	-2.7605
4	σ	2	-2.9358	-2.9091	-2.8683
5	c, ϕ	2	-2.8554	-2.8211	-2.7687
6	c, σ	2	-2.9317	-2.9012	-2.8546
7	ϕ, σ	2	-2.9272	-2.8929	-2.8404
8	c, ϕ, σ	4	-2.9369	-2.8912	-2.8213
9	c, σ	4	-2.9390	-2.9009	-2.8426
	AR(2)		-2.7632	-2.7516	-2.7338
	AR(2)-GARCH(1,1)		-2.8366	-2.8138	-2.7790

Note: See the text for the description of the cases.

Table 2: Estimated Parameters

		Case 1				Case 4	
	AR	$S_t = 1$	$S_t = 2$	C.S.	$S_t = 1$	$S_t = 2$	C.S.
c	0.037	0.008	0.155	[-0.37;0.05]	0.021		
	(0.03)	(0.03)	(0.13)		(0.03)		
ϕ_1	1.370	1.321	1.397	[-0.32;0.16]	1.345		
	(0.00)	(0.06)	(0.13)		(0.05)		
ϕ_2	-0.378	-0.323	-0.431	[-0.13;0.35]	-0.349		
	(0.01)	(0.06)	(0.13)		(0.05)		
σ^2	0.004	0.002	0.013	[-0.02;-0.01]	0.002	0.012	[-0.02;-0.01]
	(0.00)	(0.00)	(0.00)		(0.00)	(0.00)	
$p_{1,1}$		0.953		[0.91;0.98]	0.955		[0.91;0.98]
$p_{2,1}$		0.220		[0.08;0.45]	0.182		[0.07;0.36]

Note: Standard errors in parenthesis. For the parameters c, ϕ, σ^2 the column C.S. reports the confidence set of the difference of the coefficient between both states (e.g. $c_{S_t=1} - c_{S_t=2}$). For the probabilities, the column C.S. reports the confidence set of the estimated probability. The rest of the entries are the mean of the distribution. All these where obtained using the MCMC procedure described in the appendix, using 150K draws from the distribution.

Table 3: Estimated Parameters, Cont.

	Case 6			Case 8			Case 9		
	$S_t = 1$	$S_t = 2$	C.S.	$S_t = 1$	$S_t = 2$	C.S.	$S_t = 1$	$S_t = 2$	C.S.
c	0.026	0.033	[-0.03;0.02]	0.001	1.345	[-1.38;-1.32]	0.025	0.020	[-0.05;0.07]
	(0.03)	(0.04)		(0.01)	(0.02)		(0.04)	(0.04)	
ϕ_1	1.343			-0.342	0.195	[-1.59;-1.44]	1.337		
	(0.05)			(0.02)	(0.04)		(0.05)		
ϕ_2	-0.349			1.203	-0.262	[0.34;0.55]	-0.342		
	(0.05)			(0.03)	(0.04)		(0.05)		
σ^2	0.002	0.012	[-0.02;-0.01]	0.001	0.010	[-0.01;0.00]	0.002	0.011	[-0.02;-0.01]
	(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)	
$p_{1,1}$	0.955		[0.91;0.98]	0.946		[0.9;0.98]	0.955		[0.91;0.98]
$p_{2,1}$	0.189		[0.07;0.37]	0.233		[0.08;0.48]	0.183		[0.07;0.37]
$p_{1,1}^2$				0.813		[0.71;0.9]	0.589		[0.07;0.99]
$p_{2,1}^2$				0.621		[0.43;0.8]	0.454		[0.02;0.93]

Note: See Table 2. $p_{1,1}^2$ and $p_{2,1}^2$ are the transition probabilities associated with the process that governs parameters other than the variance.

Table 4: Confidence Set for difference in the Unconditional Mean Across Regimes.

	95% Confidence Set
Case 6	[-9.54;7.36]
Case 8	[-28.31;-18.09]
Case 9	[-30.03;30.03]

Note: These were computed from the outcome of the MCMC procedure.

Table 5: Information Criteria in Sub-Sample

	1975.01 - 2004.12			1975.01 - 2007.12		
Case	AIC	HQC	BIC	AIC	HQC	BIC
1	-3.098	-3.055	-2.990	-3.001	-2.961	-2.900
2	-2.999	-2.969	-2.923	-2.304	-2.276	-2.234
3	-3.066	-3.032	-2.980	-2.925	-2.893	-2.844
4	-3.104	-3.073	-3.028	-3.013	-2.985	-2.942
5	-2.970	-2.931	-2.872	-2.942	-2.906	-2.851
6	-3.100	-3.066	-3.013	-3.009	-2.977	-2.928
7	-3.094	-3.055	-2.997	-3.005	-2.969	-2.914

Note: See the text for the description of the cases.

Table 6: Parameter Estimates Case 4, Different Samples.

	Full Sample	1975.01 to 2004.12	1975.01 to 2007.12
c	0.021	0.084	0.018
ϕ_1	1.345	1.313	1.324
ϕ_2	-0.349	-0.335	-0.328
$\sigma^2_{S_t=2}$	0.002	0.002	0.002
$\sigma^2_{S_t=1}$	0.012	0.009	0.009
$p_{1,1}$	0.955	0.967	0.966
$p_{2,1}$	0.818	0.804	0.860

Note: See Table 2

17

Table 7: Information Criteria, Log-Difference of Copper Price.

Case	AIC	HQC	BIC
1	6.289	6.328	6.386
2	6.449	6.475	6.516
3	6.396	6.427	6.473
4	6.278	6.305	6.346
5	6.395	6.430	6.482
6	6.282	6.312	6.359
7	6.286	6.321	6.373

Table 8: Root-Mean-Squared Forecast Error and Tests

| | | | | Test Statistic vs. Case 4 | | | |
| Months | | | | AR | | GARCH | |
Ahead	AR	GARCH	Case 4	DM	HLN	DM	HLN
1	0.099	0.100	0.098	2.00^{**}	1.98^{*}	1.50	1.49
2	0.182	0.182	0.178	2.01^{**}	1.95^{*}	1.79^{*}	1.74^{*}
3	0.257	0.257	0.250	2.39^{**}	2.27^{**}	1.75^{*}	1.66^{*}
6	0.425	0.424	0.412	2.61^{**}	2.29^{**}	1.16	1.02
12	0.465	0.458	0.445	2.06^{**}	1.45	0.53	0.37

Note: The column DM reports the Diebold and Mariano (1995) test statistic of the null that both models have the same Root-Mean-Squared Forecast Error, while the column HLM reports the modification to the DM test suggested by Harvey et al. (1997). ** denotes rejection at 5% significance level and * at 10%.

Table 9: Forecast Coverage of a 90% Confidence Interval and Tests

| | | | | Test P-val vs. Case 4 | |
| Months | | | | | |
Ahead	AR	GARCH	Case 4	AR	CARCH
1	50.0	36.1	51.4	0.49	0.35
2	64.8	60.6	64.8	0.50	0.47
3	68.6	72.9	80.0	0.43	0.46
6	85.1	85.1	91.0	0.47	0.47
12	95.1	96.7	95.1	0.50	0.49

Note: The second to fourth column denote the coverage (in percentage) of the simulated forecast confidence bands of 90%. The last two columns show the p-value of the Giacomini and White (2006) test of predictive ability, using a cuadratic coverage accuracy loss function, of the null that both models provide the same coverage.

Table 10: Forecast Error Variance Decomposition

Months	AR		GARCH		Case 4		
Ahead	Param.	Shock	Param.	Shock	Param.	State	Shock
1	2.6	97.4	5.2	94.8	0.04	3.2	96.7
2	4.4	95.6	8.7	91.3	0.04	4.6	95.4
3	5.6	94.4	11.2	88.8	0.04	5.5	94.4
6	7.9	92.1	16.9	83.1	0.04	7.9	92.1
12	12.2	87.8	21.1	78.9	0.04	12.6	87.4

Note: Each entrance is the percentage of the forecast error variance due to each of the possible sources of uncertainty.

Figure 1: Copper Price (in logs)
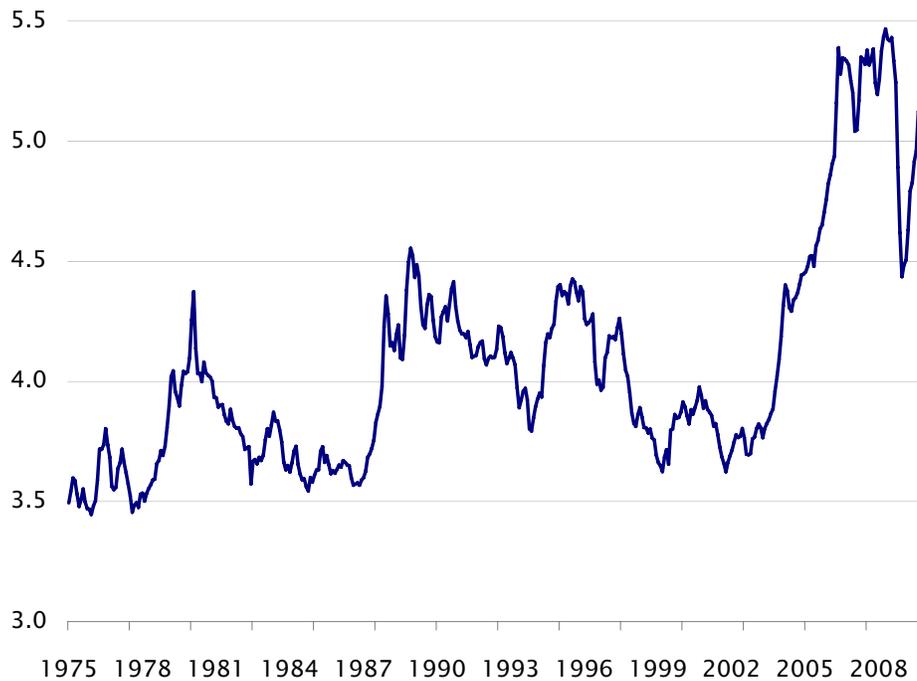
Note: The series is the log of the monthly spot price of copper (in dollars) at the London Market, from January 1975 to January 2010. The source is the International Financial Statistics database from the IMF.

Figure 2: Smoothed (two-sided) Probabilities of the Low-Variance State
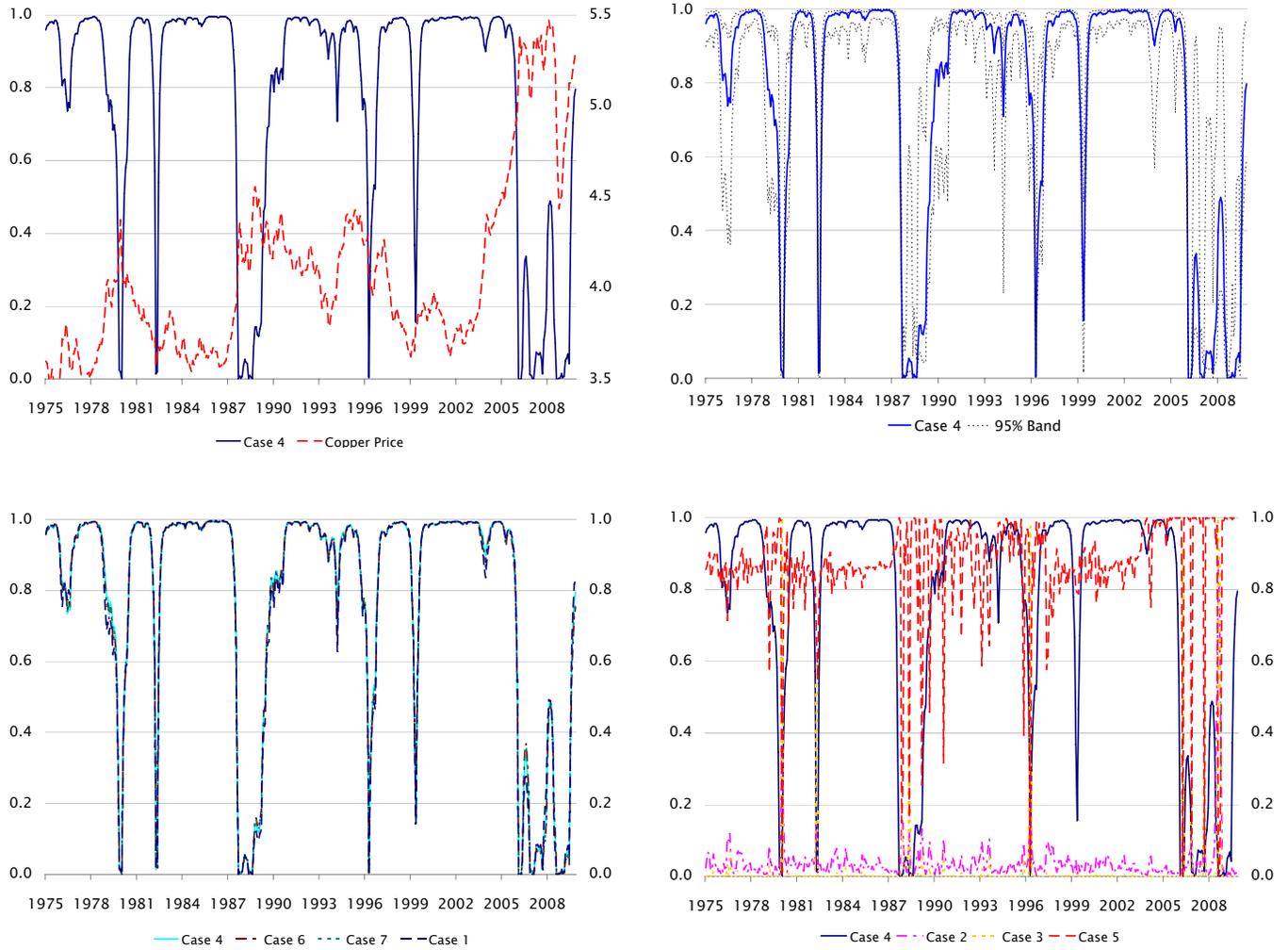
Figure 3: Smoothed (two-sided) Probabilities

Constant and Lags

Variance
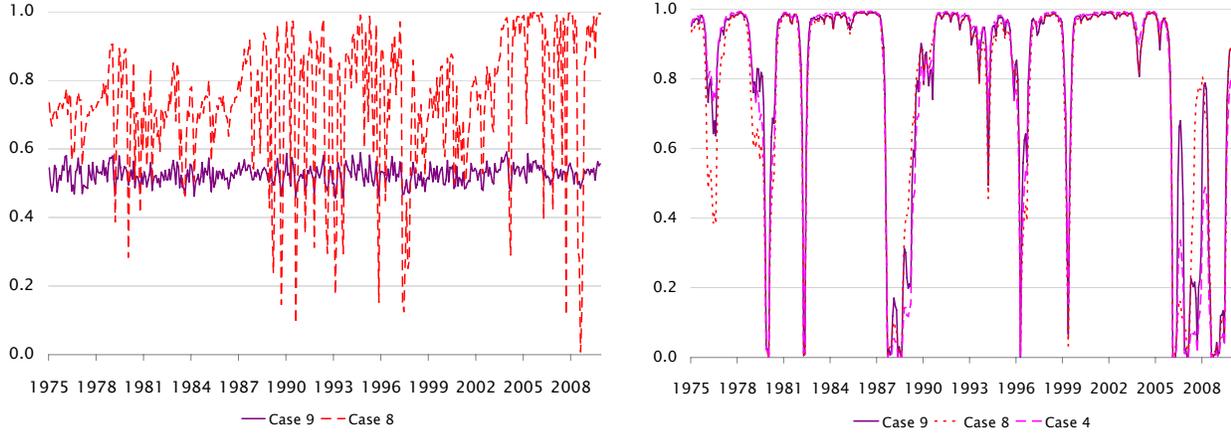
Figure 4: Smoothed (two-sided) Probabilities of the Low-Variance State, Case 4, Different Samples
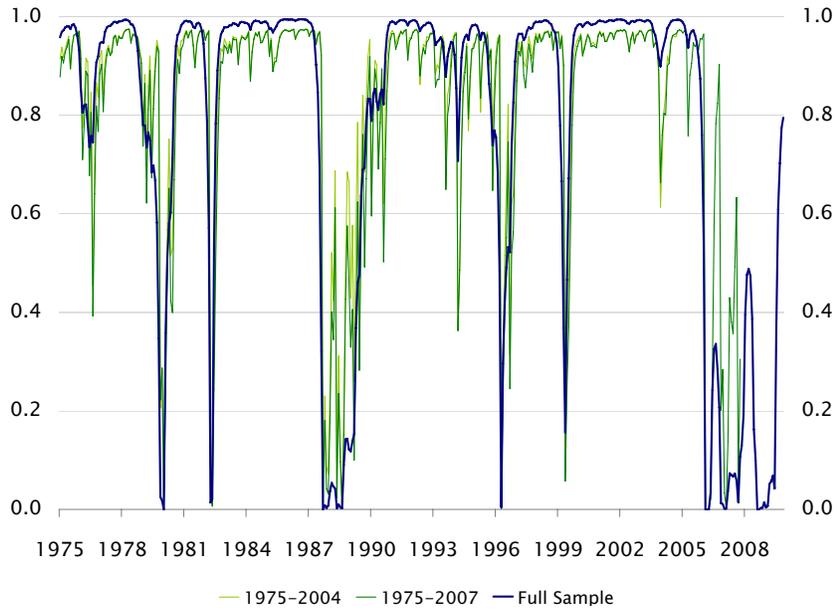
Figure 5: Smoothed (two-sided) Probabilities of the Low-Variance State, Log-Difference of Copper Price
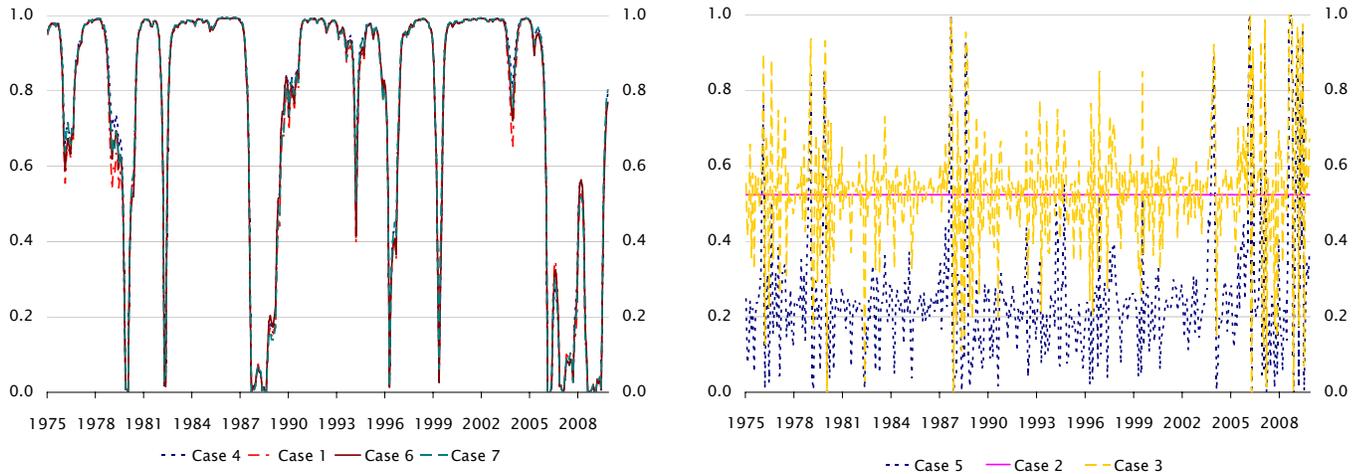
Figure 6: Forecast 90% Confidence Bands Examples

Low Variance State

High Variance State
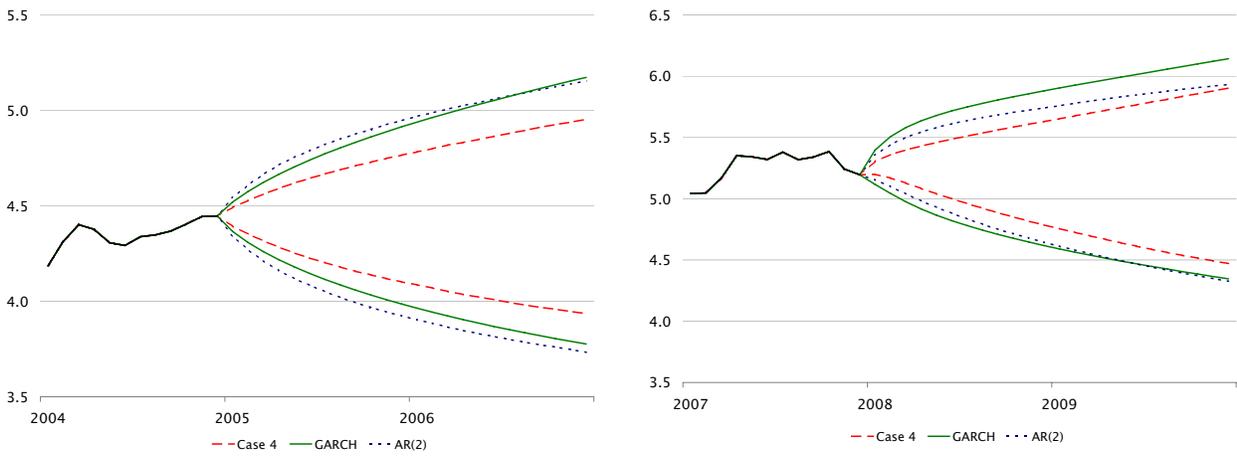
Documentos de Trabajo
Banco Central de Chile

Working Papers
Central Bank of Chile

NÚMEROS ANTERIORES

PAST ISSUES

DTBC – 606 Enero 2011
College Risk and Return
Gonzalo Castex

DTBC – 605
Determinants of Export Diversification Around The World: 1962 – 2000 Enero 2011
Manuel R. Agosin, Roberto Álvarez y Claudio Bravo-Ortega

DTBC – 604 Enero 2011
A Solution to Fiscal Procyclicality: the Structural Budget Institutions Pioneered by Chile
Jeffrey Frankel

DTBC – 603 Diciembre 2010
Eficiencia Bancaria en Chile: un Enfoque de Frontera de Beneficios
José Luis Carreño, Gino Loyola y Yolanda Portilla

DTBC – 602 Diciembre 2010
Chile's Structural Fiscal Surplus Rule: A Model – Based Evaluation
Michael Kumhof y Douglas Laxton

DTBC – 601 Diciembre 2010
Price Level Targeting and Inflation Targeting: a Review
Sofía Bauducco y Rodrigo Caputo

DTBC – 600 Noviembre 2010
Vulnerability, Crisis and Debt Maturity: Do IMF Interventions Shorten the Length of Borrowing?
Diego Saravia

DTBC – 599 Noviembre 2010
Is Previous Export Experience Important for New Exports?
Roberto Álvarez, Hasan Faruq y Ricardo A. López

DTBC – 598 Noviembre 2010
Accounting for Changes in College Attendance Profile: A Quantitative Life-cycle Analysis
Gonzalo Castex