BANCO CENTRAL DE RESERVA DEL PERÚ

Regional inflation dynamics and inflation targeting in Peru

Diego Winkelried*   José Enrique Gutierrez**

* Central Reserve Bank of Peru
** Superintendency of Banking, Insurance and Private Pension Funds

DT. N° 2012-018
Serie de Documentos de Trabajo
Working Paper series
Octubre 2012

Regional inflation dynamics and inflation targeting in Peru*

Diego Winkelried†

and

José Enrique Gutierrez‡

October 4, 2012

Abstract

The Central Reserve Bank of Peru (BCRP) has been targeting inflation for more than a decade, using Lima's inflation as the operational measure. An alternative indicator is countrywide inflation, whose quality and real-time availability have improved substantially lately. Hence, given these two somehow competing measures of inflation, two interesting policy questions arise: what have been the implications for national inflation of targeting Lima's inflation? Would shifting to a national aggregate significantly affect the workings of monetary policy in Peru? To answer these questions, we estimate an error correction model of regional inflations and investigate how shocks propagate across the country. The model incorporates (i) aggregation restrictions whereby each regional inflation is affected by an aggregate of neighboring regions, and (ii) long-run restrictions that uncover a single common trend in the system. The results indicate that a shock to Lima's inflation is transmitted fast and strongly elsewhere in the country. This constitutes supporting evidence to the view that by targeting Lima's inflation, the BCRP has effectively, albeit indirectly, targeted national inflation.

JEL Classification : C32, C50, E31, E52, R10.
Keywords : Regional inflation, inflation targeting, relative PPP, error correction model.

*A previous version of this work was circulated in Spanish under the title "Propogación de choques inflacionarios en el Perú: Evaluación del esquema de metas de inflación en la estabilidad de precios a nivel nacional". We would like to thank Gustavo Yamada, Marco Ortiz and seminar participants at the Central Reserve Bank of Peru, and an anonymous referee for helpful comments. We are also grateful to Peter Paz and Brenda Pizarro for research assistance. All remaining errors are ours. The opinions herein are those of the authors and do not necessarily reflect the views of the Central Reserve Bank of Peru or the Superintendency of Banking, Insurance and Private Pension Funds.

†Winkelried: Central Reserve Bank of Peru, Jr. Antonio Miró Quesada 441-445, Lima 1, Perú (email: diego.winkelried@bcrp.gob.pe).

‡Gutierrez: Superintendency of Banking, Insurance and Private Pension Funds (email: jgutierreza@sbs.gob.pe).

1 Introduction

The Central Reserve Bank of Peru (BCRP) has been targeting inflation for almost two decades. In the early the 90s, it embarked into a disinflation programme that brought inflation down from hyperinflation (around 140 percent in 1991) to single-digit levels (6.5 percent in 1997) and eventually to international levels (3.7 percent in 1999). The process of disinflation was gradual as the BCRP built reputation as an inflation targeter. Aimed at reinforcing its ability to anchor inflation expectations, in 1994 it started announcing yearly target ranges and in 2002 it adopted a fully-fledged inflation targeting regime, where long-run targets were announced (see Rossini and Vega, 2008).

The operational measure of inflation used by the BCRP has always been the change in the Consumer Price Index (CPI) of the Peruvian capital city, Lima. There were good reasons for this choice. Firstly, the Peruvian economy is heavily centralized: Lima concentrates about a third of the country's population and used to represent more than 70 percent of national expenditure. Thus, it seems reasonable to believe that for practical purposes Lima's inflation could serve as a proxy for national inflation (see Armas et. al., 2010). Secondly, even though national CPI data are available, its timing of publication used to pose difficulties for the real-time monitoring of the state of the economy. Historically, whereas the definitive Lima's CPI figure has been readily available at the first working day of the following month, the release of the national CPI has been subject to delays and often to revisions.

However, this situation may have changed. In the last decade, the Peruvian economy has grown at a healthy average annual rate of 6.5 percent, with many provinces growing at an even faster pace. Even though the country remains centralized with Lima amounting to 2/3 of national expenditure, the economic importance of certain provinces has increased significantly, thereby bridging the income gap with respect to Lima. On the other hand, after major improvements in survey design and sampling techniques, since early 2012 the National Statistics Office of Peru (INEI) has committed to publish national inflation as reliably and timely as that of Lima. These developments have led local academia and economic press to inquire about the suitability of national inflation, rather than Lima's, in conducting monetary policy.[1] This raises two interesting policy questions:

- What have been the implications for national inflation of the BCRP targeting Lima's inflation?
- Would shifting to a national aggregate significantly affect the workings of monetary policy in Peru?

In order to answer these questions, it is key to quantify how relevant the behavior of inflation in Lima is for the determination of inflation in the rest of the country. For this, we take the purchasing power parity (PPP), which has been suggested to be adequate when studying price differentials within regions of a country or in a monetary union (Rogers, 2007), as our conceptual framework. In its absolute version, the PPP predicts that, in the absence of trading costs and other frictions, the price levels in two regions should converge to a same level (see Parsley and Wei, 1996; Cecchetti et. al., 2002, for comprehensive surveys). For the Peruvian case, Monge and Winkelried (2009) found that absolute PPP holds in about 3/4 of all possible pairs of Peruvian cities, at a 10 percent of significance. Price level convergence is limited by the transportation costs implied by an adverse geography and the lack of well-suited communications infrastructure in some of the Peruvian regions (especially, the highlands or Sierra).

However, this notion of PPP is not completely useful for monetary policy analysis, where the interest is on the behavior of *inflation* rather than price level differentials. Thus, we rather focus on the weak, relative version of the PPP that states that once the effects of transitory regional shocks fade away, the inflation rates in two cities should converge to an equilibrium rate, regardless on whether the shocks exert permanent discrepancies on the price levels (see, for example, Busetti et. al., 2006; Beck et. al., 2009).

The goal of this paper is to answer the aforementioned questions by investigating the dynamic relationships among regional inflations in Peru, and how they may affect the design of monetary policy. In particular, we aim to study

[1] Similar experiences in the past can be found upon exploring the websites of the statistical offices and central banks of other Latin American inflation targeters. In Colombia, the CPI covered 13 cities from 1989 to 2008, and 24 cities from 2009 onwards; in Mexico, 35 cities up to February 1995 and 46 cities since March 1995; in Chile, up to January 2009 the CPI was that of Santiago de Chile, then in February 2009 the urban centers of Puente Alto and San Bernardo were incorporated, and in February 2010 the CPI became countrywide. Yet, the geographical representativeness of the CPI index used for monetary policy analysis is always taken for granted, to the extent that we have failed to find relevant documentation or research related to the implications of these changes in measuring inflation.

whether relative PPP holds among Peruvian regions and, if so, whether the adjustment towards equilibrium is fast enough. To this end, we estimate and simulate a vector error correction model for the inflation rates of 9 Peruvian economic regions, a model that shows how regional shocks propagate across the country. To achieve a parsimonious yet dynamically rich specification, we constrain the feedback effects from other regions by considering "rest of the country" aggregates rather than each region individually. These aggregates are computed using both a geographic scheme, where the influence of neighboring regions is larger than that of more distant ones, and an economic approach, where the importance of a given region is proportional to its participation in national expenditure. Once the aggregates are computed, the dynamic relationships among the regional inflations rates can be further analyzed, and in particular it can be assessed whether the data are driven by common inflation trends.

The increase in globalization over the last decades has highlighted the pervasiveness of international linkages in the world economy, and the importance of capturing those linkages in empirical models (see, among others, Elliott and Fatas, 1996). As a result, there is a large literature exploiting such interrelationships, with its most popular thread being the so-called global VAR (GVAR) advanced in Pesaran et. al. (2004), where "rest of the world" aggregates are computed using trade weights. Even though our modeling approach is related to the GVAR, there at least two important methodological differences. Firstly, our model is smaller as it includes one variable per unit (regional inflation). Although this prevents us to label shocks more adequately (for instance, supply versus demand shocks), it allows us to formally test the aggregation hypothesis that is taken for granted in the GVAR literature. Secondly, our identification strategy differs in that we are able to identify relevant shocks (for instance, a shock to Lima's inflation) from the long-run properties of the system.

We find strong evidence on the importance of shocks to Lima's inflation in influencing pricing decisions elsewhere in the country. In particular, most of the variability in individual regional inflation rates is driven by the evolution of a common trend that can be identified from the evolution of Lima's inflation. Furthermore, it is found that convergence towards this trend is fast, well within the monetary policy lag (i.e., the time it takes for a monetary policy action to affect inflation, which is believed to be between one and two years). Therefore, given such active correction mechanisms, we can conclude that by targeting Lima's inflation the BCRP has indirectly promoted inflation stability in the whole country.

The remainder of the paper is organized as follows. Section 2 motivates the discussion by briefing stylized facts in the data. Section 3 discusses methodological issues and develops a vector error correction model that allows for rich feedbacks parsimoniously. Furthermore, formal tests on the aggregation restrictions and on the relative PPP are proposed. Section 4 presents estimation results, and analyzes the effects of a shock originated in Lima. Section 5 gives closing remarks and avenues for further research.

2 A glimpse of the data

The INEI publishes CPI data for the 25 largest cities of the 24 Peruvian Departments (administrative divisions) on a monthly basis. In order to work with a manageable number of units, we use the economic classification of Peruvian cities into $n = 9$ economic regions proposed by Gonzales de Olarte (2003). This classification is based both on historical considerations (the regions are formed by contiguous Departments) and, importantly, on economic grounds, such as market articulation and integration.

The economic regions are depicted in Figure 1, together with the weights used by INEI in computing national inflation, based on expenditure surveys. As mentioned, Lima is by far the most important region in the country, with a share of 66.02 percent. It is followed by the two economic centers located at regions 4 (La Libertad and Ancash) and 6 (Arequipa, Moquegua, Tacna and Puno), each with a share slightly above 7 percent. The share of each of the remaining 6 regions varies between 2.21 (region 5) and 3.93 (region 9) percent. It is worth mentioning that these shares correlate strongly with the distribution of population across the country.

The data run from 1996 to 2011, about 190 observations after adjusting for initial conditions. Recall that throughout the sample period, the BCRP has been targeting Lima's inflation. In our empirical analysis below, inflation is defined as the annualized monthly percent change of each CPI index P_t, $1200\log(P_t/P_{t-1})$. However, to ease visualization in this section we describe the smoother year-on-year rates, $100\log(P_t/P_{t-12})$.

Figure 2(a) shows the evolution of Lima's inflation, together with the maximum and minimum rates found among the remaining 8 regions. Given the importance of Lima, at this scale national inflation is difficult to distinguish from that of Lima (with the exception of a few episodes) and hence is not included in the graph. It is notorious that the three inflation series share many common features. In particular, they tend to move in tandem (their turning points are virtually the same with small differences of about one or two months) and their variations are of comparable magnitudes. Thus, at first glance, the inflation rates seem to be driven by a common trend or, in other words, the relative PPP seems to be a feature of the data, a point that is formally addressed below.

There are, of course, episodes of discrepancies that, albeit sizeable, appear to be temporary. Two are worth highlighting. The first one occurs in 1998, when the country was hit by a particularly severe El Niño phenomenon. The economic costs of El Niño relate to disruptions in agricultural and fishing production, and to direct damages to infrastructure due to floods, and are often concentrated in the northern regions of the country (2, 3 and 4) that are closer to the equator. These regions experienced a peak in inflation that was reverted as the climate conditions returned to normality in 1999.

The second episode corresponds to the period 2007 to mid-2008, and is the consequence of two related economic phenomena. The first one is the commodities prices boom witnessed by international markets, mainly due to boosting demand in emerging markets like China. The magnitude and persistence of these shocks passed-through regional prices unevenly, mainly because the share of food in the CPI is larger in provinces than in Lima, a fact that partially explains why Lima registered the lowest inflation rate during this period.

Moreover, preliminary figures from INEI on regional output show that during the period from 2005 to 2008, Lima grew at about 7.5 percent per annum, whereas regions as Arequipa, Cusco and La Libertad grew at rates closer to 8 percent, and even more successful locations such as Ica grew above 9 percent. This strong expansion in regional output (and demand) was boosted by the extraordinarily favorable conditions the country faced as a commodity exporter. This gives an indication that many regions would have bridged the income gap with respect to Lima relatively quickly, inevitably generating inflationary pressures at the regional level. Moreover, anecdotal evidence points out to conjecture that the inflationary pressures may have been magnified by initial conditions, a *catch-up effect*. In particular, prices that were lower in provinces (because of a Balassa-Samuelson effect) would have converged to levels close to those of the capital as the provinces grew richer, thereby generating further inflation. Nonetheless, the inflationary pressures brought by these processes vanished rapidly after the financial international crisis was triggered by the collapse of Lehman-Brothers in late-2008.

Figure 2(b) depicts inflation deviations from the national rate for selected regions from both ends of the country. It can be seen that inflation deviations in northern region 2 are more closely related to those of neighboring region 3 than to those of the more distant regions, for instance 6 or 8. Similarly, the relationship between inflation in any of the southern regions 6 and 8 is much closer than that of a northern region with a southern one. Thus, an important conclusion from the visual inspection of these series is that geography seems to matter. Regional shocks, or even the responses to aggregate shocks in different regions, are likely to depend on location. This additional feature of the data will prove useful for parsimony in the empirical analysis below.

3 Methodological issues

This section presents the econometric framework used to investigate feedbacks amongst the inflation rates of n regions in a country. Two major points are considered: testing and the identification of structural shocks.

Regarding testing, aggregation restrictions are imposed into a standard, potentially large reduced form VAR of inflation rates, and we propose a test to formally assess their significance. These restrictions promote parsimony while maintaining the dynamic feedback to the inflation rate in a given region from the inflation rates in the rest of the country. Secondly, we derive a test for long-run homogeneity restrictions using the error correction form of the VAR model. In the context of a VAR model of inflation rates, long-run homogeneity can be understood as the fulfilment of the weak version of the PPP hypothesis, where the differences between inflation rates are stationary.

It is important to emphasize that we have been careful in proposing testing procedures that are robust to the stationarity properties of the data. In particular, both the aggregation and long-run homogeneity hypotheses can be tested using standard results from classical regression theory as the statistics involved are asymptotically Gaussian or chi-square, *regardless* on whether the data are stationary or not. Technical details are shown in the appendix.

The error correction form of the VAR provides a suitable formulation to infer the characteristics of the common trends driving the data. In particular, we focus on the case where the inflation rate in a single reference region does not respond to disequilibria (deviations from the relative PPP) elsewhere. Here, a shock is to be regarded as having permanent effects as long as it affects the inflation in the reference region, and is dubbed as transitory otherwise. It follows that the shock to the reference region can be easily identified using the permanent-transitory decomposition advanced in Gonzalo and Ng (2001).

3.1 The aggregation hypothesis

Our starting point is the reduced form VAR(p) model

$$\boldsymbol{y}_t = \sum_{r=1}^{p} \boldsymbol{A}_r \boldsymbol{y}_{t-r} + \boldsymbol{\varepsilon}_t \,, \tag{1}$$

where $\boldsymbol{y}_t$ is an $n \times 1$ vector of endogenous variables whose i-th element corresponds to the inflation of region i in period t, $\boldsymbol{A}_r$ ($r = 1, 2, \ldots, p$) are coefficient matrices and $\boldsymbol{\varepsilon}_t$ is the vector of mutually correlated *iid* statistical innovations. The covariance matrix of $\boldsymbol{\varepsilon}_t$ is an $n \times n$ positive define matrix $\boldsymbol{\Omega}_\varepsilon$.

It is well-documented that the usefulness of a dynamic model like (1) may be limited in finite samples due to the proliferation of parameters that need to be estimated. Indeed, each additional lag implies the estimation of n^2 coefficients, and these may be poorly estimated with the sample sizes typically encountered in applications. Thus, promoting parsimony by imposing meaningful restrictions on matrices $\boldsymbol{A}_r$ is likely to improve the inferential content of testing procedures based on the VAR system. This is the purpose of aggregation restrictions, where given weights are used to construct averages that maintain feedback effects across regions.

Consider an aggregate composed by the $n - 1$ variables in $\boldsymbol{y}_t$ other than $y_{i,t}$ ($i = 1, 2, \ldots, n$),

$$x_{i,t} = \sum_{j=1}^{n} w_{ij} y_{j,t} \qquad \text{where} \qquad \sum_{j=1}^{n} w_{ij} = 1 \qquad \text{and} \qquad w_{ii} = 0 \,. \tag{2}$$

The normalization that the weights sum to one is for algebraic convenience and involves no loss of generality. We maintain the assumption that the weights w_{ij} are *not* estimated jointly with $\boldsymbol{A}_r$, otherwise the linearity in the VAR model may be lost with aggregation. This situation corresponds to either non-random weights or stochastic weights that are predetermined, i.e. its determination is independent from $\boldsymbol{\varepsilon}_t$.

Take the i-th equation of the unrestricted VAR

$$y_{i,t} = \sum_{r=1}^{p} a_{ii}(r) y_{i,t-r} + \sum_{r=1}^{p} \sum_{j \neq i}^{n} a_{ij}(r) y_{j,t-r} + \varepsilon_{i,t} \,, \tag{3}$$

where $y_{i,t}$ is the i-th element of $\boldsymbol{y}_t$, $\varepsilon_{i,t}$ is the i-th element of $\boldsymbol{\varepsilon}_t$, and $a_{ij}(r)$ denotes the (i, j)-th element of $\boldsymbol{A}_r$. In an alternative, restricted model all dynamic feedback to $y_{i,t}$ come from its own lags and lags of the aggregate,

$$y_{i,t} = \sum_{r=1}^{p} a_{ii}(r) y_{i,t-r} + \sum_{r=1}^{p} c_i(r) x_{i,t-r} + e_{i,t} = \sum_{r=1}^{p} a_{ii}(r) y_{i,t-r} + \sum_{r=1}^{p} \sum_{j \neq i}^{n} c_i(r) w_{ij} y_{j,t-r} + e_{i,t} \,. \tag{4}$$

Clearly, if $a_{ij}(r) = c_i(r) w_{ij}$ then the restricted model (4) is equivalent to the model without restrictions (3). These $p(n - 1)$ equalities imply a total of $p(n - 1) - p = p(n - 2)$ restrictions of the form

$$a_{ij}(r) - \left[\frac{w_{ij}}{w_{ik}}\right] a_{ik}(r) = 0 \qquad \text{for} \quad j \neq k,\; k \neq i \quad \text{and} \quad r = 1, 2, \ldots, p \,. \tag{5}$$

Thus, the aggregation restrictions imply that the elements of the i-th row of $\boldsymbol{A}_r$ are proportional to each other, and the proportionality factor is given by the ratio w_{ij}/w_{ik}. In other words, $y_{j,t}$ and $y_{k,t}$ affect the expected value of future realizations of $y_{i,t}$ proportionally to their contributions to the aggregate (2).

The unrestricted model is obtained by regressing $y_{i,t}$ on the p lags of $\boldsymbol{y}_t$. This amounts to pn coefficients per equation and pn^2 in the entire VAR. On the other hand, in the restricted model $y_{i,t}$ is regressed on its p lags and the p lags of the aggregate $x_{i,t}$. Here, each equation has $2p$ coefficients and the restricted VAR has $2pn$ coefficients. Thus, the aggregation restrictions can reduce the number of coefficients to be estimated substantially, even for moderate values of n. For instance, if $p = 2$ and $n = 9$ then we have $pn^2 = 162$ coefficients in the unrestricted model, and only $2pn = 36$ in the restricted, a total of $np(n-2) = 126$ restrictions.

The aggregation restrictions can be conveniently reinterpreted as exclusion restrictions, and this is the basis for hypothesis testing. After simple manipulations, the original equation (3) can be rewritten as

$$y_{i,t} = \sum_{r=1}^{p} a_{ii}(r) y_{i,t-r} + \sum_{r=1}^{p} c_i(r) x_{i,t-r} + \sum_{r=1}^{p} \sum_{j \neq i}^{n} \delta_{ij}(r) y_{j,t-r} + \varepsilon_{i,t}\,, \tag{6}$$

where $\delta_{ij}(r) = a_{ij}(r) - c_i(r) w_{ij}$ for $r = 1, \ldots p$, $j = 1, 2, \ldots, n$ and $j \neq i$. Therefore, the restricted model has $\delta_{ij}(r) = 0$ for all r and $j \neq i$. Thus, testing the aggregation hypothesis amounts to test the joint significance restrictions $H_0 : \delta_{ij}(r) = 0$ using a standard Wald statistic. Note that H_0 has the appealing interpretation that once $x_{i,t}$ is controlled for, its constituents $y_{j,t}$ have no predictive power on $y_{i,t}$.

3.2 Long-run homogeneity

The restricted model

$$y_{i,t} = \sum_{r=1}^{p} a_{ii}(r) y_{i,t-r} + \sum_{r=1}^{p} c_i(r) x_{i,t-r} + \varepsilon_{i,t}\,, \tag{7}$$

constitutes an ARDL equation that describes the dynamic relationship between $y_{i,t}$ and $x_{i,t}$. The long-run multiplier $\text{LRM}_i = \sum_{r=1}^{p} c_i(r) / (1 - \sum_{r=1}^{p} a_{ii}(r))$ measures the effect that a unit permanent deviation of $x_{i,t}$ from its long-run level has on $y_{i,t}$. The relative PPP hypothesis between $y_{i,t}$ and $x_{i,t}$ implies the long-run homogeneity condition $\text{LRM}_i = 1$. This means that inflation in region i ($y_{i,t}$) responds to the permanent deviation on the inflation rate elsewhere ($x_{i,t}$) proportionally, such that the long-run behavior of the relative inflation ($y_{i,t} - x_{i,t}$) is not affected, i.e. $y_{i,t} - x_{i,t}$ remains stationary. Note that the fulfillment of *pairwise* relative PPP conditions, i.e. $y_{i,t} - y_{j,t}$ stationary for all $i \neq j$, requires $\text{LRM}_i = 1$ for all regions. We return to this point below.

In order to test and impose $\text{LRM}_i = 1$, the ARDL equation can be written in error correction form as

$$\Delta y_{i,t} = -\gamma_i (y_{i,t-1} - x_{i,t-1}) + \sum_{r=1}^{p-1} \pi_i^y(r) \Delta y_{i,t-r} + \sum_{r=1}^{p-1} \pi_i^x(r) \Delta x_{i,t-r} + \theta_i x_{i,t-1} + \varepsilon_{i,t}\,, \tag{8}$$

where $\gamma_i = 1 - \sum_{r=1}^{p} a_{ii}(r) \geq 0$, $\pi_i^y(r) = -\sum_{s=r+1}^{p} a_{ii}(s)$, $\pi_i^x(r) = \sum_{s=r+1}^{p} c_i(s)$ and $\theta_i = \gamma_i(\text{LRM}_i - 1)$.

The long-run homogeneity hypothesis can be written as $H_0 : \theta_i = 0$. As shown in appendix B, provided that (8) is correctly specified (i.e., $\varepsilon_{i,t}$ uncorrelated to the history of $x_{i,t}$), the t-statistic associated to H_0 has the usual asymptotic Gaussian distribution, even if the data are not stationary and so the limiting distribution of the least squares estimator of θ_i is non-standard. This provides a simple test on the relative PPP hypothesis in equation i.

If $\gamma_i > 0$ then the so-called error correction mechanism is in place and so $y_{i,t}$ responds to deviations from the aggregate $x_{i,t}$ to restore equilibria in the long-run. The "speed of adjustment" coefficient γ_i tells us the proportion of the disequilibrium $y_{i,t} - x_{i,t}$ which is corrected with each passing period. Note that if $\gamma_i = 1$, then equilibrium correction occurs in just one period. On the other hand, $\gamma_i = 0$ corresponds to the case where $y_{i,t}$ does not respond to disequilibria at all, $y_{i,t}$ is *long-run forcing*. As we will see next, this result can be exploited to identify the trends driving the dynamics of the system made from the collection of n equations like (8).

3.3 Identification of shocks from a reference region

Under the aggregation and long-run homogeneity hypotheses, the original VAR model (1) can be written as the vector error correction model (VECM)

$$\Delta \boldsymbol{y}_t = -\boldsymbol{\Pi} \boldsymbol{y}_{t-1} + \boldsymbol{\pi}_1 \Delta \boldsymbol{y}_{t-2} + \ldots + \boldsymbol{\pi}_{p-1} \Delta \boldsymbol{y}_{t-p+1} + \boldsymbol{\varepsilon}_t \,, \tag{9}$$

where the $\boldsymbol{\pi}$ matrices are constrained due to the aggregation restrictions.[2]

Recall that, by the Granger representation theorem (cf. Hamilton, 1994, p. 581), a system of n variables driven by h stochastic trends can be represented as a VECM like (9) with the long-run matrix $\boldsymbol{\Pi}$ having rank equal to $n - h$. This implies that the system contains $n - h$ cointegrating relationships.

Let $\boldsymbol{W}$ be the $n \times n$ matrix that collects the weights w_{ij} (recall that $w_{ii} = 0$). Plugging the definition in (2) into (8), it is easy to verify that matrix $\boldsymbol{\Pi}$ takes the form $\boldsymbol{\Pi} = \boldsymbol{\Gamma}_0 \boldsymbol{B}_0$ where $\boldsymbol{\Gamma}_0$ is an $n \times n$ diagonal matrix with γ_i as the i-th entry on the diagonal, and $\boldsymbol{B}_0 = \boldsymbol{I}_n - \boldsymbol{W}$ is also an $n \times n$ matrix. Explicitly,

$$\boldsymbol{\Pi} = \begin{bmatrix} \gamma_1 & 0 & 0 & \cdots & 0 \\ 0 & \gamma_2 & 0 & \cdots & 0 \\ 0 & 0 & \gamma_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \gamma_n \end{bmatrix} \begin{bmatrix} 1 & -w_{12} & -w_{13} & \cdots & -w_{1n} \\ -w_{21} & 1 & -w_{23} & \cdots & -w_{2n} \\ -w_{31} & -w_{32} & 1 & \cdots & -w_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -w_{n1} & -w_{n2} & -w_{n3} & \cdots & 1 \end{bmatrix} . \tag{10}$$

Let $\boldsymbol{1}_m$ denote an $m \times 1$ sum vector (full of ones). Then, since $\boldsymbol{W}\boldsymbol{1}_n = \boldsymbol{1}_n$ ($\boldsymbol{W}$ is *row-stochastic*), it follows that $\boldsymbol{B}_0\boldsymbol{1}_n = \boldsymbol{0}$, and thus $\boldsymbol{1}_n$ is an eigenvector of $\boldsymbol{B}_0$ associated with a zero eigenvalue. It is easy to show that the multiplicity of the zero eigenvalue is one, and thus the rank of $\boldsymbol{B}_0$ is equal to $n - 1$. Provided that $\gamma_i > 0$ for at least $n - 1$ equations, and hence the rank of $\boldsymbol{\Gamma}_0$ is either $n - 1$ or n, this result implies that $\boldsymbol{\Pi}$ is of rank $n - 1$ as well. It follows that in system (9), there must a single stochastic trend and $n - 1$ cointegrating relationships, $h = 1$.

The equality in (10) and its implied single common trend is a direct consequence of setting $\theta_i = 0$ for all i in (8): $y_{i,t}$ is cointegrated with $x_{i,t}$, and since this happens for all i, it must be the case that $y_{i,t}$ is cointegrated with each of the $n - 1$ constituents of $x_{i,t}$. In other words, $\theta_i = 0$ for all i implies pairwise cointegrating relationships of the form $y_{i,t} - y_{j,t} \sim I(0)$ for $i \neq j$, i.e. the fulfilment of the relative PPP for any pair of regions.

A more familiar VECM representation will have $\boldsymbol{\Pi}$ decomposed as the product of a full *column* rank matrix and a full *row* rank matrix, both with rank $n - 1$. We may arrived easily to such a representation if one of the adjustment coefficients γ_i happens to be equal to zero. Hence, with no loss of generality we set $\gamma_1 = 0$, so that

$$\boldsymbol{\Pi} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \gamma_2 & 0 & \cdots & 0 \\ 0 & \gamma_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \gamma_n \end{bmatrix} \begin{bmatrix} -w_{21} & 1 & -w_{23} & \cdots & -w_{2n} \\ -w_{31} & -w_{32} & 1 & \cdots & -w_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -w_{n1} & -w_{n2} & -w_{n3} & \cdots & 1 \end{bmatrix} , \tag{11}$$

or compactly $\boldsymbol{\Pi} = \boldsymbol{\Gamma}\boldsymbol{B}$, where $\boldsymbol{\Gamma}$ is the $n \times (n - 1)$ matrix obtained by removing the first column of $\boldsymbol{\Gamma}_0$, and $\boldsymbol{B}$ is the $(n - 1) \times n$ matrix resulting from removing the first row of $\boldsymbol{B}_0$. Both $\boldsymbol{\Gamma}$ and $\boldsymbol{B}$ have rank $n - 1$.

The alternative representation proves very convenient in interpreting the shocks hitting the system and especially in identifying the single common trend. The situation $\gamma_1 = 0$ and $\gamma_i > 0$ for $i > 1$ implies that $y_{1,t}$ is the only long-run forcing variable of the system and, following Gonzalo and Granger (1995), that this very variable can be used to describe the properties of the common trend. Intuitively, each $y_{i,t}$ can be thought of as the accumulation of a shock with permanent effects plus an idiosyncratic shock that deviates $y_{i,t}$ from the common trend. Since $y_{1,t}$

[2] Define $\boldsymbol{\Pi}_r^x = \mathrm{diag}(\,\pi_1^x(r), \pi_2^x(r), \ldots, \pi_n^x(r)\,)$ and $\boldsymbol{\Pi}_r^y = \mathrm{diag}(\,\pi_1^y(r), \pi_2^y(r), \ldots, \pi_n^y(r)\,)$ as the $n \times n$ diagonal matrices that collect the coefficients associated to the r-th lag effects in (8). Then, $\boldsymbol{\pi}_r = \boldsymbol{\Pi}_r^y + \boldsymbol{\Pi}_r^x \boldsymbol{W}$, where $\boldsymbol{W}$ is the matrix of weights defined below. Note that the $n \times n$ matrix $\boldsymbol{\pi}_r$ contains only $2n$ free parameters.

will not adjust to restore disequilibria, then it must be driven exclusively by the permanent shock, whereas $y_{i,t}$ for $i \neq 1$ not only responds to the permanent shock, so that it keeps $y_{i,t} - y_{1,t}$ stationary, but also to perturbations that transitorily take $y_{i,t} - y_{1,t}$ away from its constant long-term expectation.

Gonzalo and Ng (2001) propose a method to express the perturbations of the VECM $\boldsymbol{\varepsilon}_t$ as a linear combination of h permanent shocks and $n - h$ transitory shocks, the so-called PT decomposition. Thus, the above rationale provides a simple method to identify shocks that can be attributed to region 1: since only the $h = 1$ trend shock affects $y_{1,t}$, then the permanent shock in a PT decomposition can interpreted as coming from region 1.

More formally, let $\boldsymbol{u}_t$ be the $n \times 1$ vector whose first entry is the permanent shock, $u_{1,t}$, and the remaining $n - 1$ entries correspond to transitory shocks, $\boldsymbol{u}_{2,t}$. Gonzalo and Ng show that $\boldsymbol{u}_{2,t} = \boldsymbol{B}\boldsymbol{\varepsilon}_t$, whereas $u_{1,t} = \boldsymbol{\Gamma}^* \boldsymbol{\varepsilon}_t$, where $\boldsymbol{\Gamma}^*$ is a $1 \times n$ vector such that $\boldsymbol{\Gamma}^*\boldsymbol{\Gamma} = \mathbf{0}$. From (11) it is clear that $\boldsymbol{\Gamma}^*$ must be proportional to the first unit vector in $\mathbb{R}^n$, i.e. the first row of $\boldsymbol{I}_n$. More compactly, $\boldsymbol{u}_t = \boldsymbol{G}\boldsymbol{\varepsilon}_t$, where the first row of $\boldsymbol{G}$ is $\boldsymbol{\Gamma}^*$ and its lower $(n-1) \times n$ block is given by $\boldsymbol{B}$. To have a better grasp on how the structural shocks $\boldsymbol{u}_t$ relates to the innovations $\boldsymbol{\varepsilon}_t$, consider the equality $\boldsymbol{\varepsilon}_t = \boldsymbol{G}^{-1}\boldsymbol{u}_t$. Note that $\boldsymbol{B}$ can be partitioned as $\boldsymbol{B} = [-\boldsymbol{w}_{12} \ : \ \boldsymbol{I}_{n-1} - \boldsymbol{W}_{22}]$, where $\boldsymbol{w}_{12}$ and $\boldsymbol{W}_{22}$ are blocks of $\boldsymbol{W}$. Since $(\boldsymbol{I}_{n-1} - \boldsymbol{W}_{22})\mathbf{1}_{n-1} = \boldsymbol{w}_{12}$, then

$$\boldsymbol{G}^{-1} = \begin{bmatrix} 1 & \mathbf{0} \\ -\boldsymbol{w}_{12} & \boldsymbol{I} - \boldsymbol{W}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} 1 & \mathbf{0} \\ (\boldsymbol{I} - \boldsymbol{W}_{22})^{-1}\boldsymbol{w}_{12} & (\boldsymbol{I} - \boldsymbol{W}_{22})^{-1} \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{1}_{n-1} & (\boldsymbol{I} - \boldsymbol{W}_{22})^{-1} \end{bmatrix}. \tag{12}$$

The innovation to the first equation equals the permanent shock $\varepsilon_{1,t} = u_{1,t}$, and in this way we are able to identify how shocks to a reference region propagates to the rest of the country. Note also that on impact, a shock $u_{1,t} = 1$ will have a proportional effect in *all* innovations $\varepsilon_{i,t} = 1$ for $i = 1, 2, \ldots, n$, whereas a shock $u_{i,t} = 1$ for $i > 1$ will affect all innovations $(\varepsilon_{2,t}, \varepsilon_{3,t}, \ldots, \varepsilon_{n,t})$ but $\varepsilon_{1,t}$.[3,4]

4 Empirical analysis

Next we present the main results of our empirical analysis, using $y_{i,t} = 1200 \log(P_{i,t}/P_{i,t-1})$, i.e. annualized monthly inflation rates. First, we describe two approaches to measure the "rest of the country" inflation rates, one based on geographic proximity and the other based on economic importance. Under both aggregation schemes, we find supporting evidence of the aggregation hypothesis. The VECM is then estimated and its dynamic features are analyzed. It is found that relative PPP holds across the country and moreover that convergence towards a common inflation trend is fast. Furthermore, the results also point out that the influence of Lima in the behavior of regional, and consequently countrywide, inflation rates is strong.

4.1 Aggregation schemes

We entertain two alternative approaches to construct the aggregate $x_{i,t}$ which, since $\boldsymbol{x}_t = \boldsymbol{W}\boldsymbol{y}_t$, amounts to determine the weights matrix $\boldsymbol{W}$. In both cases the entries of $\boldsymbol{W}$ are given (independent from $\boldsymbol{\varepsilon}_t$), and so the theoretical considerations on linearity after aggregation discussed in the previous section apply.

The first approach corresponds to *geographic aggregation*. It is based on the basic notion that location and distance matter, and in particular that the closer region i is from region j, the stronger their mutual feedback. There are many ways to operationalize this concept and we choose one that is particularly simple (up to third-order contiguity relations among the $n = 9$ regions).

[3] The usual practice, however, is to orthogonalize $\boldsymbol{u}_t$ and consider instead $\boldsymbol{\varepsilon}_t = \boldsymbol{G}^{-1}\boldsymbol{H}\boldsymbol{\eta}_t$, where $\boldsymbol{H}$ is the lower block triangular matrix obtained by applying a Choleski decomposition to the covariance matrix of $\boldsymbol{u}_t$, and $\boldsymbol{\eta}_t$ collects the orthogonalized shocks. This makes no difference from the identification of the shock to the reference region. Just as in the case of $\boldsymbol{u}_t$, the first element of $\boldsymbol{\varepsilon}_t$ is affected only by the first element of $\boldsymbol{\eta}_t$, i.e. the orthogonalized permanent shock.

[4] These results can be generalized easily. Suppose that there are two reference regions (two common trends) such that $\gamma_1 = \gamma_2 = 0$. In this case, since the rank of $\boldsymbol{B}_0$ remains $n - 1$, $\text{rank}(\boldsymbol{\Pi}_0) = \text{rank}(\boldsymbol{\Gamma}_0) = n - 2$ so in the $\boldsymbol{\Pi} = \boldsymbol{\Gamma}\boldsymbol{B}$ decomposition $\boldsymbol{\Gamma}$ is the $n \times (n-2)$ matrix obtained by removing the first two columns of $\boldsymbol{\Gamma}_0$, and $\boldsymbol{B}$ is the $(n-2) \times n$ matrix resulting from removing the first two rows of $\boldsymbol{B}_0$. Upon examining the structure of matrix $\boldsymbol{G}$, we can reach the important conclusion that the first two innovations $(\varepsilon_{1,t}, \varepsilon_{2,t})$ are linear combinations of the two permanent shocks $(u_{1,t}, u_{2,t})$, and no transitory shock affects them. Upon orthogonalizing $\boldsymbol{u}_t$, or using another identification scheme, we could then find structural shocks coming from regions 1 and 2. For a related discussion, see Gonzalo and Granger (1995).

Define the scores

$$\tilde{w}_{ij} = \begin{cases} 0 & \text{if } i = j \text{ (regions are not considered neighbors to themselves)}, \\ 0 < a & \text{if region } j \text{ and } i \text{ are contiguos}, \\ 0 \leq b < a & \text{if region } j \text{ is located one region away from } i, \\ 0 \leq c < b & \text{otherwise (} i \text{ and } j \text{ are separated by more than two regions)}, \end{cases} \tag{13}$$

such that w_{ij} is simply the normalized version of $\tilde{w}_{ij}$, i.e $w_{ij} = \tilde{w}_{ij} / \sum_{j=1}^{n} \tilde{w}_{ij}$. It is worth-mentioning that alternative, and more sophisticated, spatial weighting schemes rendered results (available upon request) that were very similar to those reported below. From Figure 1, the symmetric matrix with raw (not normalized) scores equals

$$\tilde{\boldsymbol{W}} = \begin{bmatrix} 0 & c & b & a & a & b & a & b & b \\ c & 0 & a & b & c & c & c & c & b \\ b & a & 0 & a & c & c & b & c & a \\ a & b & a & 0 & b & c & a & b & a \\ a & c & c & b & 0 & a & a & a & b \\ b & c & c & c & a & 0 & b & a & b \\ a & c & b & a & a & b & 0 & a & a \\ b & c & c & b & a & a & a & 0 & a \\ b & b & a & a & b & b & a & a & 0 \end{bmatrix} . \tag{14}$$

We set arbitrarily $a = 1$, $b = 0.5$ and $c = 0$, and recall that our main conclusions are not sensitive to this choice.

The second approach is *economic aggregation*. Here, the strength of the feedback is determined by weights that are proportional to the participation of each region in national expenditure. The underlying assumption is that the importance of region i as a market for goods traded nationwide is well-reflected by its relative contribution to total expenditure. Under this approach, $x_{i,t}$ is a normalized measure of national inflation *without the contribution of inflation in region i*. Let Y_t denote national inflation, and let α_j be the weight region j receives in composing it (these weights are reported in Figure 1). By construction, $\sum_{j=1}^{n} \alpha_j = 1$. Then,

$$Y_t = \sum_{j=1}^{n} \alpha_j y_{j,t} = \alpha_i y_{i,t} + \sum_{j \neq i}^{n} \alpha_j y_{j,t} = \alpha_i y_{i,t} + (1 - \alpha_i) x_{i,t} , \tag{15}$$

thus Y_t is a linear combination between $y_{i,t}$ and $x_{i,t}$. It follows that $x_{i,t} = \sum_{j \neq i}^{n} \alpha_j y_{j,t} / (1 - \alpha_i)$ and, therefore,

$$w_{ij} = \frac{\alpha_j}{1 - \alpha_i} . \tag{16}$$

In this case, the matrix $\boldsymbol{W}$ is not symmetric. A large region (in economic terms) exerts more influence on a small region than the small region exerts on the larger one. More formally, provided that $\alpha_i + \alpha_j < 1$ (which is always the case), it is easy to verify that $w_{ij} < w_{ji}$ if and only if $\alpha_i > \alpha_j$. Thus, the feedback that inflation in region i receives from inflation in region j is smaller than the feedback from region i to region j, since region i constitutes, in terms of expenditure, a larger market than region j.

4.2 Aggregation hypothesis

In testing the aggregation hypothesis, an important practical issue is the determination of the lag length p. Based on the results of Breusch–Godfrey LM tests on serial correlation in the regression errors (as well as some information criteria), we could not reject the hypothesis of uncorrelated errors in any equation for $p \geq 5$. This result holds for either the unrestricted model (3) or its restricted version (4). Thus, we set $p = 5$ ($p - 1 = 4$ lags in the VECM). Then, under the null hypothesis of aggregation ($\delta_{ij}(r) = 0$ in equation (6) for all $i = 1, \ldots, n$, $j \neq i$ and $r = 1, \ldots, p$), the standard Wald statistic is asymptotically distributed as χ^2 with $p(n - 2) = 35$ degrees of freedom.

It can be seen in Table 1 that the aggregation hypothesis cannot be rejected at conventional significance levels in any case, and regardless of the aggregation scheme used to compute $x_{i,t}$ (geographic or economic). We take these results as conclusive evidence that the restricted model, which uses weighted aggregates to summarize feedback effects from the rest of the country, is capable to capture the main features of the data. The next step, thus, is to investigate the dynamics of this restricted form.

4.3 Long-run homogeneity

Unlike a standard VECM, (9) constitutes a system of seemingly unrelated regression equations (SUR) with different regressors. Thus, we choose system generalized least squares as our preferred estimation method (the results using ordinary least squares were qualitatively the same).

The results are displayed in Table 2 (for $p = 5$). The first column shows Breusch–Godfrey LM tests on serial correlation, and suggests that each dynamic equation is correctly specified. This aspect of the model is relevant to perform valid inference on θ_i (see appendix B). The second and third columns show the point estimates and p-values of γ_i and θ_i. It can be observed that in all equations, θ_i appears to be statistically insignificant, suggesting that the long-run homogeneity hypothesis ($\theta_i = 0$ for all i) cannot be rejected at conventional significance levels. On the other hand, the estimates of γ_i are statistically different from zero, except in the case of Lima. Thus, the first set of results suggests that the model described in section 3.3, with Lima as the reference region, cannot be rejected by the data. These findings hold for both the geographic and economic aggregation schemes.

The fourth column of Table 2 shows estimates and p-values of γ_i after imposing the long-run homogeneity restrictions $\theta_i = 0$ for all i. Again, all estimates but Lima's are statistically significant. Moreover, the point estimates of γ_i are considerably large for $i > 1$, ranging from about 0.65 to about 1.00 for the geographic aggregation and from about 0.75 to approximately 1.00 for the economic aggregation. This indicates that deviations from the relative PPP are corrected rather quickly. Geographic aggregation implies a somehow slower adjustment towards PPP, reflecting the influence of Lima's inflation (the common trend) in the "rest of the country" aggregates under the economic aggregation approach. All these results are confirmed in the last column of Table 2, which presents the final estimates and p-values of γ_i, after imposing both the long-run homogeneity restrictions and the fact that Lima serves as a reference region ($\gamma_1 = 0$).[5]

4.4 Regional dynamics and the importance of Lima

The estimation results reveal fast convergence of the regional inflations towards their own aggregates of external inflation, $y_{i,t} - x_{i,t}$. Since this phenomenon occurs in every region of the country, one could also expect a fast reversion towards the common trend. Figure 3 shows the responses of regional inflations relative to Lima ($y_{i,t} - y_{1,t}$), to a unit shock in Lima's inflation (the identified permanent shock of the system), under the geographic aggregation scheme (the results under the economic aggregation approach were very similar and are available upon request). As a summary, the last panel of Figure 3 shows the response of the average inflation of the $n = 9$ regions in the country. Due to the intrinsic dynamics implied by both the lags of the VECM and the sparsity of the matrix $\boldsymbol{W}$, convergence of $y_{i,t}$ towards the common trend $y_{1,t}$ is not as immediate as it is towards $x_{i,t}$. The point estimates of the impulse-response functions display hump shapes suggesting an initial overreaction of $y_{i,t}$ to the shock that eventually vanishes. Convergence is fast, nevertheless: the maximum deviation amounts to less than 0.4 percent in annual terms, and furthermore no response is statistically significant after 12 months of the occurrence of the shock (many of them are not significant even after 6 months). Thus, we may conclude that a shock to trend inflation will completely propagate to the individual regional inflations in less than a year.

On the other hand, as discussed in section 2, regional inflation rates are also affected by idiosyncratic shocks that may induce significant short-run deviations form a trend. Figure 4 shows the responses of regional inflations relative to Lima to a shock in region i that induces a deviation on impact of one percent, under the geographic aggregation scheme. Here, the shock is characterized by setting $\varepsilon_{i,0} > 0$ and $\varepsilon_{j,0} = 0$ for $i \neq j$ such that $y_{i,0} - y_{1,0} = 1$; then $\varepsilon_{i,t} = 0$ for all i and $t > 0$. It is observed that the effects of initial deviations of one percent are short-lived: in all cases, after three periods (months) of the shocks the deviation from Lima's inflation is less than 0.20 percent. Given the transitory nature of these shocks, few responses appear to be statistically significant after 12 months of the shock and none is significant after 18 months.

To assess the importance of the shock to the common trend, we perform a variance decomposition of the forecast

[5] Under geographic aggregation, the Wald statistic for the joint hypothesis that $\theta_i = 0$ for all i is $\omega_1 = 10.32$, below the 90% critical value $\chi^2_{10} = 14.68$. Also, the Wald statistic for the joint hypothesis that $\gamma_i = 0$ for $i > 1$ is $\omega_2 = 55.79$, well above the 99% critical value $\chi^2_9 = 21.66$. Under economic aggregation these statistics amount to $\omega_1 = 4.77$ and $\omega_2 = 78.69$.

errors in the VECM. Table 3 shows the contribution of the permanent shock for different forecasting horizons. The long-run $h \to \infty$ corresponds to the unconditional variance, and we first focus on these figures. The importance of the common trend cannot be exaggerated: under the geographic aggregation scheme, it contributes to at least 69.1 percent of the unconditional variance (in region 8) and up to almost 83.7 percent (in region 7). The contribution to average inflation is a stunning 89.7 percent. As expected, under economic aggregation the importance of Lima's inflation is further increased: the shock contributes to at least 74.6 percent of the unconditional variance (again in region 8), and more than 95 percent for the average inflation rate. The timing in the variance decomposition implies that the influence of the common trend shock is conspicuous one year ahead, whereas after two years the figures are close to the unconditional statistics. These findings point out that what happens in Lima serves as a benchmark in pricing decisions within the horizons that are relevant for monetary policy, say one to two years.[6]

5 Closing remarks

We have estimated and tested a multivariate dynamic model of regional inflations in Peru, in order to quantify the various feedbacks among regional inflation rates and, in particular, to assess the importance of Lima's inflation for the rest of the country. Our estimations suggest that the main features of the data can be well-described by a VECM with two important particularities. Firstly, the model is very parsimonious, as the dynamic feedback from the rest of the country to a specific regional inflation rate can be restricted to come from aggregates of neighboring regions (both in a geographic and in an economic sense) rather than from individual inflations. Secondly, the estimates of cointegrating vectors and their associated speed of adjustments provide evidence that the relative PPP holds among pairs of regional inflations, and that the only identifiable long-run forcing variable is Lima's inflation.

These findings lead us to conclude that Lima's inflation has served as a trend, an "attractor", for all remaining regional inflations. Furthermore, the typical deviation from this trend is likely to be short-lived and moderate. Since these estimates come from a sample where the BCRP has been actively and explicitly targeting Lima's inflation, it can be inferred that such practices have contributed to inflation stability in the whole country.

Looking forward, the economic importance of Lima in composing the national aggregate is expected to remain large in the medium term. Thus, the market may continue viewing Lima's inflation as a noise-free measure of national inflation, useful for price setting decisions. Therefore, in the event that the BCRP shifts to targeting national inflation, our findings indicate that the workings of monetary policy in Peru (especially the monetary policy lag) should not be significantly affected. It is unlikely that the adoption of a national inflation target would weaken the strong feedback from Lima's inflation to inflation elsewhere in the country.

Our results can be further refined. A referee suggested that, in line with findings in Monge and Winkelried (2009), the larger share of foodstuff in the CPI basket in provinces may be driving the results towards fast inflation convergence. This is mainly because food prices are subject to supply shocks that, despite their size, generally revert quickly. Hence, a natural extension of our work will be to repeat our analysis using somehow harmonized CPI across regions, for instance by using the same weights than Lima's CPIs prior to aggregation. Alternatively, the analysis can be performed using measures of *core* inflation, excluding volatile items such as foodstuff and fuels. Even though this extensions will surely render further insights to enhance our understanding of the complex transmission of regional shocks across the country, we do not expect our main qualitative results to change substantially.

[6] The case of region 8 deserves attention since it is the region where transitory shocks appear to be more important. It includes Cusco city which is by far the largest tourist destination in the country, and so its pricing dynamics (especially in services) is likely to be decoupled from the behavior in neighboring regional economies, Lima amongst them.

Technical appendix

Following theoretical results in Phillips (1986), Park and Phillips (1988) and Sims et. al. (1990), this appendix shows that both the aggregation and long-run homogeneity hypotheses can be tested using standard results from classical regression theory. In particular, it is shown that the statistics involved are asymptotically Gaussian even when the data are not stationary. This is the starting point for Wald coefficient tests that are asymptotically χ^2.

A Aggregation hypothesis

The VAR(p) model

$$\boldsymbol{y}_t = \boldsymbol{A}_1\boldsymbol{y}_{t-1} + \boldsymbol{A}_2\boldsymbol{y}_{t-2} + \ldots + \boldsymbol{A}_p\boldsymbol{y}_{t-p} + \boldsymbol{\varepsilon}_t \tag{17}$$

can be rewritten as

$$\boldsymbol{y}_t = \boldsymbol{C}_1\Delta\boldsymbol{y}_{t-1} + \boldsymbol{C}_2\Delta\boldsymbol{y}_{t-2} + \ldots + \boldsymbol{C}_{p-1}\Delta\boldsymbol{y}_{t-p+1} + \boldsymbol{H}\boldsymbol{y}_{t-1} + \boldsymbol{\varepsilon}_t\,, \tag{18}$$

where $\boldsymbol{H} = \boldsymbol{A}_1 + \boldsymbol{A}_2 + \ldots + \boldsymbol{A}_p$, $\boldsymbol{C}_1 = \boldsymbol{A}_1 - \boldsymbol{H}$ and $\boldsymbol{C}_r = \boldsymbol{C}_{r-1} + \boldsymbol{A}_r$ for $r = 2, \ldots, p-1$.

If $\boldsymbol{y}_t$ is stationary, then the least squares estimators of the $\boldsymbol{C}$ matrices and of $\boldsymbol{H}$ will be asymptotically Gaussian under standard conditions. On the other hand, following Sims et. al. (1990) and Hamilton (1994, section 18.2), when $\boldsymbol{y}_t$ is nonstationary, the least squares estimators of the $\boldsymbol{C}$ matrices in (18) are consistent and remain asymptotically Gaussian, whereas the corresponding estimator of $\boldsymbol{H}$ becomes superconsistent and have a nonstandard distribution characterized by functionals of Weiner processes. More formally, it is found that for $r = 1, 2, \ldots, p-1$

$$\sqrt{T}\text{vec}(\,\hat{\boldsymbol{C}}_r - \boldsymbol{C}_r\,) \xrightarrow{d} N(\boldsymbol{0}, \boldsymbol{V}_r) \qquad \text{whereas} \qquad \sqrt{T}\text{vec}(\,\hat{\boldsymbol{H}} - \boldsymbol{H}\,) = o_p(1)\,, \tag{19}$$

where vec($\boldsymbol{A}$) is the vector obtained by stacking the columns of matrix $\boldsymbol{A}$, and $\boldsymbol{V}_r$ denotes a positive definite covariance matrix.

The results in (19) are significant since they imply that hypothesis tests involving linear combinations of the $\boldsymbol{C}$ matrices and $\boldsymbol{H}$ are dominated asymptotically by the coefficients with the slower rate of convergence, namely those in the $\boldsymbol{C}$ matrices. In particular, tests involving linear combinations of the $\boldsymbol{A}$ matrices in (17) other than the sum $\boldsymbol{H}$ have the usual limiting distributions.

Recall that, for a given $r = 1, 2, \ldots, p$, the aggregation restrictions have the form $a_{ij}(r) - (w_{ij}/w_{ik})a_{ik}(r)$ for $j \neq k$ and $k \neq i$. These restrictions can be expressed more compactly as $\boldsymbol{R}_r\text{vec}(\boldsymbol{A}_r) = \boldsymbol{0}$, where $\boldsymbol{R}_r$ is a $p(n-2) \times n^2$ matrix whose typical row contains a 1, the ratio $-w_{ij}/w_{ik}$ and $n^2 - 2$ zeroes. In terms of the coefficientes of the reparameterized model (18), the aggregation restrictions are (consider $\boldsymbol{C}_p = \boldsymbol{0}$)

$$\begin{aligned} &\boldsymbol{R}_r\text{vec}(\,\boldsymbol{A}_r\,) \equiv \boldsymbol{R}_r\text{vec}(\,\boldsymbol{C}_r - \boldsymbol{C}_{r-1}\,) = \boldsymbol{0} \quad \text{for} \quad r = 2, \ldots, p\,, \\ &\boldsymbol{R}_1\text{vec}(\,\boldsymbol{A}_1\,) \equiv \boldsymbol{R}_1\text{vec}(\,\boldsymbol{C}_1 + \boldsymbol{H}\,) = \boldsymbol{0}\,. \end{aligned} \tag{20}$$

Therefore, using the results in (19), we have that under the null hypotheses in (20) (consider $\hat{\boldsymbol{C}}_p = \boldsymbol{0}$),

$$\begin{aligned} \sqrt{T}\boldsymbol{R}_r\text{vec}(\,\hat{\boldsymbol{C}}_r - \hat{\boldsymbol{C}}_{r-1}\,) &\xrightarrow{d} N(\boldsymbol{0}, \boldsymbol{\Omega}_r) \quad \text{for} \quad r = 2, \ldots, p\,, \\ \sqrt{T}\boldsymbol{R}_1\text{vec}(\,\hat{\boldsymbol{C}}_1 + \hat{\boldsymbol{H}}\,) = \sqrt{T}\boldsymbol{R}_1\text{vec}(\,\hat{\boldsymbol{C}}_1\,) + o_p(1) &\xrightarrow{d} N(\boldsymbol{0}, \boldsymbol{\Omega}_1)\,. \end{aligned} \tag{21}$$

where $\boldsymbol{\Omega}_r$ denote positive definite covariance matrices for $r = 1, 2, \ldots, p$. It follows that, regardless of the stationarity properties of the data, the statistics involved in testing aggregation hypothesis are asymptotically Gaussian, and thus the corresponding Wald tests have the usual asymptotic χ^2 distribution.

B Long-run homogeneity

Consider the stylized model, equation (8) with $p = 1$,

$$\Delta y_t = -\gamma(y_{t-1} - x_{t-1}) + \theta x_{t-1} + \varepsilon_t \,, \tag{22}$$

where $\Delta x_t = v_t$ and ε_t are stationary processes with zero mean and unconditional variances σ_v^2 and σ_ε^2, respectively. It is assumed that v_t and ε_t are uncorrelated at all leads and lags, which essentially implies that equation (22) is correctly specified. The arguments below are still valid for augmented equations that include lags of Δx_t and Δy_t, such that the uncorrelatedness between v_t and ε_t is guaranteed.

The purpose is to test $H_0 : \theta = 0$ in (22). For concreteness let $z_{t-1} = y_{t-1} - x_{t-1}$, which is stationary under H_0. It is not difficult to verify that, because $z_t \sim I(0)$ and $x_t \sim I(1)$,

$$\frac{1}{T}\frac{\sum_t x_{t-1}z_{t-1}\sum_t z_{t-1}\varepsilon_t}{\sum_t(z_{t-1})^2} = \frac{1}{T}\frac{O_p(T)O_p(\sqrt{T})}{O_p(T)} = o_p(1) \quad \text{and} \quad \frac{1}{T^2}\frac{(\sum_t x_{t-1}z_{t-1})^2}{\sum_t(z_{t-1})^2} = \frac{1}{T^2}\frac{O_p(T^2)}{O_p(T)} = o_p(1)\,. \tag{23}$$

On the other hand, let $W_\varepsilon(\cdot)$ and $W_v(\cdot)$ be two standard Weiner processes on $C(0,1)$, associated to the standardized series $\varepsilon_t/\sigma_\varepsilon$ and v_t/σ_v, respectively. These processes are uncorrelated and hence, due to the Gaussianity of the increments, independent. A standard result for integrated processes (cf Hamilton, 1994, section 17.5) is that

$$\frac{1}{T^2}\sum_t (x_{t-1})^2 \xrightarrow{d} \sigma_v^2 \int_0^1 W_v(r)^2 dr \equiv \mathcal{D}_2 \,, \tag{24}$$

whereas Phillips (1986, p. 327) and Park and Phillips (1988, Lemma 5.1) show that

$$\frac{1}{T}\sum_t x_{t-1}\varepsilon_t \xrightarrow{d} \sigma_v\sigma_\varepsilon \int_0^1 W_v(r)\,dW_\varepsilon(r) = N(0, \sigma_\varepsilon^2\,\mathcal{D}_2) \equiv \mathcal{D}_1 \,. \tag{25}$$

The above limiting distribution $\mathcal{D}_1$ is mixed Gaussian. This is to be understood as a normal distribution with variance proportional to $\mathcal{D}_2$ which is itself a random drawing from the space of positive scalars, in this case quadratic functionals of a Weiner process.

Finally, let s^2 denote the usual maximum likelihood estimator of the regression error variance, i.e. the average of the sum of squared residuals in regression (22), which is consistent for σ_ε^2. Upon gathering the results in (23), (24) and (25), the t statistic for testing $H_0 : \theta = 0$ satisfies, under this null hypothesis,

$$\tau = \frac{\sum_t x_{t-1}\varepsilon_t - \sum_t x_{t-1}z_{t-1}\sum_t z_{t-1}\varepsilon_t/\sum_t(z_{t-1})^2}{\sqrt{s^2\sum_t(x_{t-1})^2 - s^2(\sum_t x_{t-1}z_{t-1})^2/\sum_t(z_{t-1})^2}} = \frac{\sum_t x_{t-1}\varepsilon_t/T}{\sqrt{s^2\sum_t(x_{t-1})^2/T^2}} + o_p(1) \xrightarrow{d} \frac{\mathcal{D}_1}{\sqrt{\sigma_\varepsilon^2\,\mathcal{D}_2}} \equiv N(0,1)\,, \tag{26}$$

which follows from the continuous mapping and Cramér theorems. Therefore, the limiting distribution of τ is standard Gaussian, even though the least squares estimator of θ itself is not asymptotically Gaussian (the limiting distribution of $T\hat{\theta}$ is $\mathcal{D}_1/\mathcal{D}_2$). Similar arguments show that the asymptotic χ^2 statistics for tests for general restrictions on the coefficients θ across equations are also generated in the standard way.

References

Armas, A., L. Vallejos and M. Vega (2010), "Measurement of price indices used by the Central Bank of Peru", in Bank for International Settlements (ed.), *Monetary Policy and the Measurement of Inflation: Prices, Wages and Expectations*, Bank for International Settlements, 49, 259-283.

Beck, G. W., K. Hubrich and M. Marcellino (2009), "Regional inflation dynamics within and across Euro area countries and a comparison with the United States", *Economic Policy*, 24(57), 141–184.

Busetti, F., S. Fabiani and A. C. Harvey (2006), "Convergence of prices and rates of inflation", *Oxford Bulletin of Economics and Statistics*, 68(S1), 863-877.

Cecchetti, S. G., C. M. Nelson and R. J. Sonora (2002), "Price index convergence among United States cities", *International Economic Review*, 43(4), 1081-1099.

Elliott, G. and A. Fatas (1996), "International business cycles and the dynamics of the current account", *European Economic Review*, 40(2), 361-387.

Gonzales de Olarte, E. (2003), *Regiones Integradas: Ley de Incentivos para la Integración y Conformación de Regiones. Lineamientos Económicos y Políticos*, Lima: Fondo Editorial del Congreso del Perú.

Gonzalo, J. and C. W. J. Granger (1995), "Estimation of common long-memory components in cointegrated systems", *Journal of Business and Economic Statistics*, 13(1), 27-35.

Gonzalo, J. and S. Ng (2001), "A systematic framework for analyzing the dynamic effects of permanent and transitory shocks", *Journal of Economic Dynamics and Control*, 25(10), 1527-1546.

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

Monge, A. and D. Winkelried (2009), "PPC y convergencia de precios: Un análisis para las ciudades del Perú", Banco de la República, *Ensayos de Política Económica*, 27(58), 56-105.

Park, J. Y. and P. C. B. Phillips (1988), "Statistical inference in regressions with integrated processes: Part 1", *Econometric Theory*, 4(3), 468-497.

Parsley, D. C. and S. J. Wei (1996), "Convergence to the law of one price without trade barriers or currency fluctuations", *Quarterly Journal of Economics*, 111(4), 1211-1236.

Pesaran, M. H., T. Schuermann and S. M. Weiner (2004), "Modeling regional interdependencies using a global error-correcting macroeconometric model", *Journal of Business and Economic Statistics*, 22(2), 129-162.

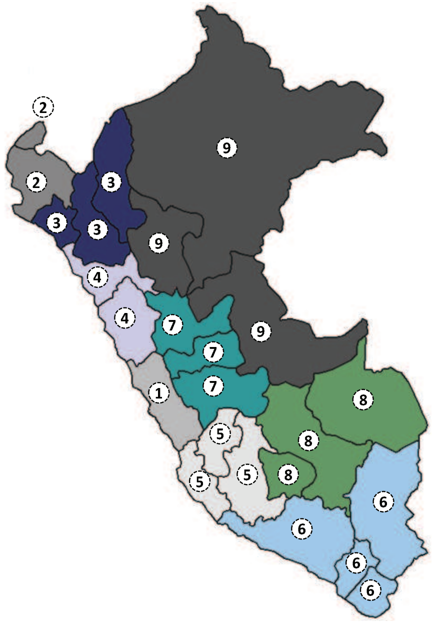
Phillips, P. C. B. (1986), "Understanding spurious regressions in econometrics", *Journal of Econometrics*, 33(3), 311-340.

Rogers, J. H. (2007), "Monetary union, price level convergence, and inflation: How close is Europe to the USA?", *Journal of Monetary Economics*, 54(3), 785–796.

Rossini, R. and M. Vega (2008), "The monetary policy transmission mechanism under financial dollarisation: The case of Peru 1996 - 2006", in Bank for International Settlements (ed.), *Transmission Mechanisms for Monetary Policy in Emerging Market Economies*, Bank for International Settlements, 35, 395-412.

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

Figure 1. *Economic regions in Peru*



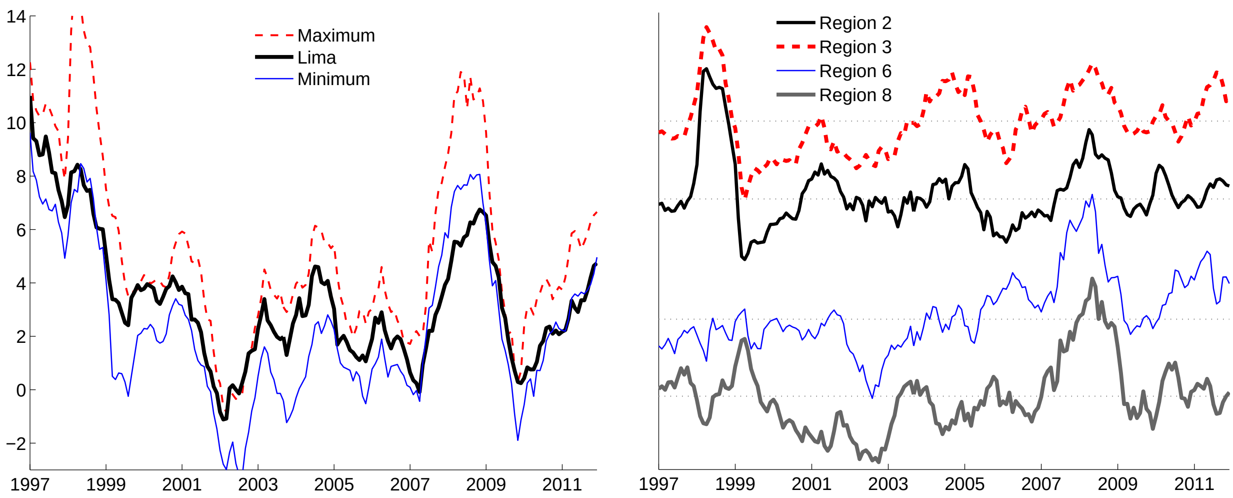
1 Lima (66.02)
2 Piura, Tumbes (2.57)
3 Lambayeque, Cajamarca, Amazonas (3.83)
4 La Libertad, Ancash (7.25)
5 Ica, Ayacucho, Huancavelica (2.21)
6 Arequipa, Moquegua, Tacna, Puno (7.57)
7 Junín, Pasco, Huánuco (3.07)
8 Cusco, Apurímac, Madre de Dios (3.55)
9 Loreto, San Martín, Ucayali (3.93)

Note: Classification of Peruvian Departments into 9 economic regions, following Gonzales de Olarte (2003, p. 41). The numbers in parentheses are the weights inflation in each region receives in composing national inflation, according to INEI.

Figure 2. *Regional inflation (year-on-year percent changes), 1997 - 2011*

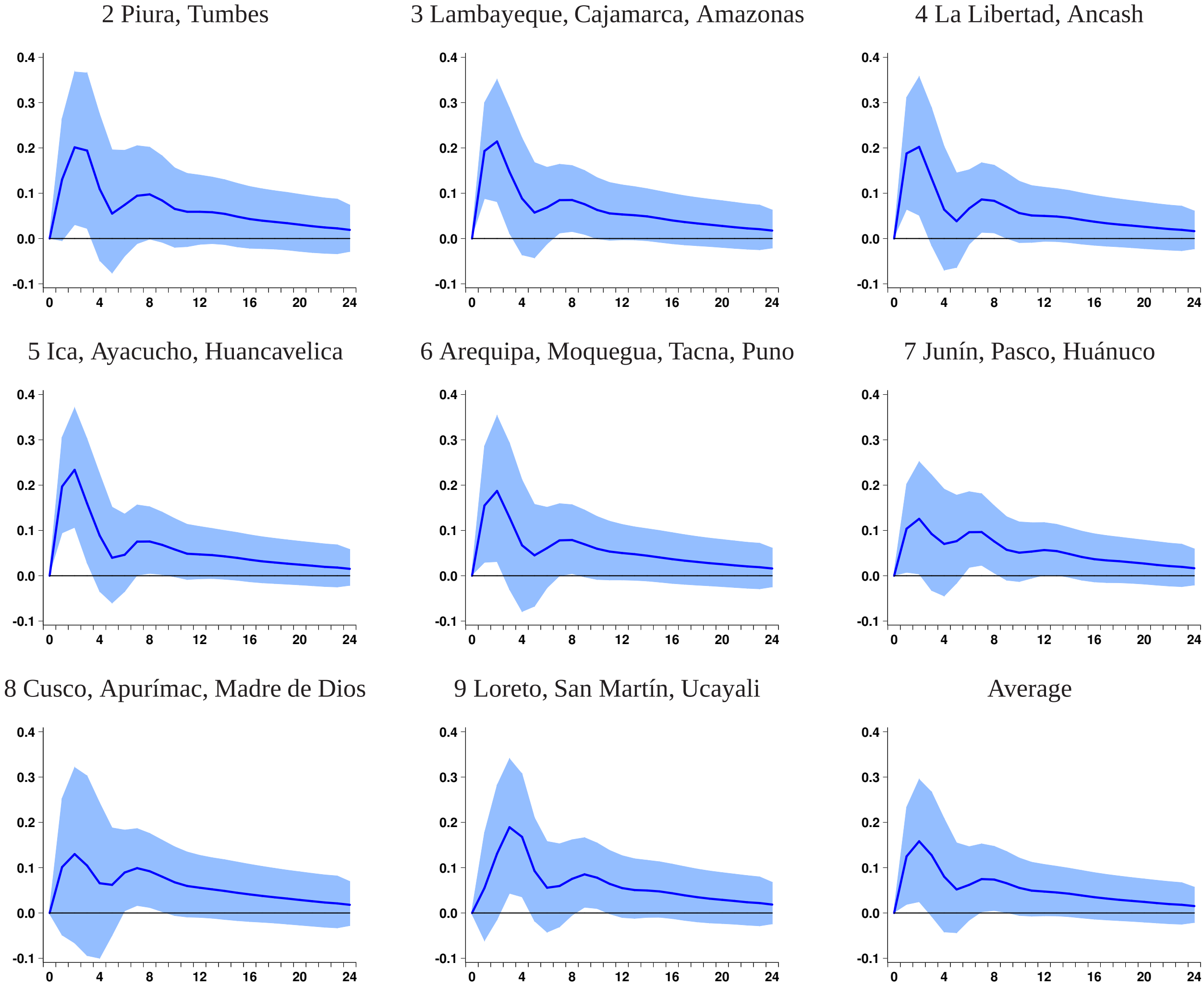
(a) Lima, maximum and minimum

(b) Deviations from national inflation



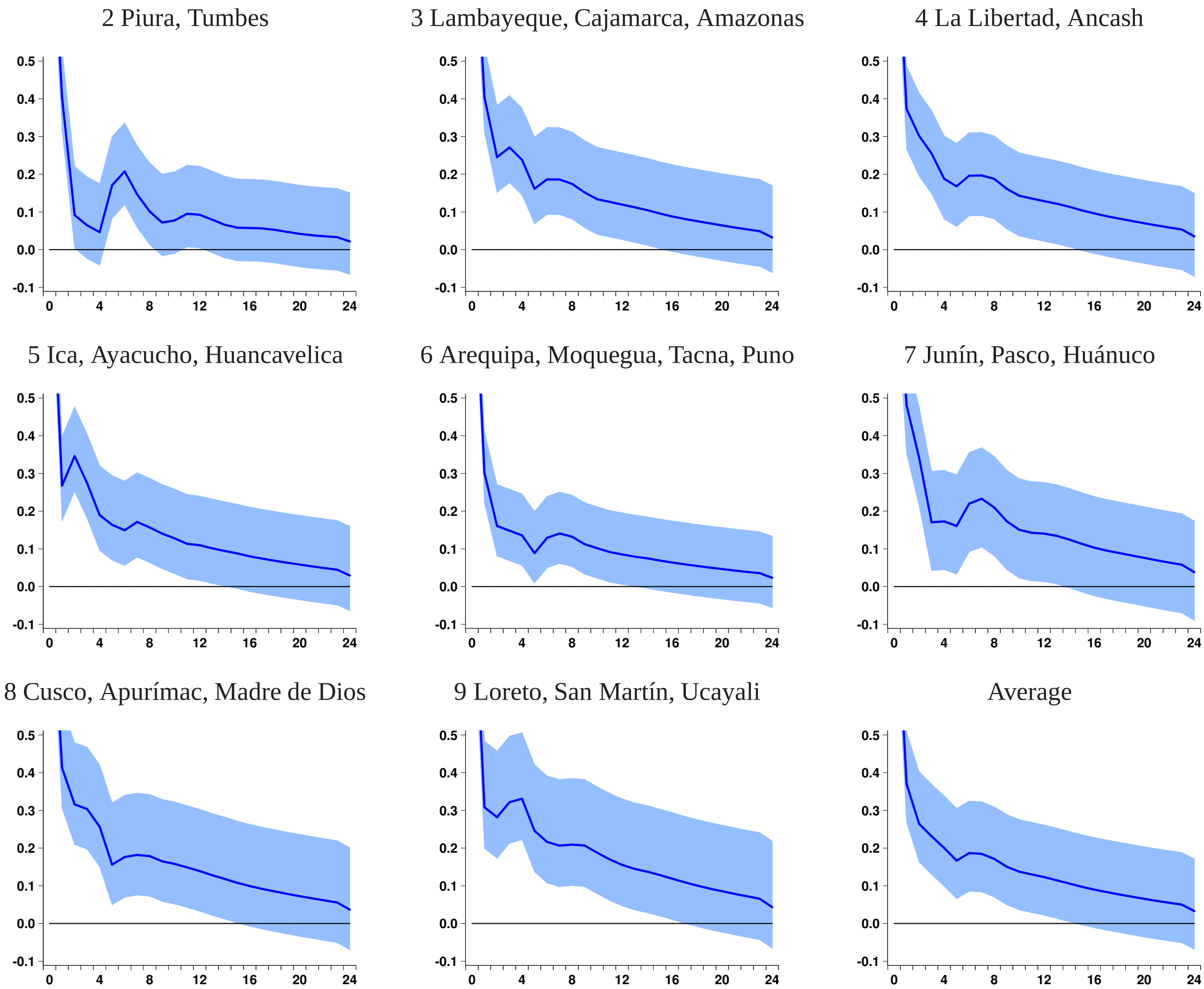
Notes: For a given period, panel (a) shows Lima's inflation and the maximum and minimum rates among the remaining 8 regions. Constants were added to the series in panel (b) to ease visualization. For each series the vertical scale ranges from −2.5 to about 4.5 percent, and the dashed horizontal lines are the corresponding zero lines.

Figure 3. *Response of regional inflations relative to Lima's to a shock in Lima (geographic aggregation)*



Notes: Each graph shows the response of $y_{i,t} - y_{1,t}$ (in annualized percent terms) to a unit shock in Lima, identified as discussed in section 3.3. Figures in the horizontal axis are months after the occurrence of the shock in period 0. Bootstrap 90 percent confidence intervals are depicted as shaded areas. The scale of the vertical axis is the same for all graphs.

Figure 4. *Response of regional inflations relative to Lima's to a regional shock (geographic aggregation)*



Notes: Each graph shows the response of $y_{i,t} - y_{1,t}$ (in annualized percent terms) to a shock in region i such that $y_{i,0} - y_{1,0} = 1$. Figures in the horizontal axis are months after the occurrence of the shock in period 0. Bootstrap 90 percent confidence intervals are depicted as shaded areas. The scale of the vertical axis is the same for all graphs.

Table 1. *Testing for aggregation*

	Geographic aggregation		*Economic aggregation*	
	χ^2 statistic	p-value	χ^2 statistic	p-value
1 Lima	31.654	0.630	32.295	0.599
2 Piura, Tumbes	35.446	0.447	38.924	0.297
3 Lambayeque, Cajamarca, Amazonas	27.838	0.800	31.513	0.637
4 La Libertad, Ancash	28.050	0.792	30.108	0.703
5 Ica, Ayacucho, Huancavelica	39.261	0.285	39.527	0.275
6 Arequipa, Moquegua, Tacna, Puno	28.221	0.785	31.129	0.656
7 Junín, Pasco, Huánuco	29.078	0.749	26.446	0.850
8 Cusco, Apurímac, Madre de Dios	37.764	0.344	36.229	0.411
9 Loreto, San Martín, Ucayali	35.291	0.454	38.136	0.329

Notes: Wald tests (asymptotically χ^2) for $H_0 : \delta_{ij}(r) = 0$ in equation (6), for all $i = 1, \ldots, n$, $j \neq i$ and $r = 1, \ldots, p$, and for different aggregation schemes. The lag length was set to $p = 5$, which was found to be the minimum value to render serially uncorrelated residuals in all equations (both in the unrestricted and in the restricted models). In all tests, the number of degrees of freedom is $p(n-2) = 35$.

Table 2. *Estimation results*

	Breusch–Godfrey LM test	γ_i	θ_i	γ_i ($\theta_i = 0$)	γ_i ($\theta_i = \gamma_1 = 0$)
			Geographic aggregation		
1 Lima	0.074 (0.929)	0.548 (0.122)	0.274 (0.109)	0.399 (0.084)	–
2 Piura, Tumbes	2.036 (0.134)	1.128 (0.014)	0.194 (0.449)	1.056 (0.015)	1.059 (0.016)
3 Lambayeque, Cajamarca, Amazonas	0.474 (0.623)	0.769 (0.011)	0.224 (0.107)	0.743 (0.010)	0.787 (0.005)
4 La Libertad, Ancash	0.038 (0.963)	0.855 (0.034)	0.148 (0.545)	0.755 (0.028)	0.763 (0.022)
5 Ica, Ayacucho, Huancavelica	0.177 (0.838)	0.800 (0.058)	0.255 (0.108)	0.718 (0.041)	0.729 (0.032)
6 Arequipa, Moquegua, Tacna, Puno	1.478 (0.231)	0.997 (0.007)	0.197 (0.311)	0.982 (0.003)	1.000 (0.003)
7 Junín, Pasco, Huánuco	1.044 (0.354)	0.681 (0.023)	0.165 (0.211)	0.645 (0.020)	0.647 (0.025)
8 Cusco, Apurímac, Madre de Dios	2.046 (0.133)	0.692 (0.016)	0.057 (0.778)	0.635 (0.015)	0.625 (0.018)
9 Loreto, San Martín, Ucayali	3.390 (0.076)	0.806 (0.001)	0.245 (0.216)	0.704 (0.002)	0.647 (0.004)
			Economic aggregation		
1 Lima	0.084 (0.920)	0.125 (0.696)	0.178 (0.210)	0.071 (0.756)	–
2 Piura, Tumbes	1.318 (0.271)	1.044 (0.022)	0.043 (0.885)	0.969 (0.008)	0.960 (0.006)
3 Lambayeque, Cajamarca, Amazonas	0.238 (0.788)	0.900 (0.008)	0.099 (0.607)	0.890 (0.004)	0.873 (0.005)
4 La Libertad, Ancash	0.121 (0.886)	0.873 (0.008)	0.037 (0.876)	0.800 (0.005)	0.793 (0.005)
5 Ica, Ayacucho, Huancavelica	2.157 (0.119)	0.969 (0.008)	0.153 (0.382)	0.985 (0.002)	0.968 (0.002)
6 Arequipa, Moquegua, Tacna, Puno	1.242 (0.292)	0.848 (0.003)	0.112 (0.601)	0.842 (0.001)	0.834 (0.001)
7 Junín, Pasco, Huánuco	0.551 (0.577)	0.873 (0.001)	0.084 (0.510)	0.878 (0.000)	0.866 (0.000)
8 Cusco, Apurímac, Madre de Dios	1.351 (0.262)	0.787 (0.002)	0.021 (0.913)	0.750 (0.000)	0.749 (0.000)
9 Loreto, San Martín, Ucayali	1.707 (0.185)	0.806 (0.013)	0.166 (0.330)	0.790 (0.007)	0.787 (0.006)

Notes: SUR estimates. The Breusch–Godfrey is asymptotically distributed as a χ^2 variate with one degree of freedom under the null hypothesis of no first-order serial correlation in the regression errors. Figures in parentheses are p-values. In the case of coefficient estimates, the p-values are for the null hypothesis that the corresponding coefficient is equal to zero.

Table 3. *Contribution of a shock in Lima (permanent) to the forecast error variance*

	$h = 0$	$h = 1$	$h = 6$	$h = 12$	$h = 18$	$h = 24$	$h \to \infty$
				Geographic aggregation			
2 Piura, Tumbes	51.8	55.5	60.3	65.9	70.3	73.7	77.2
3 Lambayeque, Cajamarca, Amazonas	56.5	63.2	68.8	73.9	77.6	80.6	83.3
4 La Libertad, Ancash	49.3	57.3	64.3	70.6	75.1	78.4	80.9
5 Ica, Ayacucho, Huancavelica	56.5	61.8	67.8	74.0	78.2	81.3	83.3
6 Arequipa, Moquegua, Tacna, Puno	44.4	50.0	57.9	66.4	72.0	76.0	78.2
7 Junín, Pasco, Huánuco	53.7	56.7	66.3	73.9	78.8	82.2	83.7
8 Cusco, Apurímac, Madre de Dios	18.2	25.1	39.9	53.2	61.2	66.8	69.1
9 Loreto, San Martín, Ucayali	40.0	47.3	60.3	68.7	74.1	77.9	80.1
Average	76.1	76.3	79.0	82.7	85.5	87.6	89.7
				Economic aggregation			
2 Piura, Tumbes	50.5	57.3	64.4	71.4	76.1	79.4	81.2
3 Lambayeque, Cajamarca, Amazonas	56.4	66.2	74.8	81.1	84.8	87.3	88.3
4 La Libertad, Ancash	49.7	59.3	68.3	75.7	80.3	83.3	84.6
5 Ica, Ayacucho, Huancavelica	58.5	68.4	77.1	83.3	86.8	89.1	89.8
6 Arequipa, Moquegua, Tacna, Puno	46.2	53.8	65.1	73.6	78.9	82.3	83.6
7 Junín, Pasco, Huánuco	58.2	62.7	77.6	84.9	88.7	91.0	91.4
8 Cusco, Apurímac, Madre de Dios	21.0	28.0	46.0	59.5	67.6	73.0	74.6
9 Loreto, San Martín, Ucayali	41.3	50.8	67.7	77.0	82.2	85.4	86.5
Average	78.2	82.2	88.0	91.8	93.7	94.9	95.3

Notes: The shock to Lima's inflation is identified as discussed in section 3.3. The figures show the percentage of the forecast error variance that is attributable to this shock, for various forecasting horizons h. The case $h \to \infty$ corresponds to the contribution to the unconditional variable of the forecast error.