The long-run relationship between money and prices in Mexico: 1969-2010.

Gomez-Ruano, Gerardo

Universidad Iberoamericana

2014

The long-run relationship between money and prices in Mexico: 1969-2010

Gerardo Gomez-Ruano
Department of Economics, Universidad Iberoamericana, DF, Mexico
E-mail: ggomezr@alum.bu.edu

This paper performs a spectral analysis (univariate and bivariate) of monthly series of growth in money (a narrow and a broad aggregate) and in prices for Mexico. This analysis allows the identification of the most important frequencies for each series, as well as of some measures of association between the series, at different frequencies. In particular, zero frequency measures, typically used for identifying the long-run relationship between money and prices, are obtained. In addition to the analysis of the entire series (1969-2010), a rolling sample analysis for the zero frequency is also carried out to allow for changes in the long-run relationships.

I. Introduction

The quantitative theory of money and prices is probably one of the oldest theories in Economics. Various historians attribute it to David Hume, but there are hints of it for earlier times (see, for example, Robbins (1998)).

In more recent times, the study by Vogel (1974) and a well-known article by Lucas (1980) provided empirical evidence that supported this theory. Nevertheless, the results by Lucas have been the target of major scrutiny, showing certain instability in the relationship between money and prices.

Some economists, like Sargent and Surico (2011), have proposed models to explain the aforementioned instability, while others, like Benati (2009), have spent time documenting the long-run behavior of this relationship for a good many of countries and time periods. These studies of long-run behavior are mainly done with tools from spectral analysis, which determine the behavior of time-series at the lowest frequencies. Such a study has never been done for Mexico, partly because of the complexity of this technique.

The main purpose of this paper is to briefly introduce the methodology of spectral analysis and the resulting estimations for Mexico, with monthly data from 1969 to 2010.

Section II introduces basic concepts from spectral analysis, which are needed to understand the results; section III presents the data; section IV presents the results obtained for the entire sample; and section V presents the results from a rolling sample. The paper ends with a brief discussion and conclusion; and a small appendix contains the estimators that were employed.

II. Elements of Spectral Analysis[1]

There are generally two approaches for analyzing stationary stochastic processes: the first and most common is the "time domain" approach, the second and least common is the "frequency domain" approach.

The "frequency domain" approach is also known as spectral analysis, due to the fact that light is composed of electromagnetic waves of different frequencies and the continuum of such frequencies is called the spectrum.

Stationary processes

It is assumed that the reader is familiar with the concept of random variable. It is then needed to define what a stationary process is. For practical purposes, we will limit ourselves to univariate processes (those of only one variable) which occur in discrete time (integer units of time) and whose realizations are real numbers (and not, say, complex numbers).

We define a stochastic process, $\{X_t\}$, as a family of random variables indexed by the symbol t. We assume that t belongs to the set of integers and that, for all t, X_{t} is a random variable with density function $f_t(\cdot)$.

We define the mean of X_t, μ_t, as

$$\mu_t = E[X_t] = \int_{-\infty}^{\infty} x f_t(x) dx$$

[1] This is an informal presentation, which follows Priestley (1981), Fuller (1996), Brockwell and Davis (1991), and Box et al. (2008).

and the autocovariance of X_t, $cov(X_t, X_s)$, as

$$cov(X_t, X_s) = E[(X_t - \mu_t)(X_s - \mu_s)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x_1 - \mu_t)(x_2 - \mu_s)f_{t,s}(x_1, x_2)dx_1dx_2$$

where $f_{t,s}(\cdot,\cdot)$ is the joint density function of (X_t, X_s).

We define a stationary stochastic process, or stationary process, as a stochastic process, $\{X_t\}$, with $\mu_t = \mu$, and $cov(X_t, X_{t+d}) = cov(X_{t+k}, X_{t+k+d})$ for all t, d and k integer numbers.

It follows from the definition of stationary process that the mean of a stationary process is a constant μ, and that the autocovariance can be expressed as a function that depends solely on the distance d.

Examples:

i. White Noise

   A stochastic process $\{\varepsilon_t\}$ is called white noise when it is composed of a sequence of uncorrelated random variables whose mean and variance are the same:

$$E[\varepsilon_t] = \mu$$

$$covar(\varepsilon_t, \varepsilon_{t+d}) = \begin{cases} 0, & d \neq 0 \\ \sigma^2, & d = 0 \end{cases}$$

   Clearly, white noise is a stationary process.

ii. Autoregressive process of order 1, also known as AR(1)

   A stochastic process $\{U_t\}$ is called autoregressive of order 1 if

$$U_t = aU_{t-1} + \varepsilon_t$$

   where a is a constant and $\{\varepsilon_t\}$ is white noise.

   If $|a| < 1$ and the white noise has a zero mean, $E[\varepsilon_t] = 0$, then $\{U_t\}$ is a stationary process:

$$U_t = \sum_{s=0}^{\infty} a^s \varepsilon_{t-s}$$

$$E[U_t] = E\left[\sum_{s=0}^{\infty} a^s \varepsilon_{t-s}\right]$$

$$= \sum_{s=0}^{\infty} a^s E[\varepsilon_{t-s}]$$

$$= \sum_{s=0}^{\infty} a^s \cdot 0$$

$$= 0$$

$$covar(U_t, U_{t+d}) = E[U_t, U_{t+d}]$$

$$= E\left[\sum_{s=0}^{\infty}\sum_{r=0}^{\infty} a^s a^r \varepsilon_{t-s}\varepsilon_{t+d-r}\right]$$

We can simplify the double summation since all the terms $a^s a^r \varepsilon_{t-s}\varepsilon_{t+d-r}$ with $t - s \neq t + d - r$ have a zero mean (recall that $covar(\varepsilon_t, \varepsilon_{t+d}) = 0$ for $d \neq 0$). Hence, we only need to consider terms with $t - s = t + d - r$, or equivalently $s = r - d$. We can therefore substitute s in terms of r or vice versa. For convenience, we will keep only one summation going from zero to infinity; so we will make the substitution depending on whether d is positive or negative. If $d \leq 0$ we have that $d = -|d|$. Substituting s according to $s = r - d$ we remain with a single summation from zero to infinity:

$$E\left[\sum_{s=0}^{\infty}\sum_{r=0}^{\infty} a^s a^r \varepsilon_{t-s}\varepsilon_{t+d-r}\right] = E\left[\sum_{r=0}^{\infty} a^{r+|d|} a^r \varepsilon_{t-|d|-r}\varepsilon_{t-d-r}\right]$$

$$= E\left[\sum_{r=0}^{\infty} a^{2r+|d|} \varepsilon_{t-|d|-r}^2\right]$$

$$= \sigma^2 \sum_{r=0}^{\infty} a^{2r+|d|}$$

$$= \sigma^2 \frac{a^{|d|}}{1 - a^2}$$

If $d > 0$, the result is the same except for the use of the absolute value; hence the expression using the absolute value of d is correct.

Frequencies

So far we have considered processes that are defined on the time domain. However, it also possible to define processes on the frequency domain.

When we speak of frequencies, we are implicitly talking about a repeating event (a periodic event). The frequency is nothing more than the amount of times that such an event happens per unit of time.

In mathematics, the most popular periodic functions are the trigonometric functions sine and cosine, $\sin(x)$ and $\cos(x)$ respectively. We will later see how these two functions constitute the building blocks for any stationary process.

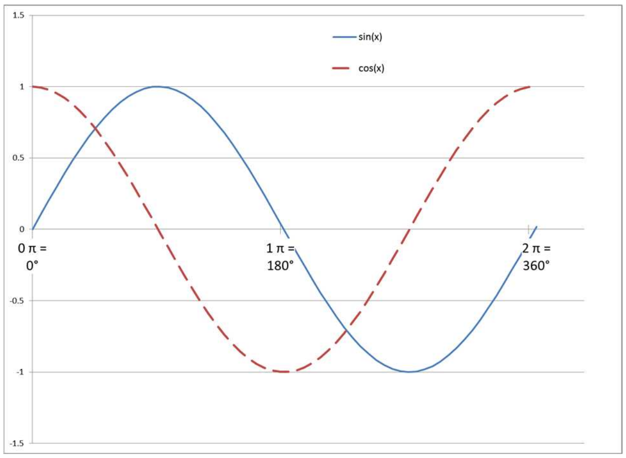

Fig. 1. Sine and cosine functions

Although trigonometric functions can be computed for angular degrees (for example, the sine of 45 degrees or 45°), the common practice is to use "radians" as their argument. The conversion from one unit to the other is fairly straightforward since 2π radians equal 360 angular degrees (a complete turn) and both units are directly proportional (see the accompanying graph, which has the unit conversion on the axis).

Sinusoidal processes

Let us begin by considering a stochastic process with sine and cosine functions:

$$Y_t = u\cos(t) + v\sin(t)$$

where u and v are random variables with the same variance σ^2, zero mean, and zero covariance. $\{Y_t\}$ is a stationary process since

$$\begin{aligned} E[Y_t] &= E[u\cos(t) + v\sin(t)] \\ &= E[u]\cos(t) + E[v]\sin(t) \\ &= 0 \cdot \cos(t) + 0 \cdot \sin(t) \\ &= 0 \end{aligned}$$

That is, the unconditional mean is constant. And using the trigonometric identity $\cos(x-y) = \cos(x)\cos(y) + \sin(x)\sin(y)$ together with the zero covariance assumption (i.e., $E[uv] = 0$), we have that

$$E[Y_t Y_{t+d}] = E[(u\sin(t) + v\cos(t))(u\sin(t+d) + v\cos(t+d))]$$

$$= E[u^2\sin(t)\sin(t+d)$$
$$+ uv(\sin(t)\cos(t+d) + \cos(t)\sin(t+d)) + v^2\cos(t)\cos(t+d)]$$

$$= E[u^2]\sin(t)\sin(t+d) + E[uv](\sin(t)\cos(t+d) + \cos(t)\sin(t+d))$$
$$+ E[v^2]\cos(t)\cos(t+d)$$

$$= \sigma^2\sin(t)\sin(t+d) + \sigma^2\cos(t)\cos(t+d)$$

$$= \sigma^2 \cos(d)$$

In other words, the autocovariance depends only on the distance d. Notice that for $d = 0$ we get the variance of Y_t, which is σ^2 (and not $2\sigma^2$) since $\sin^2(x) + \cos^2(x) = 1$ for all x. That is, whenever the sine grows, the cosine shrinks such that the sum of both variances is constant and equal to one.

In order to convert the time units into radians, one multiplies t by ω, where $\omega \equiv 2\pi f$ and f is the amount of cycles per time unit:

$$Y_t = u\cos(\omega t) + v\sin(\omega t)$$

$$= u\cos(2\pi f t) + v\sin(2\pi f t)$$

In this way we see that Y_t has a frequency of f cycles per time unit, and our proof for stationarity remains valid since we did not require t or d to be integers.

Let $\{\omega_n\}_{n=1}^{\infty}$ be a sequence of positive distinct constants, and let $\{a_n\}_{n=1}^{\infty}$, $\{b_n\}_{n=1}^{\infty}$ be two real-valued stochastic processes with zero mean, zero covariance, and variance σ_n^2. Then, subject to some regularity conditions, the process

$$W_t = \sum_{n=1}^{\infty} (a_n \cos(\omega_n t) + b_n \sin(\omega_n t))$$

is a stationary process. Thus, the stationary process $\{W_t\}$ is basically a sum of sinusoidal functions whose coefficients form a bivariate stochastic process.

In what follows, it will prove convenient to use the following complex exponential notation (this notation is based upon the polar form of complex numbers[2]), which is equivalent to the long sine and cosine notation:

[2] That is, $z = a + ib = r(\cos\theta + i\sin\theta) = re^{i\theta}$, where $r = |z| = \sqrt{a^2 + b^2}$, $\cos\theta = \frac{a}{r}$, and $\sin\theta = \frac{b}{r}$.

$$W_t = \sum_{k=-\infty}^{\infty} A_k e^{i\, sgn(k)\omega_{|k|} t}$$

with

$$A_k = \begin{cases} \frac{1}{2}(a_k - ib_k), & k \geq 1 \\ 0, & k = 0 \\ \frac{1}{2}(a_{|k|} + ib_{|k|}), & k \leq -1 \end{cases}$$

and where i is the imaginary unit with $i^2 = -1$, and $sgn(x)$ is the sign of x. Thus, we simply construct the complex-valued stochastic process$\{A_k\}_{k=-\infty}^{\infty}$ from our previous sequence $\{\omega_n\}_{n=1}^{\infty}$ and our previous real-valued stochastic processes $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$. The mean of a complex-valued random variable $Z = a + ib$ is $\mu_Z = E[Z] = E[a] + iE[b]$. While the variance is $var(Z) = E[(Z - \mu_Z)(Z^* - \mu_Z)]$, where Z^* is the complex conjugate of Z; that is, if $Z = a + ib$ then $Z^* = a - ib$.

It follows that $\{A_k\}_{k=-\infty}^{\infty}$ is a complex-valued stochastic processes with zero mean and variance $var(A_k) = \frac{1}{2}\sigma_{|k|}^2$ for all $k \neq 0$. Moreover, since the summation is symmetric, the amount of variance contributed by the frequency ω_n is $var(A_n) + var(A_{-n}) = \sigma_n^2$. The representations are thus equivalent.

An important characteristic of the sinusoidal process, which is inherited by the spectral representation that will be seen later, is that the variance of the stochastic process W_t equals the sum of the variances at each distinct frequency:

$$var(W_t) = var(W_0) = var\left(\sum_{k=-\infty}^{\infty} A_k\right) = var\left(\sum_{n=1}^{\infty} a_n\right) = \sum_{n=1}^{\infty} \sigma_n^2$$

Notice that we use the property whereby the variance of a stationary process is the same for all t and therefore the same as that for $t = 0$. Since $\sin(0) = 0$, only the coefficients of the cosine functions remain. And because their covariance is zero, we get only the sum of their variances as the variance of the stationary process W_t.

The alias effect and the frequency continuum

Whenever a stationary process is observed exclusively for discrete units of time, it is impossible to distinguish some frequencies (high ones) from others (lower ones). This is known as the "alias" effect. An intuitive way to understand the alias effect is by following this line of reason: A necessary condition to identify a periodic movement is the ability to observe at least two different states of this movement, say, a low and high state. Unfortunately, this is not a sufficient condition, since there might be movements with a higher frequency (for example, twice the frequency) which coincide for both the only states that we are able to observe.

A clear example of the alias effect can be seen in Fig. 2:

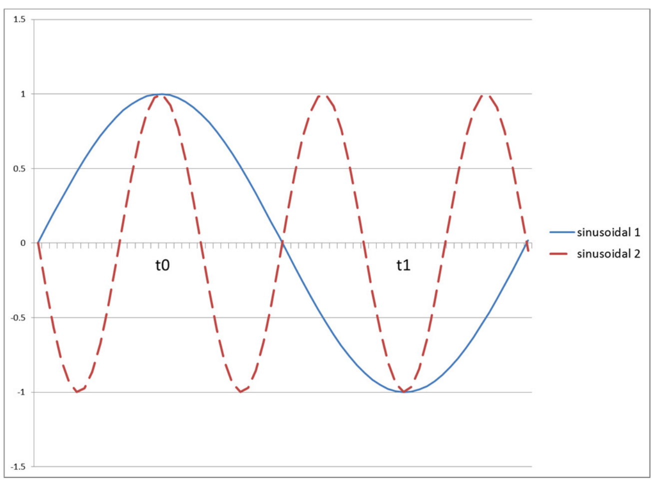

Fig. 2. An example of the alias effect

In this graph, we notice that eliciting observations at t_0 and t_1 is not enough to distinguish the sinusoidal 1 from the sinusoidal 2, because both take the same values at those points. In other words, the frequency of sinusoidal 1 is an alias of the frequency of sinusoidal 2.

It follows from the aforementioned observations that whenever we are working with discrete processes, we are limited to some maximum observable frequency ω^*. If the time units employed are the same as those of the observations, as is common practice, then this maximum observable frequency is $\omega^* = \pi$. In other words, the shortest cycles that we are able to tell are those which take at least two time units.[3]

One might be tempted to conclude that, since it is hard to identify high frequencies, it should be equally hard to identify non-integer frequencies: Is it possible to identify cycles of length 2.4 or $\sqrt{10}$ months? The answer is that *it is* possible.

In summary, we are in principle able to identify any frequency ω in the interval $(0, \pi]$. This is not only possible but desirable because, as we shall see, it turns out that every discrete-time stationary process can be described in terms of this continuum of frequencies.

Spectral representation of stationary processes

An important result of the field of stochastic processes is the spectral representation theorem for stationary processes.

Starting from our previous representation,

$$W_t = \sum_{k=-\infty}^{\infty} A_k e^{i\, sgn(k) \omega_{|k|} t}$$

With $\{A_k\}_{k=-\infty}^{\infty}$ being a complex-valued stochastic process, we can imagine a generalized version where every frequency ω in the interval $(0, \pi]$ is taken into account. The spectral representation theorem states that every discrete-time stationary process with zero mean can be represented in this way. In other words, it can be represented as

[3] This maximum observable frequency is also known as the Nyquist frequency.

$$X_t = \int_{-\pi}^{\pi} e^{it\omega} dZ(\omega)$$

Where $Z(\omega)$ is an orthogonal complex-valued stochastic process satisfying $dZ^*(\omega) = dZ(-\omega)$, and $E[dZ(\omega)] = 0$. Notice that $\omega < 0$ is only an artifact so that, together with the symmetry of dZ, we can obtain real coefficients (just as in our last representation of W_t).

The similarity between the sinusoidal process and the spectral representation is evident; $dZ(\omega)$ works as the coefficient A_n, and both are complex-valued stochastic processes (one is defined over the set of integers and the other over a continuous interval from $-\pi$ to π).

Just as we did earlier for W_t, we can also compute the variance for X_t:

$$var(X_t) = var(X_0) = var\left(\int_{-\pi}^{\pi} dZ(\omega)\right) = \int_{-\pi}^{\pi} E[|dZ(\omega)|^2]$$

Defining

$$dH(\omega) = E[|dZ(\omega)|^2]$$

we have that $dH(\omega)$ is the variance due to the frequency ω. If $H(\omega)$ is differentiable, then $dH(\omega) = h(\omega)d\omega$, and $h(\omega)$ is called "the power". Thus, $h(\omega_0)$ is interpreted as the variance of the random coefficients in the neighborhood of ω_0.[4]

Example:

i. The power of a white noise process

   The stochastic process $\{\varepsilon_t\}$ gets the name "white noise" from the fact that it is made of all the frequencies in the same intensity. That is,

---

[4] Recall that if a function is differentiable, it means its derivative is continuous. And continuity of a function $f(\cdot)$ implies that for every $\epsilon > 0$ and x_0, there is a neighborhood of x_0, such that $|f(x) - f(x_0)| < \epsilon$ for all x in that neighborhood.

$h(\omega) = \frac{\sigma^2}{2\pi}$, where $-\pi \leq \omega \leq \pi$, and σ^2 is the variance of ε_t.

Every frequency ω contributes in the same amount, just as every color (which is just an electromagnetic wave at a specific frequency) contributes to the white color in the same amount. For purposes of displaying the power we use $\sigma^2 = 1$, so that the area under the "curve" (meaning, the total variance) equals one.

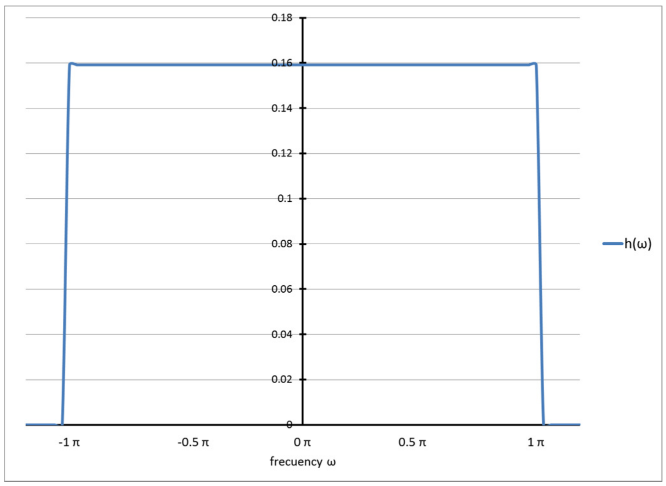

Fig. 3. The power of a white noise process

ii. The power of an autoregressive of order 1

Assuming that the process is stationary ($|a| < 1$ and $E[\varepsilon_t] = 0$), the power of an AR(1) depends on the value of a and the frequency ω:

$h(\omega) = \frac{\sigma^2}{2\pi(1-2a\cos(\omega)+a^2)}$, where $-\pi \leq \omega \leq \pi$, and σ^2 is the variance of ε_t.

Whenever $a > 0$, the power of the low frequencies is greater than the power of the high frequencies; the opposite is true for $a < 0$. For purposes of displaying the power we use $\sigma^2 = 1$ and $a^2 = 0.5^2 = 0.25$. This implies an area (total variance) of $var(U_t) = \frac{\sigma^2}{1-a^2} = \frac{3}{4}$.

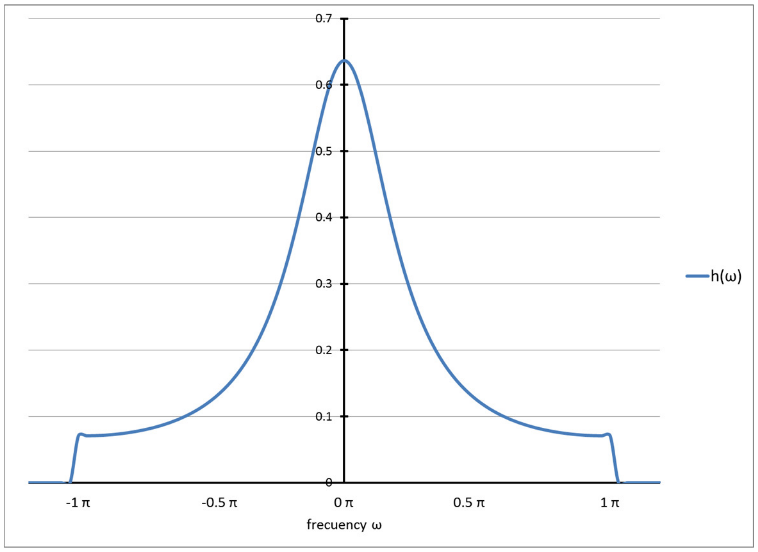

Fig. 4. An example of the power for an AR(1) process

Since the power (sometimes called "power spectrum'', or simply "spectrum") of every real-valued process is symmetrical, we might just as well display only the positive frequencies. Another common practice, to be employed later on, is to display the natural logarithm of the power, instead of the power itself. This is done mainly because oftentimes the power of some frequencies is several orders of magnitude greater than that of others. Finally, for ease of interpretation, we will use the length of the cycles, instead of their frequency, on the horizontal axis.

Coherency and gain

Starting off from the spectral representation, one is able to derive different measures of association between two stochastic processes. Two of those measures which will be used throughout are the coherency and the gain.

Suppose we have two stationary processes $\{X_{1,t}\}$ and $\{X_{2,t}\}$; then, by the spectral representation theorem, we can express them as

$$X_{1,t} = \int_{-\pi}^{\pi} e^{it\omega} dZ_1(\omega)$$

$$X_{2,t} = \int_{-\pi}^{\pi} e^{it\omega} dZ_2(\omega)$$

The coherency $\mathcal{C}_{12}(\omega)$ at a frequency ω between two stationary processes $\{X_{1,t}\}$ and $\{X_{2,t}\}$; is defined as

$$\mathcal{C}_{12}(\omega) = \left|\frac{cov\big(dZ_1(\omega), dZ_2(\omega)\big)}{[var(dZ_1(\omega))var(dZ_2(\omega))]^{1/2}}\right|$$

The use of the absolute value\norm is due to the fact that the covariance between two complex-valued random variables is, in general, complex as well. The coherency $\mathcal{C}_{12}(\omega)$ can be interpreted, grossly speaking, as the correlation between the random coefficients of the processes $\{X_{1,t}\}$ and $\{X_{2,t}\}$, at the frequency ω. In fact, the coherency satisfies the inequality $0 \leq \mathcal{C}_{12}(\omega) \leq 1$. It is, grossly speaking, a measure of linear association.

The gain $\mathcal{G}_{12}(\omega)$ of the process $\{X_{1,t}\}$ over the process $\{X_{2,t}\}$ at frequency ω is defined by

$$\mathcal{G}_{12}(\omega) = \left|\frac{cov(dZ_1(\omega), dZ_2(\omega))}{var(dZ_2(\omega))}\right|$$

Once again, we use the norm because this covariance is, in general, complex-valued. The gain $\mathcal{G}_{12}(\omega)$ can be interpreted, grossly speaking, as the least-squares estimate for the slope of a linear regression where $dZ_1(\omega)$ is the dependent variable and $dZ_2(\omega)$ is the independent variable.

Unlike the coherency, the gain is---in general---asymmetric. That is, $\mathcal{G}_{12}(\omega) \neq \mathcal{G}_{21}(\omega)$.

III. Data

We briefly introduce the data in this section. The series of the two monetary aggregates for Mexico are available online from the 1980s onwards.[5] The author requested the earlier

[5] At the mexican central bank's website.

observations directly from the central bank. The monetary aggregates are called M1a and M4a. Both of them have the letter "a" because they include the public sector. M1a includes (1) coins and bills in the hands of the public, (2) checkable deposits at domestic banks/subsidiaries, (3) demand deposits at domestic banks/subsidiaries, and (4) demand deposits at credit unions. M4a includes M1a plus (1) domestic securities in hands of residents, (2) domestic securities in hands of non-residents, (3) deposits made at mexican branches/subsidiaries abroad. Both monetary aggregates are available from 1960 onwards and have a monthly frequency.

The Consumer Price Index (CPI) for Mexico is available online from the mexican statistics bureau.[6] The series begins around 1969, and is available at a monthly frequency as well. Hence our data is restricted to 1969 through 2010.[7]

Inflation

Starting off from the CPI series, we obtain the monthly logarithmic change from 1969 through 2010. These data are presented in the following figure.

[6] This bureau is called the Instituto Nacional de Geografía y Estadística (INEGI).
[7] The CPI is known in Mexico as the *Indice Nacional de Precios al Consumidor* (INPC).

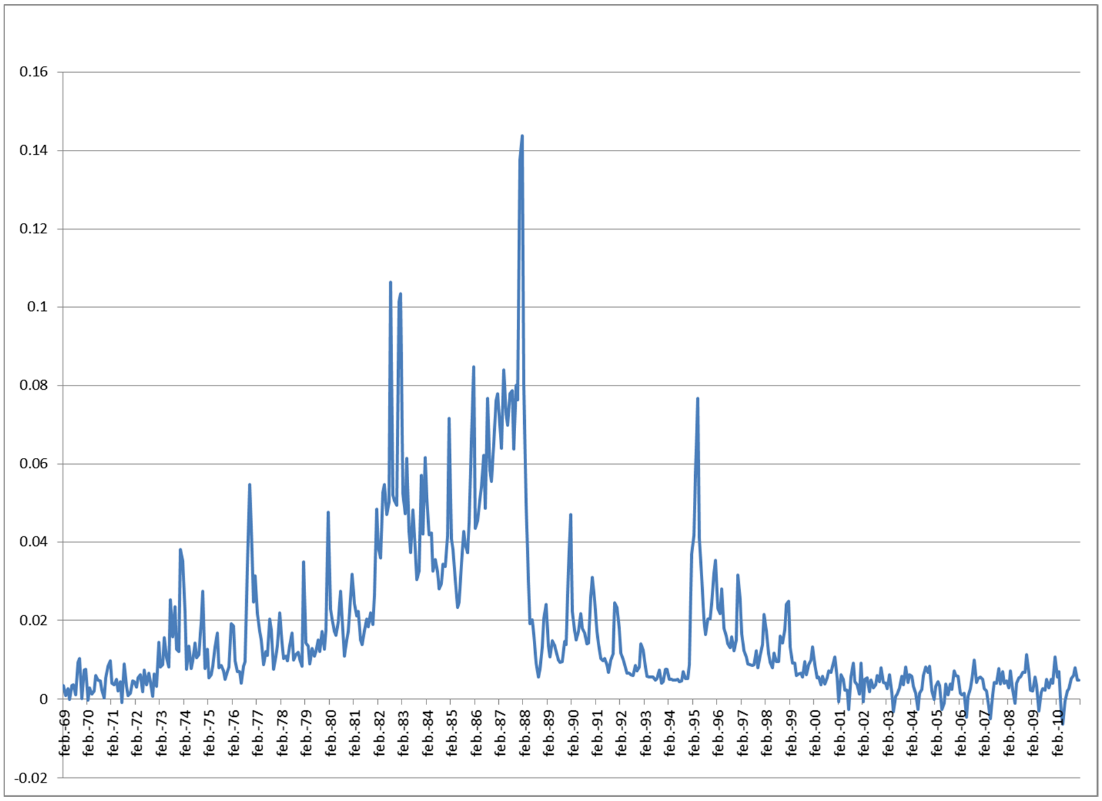

Fig. 5. CPI monthly log-change

The behavior of inflation over the last four decades has been analyzed by different economic historians (see, for example, Moreno-Brid and Bosch (2010), or Kuntz (2010)). An interesting feature, hitherto ignored by other authors, is that while the annual inflation rates of the last years is similar to that of 1969-1972, the seasonality of the monthly inflation is much greater. This is likely the result of the stability in the (implicit/explicit) target for the interest rate that the mexican central bank has kept throughout the last years.

The growth of M1a

Both monetary aggregates are transformed using a monthly logarithmic change, as was done for the price index.

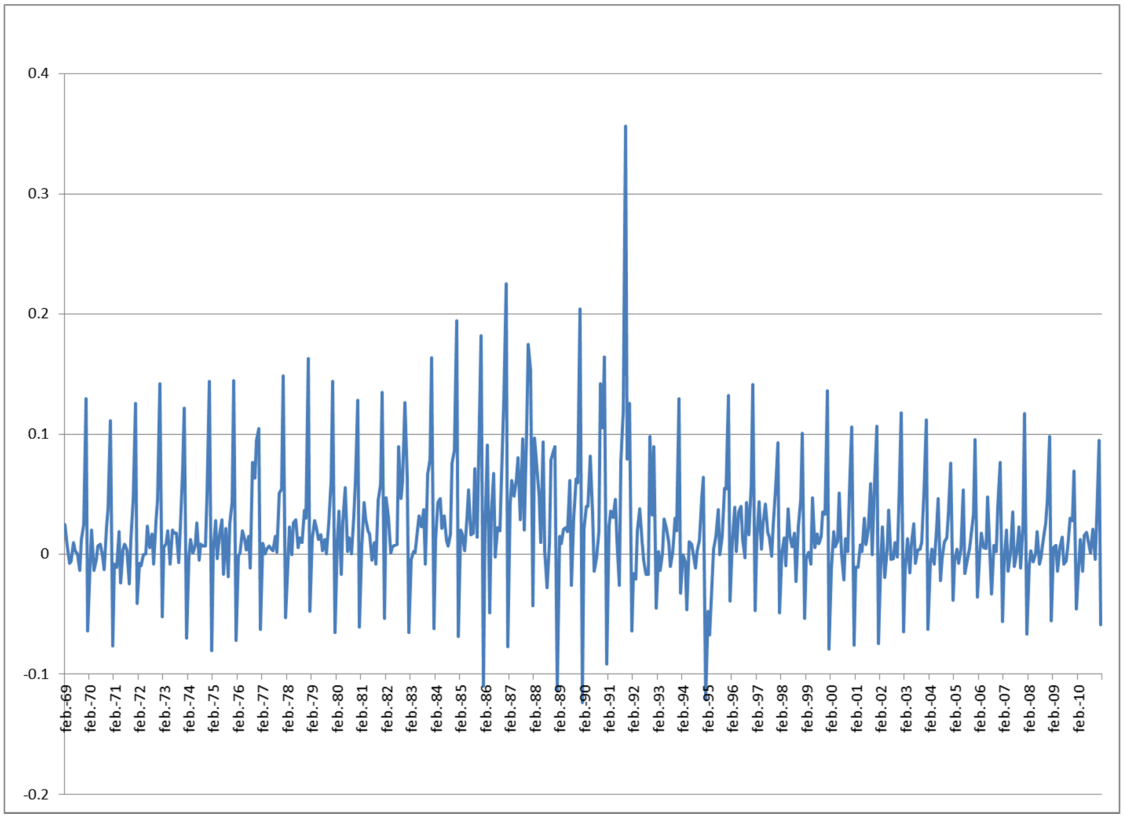

Fig. 6. M1a monthly log-change

The growth of M4a

The series for M4a is visibly "heavier" (i.e., it has more inertia, it takes longer for it to get back to the mean), and it is closer in behavior to the price series.

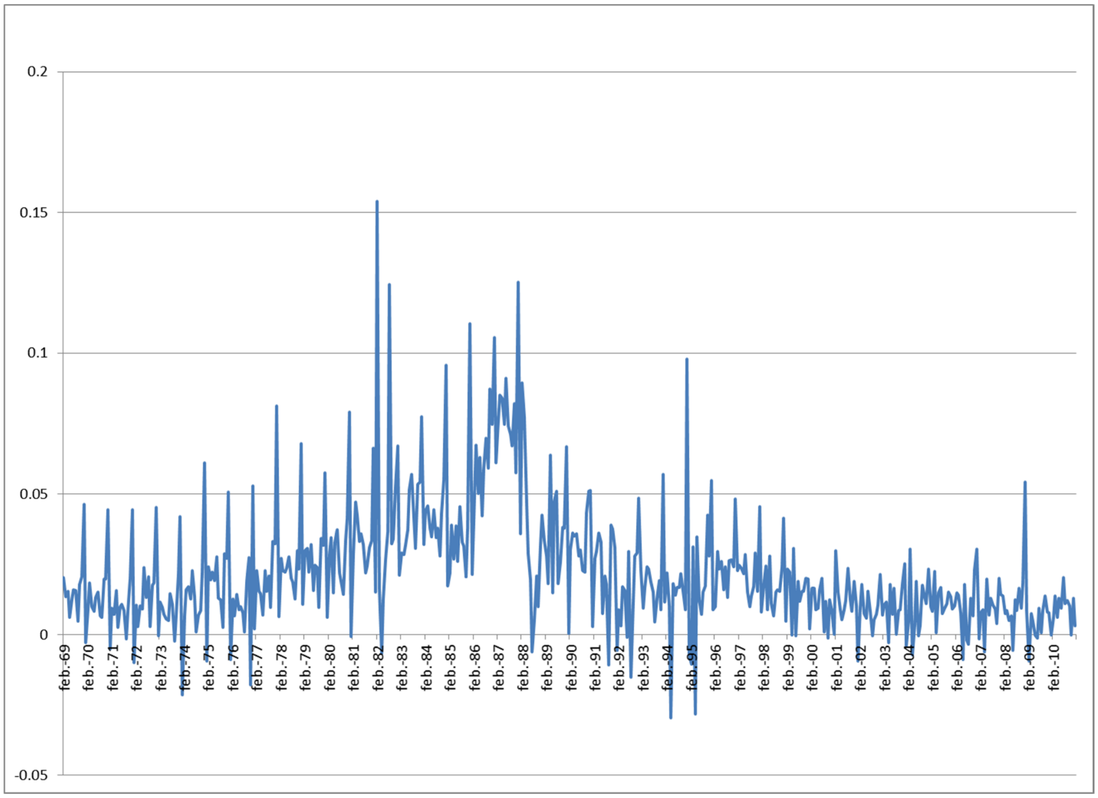

Fig. 7. M4a monthly log-change

Indeed, M4a appears to be an hybrid between the price index and M1a. Not surprisingly, this crude characterization will become apparent in our quantitative analysis as well.

IV. Results

For the sake of exposition, the rest of the paper (sections IV and V) will employ the terms "prices" and "monetary aggregates" when referring to the respective monthly log-changes. Also, keep in mind that what follows are all estimates, even if we do not always explicitly mention it.

We now present the three main results of the quantitative analysis: the spectra for each series (prices, M1a, M4a), and the coherencies and gains for the duplets (prices, M1a) and (prices, M4a).

The specifics of our methodology are somewhat technical, and therefore found at the end.

The power

The power of a time series summarizes a great amount of information in a non-parametric fashion.

We have chosen to display the logarithm of the power in our figures, for the sake of exposition.

a) *The power of prices.* The power estimate for prices has an expected shape: that of a heavy series whose lowest frequencies embody most of the total variance. "the typical spectral shape of an economic variable" in the words of Granger (1966).

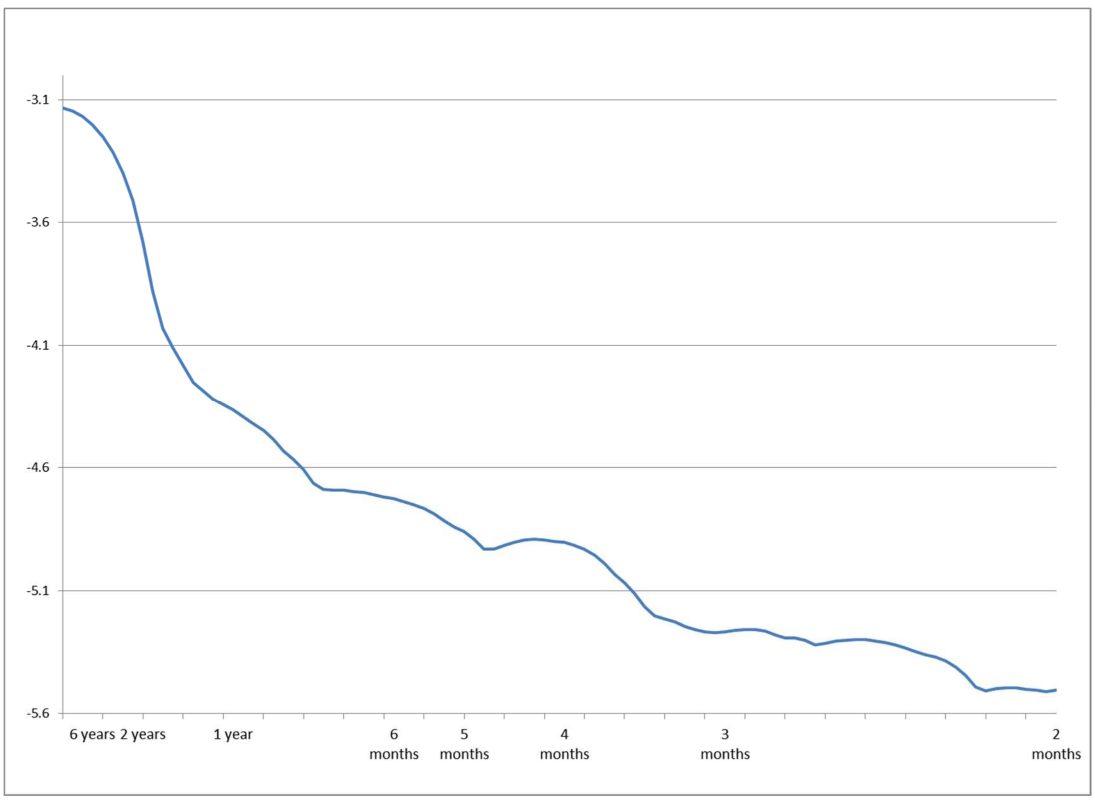

Fig. 8. Log of the power estimate for prices

Thanks to the properties of logarithms, we can see that the cycles with a length of around 6 years explain 200% more of the total variance than those with a length of around 3 months.

b) *The power of M1a.* The power of M1a is rather different from that of the other two series: it does not show the typical spectral shape of an economic variable.

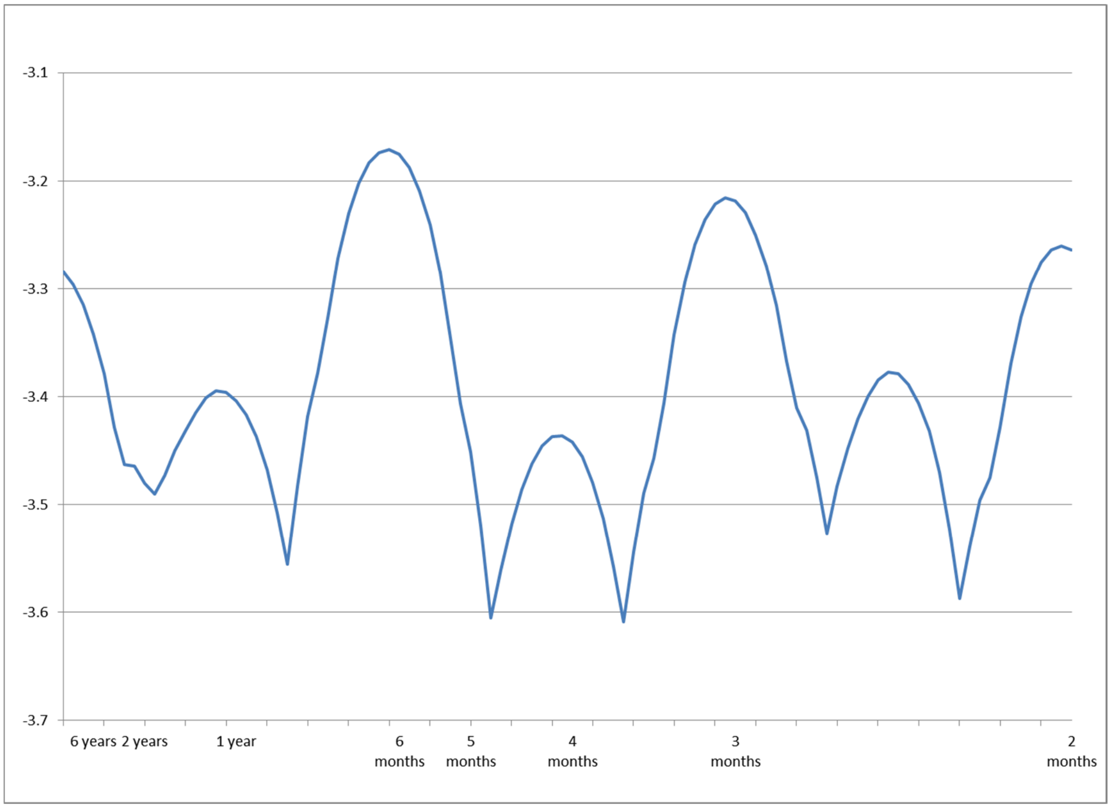

Fig. 9. Log of the power estimate for M1a

On the contrary, it is a much "lighter" series, with a heavy seasonal behavior. Cycles with length of around 6, 3, and 2 months explain more than those close to one or more years.

c) *The power of M4a.* Finally, the estimated spectral shape of M4a is somewhere in between that of M1a and prices: it a very "heavy" series, just like prices, but with a certain seasonality around 6, 3, 2.4, and 2 months.[8]

[8] A seasonality around 2.4 months may seem awkward. It is actually not because 2.4 months are one fifth of a year. In other words, this is a harmonic frequency of the yearly frequency. Together, the cycles of length 6, 3, 2.4, and 2 months, help describe common yet not uniform seasonalities.

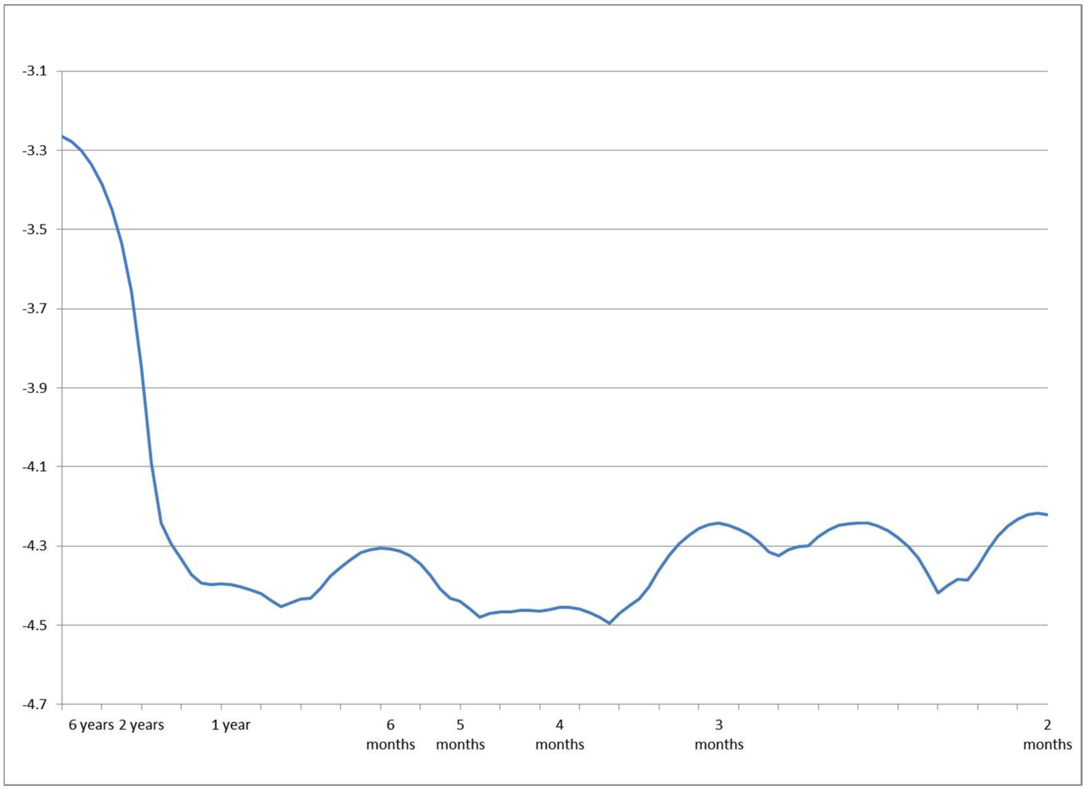

Fig. 10. Log of the power estimate for M4a

Unlike the power estimate for prices though, the power estimate for M4a shows a greater variance for cycles with length of 3 months or less than for cycles with length around 4 months or a year.

The coherency

The coherency was estimated for the duplets (M1a, prices) and (M4a, prices). Just like the correlation coefficient, the coherency is a symmetric measure: the coherency between M1a and prices is the same as that between prices and M1a. However, the coherency is not negative because it takes the absolute value to turn complex numbers into real ones.

a) *Coherency between M1a and prices.* The coherency estimate is statistically not different from zero for the highest frequencies, modest for intermediate frequencies, and considerable for very low frequencies.

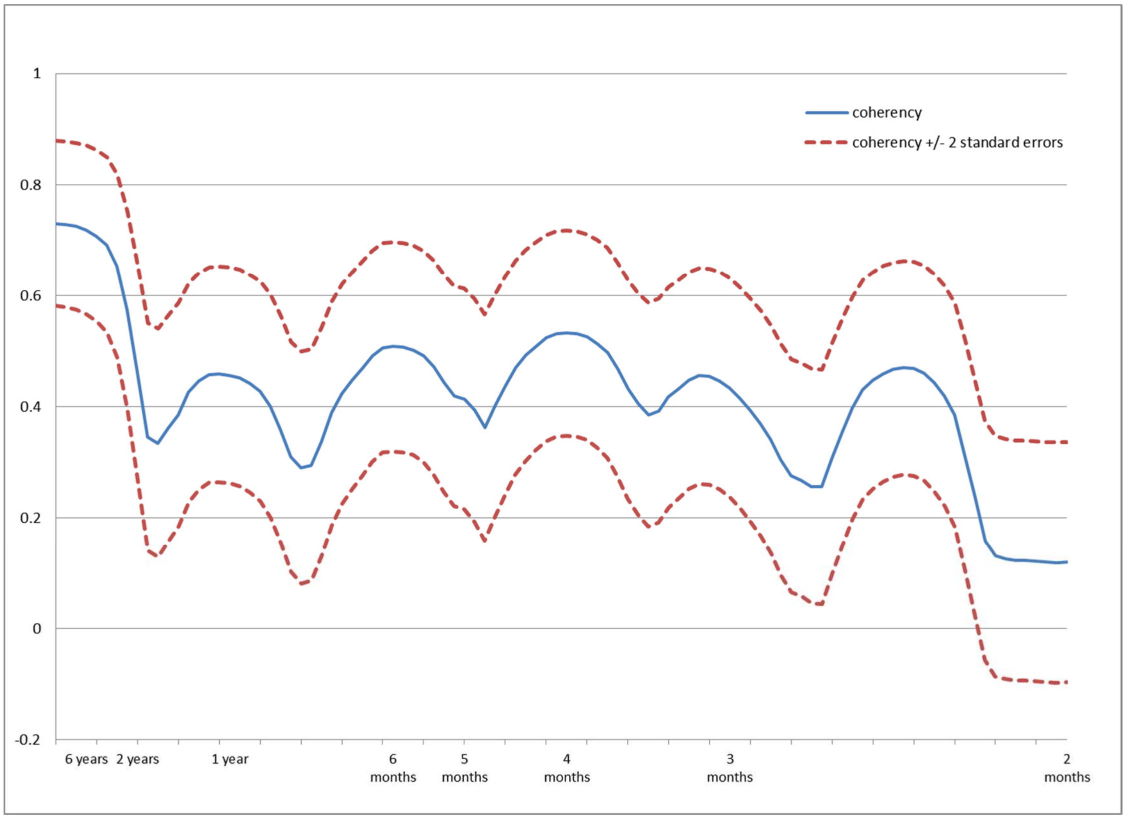

Fig. 11. Coherency estimate for (M1a, prices)

We conclude that the relationship between M1a and prices is close to linear only for cycles of length greater than or equal to 2 years.

b) *Coherency between M4a and prices.* For the coherency estimate of (M4a, prices), the relationship for cycles with length of 2 or more years is almost linear: the coherency estimate is statistically greater than 0.8. Moreover, for cycles with lengths of 6 years or more, the coherency is statistically not different from 1.

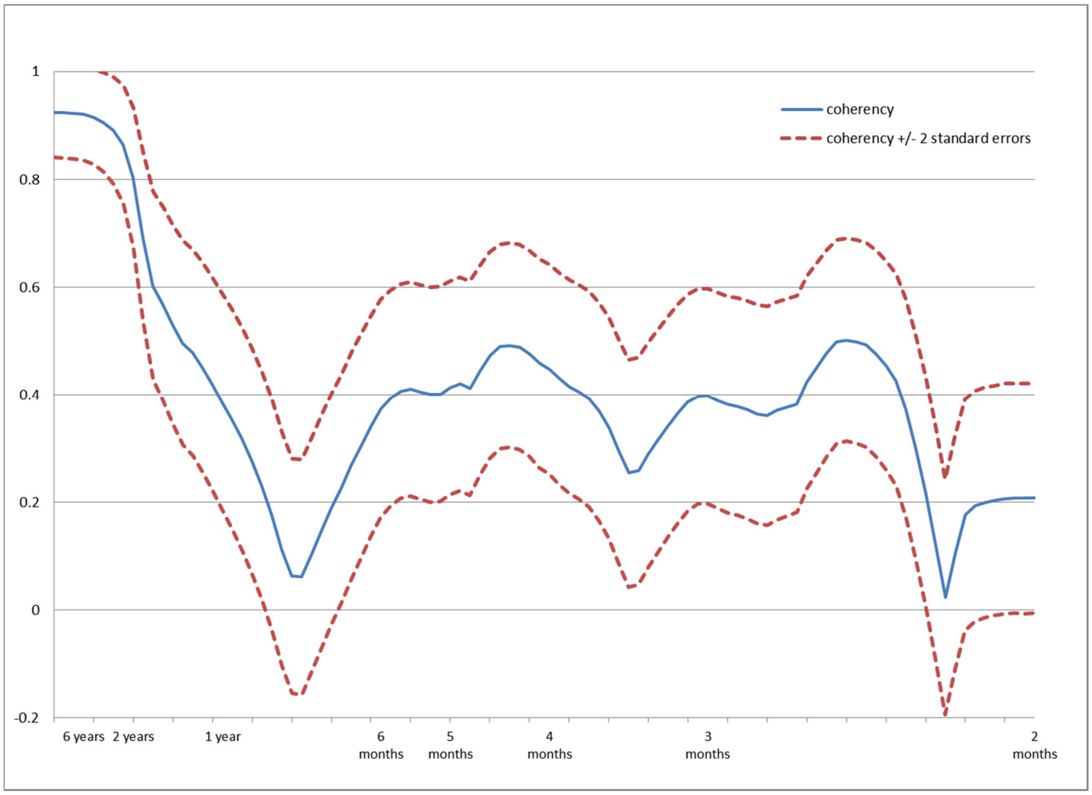

Fig. 12. Coherency estimate for (M4a, prices)

Other cycles have a modest or statistically insignificant coherency.

The gain

As previously said, the gain can be interpreted---grossly speaking---as the slope of a linear regression from the random coefficients of one process over those of the other. In our two following estimations, we have proceeded as if prices were the left-hand-side (or independent) variable.

a) *The gain of prices over M1a.* As the figure shows, our gain estimate has a very clear shape in the case of M1a:

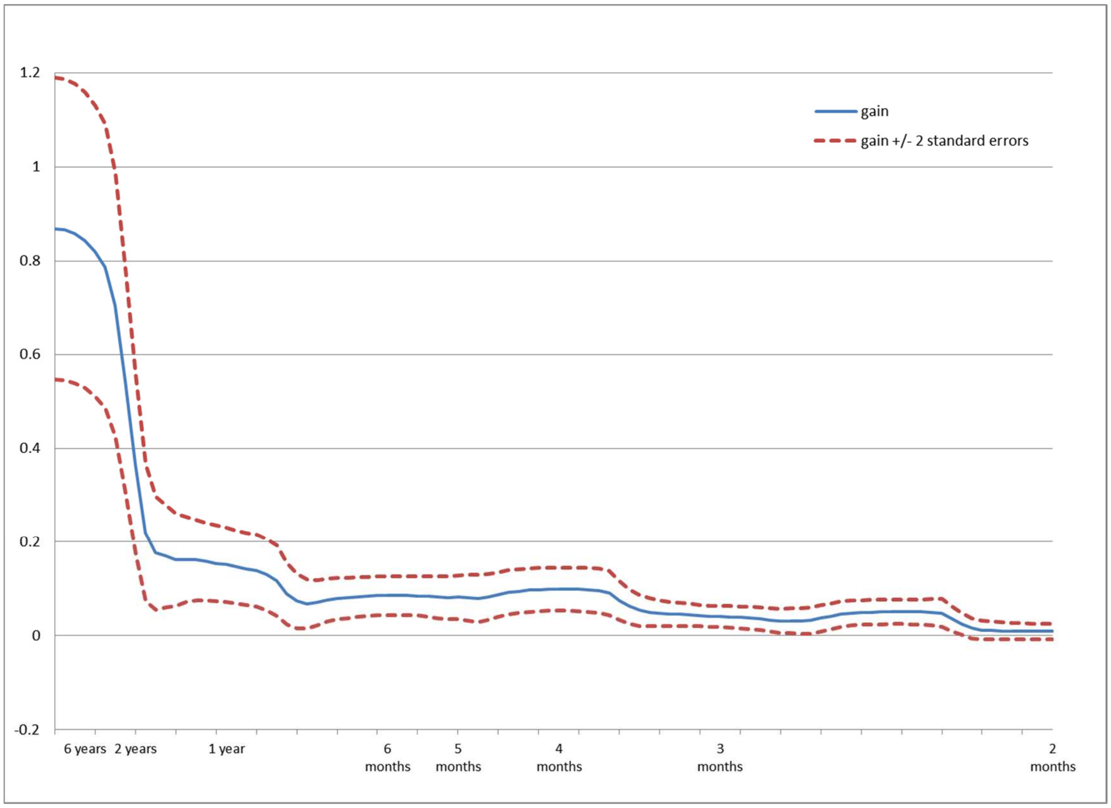

Fig. 13. Gain estimate of prices over M1a

For cycles with a length of less than 2 years, the "slope" is less than 0.2; while for cycles with a length of 6 or more years, the "slope" is around 0.85 and statistically not different from 1.

b) *The gain of prices over M4a.* In the case of M4a, the "slope" is around 0.2 for cycles with length of less than 1 year; and the "slope" is around 1 for cycles with length of more than 2 years.

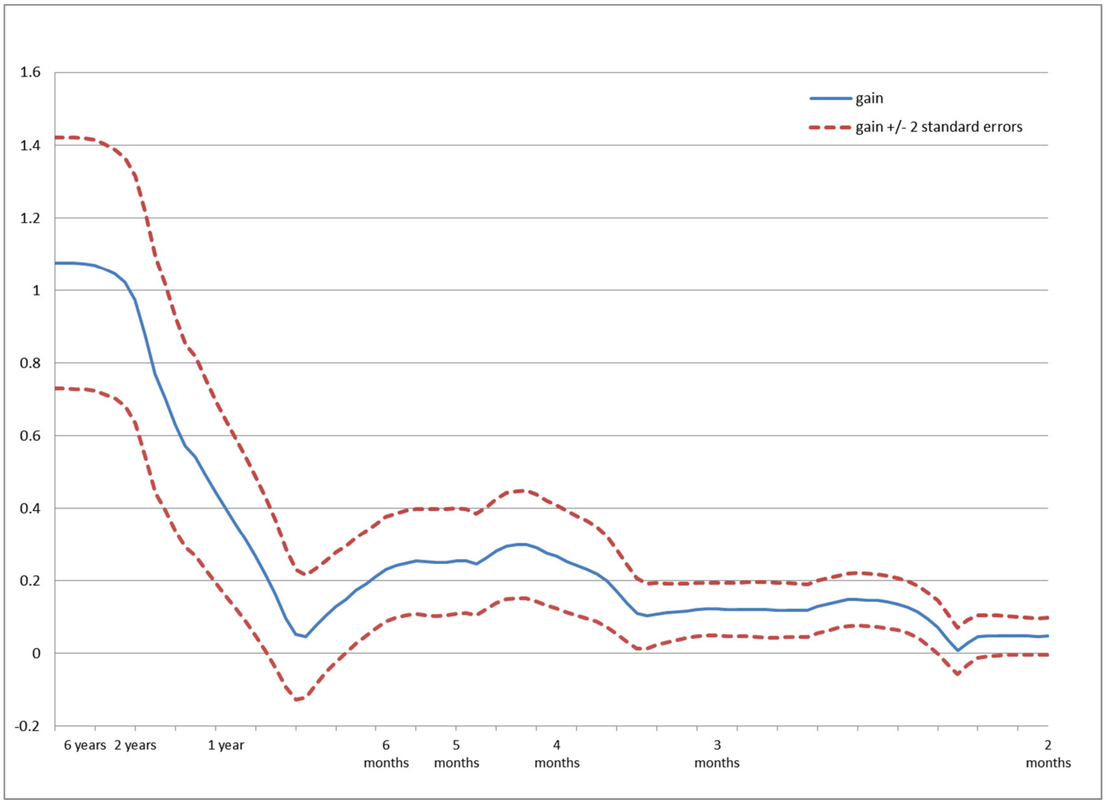

Fig. 14. Gain estimate of prices over M4a

Notice that the gain over M4a is greater than that over M1a, even though there is no "a priori" reason for the gains to be different. In fact, if M1a was a constant proportion of M4a, the gains would be (trivially) identical. There are some reasons to believe that the gain of a narrow aggregate should be smaller than that of a broader one. But the results of Benati (2009) show that this is not always the case.

V. Analysis for Rolling Samples

In this section, we include an analysis of rolling samples analogous to the one done by Benati for various OECD countries.

The main reason for performing this type of analysis is the concern about the stability of the spectrum. Since it is impossible to simultaneously display three surfaces (the spectrum and the confidence bands), we only show the results for the zero frequency, just like Benati does.

A window length of 15 years was chosen, since it is about the minimum length required to get discernible patterns from this dataset.

The coherency at the zero frequency

As mentioned before, the importance of the zero frequency is due to its interpretation as the "long run". This in turn is because it is the frequency whose neighborhood contains the lowest frequencies (and hence, arbitrarily large cycles).

a) *The coherency between M1a and prices.* The coherency between M1a and prices at frequency zero seems quite unstable. The figure shows what appears to be 4 different episodes classifiable into 3 different regimes.

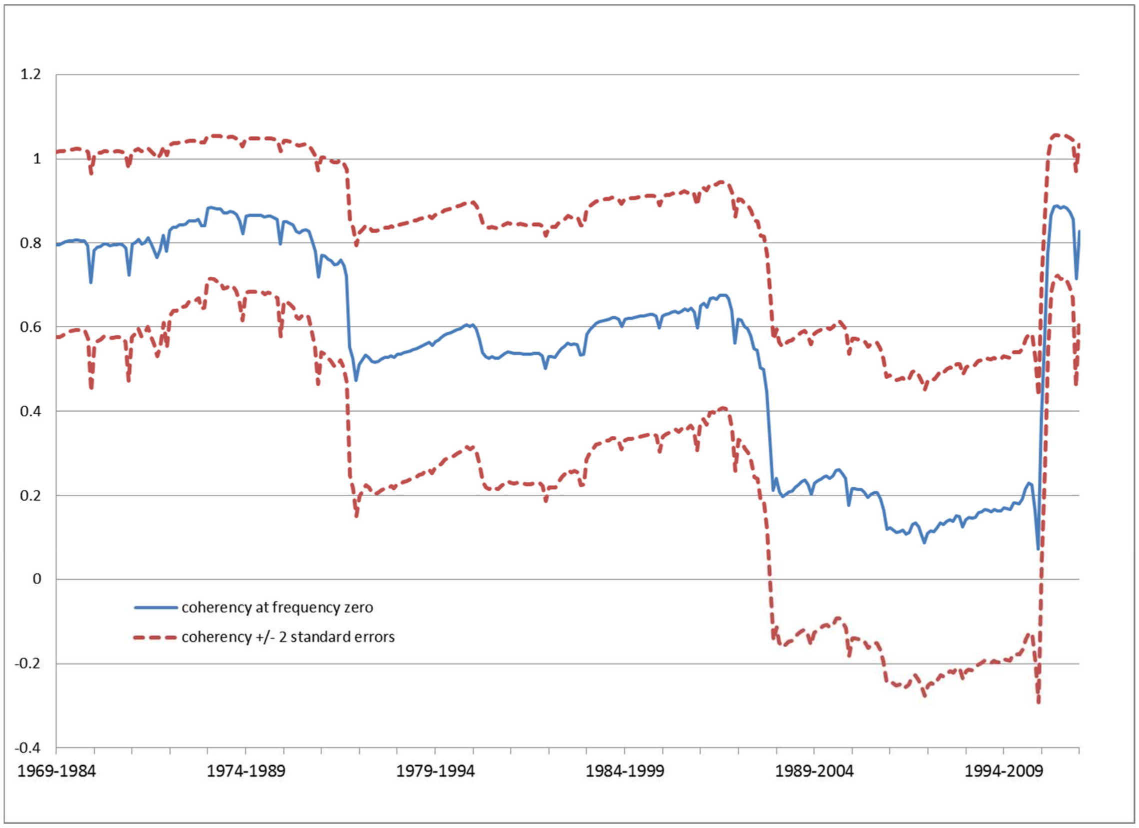

Fig. 15. Coherency at frequency zero for (M1a, prices), rolling sample

The first regime has a high coherency and matches two episodes of relatively healthy public finances (their start points being 1969-1976 and 1995 onwards). The second regime has a moderate coherency and matches an episode with escalating inflation rates (start points being1976-1987). The third regime has a very low coherency, and matches an episode of low inflation rates that were achieved through national agreements and a fixed exchange rate (1988-1994).[9]

b) *The coherency between M4a and prices.* The coherency between M4a and prices is fairly stable. Despite the shock of 1988 (start point), it slowly recovered to previous levels.

[9] Strictly speaking, it was a crawling-peg regime. For detailed accounts of mexican economic history see, for example, Moreno-Brid and Bosch (2010), or Kuntz (2010).

Fig. 16. Coherency at frequency zero for (M4a, prices), rolling sample

The confidence bands provide support to our previous assertion as well: the upper band never descends to the levels of the lower band, and both bands slowly return to their previous after the 1988 shock.

The gain at frequency zero

Just like the coherency at frequency zero, the gain at frequency zero is used to describe long-run behavior. In particular, the gain is like "the slope" of a linear relationship.

a) *The gain or prices over M1a.* The gain of prices over M1a is almost a mirror image of coherency: it is also possible to distinguish the previously mentioned episodes, although this time the last episode is more akin to the second than to the first.

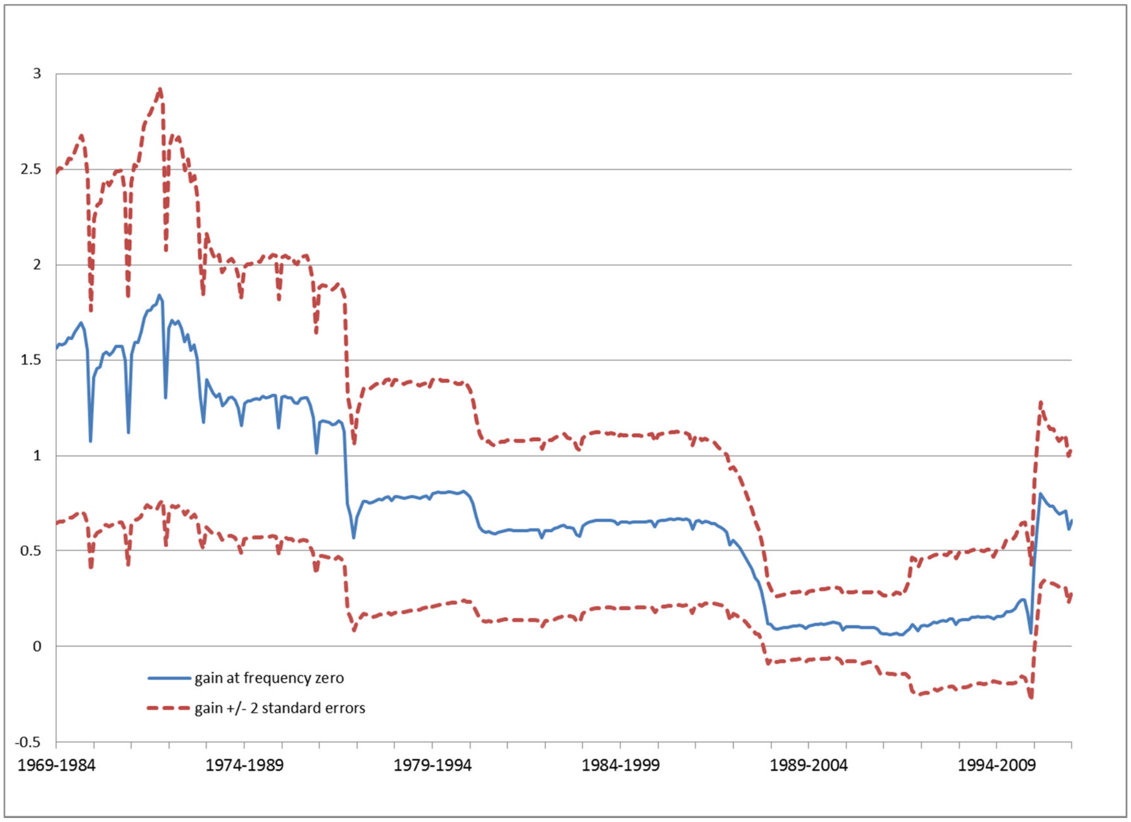

Fig. 17. Gain at frequency zero from prices over M1a, rolling sample

Notice that, despite both being low inflation episodes, the very last episode has only about half the gain of the first episode.

b) *The gain of prices over M4a.* The gain of prices over M4a is far more stable than that over M1a. In fact, it is even more stable than the corresponding coherency: it took less than three years for the gain to recover from the 1988 shock, and go back to previous levels.

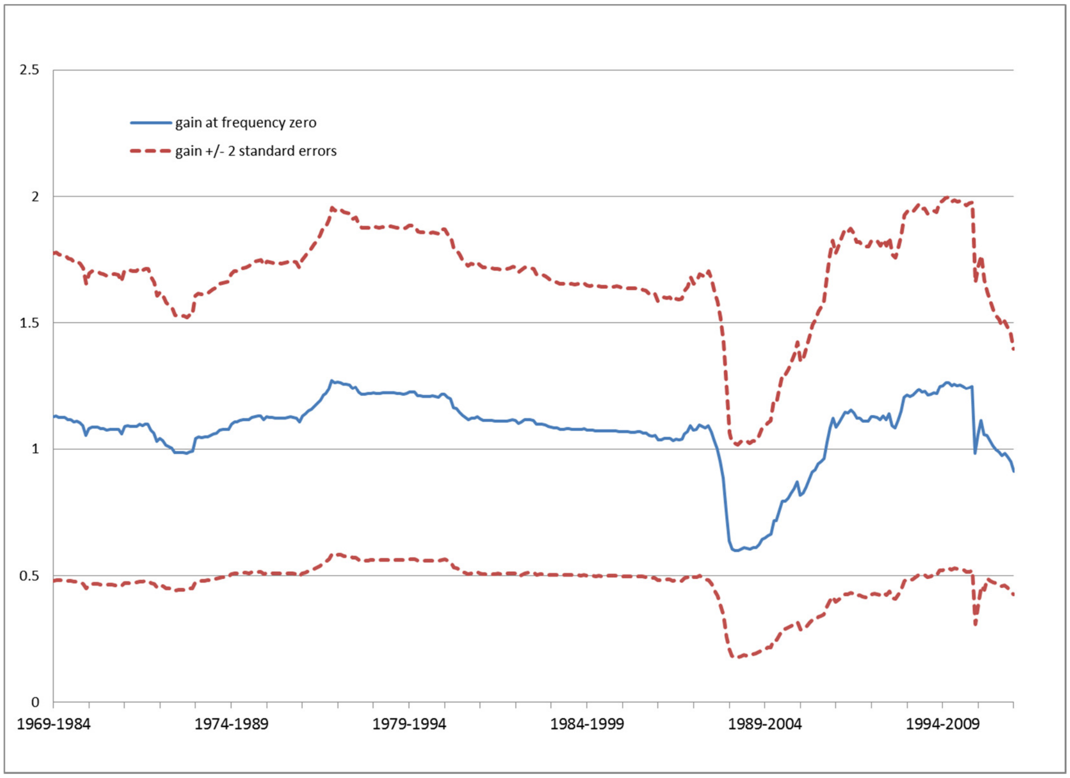

Fig. 18. Gain at frequency zero from prices over M4a, rolling sample

It may seem as if the gain has decreased in the last years, but it is far from a statistically significant change.

VI. Discussion

We now consider the results for Mexico, that we just presented, with those obtained by Benati for other countries as well as the observations done by Sargent y Surico (2011) for the US.

As previously mentioned, one of the present-day puzzles of monetary economics is the apparent instability of the long-run money-prices relationship. Benati documents this instability through his zero-frequency rolling-sample estimations.[10]

[10] Benati employs 25-year-width windows. We chose 15-year-width windows, given the smaller span of our sample. Our results are fairly comparable though.

It might be comforting for the mexican policymakers to know that Mexico is not the only country exhibiting instabilities for the zero frequency coherency and gain between prices and a narrow monetary aggregate: the USA, the UK, Canada, Japan, Norway, and the Euro Zone, all show unstable gains which take both values above and below one. Nevertheless, all of these economies (except for USA, UK, and Norway) show coherencies that are consistently close to one (greater than 0.9).

Another piece of information that is worth mentioning is that, according to Benati, the zero-frequency gains for most other countries oscillate around 0.75: just like it does for Mexico for the last 15 years.

Sargent and Surico (2011) propose a model with rigidities, habit formation, price-indexing, and unit-root technological shocks, in order to explain the instabilities for the USA. They argue that this model is able to explain the variations in (a statistic that is very similar to) the gain.

VII. Conclusions

Section IV presented the results for the entire sample (1969-2010).

Considering the behavior of the aggregates and prices during this entire time period, a critical result is that only cycles in money-growth that lasted for more than two years were associated with cycles in inflation.

For the growth in M1a, a lasting deviation of 1% is related to a lasting deviation of 0.8% in the inflation rate.[11] On the other hand, a lasting deviation of 1% in the growth rate of M4a translates into a lasting deviation of 1.1% in the inflation rate.

Section V presented results for 15-year rolling sample.

An important result from this exercise was that, indeed, both the “long-run” coherency and the “long-run” gain of (prices,M1a) are unstable. It seems as if the monetary/fiscal regime plays a crucial role in shaping their relationship.

A comparison of our results with those of Benati let us conclude that Mexico is not as different from other countries like the US and the UK. They too exhibit “long-run” instabilities in their relationships between prices and narrow monetary aggregates.

Finally, we showed that the relationship between prices and M4a is a fairly stable one, with a gain that is not statistically different from one throughout the whole period.

Acknowledgements

I thank Carlos Capistrán Carmona and Pierre Perron. The usual disclaimer applies.

References

Andrews, D. W. K. (1991), “Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation”, *Econometrica*, vol. 59, no.3, pp. 817-858.

Benati, L. (2009), “Long Run Evidence on Money Growth and Inflation”, *Working Paper Series*, no. 1027, European Central Bank.

[11] By a “lasting deviation” we mean what happens at the zero frequency or the “long-run”. Recall that the zero frequency measure tells us what happens for arbitrarily long cycles.

Box, G. E. P., G. M. Jenkins, y G. C. Reinsel (2008), *Time Series Analysis: Forecasting and Control*, fourth edition, John Wiley and Sons.

Brockwell, P. J. y R. A. Davis (1991), *Time Series: Theory and Methods*, second edition, Springer.

Fuller, W. A. (1996), *Introduction to Statistical Time Series*, second edition, John Wiley and Sons.

Granger, C. W. J. (1966), "The Typical Spectral Shape of an Economic Variable", *Econometrica*, vol. 34, no. 1, pp. 150-161.

Kuntz, S. F. coordinator (2010), *Historia económica general de México: De la Colonia a nuestros días*, El Colegio de México and the Secretaría de Hacienda.

Lucas, R. E. (1980), "Two Illustrations of the Quantity Theory of Money", *American Economic Review*, vol. 70, no. 5, pp. 1005-1014.

Moreno-Brid, J. C. y J. R. Bosch (2010), *Desarrollo y crecimiento en la economía Mexicana*, Fondo de Cultura Económica.

Priestley, M. B. (1981), *Spectral Analysis and Time Series*, Academic Press.

Robbins, L. (1998), *A History of Economic Thought*, Princeton University Press.

Sargent, T. J. y P. Surico (2011), "Two Illustrations of the Quantity Theory of Money: Breakdowns and Revivals", *American Economic Review*, vol. 101, no. 1, pp. 109-128.

Appendix

All the estimators follow Priestley (1981). The power estimate $h(\omega)$ of the process $\{X_t\}$ is

$$\widehat{h(\omega)} = \frac{1}{2\pi}\sum_{s=-(N-1)}^{(N-1)} \lambda(s)\widehat{R(s)}\cos(s\omega)$$

where N is the number of observations, $\lambda(s)$ is the lag window, and $\widehat{R(s)}$ is a consistent estimator of the autocovariance function at a given distance s. This estimator, biased but with a low mean square error, is

$$\widehat{R(s)} = \frac{1}{N}\sum_{t=1}^{N-|r|}(x_t - \bar{x})(x_{t+|r|} - \bar{x})$$

where $\bar{x} = \frac{1}{N}\sum_{t=1}^{N} x_t$, and $\{x_t\}$ are observations, that is, realizations of the process $\{X_t\}$.

The lag window that was employed for all our estimations was the quadratic spectral window, with a scale parameter $m = 12$.[12] That is

$\lambda(s) = \frac{3}{\delta^2}\left(\frac{\sin(\delta)}{\delta} - \cos(\delta)\right)$, where $\delta = \frac{6\pi|s|}{5m} = \frac{\pi|s|}{10}$.

In order to estimate the coherency between processes $\{X_{1,t}\}$ and $\{X_{2,t}\}$, we can use the fact that

$$\mathfrak{C}_{12}(\omega) = \left[\frac{(c_{12}(\omega))^2 + (q_{12}(\omega))^2}{h_{11}(\omega)h_{22}(\omega)}\right]^{1/2}$$

where $c_{12}(\omega)$ is the co-spectrum, and $q_{12}(\omega)$ is the quadrature spectrum of processes $\{X_{1,t}\}$ and $\{X_{2,t}\}$ at frequency ω, and $h_{ii}(\omega)$ is the power spectrum of process $i = 1,2$ at frequency ω.

Thus, the employed estimator is

[12] See Priestley (1981) and Andrews (1991).

$$\widehat{\mathfrak{C}_{12}(\omega)} = \left[\frac{\left(\widehat{c_{12}(\omega)}\right)^2 + \left(\widehat{q_{12}(\omega)}\right)^2}{\widehat{h_{11}(\omega)}\widehat{h_{22}(\omega)}}\right]^{1/2}$$

where the estimator $\widehat{h_{\iota\iota}(\omega)}$ is the same as before (without any subindex), and the co-spectrum and quadrature spectrum estimators are

$\widehat{c_{12}(\omega)} = \frac{1}{4\pi}\sum_{s=-(N-1)}^{(N-1)} \lambda(s)\left[\widehat{R_{12}(s)} + \widehat{R_{12}(-s)}\right]\cos(s\omega)$, and

$\widehat{q_{12}(\omega)} = \frac{1}{4\pi}\sum_{s=-(N-1)}^{(N-1)} \lambda(s)\left[\widehat{R_{12}(s)} - \widehat{R_{12}(-s)}\right]\sin(s\omega)$,

where $\widehat{R_{12}(s)} = \frac{1}{N}\sum_t \left(x_{2,t} - \overline{x_2}\right)\left(x_{1,t+s} - \overline{x_1}\right)$.

The asymptotic variance of the coherency estimator is given by

$$avar\left(\widehat{\mathfrak{C}_{12}(\omega)}\right) = \frac{m}{2N}\left(1 - \mathfrak{C}_{12}(\omega)^2\right)$$

our confidence bands are calculated using the asymptotic standard error that results from this expression.

Finally, for the gain estimate between processes $\{X_{1,t}\}$ and $\{X_{2,t}\}$, we use the fact that

$$\mathcal{G}_{12}(\omega) = \frac{\left[(c_{12}(\omega))^2 + (q_{12}(\omega))^2\right]^{1/2}}{h_{22}(\omega)}$$

Therefore using the estimator

$$\widehat{\mathcal{G}_{12}(\omega)} = \frac{\left[\left(\widehat{c_{12}(\omega)}\right)^2 + \left(\widehat{q_{12}(\omega)}\right)^2\right]^{1/2}}{\widehat{h_{22}(\omega)}}$$

where the estimators for the co-spectrum, the quadrature spectrum, and the power, are the same as before. The asymptotic variance of the gain estimator is given by

$$avar\left(\widehat{\mathcal{G}_{12}(\omega)}\right) = \frac{m}{2N}\mathcal{G}_{12}(\omega)^2\left[1 + \frac{1}{\mathfrak{C}_{12}(\omega)^2}\right]$$

The (asymptotic) standard errors are calculated from this expression.