BANCO CENTRAL DE RESERVA DEL PERÚ

Constructing a real-time coincident recession index: an application to the Peruvian economy

Liu Mendoza Pérez*   Daniel Morales Vásquez**

* Universidad Peruana de Ciencias Aplicadas
** Rimac Seguros

DT. N° 2012-020
Serie de Documentos de Trabajo
Working Paper series
Noviembre 2012

Constructing a real-time coincident recession index: an application to the Peruvian economy *

Liu Mendoza Pérez† **Daniel Morales Vásquez‡**

November, 2012

Abstract

In everyday macroeconomic analysis, businessmen and policymakers monitor many variables in order to assess the current situation of a country's business cycle. However, making this assessment is extremely difficult, especially on the verge of recessions: does a drop in one or more of these series reveal the beginning of a recession? Or is it a signal of a temporal deceleration? To answer these questions we have constructed a monthly probabilistic coincident index to detect how close we are of a recession in the Peruvian economy using a non-linear Markov-switching model. In the construction of this index, we have explored the informational content of tendency surveys and international economic variables. We find that the index detected with promptness and reliability the recent recession period associated with the international financial crisis even in real-time analysis. However, since it has been developed with information comprising eight years due to limited data availability, its future recession detection capability has yet to endure the test of time.

JEL Classification : C32, E32, E44.
Keywords : business cycles, Markov-switching models, recession index, business tendency surveys.

* We want to thank Diego Winkelried, César Carrera, Leonidas Cuenca, Javier Kapsoli, and Freddy Espino for their helpful comments. Especial thanks to Henry Pérez for his excellent computational assistance. Any remaining errors are the authors' sole responsibility.
† Mendoza: Universidad Peruana de Ciencias Aplicadas. Email: pceflmen@upc.edu.pe.
‡ Morales: Rimac Seguros. Email: dmoralesv@rimac.com.pe, daniel.morales@pucp.edu.pe.

1. Introduction

One key issue for policymakers and businessmen is to determine the onset of a recession or a recovery phase. To the former, this knowledge is of assistance in making prompt decisions on the implementation or withdrawal of measures to boost economic activity. To the latter, it provides guidance in anticipating market movements and allows the preventive adjustment of business strategies.

But the acquisition of this knowledge on time is a difficult task, particularly at the beginning of a recession or a recovery phase. First, because these phases are not directly observable phenomena and, therefore, they should be inferred from a set of economic variables. And second, because some of the most informative variables on these phases are published with delay, being GDP the most prominent case.

To overcome these issues, many methods have been developed to detect in a timely and systematic manner the beginning of such phases, a process that is called turning point (TP) detection. One of the most used methods is that of leading indexes[1], which estimate the most recent values of a reference series that reflects economic activity (usually GDP) from a series of economic variables that lead its evolution. With this information, it is possible to grasp an idea regarding the beginning of recessions or recoveries.

However, if one of these indexes shows that economic activity is declining, can we state that a recession phase has just started or this is just a temporary deceleration? No direct answer can be given to these questions based on leading indexes, since they answer the question "what is or will be the level or growth rate of economic activity?" and not the question "have we entered a recession or has a recovery phase just begun?". From this example, it is clear that leading indexes provide quantitative, not qualitative information about the business cycle, the latter being perhaps even more important for decision-making of policy-makers and businessmen alike.

In order to complement everyday business cycle analysis, probability-based models have been used to monitor recessions. A first line of research on this matter started with Neftçi (1982), who generates binary variables which reflect the states of recession and stable growth from a set of variables highly correlated with GDP. This methodology was used by Moron et al. (2002) and Ochoa and Lladó (2002), who managed to identify the probability of occurrence of TP for the Peruvian economy. Although this methodology represents an important step forward towards obtaining qualitative information on business cycles, its TP detection power is limited, as suggested by Anas and Ferrara (2002).

That is why a second line of probability-based models was developed with the use of Markov-switching (MS) models initially proposed by Hamilton (1989). These models assume that the economy operates within different regimes or states that change the relationships among variables; even though these regimes are non-observable, they can be inferred via probabilities from the data. Therefore, when applied to business cycle analysis, these models allow us to estimate the probability of a state of recession or recovery in each moment in time. Clearly, this represents a sign of qualitative information to detect recessions.

[1] See Mendoza and Morales (2011) for a discussion on the different types of leading indexes developed so far, and their applications to the Peruvian economy in previous studies.

Based on MS models, Mendoza and Morales (2011) make the first attempt to elaborate a real-time coincident recession index for the Peruvian economy. In this effort, the authors state the need to incorporate variables extracted from business and consumer tendency surveys to elaborate recession indexes in line with Bardaji et al. (2009). These variables constitute a potentially valuable source of information to detect recessions with promptness and reliability since they are published in a timely manner and are not subject to backward revisions. Moreover, the authors state the importance of constructing a recession index with a minimum optimal number of variables, selecting only those which give few “false alarms” in recession detection. To implement this, the authors deploy the backward elimination process proposed by Bardaji et al. (2009). Finally, in line with Choy and Chow (2009), the authors incorporate real and financial international variables, which are key to constructing recession indexes in the case of small open economies such as the Peruvian economy.

Although the index proposed by Mendoza and Morales (2011) shows great reliability in identifying recessions, it was elaborated by aggregating recession probabilities from different variables through averaging. This process of aggregation is suitable for numeric series, but not for probabilities which are a different mathematical entity. Consequently, it is necessary to find a new way to aggregate information depicted in individual series while keeping consistency with the theory on probabilities.

Hence, this paper seeks to construct a real-time coincident recession index for monthly surveillance of recessions that can be used in everyday macroeconomic analysis. In line with Mendoza and Morales (2011), we will aim at exploiting the informational content of business and consumer tendency surveys, as well as real and financial international variables. Although these variables contain potentially valuable information for business cycle analysis, they have been underexploited in previous studies applied to the Peruvian economy[2]. But in contrast to Mendoza and Morales (2011), we will adopt a new methodology to aggregate the informational content of individual series respecting the theory on probabilities. Additionally, we will apply Bardaji et al. (2009)’s methodology to determine the minimum optimal number of variables to elaborate our recession index. Finally, our recession index will undergo tests in order to assess the reliability of its recession detection capability in real time.

We find that indexes elaborated with combinations of very informative series have a better recession detection capability with respect to univariate indexes. This is due to the fact that data aggregation not only reflects the notion of co-movements of series within business cycles, but also allows the “smoothing” of idiosyncratic shocks that affect each series, reducing “false alarms”. Moreover, we verify that the insertion of international variables as well as variables extracted from tendency surveys substantially improves the recession detection capability of our probabilistic index. The final recession index is composed of three variables: manufacturing and trade sales in the United Sates, the sector confidence

[2] Although some studies applied to the Peruvian economy have used information extracted from business or consumer tendency surveys, its usage has not been thorough since they considered only the aggregate confidence indexes (Escobal and Torres (2002), and Ochoa and Lladó (2002)). As suggested by Bardaji et al. (2009), the many questions asked in surveys could convey divergent information on the business cycle. Therefore, it is necessary to test the usefulness of all the survey components. In this line, Etter and Graff (2011) find that business surveys conducted monthly by the central bank of Peru contain important information which leads the evolution of annual GDP growth. See Carrera (2012) for an accurate description of these surveys.

index, and the G7 leading index. This index detected with promptness and reliability the recent recession period experienced by the Peruvian economy arising from the international financial crisis, even in real-time analysis. Furthermore, the index comprises variables of different nature, that is, international and tendency survey variables. This heterogeneity may enhance the recession detection capability of our index going forward. The manufacturing and trade sales in the United States and the G7 leading index, for example, may capture recessions caused by external shocks. And the sector confidence index may also capture recessions caused by international crises, as well as recessions caused by political crises and even climatic factors since it contains information from fishing and agriculture companies. Therefore, the final index represents a potentially useful tool for policymakers and businessmen decision-making regarding TP detection. However, since the index was constructed with information comprising 8 years due to limited data availability, its future recession detection capability has yet to endure the test of time.

The structure of this paper is as follows. Section 2 deals with the stylized facts of the Peruvian business cycle in the period 2000-2011. Section 3 presents the econometric methodology and the main results. Finally, section 4 shows the main conclusions.

2. Stylized Facts

With the aim of reversing the economic recession that started in 2001 after the “dot-com crisis” and the 9/11 attack on the Twin Towers, monetary authorities in the US reduced their main reference rate to levels close to historical lows. The resulting abundance of liquidity not only reflected an increase in credit supply to the public which enabled the recovery of the real economy, but also succeeded in raising speculative demand for commodities whose prices soared (Taylor, 2009) and led to larger FDI inflows to emerging markets (Skovgaard and Clyde, 2011).

Figure 1: Peruvian GDP
(YoY % change)

Source: Peruvian Central Bank.

Thus, in this context of high commodity prices (especially copper, zinc and gold, the main export products of Peru), strong FDI inflows and a recovering world economy, the Peruvian economy started a process of accelerating and sustained growth. Peruvian GDP, which had barely grown 0.2% in 2001, grew 6.8% in 2005 and 8.9% in 2007 (see Figure 1). This process continued up to mid-2008, the year in which the international financial crisis unfolded amid falling commodity prices, suspensions of local and foreign investment decisions, and short-term capital outflows (Dancourt and Mendoza, 2009).

However, unlike the 1998-1999 crisis, there was not a sizeable increase in the exchange rate due to the significant amount of international reserves accumulated during the period of high commodity prices between 2003 and 2008 (Dancourt and Mendoza, 2009). The minor exchange rate depreciation resulted in a diminished impact of the balance sheet effect on the disposable income of families and on the profits of companies; given the fact that financial dollarization reaches 50%. Furthermore, the existence of fiscal savings and a low public debt to GDP ratio (close to 20%) enabled the government to implement an economic stimulus plan which accelerated the Peruvian economy's recovery (Dancourt and Mendoza, 2009). Consequently, after growing -1.2% in the second quarter of 2009, the Peruvian economy moved on to growth rates of 3.4% in the fourth quarter of 2009 and 10% in the second quarter of 2010. This strong recovery was also reflected in job creation, which moved from a 1.2% growth rate in 2009 to 5% in 2010, and in the subsequent acceleration of household consumption which moved from 2.6% growth in 2009 to 6% growth in 2010.

In 2010, the Peruvian economy grew at an elevated pace of 8.8%, due to: (i) the inventory accumulation effect after the inventory depletion generated by the international financial crisis; (ii) the sharp recovery of private investment and consumption; (iii) and the strong economic impulse arising from the government's economic stimulus plan. Similar growth rates were observed up to early 2011, when economic deceleration signs were spotted after a sharp fiscal expenditure adjustment implemented by the government and a rise in political uncertainty generated by the period of presidential elections. Both factors led to a deceleration in the Peruvian economy towards growth rates around 5% in late 2011 (APOYO Consultoría, 2012).

We can see, then, that during the 2002-2011 period, the Peruvian economy went through one recession period (2008-2009) and one deceleration period (2010-2011). Furthermore, this recession period showed a different duration with respect to period of positive economic growth. This suggests that the occurrence of these phases should be modeled as a non-linear phenomenon. Therefore, it is necessary to implement a non-linear methodology which allows to capture the non-linear nature of business cycles in general and, in Peru, in line with the non-linear nature of the Peruvian business cycle found by Rodriguez (2010).

3. Econometric methodology and main results

To elaborate the recession index (RI), we will initially consider variables that previous work has found to be highly correlated to GDP. Additionally, we will propose a set of variables extracted from business and consumer tendency surveys, as well as a set of real and financial international variables that could be potentially useful to construct our RI. As pointed out in Carrera (2012), central tendency indicators of these series reflect reasonably well the distribution of answers in these surveys. The series proposed will be evaluated in the period January 2004 to September 2011, given the limited availability of variables extracted from business and consumer tendency surveys. A more detailed description of all the series to be considered with their respective starting dates, units and sources can be found in Table 1 in Annex 1 of this paper.

3.1. Selection of initial candidates

Initially, our database is composed of 71 variables. In order to reduce it, we will conduct an exploratory analysis calculating three indicators which will allow us to assess the relationship of these initial candidates with the business cycle, which will be proxied by Peruvian GDP[3]. In particular, we will seek to find leading and coincident variables to construct our RI , which means that lagging variables will be discarded in this first step. To make this objective operational, we calculated the following three statistics applied to the cyclical component of variables estimated using the asymmetric Christiano and Fitzgerald filter following Mendoza and Morales (2011)[4]:

(a) Spectral coherence[5], which will allow us to assess if each of the proposed series relates to GDP at different frequencies. We will select series which show a spectral coherence higher than 0.15 for a 1.5-to-8 year period, which suggests that the analyzed series contains relevant information on the cyclical component of GDP.

(b) Mean delay[6], which will allow us to measure the degree of lead and lag in the movements of each series with respect to GDP and thus evaluate their predictive power on the business cycle. Series with a positive mean delay greater that 0.5 will be considered leading variables. Series with a negative mean delay lower than 0.5 will be considered lagging variables.

(c) Dynamic cross-correlations with respect to GDP, which will allow us to evaluate the lead or lag of analyzed variables' cyclical component with respect to GDP's cyclical component, thus complementing the information gathered in the two previous steps. We will evaluate the maximum cross-correlation in a period of +/-8 months. Series with a maximum correlation in absolute value lower than 0.5 will be discarded due to a low correlation with the economic cycle. Additionally, series with a maximum correlation in absolute value greater than 0.5, but depicted in the +8 months horizon, will be discarded for being lagging variables.

The results of these procedures can be found in Table 2 in the Annex 1. In general, variables which have a spectral coherence greater than 0.15 also show a maximum cross-correlation in absolute value greater than 0.5. Additionally, variables classified as lagging for having their maximum cross-correlation in absolute value in the +8 months horizon also showed a negative mean delay. As a consequence, they were discarded from the analysis. Thus, after the elimination process of series with poor informational content and lagging series, the database was reduced from an initial number of 71 variables to 39 candidates. It is worth noting that the majority of variables extracted from business tendency surveys and international variables survived this elimination process with very

[3] Previously, some series were transformed to logs and were seasonally and labor-day adjusted if necessary. The details of these procedures can be found in Table 1 in Annex 1.
[4] We used this filter to extract the cyclical component of GDP since, unlike the Baxter and King or the symmetric Christiano and Fitzgerald, we do not lose information at the beginning and in the end of the sample. This is particularly important for our work due to the fact that the analyzed period of our sample is relatively short.
[5] See Annex 2 for further details.
[6] Ídem.

high cross-correlations and mean delays, which make them potential candidates to construct our final RI . Additionally, although the mean delay and cross-correlation criteria give us an idea on the temporal relationship between the analyzed series and the business cycle, they do not provide exact information on the leading horizon. Therefore, we will apply an additional methodology to determine this information for each of the remaining variables.

3.2. The univariate Markov-switching model

The next step within our methodology is the application of the MS model to each of the 39 candidates. But before proceeding with this estimation, it is convenient to conduct a theoretical revision of these models.

MS models assume that the analyzed variables depict different data generating processes which, in turn, depend on changing regimes in which the economy evolves. These changing regimes are reflected in a discrete variable, s_t . Although this variable cannot be observed directly, it can be inferred via probabilities from the data. In particular, the variable s_t will take the values 1, 2 or 3, ($s_t \in (1,2,3)$) for domestic data series, which represent the states of recession, recovery and expansion, respectively[7]. For international data series, s_t will take the values 1 and 2, ($s_t \in (1,2)$), in line with the findings of the international applied literature reviewed in the introductory part of this document.

The MS model for a variable y_t is as follows: $y_t = m(s_t) + e_t$. That is, y_t is modeled as a function of its mean, $m(s_t)$, which is not constant over time, but changes according to the values taken by the variable s_t . This allows us to assess whether the dynamic of y_t fits in the pattern seen during expansions, recoveries or recessions. The variable s_t changes from one regime to another through transition probabilities that follow a first-order Markov chain: $\Pr(s_t = j | s_{t-1} = k, s_{t-2} = l, ...) = \Pr(s_t = j | s_{t-1} = k) = p_{kj}$. That is, the probability to be in regime j in period t depends only on the regime in which the economy was in period $t-1$ and not on preceding periods. Besides, it is assumed that such probability is invariant in time.

Transition probabilities must add up to one: $\sum_{j=1}^{3} p_{kj} = 1; \forall j,k \in \{1,2,3\}$; that is, the probabilities of moving from a regime k towards any other possible regime must add up to one. Additionally, there can be persistent states (with transition probabilities close to unity), but not absorbent ones (equal to unity).
From this explanation, we can observe that MS models characterize the main features of the theory on business cycles: (i) non-observability of the business cycle, and (ii) non-linearity in the evolution of the business cycle. That is to say, s_t depicts the changing but unobservable behavior of the data, which behaves differently during recessions, recoveries and expansions.

---

[7] This in line with the results of Rodríguez (2010), who finds that the best specification of a MS model for the Peruvian economy is a three-regime model.

In the case of MS models, the non-observable variable s_t can be inferred from the data through the iterative algorithm of Expectation Maximization[8], which allows us to obtain the probabilities of being in each regime at every point in time. When all the data available in the sample is used, we obtain smoothed probabilities represented by SP_{s_t}. When only the data available in each moment in time is used, we obtain filtered probabilities, represented by FP_{s_t}. Given the fact that in this study we are interested in obtaining recession probabilities, SP_t and FP_t will refer to smoothed and filtered probabilities of recessions from now on.

Since the proposed model does not include lags of the endogenous variables, it may be affected by error autocorrelation which could potentially affect the TP detection capability of the model. However, as suggested by Bellone and Saint-Martin (2003), the effect of the autoregressive parameters will be largely captured by transition probabilities forming the Markov chain[9].

3.3. Application of the univariate MS model to the candidate series

As an initial step, in order to express variables of diverse nature (quantitative, surveys, rates, etc.) in the same units, we proceed to calculate the standardized growth rates for each of the 39 variables selected in the process described in 3.1[10]. Then, we will estimate a MS model for each of these transformed variables. In particular, the estimated model applied to each of the candidates is: $y^*_{i,t} = m(s_{i,t}) + e_{i,t}$, where y^* refers to standardized growth rates and the subscript i refers to the i-eth variable analyzed in each model from the total number of candidates.

We will compare the smoothed probabilities of each of the candidate series (SP_t^i) with a dichotomous variable which will serve as a reference in order to evaluate its predictive power. Unlike the United States, where an official committee determines the beginning and the end of recessions (National Bureau of Economic Research, NBER), we will use the technical recession criterion applied to Peruvian GDP. That is, the dichotomous variable will take the value of 1 indicating recession when seasonally adjusted Peruvian GDP has accumulated six months of negative monthly growth rates, and the value 0 in any other case.

To dismiss series with low predictive power of recessions, we will calculate the Quadratic Probabilistic Score (QPS) proposed by Brier (1950) associated with each variable i: $QPS(i,R) = \frac{1}{T}\sum_{t=1}^{T}\left(SP_t^i - R_t\right)^2$, where SP_t^i corresponds to smoothed probabilities of being in recession obtained from the i-eth series, and $R_t = \{1,0\}$ is the dichotomous variable which indicates recession according to our reference criterion. Following Bellone and

[8] See Krolzig (1997).

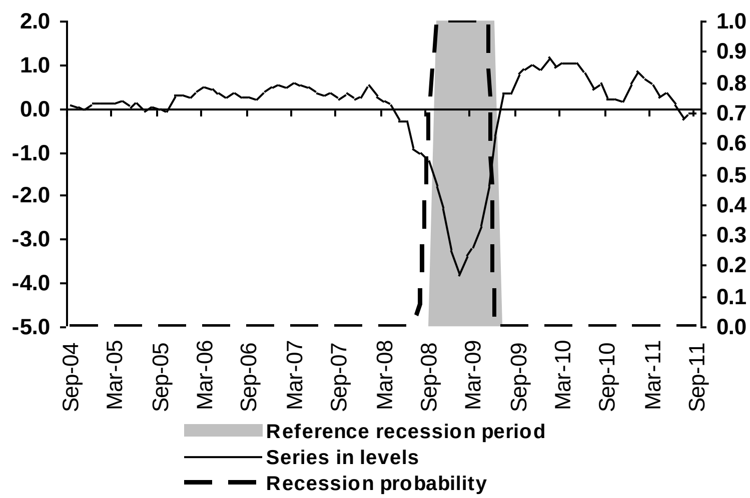
[9] Harding and Pagan (2003) show a simple linear approximation of Hamilton (1989)'s models which clearly depicts the "auto-correlation" effect in MS models due to the presence of Markov chains.

[10] In the construction of these standardized series, we take into account different degrees of smoothing. The standardization process and the selection of the optimal degree of smoothing are described in the Annex 2.

Saint-Martin (2003), the series with a QPS greater than 0.15 will be discarded from the analysis[11].

The results are shown in Table 3 in the Annex 1[12], where we observe that the Bellone and Saint-Martin (2003)'s criterion allows us to discard 14 variables. Moreover, we can observe that the variable with the lowest QPS is the leading index of the six major Asian economies whose level and recession probabilities are shown in Figure 2. We observe that the SP_t of this variable detected clearly the onset of the recession period of the Peruvian economy associated with the international financial crisis with a lead of one month (September 2008). Furthermore, their SP_t went below the threshold that defines recessions (0.5) in June 2009, the same month that our reference recession series indicates the end of the recession in Peru. With that, we can conclude that this indicator marked successfully the chronology of recessions in the analyzed period, indicating with a certain lead the beginning of the recessionary phase of the Peruvian economy.

Figure 2: Recession probabilities estimated from the leading index of the six major Asian economies



Source: own calculations.

3.4. Construction of the recession probabilistic index

Although useful, detecting recessions from univariate MS models does not capture the co-movement between variables that is implicit in the concept of business cycle. A strategy to overcome this issue is to combine different variables into an aggregate RI in order to optimize the recession signals extracted from individual series. So, we follow Stock and Watson (2003) and Bellone and Saint-Martin (2003), who found that indexes arising from the aggregation of multiple variables have a better recession detection capability than univariate indicators. According to the authors, this is a consequence of possibly unstable

[11] Note that the QPS criterion allows us to find the exact number of months which a variable leads or lag the GDP. We dismissed the lagging variables detected this way. See Annex 2 for further details.

[12] We verified that using standardized percentage variations yields the same results (recession probabilities) compared to using simple percentage variations.

and noisy signals sent from individual variables that can be reduced when combined with other series.

3.3.1 Aggregation of variables in a single recession indicator

Up to now, a MS model has been estimated for each candidate variable and their recession detection capability has been evaluated individually, obtaining as a result a reduced subset of candidates. However, our aim is to obtain a single probabilistic recession indicator, RI. As is pointed out by Carnot and Tisson (2002), the absence of theoretical arguments for combining the obtained probabilities might be a disadvantage for the adoption of MS models for detecting recessions.

A model which naturally adapts to our purpose to generate a single recession indicator from a set of variables is the Markov-switching dynamic factor model (Chauvet 1998). This model assumes that the analyzed variables share an autoregressive common factor which changes from one regime to another following a Markov chain, and thus the probability of being in a recession state can be obtained from this set of variables. However, as stated in Bellone and Saint-Martin (2003), dynamic factor models have some serious convergence problems, which is a consequence of the huge amount of parameters to be estimated even in the most parsimonious specifications.

Another approach is to aggregate the smoothed probabilities calculated in each univariate MS model using simple and weighted averages[13], and then take the final result as a RI. This is known as aggregation in «probability space» (Bellone and Saint-Martin, 2003). However, this aggregation method, used in Mendoza and Morales (2011), does not take into account the fact that probabilities, being different mathematical entity than numbers, cannot be aggregated using simple or weighted averages. In turn, different methods should be used for this purpose, such as the Fisher average or the Bonferroni inequality (Padgett and Tomlinson, 2003). However, in both cases the obtained aggregate probabilities tended to be higher than the initial probabilities, a counterintuitive result for the purpose of this study.

As a consequence, a third strategy will be applied, in which the series are first aggregated through a simple average into one series[14], and then a MS model is estimated from it. The resulting smoothed probabilities will be taken as a RI. This process is known as aggregation in "temporal space" (Bellone and Saint-Martin, 2003). Regarding the timing of the aggregation of each variable into the index, two possibilities arise: to take into account either the leading degree or the publication delay of each variable. Specifically, suppose that we have a 10-month leading variable which is published with a two-month delay; then, we have the options to calculate the RI on period t using that variable in $t-10$ or in $t-2$. Even though aggregating series based on the publication delay implies penalizing leading variables with a lower QPS, this choice would allow an earlier detection of recessions. For

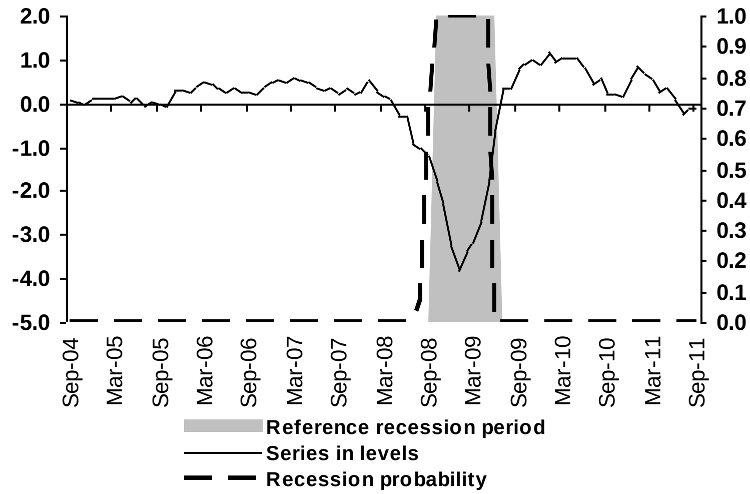
[13] The weights to be used can be based on the QPS of each individual series as follows:

$$w_i = \frac{\dfrac{1}{QPS(i,R)}}{\displaystyle\sum_{k=1}^{n} \frac{1}{QPS(k,R)}}$$

[14] As long as the variables are expressed in standardized growth rates, they can be aggregated this way.

instance, if the 10-month leading variable is incorporated in the index with a lead of 10 months, the recession alarm will arrive late. However, if we incorporate using the last available observation, the recession alarm would be triggered eight months earlier. For this reason, the second aggregation strategy will be used.

Figure 3: Recession probabilities of the aggregate RI



Source: own calculations.

However, as shown in Figure 3, an index constructed from the aggregation of the best 25 variables altogether is not significantly superior in recession detection capacity than the RI of the best individual variable (leading index of the six major Asian economies). For this reason, a valid question may arise: if from this pool of 25 variables we choose only the most informative ones, may the recession detection capacity of the RI be enhanced?

3.3.2 Choosing the optimal set of variables for the recession index

Although the previous RI does not exceed the recession detection capacity of the best variable alone, aggregation should not be discarded yet. It is necessary to consider that this aggregate RI incorporates some variables that may add little information to determine the recession state, others that may add noisy instead of valuable information, or even others that may add redundant information which increases the number of variables for the construction of the RI and thus needlessly complicates the subsequent updating process. For this reason, it is important to set out a strategy aimed at eliminating the less informative variables of the index.

For this purpose, a method based on the iterative process of Bardaji et al. (2009) is proposed. First, n recession indicators will be estimated which includes the n candidates less the $j-th$ one, where $j=1,...,n$; each RI estimated this way is denominated RI_{nj}. Then, the QPS for each RI_{nj} will be calculated, and finally, the variable associated to the RI_{nj} with the highest QPS will be discarded from the pool of candidates, thus eliminating the least informative variable. These three steps will be repeated until the index with the lowest QPS is obtained. Then, we will get the minimum optimal number $n*$ for the construction of the index[15].

Figure 4: QPS obtained in the backward elimination process

Source: own calculations.

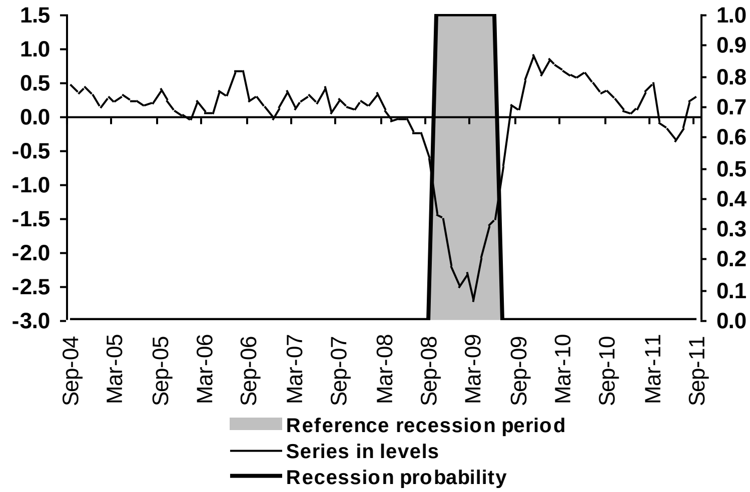
The elimination process is depicted in Figure 4. The Y-axis contains the QPS of each index after the removal of the least informative variable, while the X-axis contains the number of variables of the index. The iterative process allows a progressively better adjustment of the proposed index in each iteration until a minimum QPS is reached. In particular, the adjustment of the index to the reference variable significantly improves until the extraction of the 17th variable. This adjustment remains almost constant until the 3rd variable is removed, when it suddenly worsens.

The three-variable index was chosen as the final RI, as long as it is the easiest to update and has the lowest QPS. The final index is composed of three variables: manufacturing and trade sales in the United States, the sector confidence index and the G7 leading index[16]. It can be concluded that the proposed elimination process allows selecting variables of different nature, in this case, real international variables and survey-based variables. This guarantees a heterogeneous composition of the index, in line with the generalized nature of a recession. Moreover, this feature makes the RI more capable to identify future recessions originated by different sources, such as international crisis, political crisis, or even weather factors which are captured in the sector confidence index which summarizes information from different economic sectors.

[15] Anas and Ferrara (2002) found that the marginal benefit of incorporating more than six variables is null, which is an additional advantage over the use of leading indicators which are usually constructed using between 10 to 20 variables.

[16] The final index is composed of variables with a different number of regimes: while international variables have two regimes, survey-based variables have three. In order to determine the number of regimes of the final index we followed Humala (2005), who proposes the use of goodness-of-fit indicators to obtain the best specification since statistic tests which contrast the number of regimes do not have a standard distribution (Hansen, 1992). With this proposal, we verified that the optimal number of regimes for our final index is three.

Figure 5: Recession probabilities of the optimal RI

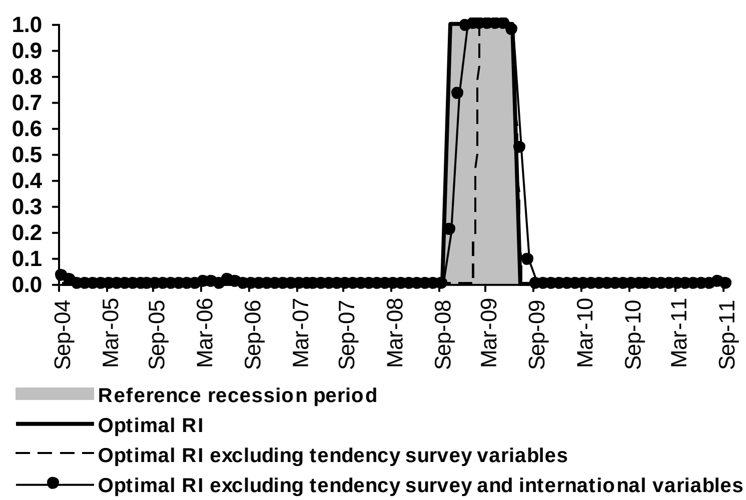


Source: own calculations.

The recession probabilities extracted from the optimal RI are shown on Figure 5. It can be observed that the RI performs well detecting the 2008 recession, provided that it crossed the recession threshold of 0.5 just a month after the bankruptcy of Lehman Brothers, event which marked the beginning of the international crisis. It is worth mentioning that the series in levels drops at the end of the sample as a consequence of an economic deceleration caused by a reduction in public expenditure and political uncertainty. However, this drop did not translate into an increase of the recession probability on that period, so the indicator did not trigger a false alarm considering it as a recession.

Finally, it is worthwhile to evaluate the quality of the information supplied by international and survey-based variables; in other words, whether their inclusion in the analysis enhances the recession detection capacity of the index. For this purpose, the iterative process will take place taking into account only domestic financial and real variables (9), then international real variables (15), and finally the survey-based variables (25).

Figure 6: Recession probabilities of the optimal RI



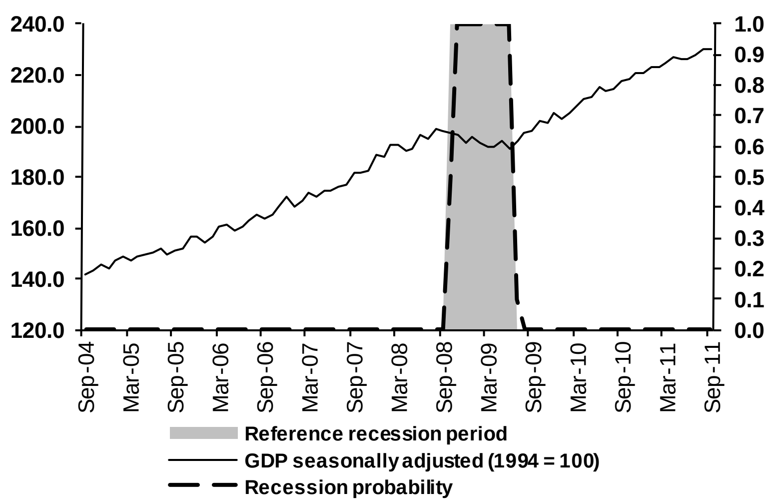
Source: own calculations.

As is shown in Figure 6, the inclusion of international and survey-based variables enhances the detection power of the index. In fact, the QPS of the final optimal index is 0.56 10^{-5}, but the QPS increases to 0.05 when the survey-based variables are excluded from the analysis and to 0.01 when the international real series are excluded. For this reason, we can conclude that these variables contain relevant and reliable information which complements the predictive power of local financial and real series to detect recessions promptly.

3.4 Real-time detection power

It has been showed that the RI composed by the manufacturing and trade sales in the United States, the sector confidence index and the G7 leading index shows the best recession detection capacity from a historical point of view. However, does it have the same detection power in real-time? To answer this question, a real-time test will take place at the beginning of a recessive period; in this case, at the beginning of the international crisis of 2008-2009. The following scenario is set out: if the indicator had been used on the verge of the international crisis, would it have detected the incoming recession in the Peruvian economy with a steady recession signal after subsequent updates? It is worth mentioning first that, as long as the construction of the index was based on the publication delay of each individual variable, a potential problem on real-time detection is solved. A second issue is that the RI must be calculated using only the available information until the evaluation period. To deal with this issue, the RI will be constructed using the filtered probability FP_t, so that we guarantee that the estimation of the recession probabilities takes into account only the information available at each period.

Figure 7: Real-time recession probabilities of the optimal RI



Source: own calculations.

The real time detection test of the optimal recession index shows that the RI crossed the threshold of 0.5 indicating the onset of a recession in October 2008, only one month after the bankruptcy of Lehman Brothers which marked the beginning of the international crisis (see Figure 7). According to the index, the recession lasted until June 2009, when the RI went below the 0.5 threshold and seasonally adjusted GDP began to show positive growth rates. It is worth highlighting the strength of the subsequent recession signals, which reduces the probability of getting false alarms. The stability of the recession signal would be related to some features of MS models, which tend to satisfactorily represent very persistent, fat-tailed series, in line with the findings of Bellone and Saint-Martin (2003).

4. Conclusions

Spotting the beginning of a recession or a recovery is useful for policy-makers and businessmen alike. It allows the former to adopt or withdraw economic stimulus measures in a timely fashion. And it allows the latter to adapt their strategies on time to minimize losses when facing sudden falls in demand during recessions, or to increase profits on the verge of an economic recovery.

In that sense, a probabilistic recession index meets entirely the objective described above by conveying qualitative information on the business cycle to complement the quantitative information provided by leading indexes. Therefore, in this paper we aim at constructing a probabilistic index for monthly recession surveillance applied to the Peruvian economy. This index was constructed using a non-linear Markov-switching model.

In the construction of our index, we evaluated the recession detection capability of variables extracted from business and consumer tendency surveys, as well as real and financial international variables. Carrera (2012) points out that survey data in Peru covers a period of relatively low inflation and there has been an increasing interest in explaining the process by which agents form their expectations. We depart and find that these series contain also valuable information to detect recessions. Moreover, we used a methodology to combine the information contained in different variables in order to obtain an optimal recession index.

We find that indexes elaborated with combinations of very informative series have a better recession detection capability when compared to univariate indexes. This is due to the fact that data aggregation not only reflects the notion of co-movements of series within business cycles, but also allows the "smoothing" of idiosyncratic shocks that affect each series, reducing "false alarms". Moreover, we verify that the insertion of international variables as well as variables extracted from tendency surveys substantially improves the recession detection capability of our probabilistic index.

The final recession index is composed of three variables: manufacturing and trade sales in the United Sates, the sector confidence index and the G7 leading index. The final index detected with promptness and reliability the recent recession period experienced by the Peruvian economy associated with the international financial crisis, even in real-time analysis. Furthermore, the index is formed by variables of different nature, that is, international and tendency survey variables. This heterogeneity may enhance the recession detection capability of our index going forward. The manufacturing and trade sale in the United States and the G7 leading index, for example, may capture recessions caused by external shocks. And the sector confidence index may also capture recessions caused by international crises, as well as those caused by political crises and even climatic factors since it contains information from fishing and agriculture companies. Therefore, the final index represents a potentially useful tool for policymakers and businessmen decision-making regarding TP detection. However, since the index was constructed with information comprising eight years due to limited data availability, its future recession detection capability has still to endure the test of time.

REFERENCES

ANAS, Jaques and FERRARA, Laurent (2002) Detecting Cyclical Turning Points: The ABCD Approach and Two Probabilistic Indicators. XXVI CIRET Conference, Taipei, October 2002.

BARCLAYS CAPITAL (2009) A Real Stress Test. Working Paper February 27th 2009.

BARDAJI, Jose, CLAVEL, Laurent and TALLET, Frederic (2009) Constructing a Markov Switching Turning Point Index Using Mixed Frequencies with an Application to French Business Survey Data, pp. 111-132. Journal of Business Cycle Measurement and Analysis No. 2009/2.

BELLONE, Benoit and SAINT-MARTIN, David (2003) Detecting Turning Points with Many Predictors through Hidden Markov Models. Working Paper, Séminaire Fourgeaud, December 2003.

BRIER, Glenn (1950) Verification of Forecasts Expressed in Terms of Probability, pp. 1-3. Monthly Weather Review No. 75.

CAMACHO, Máximo, PEREZ-QUIROZ, Gabriel and PONCELA, Pilar (2010) Green Shoots in the Euro Area: A Real-Time Measure. Madrid: Bank of Spain.

CARRERA, Cesar (2012) Estimating Information Rigidity Using Firms' Survey Data, The B.E. Journal of Macroeconomics: Vol. 12: Iss. 1 (Topics), Article 13.

CARNOT, Nicolas and TISSOT, Bruno (2002) La Prévision Economique, 120-128. Economica.

CHOW, Hwee and CHOY, Kuan (2009) Analyzing and Forecasting Business Cycles in a Small Open Economy: A Dynamic Factor Model for Singapore. Working Paper No. 05-2009. Singapur: Business University of Singapore.

CROUX, Christophe, FORNI, Mario and REICHLIN, Lucrezia (2001) A measure of comovement for Economic Indicators: Theory and Empirics, pp. 430-450. The Review of Economics and Statistics No. 83

DANCOURT, Oscar and Mendoza, Waldo (2009) Perú 2008-2009: del auge a la Recesión, Choque Externo y Respuestas de Política Macroeconómica. Crisis Internacional: Impacto y Respuestas de de Política Económica en el Perú. Fondo Editorial de la Pontifica Universidad Católica del Perú. Lima.

ESCOBAL, Javier and TORRES, Javier (2002) Un Sistema de Indicadores Líderes del Nivel de Actividad para la Economía Peruana. Working Paper No. 39. Lima: GRADE.

ETTER, Richard and GRAFF, Michael (2011) A Composite Leading Indicator for the Peruvian economy based on the BCRP's monthly business tendency surveys. Working Paper No. 6. Lima: BCRP.

HAMILTON, James (1989) A New Approach to the Economic Analysis of Non-stationary Time Series and the Business Cycle, pp. 357-384. Econometrica No. 57.

HARDING, Don and PAGAN, Adrian (2003) A Comparison of Two Business Cycle Dating Methods, pp. 1681-1690. Journal of Economic Dynamics and Control N° 27 (9).

HANSEN, Bruce (1992) The Likelihood Ratio Test Under Non-Standard Conditions: Testing the Markov Switching Model of GNP. Journal of Applied Econometrics N° 7(8), pp 61-82.

HUMALA, Alberto (2005) Modelling interest rate pass-through with endogenous switching regimes in Argentina. Doctoral Thesis. Unpublished.

KAPSOLI, Javier and BENCICH, Brigitte (2002) Indicadores Líderes, Redes Neuronales y Predicción de Corto Plazo. Working Paper No. 213. Lima: CISEPA.

KROLZIG, Hans-Martin (1997) Markov-Switching Vector Autoregressions. Statistical Inference and Application to Business Cycle Analysis. Lecture Notes in Economics and Mathematical Systems No. 454.

MENDOZA, Liu and MORALES, Daniel (2011) Construyendo un Índice Coincidente de Recesión para la Economía Peruana. Centro de Investigación Económico y Social. Available at: http://cies.org.pe/files/documents/investigaciones/politica-macroeconomica-y crecimiento/construyendo_un_indice_coincidente_de_recesion.pdf

MORÓN, Eduardo, CASAS, Carlos and CARRANZA, Eliana (2002) Indicadores Líderes para la Economía Peruana. Working Paper No. 49. Lima: Universidad del Pacífico.

NEFTCI, Salih (1982) Optimal predictions of cyclical downturns, pp. 307-327. Journal of Economic Dynamics and Control No. 4.

PADGETT, W. and TOMLINSON, Meredith (2003) Lower Confidence Bounds for Percentiles of Weibull and Birnbaum-Saunders Distributions, pp. 429-433. Journal of Statistical Computation and Simulation. N° 73(6).

OCHOA, Enrique and LLADÓ, Jorge (2002) Modelos de Indicadores Líderes de Actividad Económica para el Perú. Estudios Económicos No. 10. Lima: BCRP.

RODRIGUEZ, Gabriel (2010) Application of three Non-Linear Econometric Approaches to Identify Business cycles in Peru. Working Paper N° 284. Pontificia Universidad Católica del Perú.

SKOVGAARD, Lauge and CLYDE, Gary (2011) Foreign Direct Investment in Time of Crises. Working Paper 11-3. Peterson Institute of International Economics.

STOCK, James and WATSON, Mark (2003) Forecasting Output and Inflation: The Role of Asset Prices. Working Paper N° 8180. Cambridge: NBER.

TAYLOR, John (2009) The Financial Crisis and the Policy Responses: An Empirical Analysis of What Went Wrong. Working Paper N° 14631. NBER.

ANNEX 1

Table 1: Initial database[17]

	Units	Source	Log.	Seasonal Adj.	Labor-day Adj.
Domestic real variables					
Current account savings of the Central Government	S/. Millions	BCRP	Yes	Yes	Yes
Water consumption in Lima and Callao	Thousands of m^3	INEI	Yes	Yes	Yes
Banking sector credit to the private sector	S/. Millions	BCRP	Yes	Yes	Yes
Total credit/banking system liquidity	%	BCRP	Yes	Yes	Yes
Money	S/. Millions	BCRP	Yes	Yes	Yes
Traditional fishing exports	US$ Millions	BCRP	Yes	Yes	No
Current non-financial expenditure of the Central Government	S/. Millions	BCRP	Yes	Yes	Yes
Capital expenditure of the Central Government	S/. Millions	BCRP	Yes	Yes	Yes
Non-financial expenditure of the Central Government	S/. Millions	BCRP	Yes	Yes	Yes
Imports of intermediate goods	US$ Millions	SUNAT	Yes	Yes	No
Domestic value-dded tax	S/. Millions	SUNAT	Yes	Yes	Yes
Value-added tax	S/. Millions	SUNAT	Yes	Yes	Yes
Machinery and equipment PPI	Index 1994=100	INEI	Yes	Yes	No
Domestic products PPI	Index 1994=100	INEI	Yes	Yes	No
Volume index of intermediate goods production	Index 1994=100	INEI	Yes	Yes	Yes
Volumen index of diverse manufacturing production	Index 1994=100	INEI	Yes	Yes	Yes
Transport and communications CPI	Index 2009=100	INEI	Yes	Yes	No
GDP construction	Index 1994=100	INEI	Yes	Yes	Yes
GDP fishing	Index 1994=100	INEI	Yes	Yes	Yes
Primary GDP	Index 1994=100	INEI	Yes	Yes	Yes
Electricity production	Gigawatts/hour	COES-SINAC	Yes	Yes	Yes
Selective consumption tax	US$ Millions	INEI	Yes	Yes	Yes
Terms of trade	Index 2001=100	BCRP	Yes	Yes	No
Gross value added of chemical, rubber and plastic products	Index 1994=100	INEI	Yes	Yes	Yes
Domestic financial variables					
Net international reserves	US$ Millions	BCRP	Yes	Yes	No
Country risk	Basic points	Bloomberg	No	No	No
Active interest rate in foreign currency up to 360 days	%	SBS	No	No	No
Effective reserve requirement rate in foreign currency	%	BCRP	Yes	Yes	No
Effective reserve requirement rate in domestic currency	%	BCRP	Yes	Yes	No
Interest rate on deposit certificates of the BCRP	%	BCRP	No	No	No
Business tendency survey variables					
Economic activity index	Difussion Index	BCRP	No	No	No
Business confidence index	Difussion Index	BCRP	No	No	No
3-months ahead expectations index	Difussion Index	BCRP	No	No	No
Inventory index	Difussion Index	BCRP	No	No	No
New orders index	Difussion Index	BCRP	No	No	No
Sales index	Difussion Index	BCRP	No	No	No
Consumer tendency survey variables					
Consumer confidence index	Difussion Index	APOYO Consultoría	Yes	No	No
Consumer confidence index - Family component	Difussion Index	APOYO Consultoría	Yes	No	No
Consumer confidence index - Future situation	Difussion Index	APOYO Consultoría	Yes	No	No
Consumer confidence index - Country component	Difussion Index	APOYO Consultoría	Yes	No	No
Consumer confidence index - Prices component	Difussion Index	APOYO Consultoría	Yes	No	No
Consumer confidence index - Current situation	Difussion Index	APOYO Consultoría	Yes	No	No
Consumer confidence index - Jobs component	Difussion Index	APOYO Consultoría	Yes	No	No
Consumer confidence index - Housing component	Difussion Index	APOYO Consultoría	Yes	No	No
International real variables					
Asia: leading index of 5 major economies	Index	OECD	No	No	No
China: leading index	Index 2009=100	OECD	No	No	No
China: retail sales of consumer goods	Yuan Hundreds of Millions	National Bureau of Statistics of China	Yes	Yes	no
US: leading index	Index	OECD	No	No	No
US: consumer confidence index (current situation)	Difussion Index	The Conference Board	Yes	Yes	Yes
US: unemployment rate	%	Bureau of Labor Statistics	Yes	Yes	No
US: manufacturing and trade sales	US$ del 2005	Bureau of Economic Analysis	Yes	Yes	No
Europe: business climate for Germany	2000=100	IFO	No	No	No
Europe: extra-Eurozone exports	Euro Millions	Eurostat	Yes	Yes	No
Europe: leading index (EU 17)	Index	OECD	No	No	No
Europe: new orders in manufacturing	Difussion Index	CEIC	Yes	Yes	No
Europe: industrial production (EU 16, excl. construction)	Index	Eurostat	No	No	No
Europe: industrial production (EU 27, excl. construction)	Index	Eurostat	No	No	No
Europe: economic sentiment index (EU 16)	Index	Eurostat	No	No	No
Europe: economic sentiment index (EU 27)	Index	Eurostat	No	No	No
Europe: retail sales index (EU 16)	Index	Eurostat	Yes	No	No
Europe: retail sales index (EU 27)	Index	Eurostat	Yes	No	No
G7: leading index	Index	OECD	No	No	No
NAFTA: leading index	Index	OECD	No	No	No
OECD: leading index (Europe)	Index	OECD	No	No	No
OECD: leading index (plus 6 major non-member economies)	Index	OECD	No	No	No
OECD: leading index (total)	Index	OECD	No	No	No
International financial variables					
Dow Jones	Index	CEIC	No	No	No
Vix index	Index	CBOE	No	No	No
Country risk - Latam	Basic points	Bloomberg	No	No	No
TED spread	%	CEIC	No	No	No
Yen-dollar exchange rate	Yen per US$	CEIC	No	No	No

[17] Domestics real and financial variables were taken from Escobal and Torres (2002), Kapsoli and Bencich (2002), Morón et al. (2002), and Ochoa and Lladó (2002). International financial variables were taken from Barclays (2009). International real variables were taken from Anas and Ferrara (2002), Bellone and Saint-Martin (2003), and Camacho et al. (2010).

Table 2: Selection of candidate variables

	Spectral coherence	Mean delay	Maximum correlation 2/	Period of maximum correlation 3/
Domestic real variables				
Current account savings of the Central Government	0.24	-0.01	0.49	0
Water consumption in Lima and Callao	0.17	0.32	0.50	-4
Banking sector credit to the private sector	0.11	-0.54	0.52	7
Total credit/banking system liquidity	0.21	-21.74	0.46	10
Money	0.62	-0.11	0.79	0
Traditional fishing exports	0.01	0.81	0.20	-6
Current non-financial expenditure of the Central Government	0.39	0.02	0.63	0
Capital expenditure of the Central Government	0.09	0.46	0.37	-3
Non-financial expenditure of the Central Government	0.33	0.12	0.58	0
Imports of intermediate goods	0.52	-0.08	0.73	1
Domestic value-dded tax	0.31	-0.18	0.58	1
Value-added tax	0.59	-0.21	0.81	2
Machinery and equipment PPI	0.03	21.88	-0.02	-9
Domestic products PPI	0.08	-0.74	0.47	5
Volume index of intermediate goods production	0.69	-0.01	0.83	0
Volumen index of diverse manufacturing production	0.00	0.97	0.20	-8
Transport and communications CPI	0.00	-2.01	0.06	3
GDP construction	0.48	-0.01	0.70	0
GDP fishing	0.01	-22.03	-0.09	4
Primary GDP	0.11	0.14	0.35	-1
Electricity production	0.20	-0.15	0.47	1
Selective consumption tax	0.02	0.82	0.31	-7
Terms of trade	0.21	-0.14	0.47	1
Gross value added of chemical, rubber and plastic products	0.62	0.21	0.85	-2
Domestic financial variables				
Net international reserves	0.53	-0.03	0.73	0
Country risk	0.00	-1.35	0.20	7
Active interest rate in foreign currency up to 360 days	0.07	-21.93	0.09	10
Effective reserve requirement rate in foreign currency	0.08	-0.82	0.56	6
Effective reserve requirement rate in domestic currency	0.12	-0.45	0.45	4
Interest rate on deposit certificates of the BCRP	0.06	-21.50	0.07	9
Business tendency survey variables				
Economic activity index	0.17	0.74	0.77	-6
Business confidence index	0.08	1.30	0.80	-7
3-months ahead expectations index	0.12	1.02	0.83	-6
Inventory index	0.15	0.52	0.57	-5
New orders index	0.15	0.82	0.78	-6
Sales index	0.09	0.97	0.70	-6
Consumer tendency survey variables				
Consumer confidence index	0.00	3.98	0.33	-10
Consumer confidence index - Family component	0.00	5.34	0.30	-10
Consumer confidence index - Future situation	0.00	2.67	0.28	-10
Consumer confidence index - Country component	0.00	2.25	0.28	-10
Consumer confidence index - Prices component	0.03	21.93	0.44	-10
Consumer confidence index - Current situation	0.00	12.80	0.41	-10
Consumer confidence index - Jobs component	0.03	0.65	0.27	-5
Consumer confidence index - Housing component	0.00	2.86	0.39	-10
International real variables				
Asia: leading index of 5 major economies	0.31	0.19	0.59	-2
China: leading index	0.01	2.27	0.58	-8
China: retail sales of consumer goods	0.23	-0.12	0.49	0
US: leading index	0.11	0.05	0.33	0
US: consumer confidence index (current situation)	0.03	0.33	0.19	-2
US: unemployment rate	0.01	-22.11	0.15	10
US: manufacturing and trade sales	0.18	0.03	0.44	0
Europe: business climate for Germany	0.13	0.32	0.42	-3
Europe: extra-Eurozone exports	0.14	0.05	0.38	0
Europe: leading index (EU 17)	0.18	0.08	0.43	-1
Europe: new orders in manufacturing	0.18	0.13	0.44	-1
Europe: industrial production (EU 16, excl. construction)	0.20	0.07	0.45	-1
Europe: industrial production (EU 27, excl. construction)	0.20	0.07	0.46	-1
Europe: economic sentiment index (EU 16)	0.10	0.35	0.38	-3
Europe: economic sentiment index (EU 27)	0.12	0.39	0.42	-3
Europe: retail sales index (EU 16)	0.01	0.49	0.12	-4
Europe: retail sales index (EU 27)	0.01	0.51	0.15	-6
G7: leading index	0.17	0.10	0.42	-1
NAFTA: leading index	0.10	0.07	0.32	0
OECD: leading index (Europe)	0.19	0.06	0.45	0
OECD: leading index (plus 6 major non-member economies)	0.23	0.12	0.49	-1
OECD: leading index (total)	0.18	0.11	0.43	-1
International financial variables				
Dow Jones	0.02	0.59	0.21	-4
Vix index	0.01	-20.74	0.39	10
Country risk - Latam	0.00	-12.54	0.27	8
TED spread	0.03	-21.09	0.16	8
Yen-dollar exchange rate	0.02	21.94	0.14	-10

Table 3: Final selected variables

	Optimal k	Optimal lead	QPS of optimal lead	Lead used to construct the RI
Domestic real variables				
Current account savings of the Central Government	3	4	0.09639	
Water consumption in Lima and Callao	4	5	0.10955	
Banking sector credit to the private sector	3	-5	0.10976	-2
Total credit/banking system liquidity	3	-4	0.15975	
Money	4	0	0.06500	-2
Current non-financial expenditure of the Central Government	3	-1	0.09965	-2
Non-financial expenditure of the Central Government	3	0	0.11441	-2
Imports of intermediate goods	6	0	0.03448	-1
Domestic value-dded tax	5	1	0.11711	
Value-added tax	5	1	0.00044	
Volume index of intermediate goods production	5	1	0.08126	
GDP construction	3	-1	0.10465	-2
Primary GDP	5	2	0.10621	
Electricity production	6	1	0.00097	
Terms of trade	6	1	0.00023	
Gross value added of chemical, rubber and plastic products	5	1	0.06554	
Domestic financial variables				
Net international reserves	6	0	0.02611	0
Country risk	6	0	0.04479	0
Effective reserve requirement rate in foreign currency	6	-4	0.01561	0
Effective reserve requirement rate in domestic currency	6	1	0.00766	
Business tendency survey variables				
Economic activity index	3	-1	0.06818	0
Business confidence index	6	-3	0.05643	0
3-months ahead expectations index	4	-3	0.05410	0
Inventory index	5	-1	0.11185	0
New orders index	5	-2	0.01070	0
Sales index	6	-1	0.02070	0
International real variables				
Asia: leading index of 5 major economies	6	-1	0.02252	-2
China: leading index	3	4	0.00185	
China: retail sales of consumer goods	3	10	0.10419	
US: manufacturing and trade sales	6	0	0.01150	-1
Europe: business climate for Germany	5	-3	0.14961	
Europe: extra-Eurozone exports	6	0	0.00941	-2
Europe: leading index (EU 17)	6	0	0.00622	-2
Europe: retail sales index (EU 27)	6	0	0.00772	-2
G7: leading index	6	0	0.01196	-2
OECD: leading index (Europe)	6	0	0.01140	-2
OECD: leading index (plus 6 major non-member economies)	6	-1	0.00134	-2
OECD: leading index (total)	6	-1	0.00262	-2
International financial variables				
Country risk - Latam	4	0	0.09195	0

ANNEX 2

Spectral coherence and mean delay

Coherence is a measure of the degree of relationship, as a function of frequency, between two time series. Following Croux et. al (2001), we define spectral coherence between variables X and Y as the correlation between them for each frequency (ω): $C_{XY}(\omega) = \frac{f_{XY}(\omega)}{\sqrt{f_X(\omega) f_Y(\omega)}}$. For business cycle analysis, economists are usually interested in the periodicity range 1.5 to 8 years, so high coherences within this period is an evidence that the series contains important information about the cyclical behaviour of GDP.

The mean delay measures the lags in the movements of a series with respect to another one. Following Croux et. al (2001), mean delay can be defined as the ratio between the phase and the frequency: $M = \Phi(\omega)/\omega$. A positive mean delay for a given frequency points to the fact that the cyclical component of Y leads the cyclical component of X .

Standardization process, optimal degree of smoothing selection and optimal lead/lag classification of the selected series

Let us call $Y_t = (Y_{1,t},...,Y_{N,t})'$ the $Nx1$ vector composed of N time series that were selected in the process described in 3.1. And let us call $y_t = (y_{1,t},...,y_{N,t})'$ the vector composed by the growth rates of Y_t which depend on $k = (1,...,N)$, which is the vector that contains the degree of smoothing of each series. Hence, $y_{i,t} = (1 - L^{k_i})*\log(Y_{i,t})*100 = (\log(Y_{i,t}) - \log(Y_{i,t-k_i}))*100$, where L^{k_i} is the lag operator. Then we construct the vector containing the standardized growth rates $y^*_t = (y^*_{1,t},...,y^*_{N,t})'$, where: $y^*_{i,t} = \frac{(y_{i,t} - \mu_{y_i})}{\sigma_{y_i}}$, with μ_{y_i} and σ_{y_i}, the respective sample mean and standard deviation of the growth rates of the i-eth series. Through this process of standardization we can aggregate series of diverse nature (quantitative, surveys, rates, etc.), since they are expressed in the same units.

The optimal value of k_i (the degree of smoothing of each series) will be that which minimizes the QPS applied to each variable i and degree of smoothing k: $QPS(i,k,R) = \frac{1}{T}\sum_{t=1}^{T}\left(SP_t^i(k) - R_t\right)^2$, where $SP_t^i(k)$ corresponds to smoothed probabilities of being in recession obtained from the i-eth series and k degree of smoothing, and $R_t = \{1,0\}$ is the dichotomous variable which indicates recession according to our reference criterion. Regarding the value of k_i, it is necessary to take into consideration that the lower the degree of smoothing, the faster the recession detection capability. However, the probability to obtain false alarms also hereby increases. In contrast, if the degree of smoothing is large, the probability to obtain false alarms decreases, as well as the velocity in recession detection. Consequently, we will narrow the search of the optimal k_i in the $k_i \in \{3,...,6\}$ space, which will allow us to balance the before-mentioned risks.

As discussed in Section 3.1, the classification methodology based on the estimation of spectral coherence, mean delay and cross-correlations allowed us to establish whether a variable is coincident, leading or lagging, but not to determine the exact months of such lead or lag. Therefore, to gather this information we will use the QPS criterion applied to recession probabilities of each series with a lead and lag horizon of 10 months. In this process, we will be able to determine the lead or lag months which minimize the QPS criterion, enabling us to eliminate again lagging variables[18].

[18] Note that the proposed methodology estimates the optimal k and then the optimal lead/lag. However, it is possible to find both elements simultaneously.