ISSN 1518-3548

BANCO CENTRAL DO BRASIL

Working Paper Series 69

r-filters: a Hodrick-Prescott Filter Generalization

Fabio Araujo, Marta Baltar Moreira Areosa and José Alvaro Rodrigues Neto

February, 2003

ISSN 1518-3548
CGC 00.038.166/0001-05

Working Paper Series	Brasília	n. 69	Feb	2003	P. 1-37

Working Paper Series

Edited by:

Research Department (Depep)
(E-mail: workingpaper@bcb.gov.br)

Authorized by Afonso Sant'Anna Bevilaqua (Deputy Governor for Economic Policy).

General Control of Publications

Banco Central do Brasil
Secre/Surel/Dimep
SBS – Quadra 3 – Bloco B – Ed.-Sede – M1
Caixa Postal 8.670
70074-900 Brasília – DF – Brazil
Phones: (5561) 414-3710 and 414-3711
Fax: (5561) 414-3626
E-mail: editor@bcb.gov.br

Banco Central do Brasil Information Bureau

Address:	Secre/Surel/Diate
	SBS – Quadra 3 – Bloco B
	Edifício-Sede – 2º subsolo
	70074-900 Brasília – DF – Brazil
Phones:	(5561) 414 (....) 2401, 2402, 2403, 2404, 2405, 2406
DDG:	0800 992345
Fax:	(5561) 321-9453
Internet:	http://www.bcb.gov.br
E-mails:	cap.secre@bcb.gov.br
	dinfo.secre@bcb.gov.br

r-filters: a Hodrick-Prescott Filter Generalization

Fabio Araujo*

Marta Baltar Moreira Areosa

José Alvaro Rodrigues Neto

Abstract

A two-parameter family of filters is proposed in which the HP filter is considered as the lowest order member. While the HP filter converges to linear time trend as the smoothing factor grows, the higher order members of the proposed family converge to higher order polynomial time trends. The filter order – the new parameter introduced – allows to set the filter selectivity. Furthermore, two different methods to implement these filters are presented.

Keywords: Economic cycles, Low-pass filter, Fourier transform
JEL Classification: E32, C22, C52

* Central Bank of Brazil. E-mail address: workingpaper@bcb.gov.br

** The authors are grateful to Rossana Bastos Pinto and CAPES for their support.

1. Introduction

The filter proposed by Hodrick and Prescott **[1]**, the so-called HP filter, has been very useful in economic times series analysis. The main idea is to decompose a time series into its high and low frequency components. In this sense, it is widely applied to those series generated by the sum of two unobservable parcels with different spectral components.

There are in the literature some theoretical articles and many applications for such filter; being potential GDP estimation the most widely discussed application. Actually, this application involves the largest amount of empirical works using HP. Even when a production function is considered, this technique is useful as seen in Apel *et al* **[9]**. Additionally, Röger and Ongena **[11]** compared some potential GDP estimation methods – such as linear time trend, production function and HP filtering – and concluded that the HP combined with ARIMA[1] projection results in the most coherent output gap. HP filtering has also been used in many other applications.

A more theoretical approach can be found in a few papers. Hodrick and Prescott **[1]**, besides proposing the filter, mention some properties of the filtered series, although their main focus was on empirical analyses. Complementing this work, King and Rebelo **[4]** used time and frequency domain analysis to establish the main HP filter properties.

Cogley and Nason **[3]** argues that the HP filter is incompatible with a business cycle analysis because it artificially introduces cycles in the time series under analysis. Furthermore, the generated economic cycles are complex since the HP filter is not a time invariant procedure. On the other hand, Razzak **[7]**, based on an empirical study, argues that the filter corresponding to the border of the HP procedure is consistent with policy formulation, in spite of the fact that a spectral analysis indicates undesirable properties, such as phase distortion and poor frequency selectivity. This indicates that there is some divergence about the HP filter usage.

It is important to highlight that any signal extraction procedure will distort its resulting components. Jürgen **[6]** argues that even for the optimal linear extraction with known data generation processes (DGP) the distortion cannot be avoided. He also suggests,

[1] The ARIMA projection is used in order to reduce the border distortion.

along with Giorno *et al* **[10]**, that optimal λ value depends on the DGP. In this work, any hypothesis about the DGP will not be proposed: the main focus is on the implications of the λ choice.

In this paper, the initial idea was not only to extend the analysis presented in **[4]** using a linear algebra method but also to give a geometric interpretation. Such an analysis, however, results in an HP generalization, where higher order filters are found as other members of the HP filter family and so forth called *r*-filters. In addition to its possible time series applications, the *r*-filter analysis, using spectral and linear algebra techniques, allows a deeper understanding of the HP problem. Other generalization procedures are proposed in the literature. For instance, Reeves *et al* **[5]** proposes a generalization with some assumptions about the data generation process. This proposed filter is optimal as a fourth order filter and depends on a specific application (DGP). On the other hand, *r*-filters keep the flexibility found in the HP, given by the choice of smoothing degree (λ), and adding one more parameter related to filter selectivity (r).

Baxter and King **[8]** propose the use of a band-pass filter, instead of low pass filters, for business cycles studies. Their main idea for the filter design[2] is to take the impulse response of an ideal band-pass filter and restrict it to limited domain in order to make it feasible, since there are only finite samples in practical use. This procedure is equivalent to truncating the Laurent expansion of the frequency response, so the approximation is better[3] as more terms of the expansion are considered. Furthermore, as the resulting impulse response is a two-sided sequence, this filter cannot be applied at the series border. Following this idea, *r*-filters can be used to design band-pass filters since higher frequency selectivity can be reached without losing the results at the series border. In fact, only minor distortion is introduced to the results.

In Section 2 the *r*-filter family is proposed. Its spectral properties are studied; the geometric interpretation given to the HP is extended to the whole family such as its relation with least squares adjustment. In Section 3, an application is presented in order to compare different filters. The conclusions are stated in Section 4. The proofs of results are presented in the appendices.

[2] An optimization process is used to design the filter.
[3] Better means higher frequency selectivity.

2. A more general family of filters: *r*-filters

In this section a family of filters called *r*-filters is proposed, derived from a generalized minimization problem as stated in the following equation[4].

$$\underset{\{y_j\}_{j=1}^{N}}{\text{Min}} \quad F_{\lambda,r}$$

$$F_{\lambda,r} = \underbrace{\sum_{i=1}^{N}(y_i - x_i)^2}_{F_1} + \lambda \underbrace{\sum_{i=\frac{r}{2}+1}^{N-\frac{r}{2}}\left(\Delta^r y_i\right)^2}_{F_2} \qquad \textbf{(1)}$$

Before solving this problem, it is important to offer some insights into it. Firstly, certain notations should be introduced. In this paper, any economic time series will be expressed as a sequence of real numbers, where each observation is an element of the sequence. The series to be filtered will be called the input sequence, represented by $\{x_i\}$. Analogously, the filtered series will be the output sequence, represented by $\{y_i\}$. Alternatively, a point (or a vector) may represent these sequences.

The loss function $F_{\lambda,r}$ (Equation 1) expresses a trade-off between the output sequence element alignment according to $(r\text{-}1)^{\text{st}}$ degree polynomial (*F*$_2$ in Equation 1) and their fit to the input sequence (*F*$_1$ in Equation 1). λ is the smoothing parameter that represents this trade-off. In extreme cases, when $\lambda \to +\infty$ the resulting output sequence converges to a polynomial. On the other hand, when $\lambda \to 0$ it converges to $\{x_t\}$. In fact, it will be shown that, when $\lambda = +\infty$, the resultant sequence is the same as the one that would be obtained if Least Square Polynomial Adjustment (LSPA) were applied.

It is important to observe that the HP filter corresponds to the case $r = 2$. As a particular case, the result mentioned above also holds. Thus, when $\lambda = +\infty$, the output sequence converges to a first order polynomial that is the same obtained when the Least Square Linear Adjustment (LSLA) is applied. This result is well known in the literature.

---

[4] $\Delta^n y_i$ is the n^{th} difference centered in y_i and it is given by $(L\text{-}1)^n$. $L^{-n/2}$ where *L* represents the lag operator .For instance, $\Delta^4 y_i = y_{i+2} - 4y_{i+1} + 6y_i - 4y_{i-1} + y_{i-2}$.

2.1. Spectral properties of *r*-filters

In order to derive the main spectral properties of r-filters, a problem simpler than the one expressed in equation 1 is considered. It is supposed that the input sequence and, consequently, the output sequence have infinite length. Incorporating this hypothesis, the previous problem is modified to give:

$$\underset{\{y_i\}}{\text{Min}}\ F_{r,\lambda}$$
$$F_{r,\lambda} = \underbrace{\sum_{i=-\infty}^{+\infty} (y_i - x_i)^2}_{F_1} + \lambda \underbrace{\sum_{i=-\infty}^{+\infty} (\Delta^r y_i)^2}_{F_2} \qquad \textbf{(2)}$$

A difference equation $y_k + \lambda \Delta^{2r} y_k = x_k$ is obtained from the first order condition[5], where $\Delta^{2r} y_i$ symbolizes the $2r^{\text{th}}$ difference centered in y_i. So, $P(L) = \left[(L-1)^{2r} + (1/\lambda) L^r\right] \cdot L^{-r}$ is the characteristic polynomial of the resulting difference equation[6]. Using the Fourier transform, it can be seen that the frequency response of this filter family is given by:

$$H_{r,\lambda}(\omega) = \frac{e^{-i\omega r}/\lambda}{(e^{-i\omega}-1)^{2r} + e^{-i\omega r}/\lambda} = \frac{1}{1 + 2^r \lambda (\cos\omega - 1)^r} \qquad \textbf{(3)}$$

This function gives real values for every $\omega \in \Re$, meaning that this is a zero-phased filter for every pair (r, λ). Unfortunately, $|H_{r,\lambda}(\omega)|$ has two undesirable features for odd values of r. This function has an infinite peak at frequency $\omega_p = 1 - \frac{1}{2\sqrt[r]{\lambda}}$, besides not being necessarily convex even in the finite case. In this way, only filters defined for even values of r will be considered from here on, thus the HP is the first member of this family.

It is important to remember that $\omega = \dfrac{2\pi\, T_s}{T}$ where T_s is the sampling period and T is the period of the cycle to be filtered, both in the same time unit. Thus, to filter an eight-year cycle on a quarterly data series, for instance, T_s and T would respectively be equal to ¼ and 8, making $\omega = \frac{\pi}{16}$.

---

[5] This condition is necessary to ensure that the solution is a global minimum.

2.2. Equivalence between filters of different orders

In order to establish the equivalence between different filters, an equivalence parameter must be defined. Borrowing terminology from electrical engineering, a natural parameter could be the cut-off frequency, defined as the point at which the frequency response magnitude is $\sqrt{2}/2$. However, r-filters have an odd feature described in Equation 4. This equation states that the only inflection point for $\omega \in (0,\pi)$ occurs at the frequency in which $H_{r,\lambda}(\omega) = 0.5$, for every pair ($r$, λ). So, as the frequency response is monotone over this interval, filters are more selective at this frequency. Based on this fact, a “modified” cut-off frequency, defined as the point where $H_{r,\lambda}(\omega) = 0.5$, will be used throughout the paper as an equivalence parameter.

$$\frac{\partial^2 H_{r,\lambda}}{\partial \omega^2}(\omega_0) = 0 \Rightarrow H_{r,\lambda}(\omega_0) = 1/2 \qquad \textbf{(4)}$$

Looking for an equivalence expression, it can be noted that for a given cut-off frequency, $\omega_0 \in (0,\pi)$, there are an infinite number of pairs (r, λ) that satisfy Equation 5.

$$H_{r,\lambda}(\omega_0) = 1/2 \Leftrightarrow \cos(\omega_0) = 1 - \frac{1}{2\lambda^{1/r}} \qquad \textbf{(5)}$$

Equation 6 shows the relation of any two solutions, (r_1, λ_1) and (r_2, λ_2), for Equation 5 given ω_0.

$$\lambda_2^{r_1} = \lambda_1^{r_2} \qquad \textbf{(6)}$$

---

[6] It is important to note that if $P(L) = 0$ then $P(\frac{1}{L}) = 0$

Figure 1 shows various equivalent filters with different values of *r*.

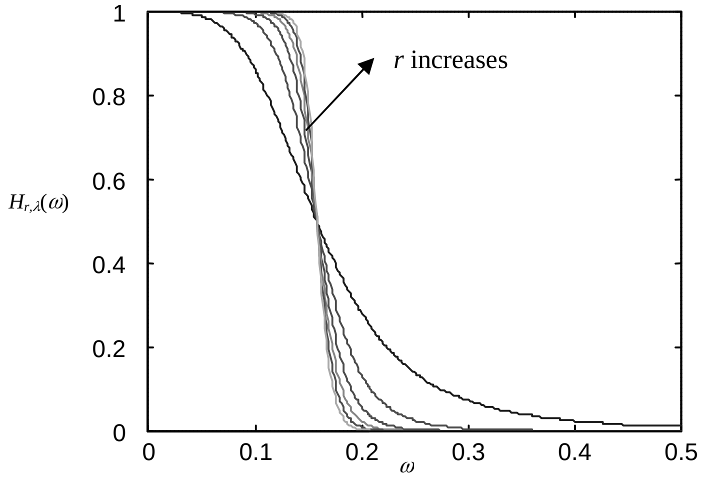


Figure 1: *r*-filter frequency responses for *r* = 2, 4, 6, 8, 10 and 12

So, for a given cut-off frequency filters defined by $H_{r_1,\lambda_1}(\omega_c)$ and $H_{r_2,\lambda_2}(\omega_c)$ will be understood as equivalent if Equation 6 holds. Furthermore, if two filters are equivalent and $r_2 > r_1$, $H_{r_2,\lambda_2}(\omega_c)$ is more selective[7] than $H_{r_1,\lambda_1}(\omega_c)$. As the filter order is defined by *r*, a direct relation between order and selectivity can be established. A similar result has already been observed in **[4]** when the exponential smoothing filter (ES) is compared to the HP filter. It is stated that the HP filter looks more like an ideal filter, i.e., it is more selective than the ES filter.

Equation 6 provides an easy way to calculate equivalent filters. For instance, all filters on the same column of the following table are equivalent.

r = 2, λ = 100	r = 2, λ = 1600	r = 2, λ = 14400
r = 4, $\lambda = 100^2$	r = 4, $\lambda = 1600^2$	r = 4, $\lambda = 14400^2$
r = 6, $\lambda = 100^3$	r = 6, $\lambda = 1600^3$	r = 6, $\lambda = 14400^3$
r = 8, $\lambda = 100^4$	r = 8, $\lambda = 1600^4$	r = 8, $\lambda = 14400^4$

[7] In the present case, i.e., a low-pass filter, selectivity denotes the capability of separating low from high frequencies. As the filter becomes more selective, cycles whose frequency is above the cut-off frequency are strongly attenuated. Geometrically, the filter frequency response approximates a step function.

All these properties could be derived due to the infinite length sample hypothesis. Unfortunately, sample is always finite. So the results found so far are only approximations of the behavior of the central observations when the sample size (*N*) is large enough.

2.3. The *r*-filters calculation

In order to highlight the dynamical characteristics of *r*-filters, only the infinite length sample case has been analyzed in previous sub-sections. This sub-section focuses on the calculation of *r*-filters given a finite length sample. Thus, the original problem (Equation 1) should be solved.

It will be proved in Appendix B that first order conditions are sufficient to find the solution and generate a linear system of *N* equations and *N* variables, described by:

$$y = T_{\lambda,r}(x) = \left(I + \lambda.B(r)\right)^{-1} x \qquad \textbf{(7)}$$

where x and y are vectors in the $\Re^N$, which respectively represent, as previously defined, an input sequence and an output sequence, both with *N* observations. *I* and *B*(*r*) respectively correspond to the identity matrix and a symmetric square matrix, both of order *N*, obtained from the first order condition. While *I* is derived from the first parcel of the loss function (F_1 in Equation 1), *B*(*r*) comes from the second parcel (F_2 in Equation 1). This filtering process is equivalent to a linear transformation. Moreover, this transformation is $\Re^N$ isomorphism.

Building *B*(*r*) without evaluating the derivative is one of the most important steps of the filtering algorithm[8]. The general term of the *B*(*r*) matrix, $b_{i,j}^{(r)}$, is given by:

$$b_{i,j}^{(r)} = \sum_{k\in\Gamma} (-1)^{i-j} C_r^{i-k+\frac{r}{2}} C_r^{j-k+\frac{r}{2}} \qquad \textbf{(8)}$$

where $\Gamma(N,r,i,j) = \left\{k \in \mathrm{Z} \;\middle|\; \max\left(1+\tfrac{r}{2}, i-\tfrac{r}{2}, j-\tfrac{r}{2}\right) \le k \le \min\left(N-\tfrac{r}{2}, i+\tfrac{r}{2}, j+\tfrac{r}{2}\right)\right\}$

[8] Matrix B(*r*) is derived in Appendix A, proving Equation 6.

Fortunately, *B(r)* presents a particular structure that allows the non use of Equation 8 in the calculation of all its N^2 elements, reducing the necessary computation. Figure 2 illustrates the $B(r)$ matrix structure.

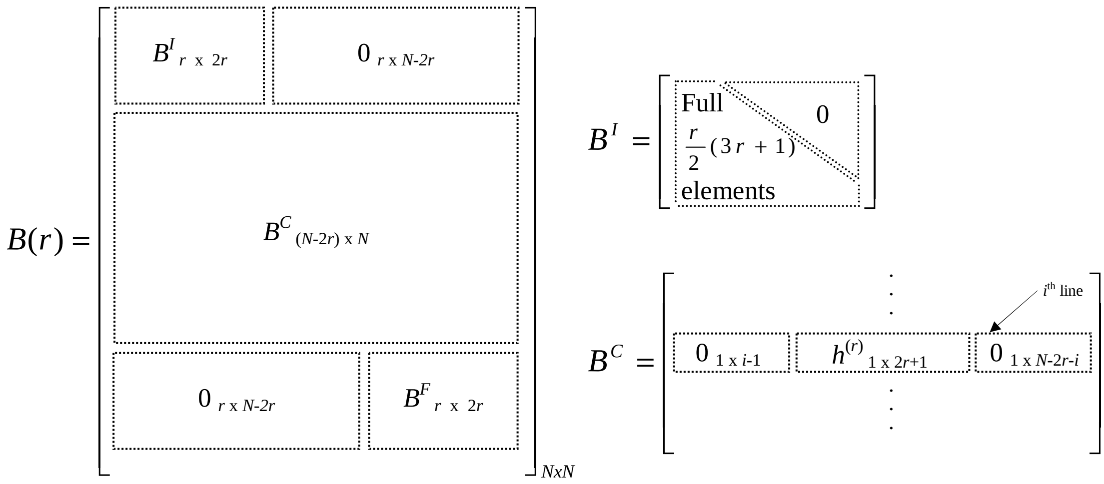


Figure 2: $B(r)$ matrix structure

where *h(r)* is a vector whose elements are Newton's binomial coefficients of order 2r with alternate signals and matrix B_F is equal to a vertical reflection of B_I[9] . Since r (filter order) is usually much smaller than N (sample size)[10], the number of elements in matrix B^I (or B^F) to be calculated is much smaller than the total number of $B(r)$ elements. For example, considering a fourth-order filter (r = 4) applied to a sample with 40 observations[11], matrix B^I has 26 non-null elements, while matrix $B(r)$ has 1600 elements.

There are two other methods to construct *B(r)* matrix. The simplest one consists of decomposing matrix *B(r)* as the product of three matrices with simple structures, as shown in Appendix A.

2.4. *r*-filters as a generalization of the least square polynomial adjustment

This sub-section aims at analyzing[12] the fact that r-filters can be interpreted as a generalization of the least square polynomial adjustment (LSPA) using an $(r\text{-}1)^{st}$ degree

[9] This statement is proven in Appendix A (property A.2)
[10] In fact, this is a necessary condition for using the r-filters. Otherwise, there would not be enough data to apply this kind of filtering.
[11] Actually this sample size is small for econometric uses.
[12] A rigorous proof is available in Appendix C

polynomial[13]. An important technical step to meet this purpose consists of building a surface in $\Re^N$, henceforth called M^r, that encompasses all (r-1)st degree polynomials.

$M^r = \{a_1.\bar{0} + a_2.\bar{1} + \ldots + a_r.\overline{r-1} | a_1, a_2, \ldots, a_r \in \Re\}$ where $\{\bar{0}, \bar{1}, \ldots, \overline{r-1}\}$ represents a set of linearly independent vectors given by the formula:

$$\bar{k} = (1^k,\ 2^k, \ldots,\ N^k),\ \forall\ k = 0, 1,\ \ldots,\ r-1$$

In order to provide some insight into M^r, an example should be presented. In case r = 2, M^2 is the sub-vector space of dimension 2 in $\Re^N$ and it is given by:

$$M^2 = \{a_1.\bar{0} + a_2.\bar{1} | a_1, a_2 \in \Re\} = \{(a_1 + a_2, a_1 + 2a_2, a_1 + 3a_2, \cdots, a_1 + Na_2) | a_1, a_2 \in \Re\}.$$

It should be noted that each pair (a_1, a_2) defines a point in $M^2 \subset \Re^N$. This point corresponds to a finite time series whose observations lie in a straight line as shown in Figure 3. In this particular case, a_1 and a_2 would respectively be the intercept and angular coefficient of this straight line.

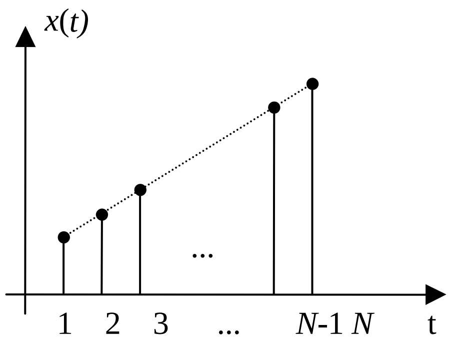


Figure 3: *A point in surface M^2*

Similar to the previous case, M^r is a hyper-plane of dimension r in $\Re^N$, spanned by vectors $\bar{0}, \bar{1}, \ldots, \overline{r-1}$, that encompasses all points whose coordinates are aligned according to an (r-1)st degree polynomial. Therefore, coordinates of a point $(a_1, a_2, \ldots, a_r)$ in M^r in the base $\Psi = \{\bar{0},\ \bar{1},\ \cdots,\ \overline{r-1}\}$ be can interpreted as coefficients of an (r-1)st degree polynomial. Analogously to Figure 3, Figure 4 shows the general case.

[13] In fact, any coefficient of the adjusted polynomial can be zero, even the highest order coefficient. Therefore, the adjusted polynomial degree might be smaller than r-1.

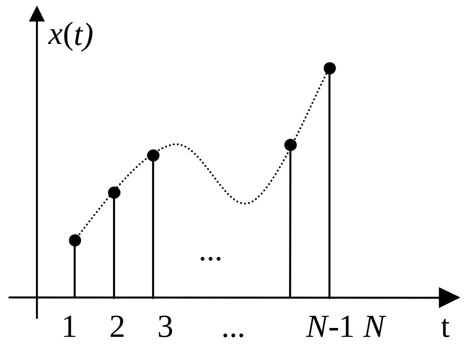
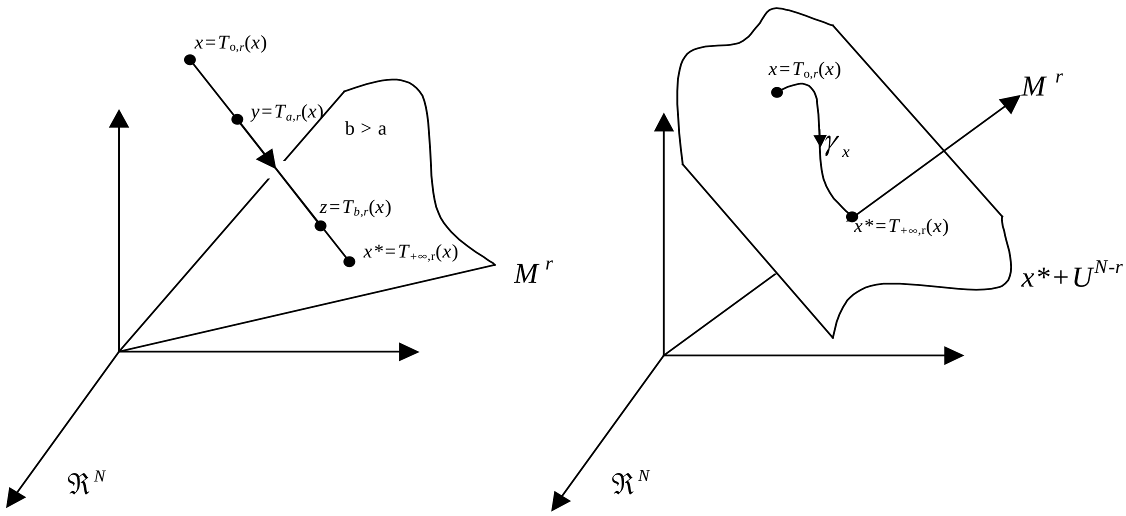


Figure 4: a time series that lies on a polynomial

Now, let $x^* = (x_1^*, x_2^*, ..., x_N^*)$ representing the sequence obtained by applying LSPA to x. It can be understood as the point in $\Re^N$ that corresponds to the orthogonal projection of x on hyper-plane M^r, as shown in Figure 5.a. Furthermore, if y is the result when $T_{\lambda,r}$ (Equation 7) is evaluated at x, then y belongs to space $x^* + U^{N-r} = \{x^* + u \mid u \in U^{N-r}\}$, where U^{N-r} is the orthogonal complement of M^r.



(a) (b)

Figure 5: The orthogonal projection given by HP
(a) focus on plane M^r (b) focus on $x^* + U^{N-r}$ space

In Figure 5.a, N dimension Euclidean space and plane M^r are shown. A point x is projected on x^* using LSPA, which also corresponds to applying the r-filter with $\lambda = +\infty$. Points y and z are in space $x^* + U^{N-r}$, that is orthogonal to M^r. Moreover, as λ grows $T_{\lambda,r}(x)$ gets closer to x^*. In fact, by varying λ, it is possible to define a continuous path starting at x and ending at x^*. Again, space $x^* + U^{N-r}$, represented by a straight segment in Figure 5.a, is in fact a space of dimension N-r. This is clearer in Figure 5.b, in which hyper-plane M^r is represented as a line.

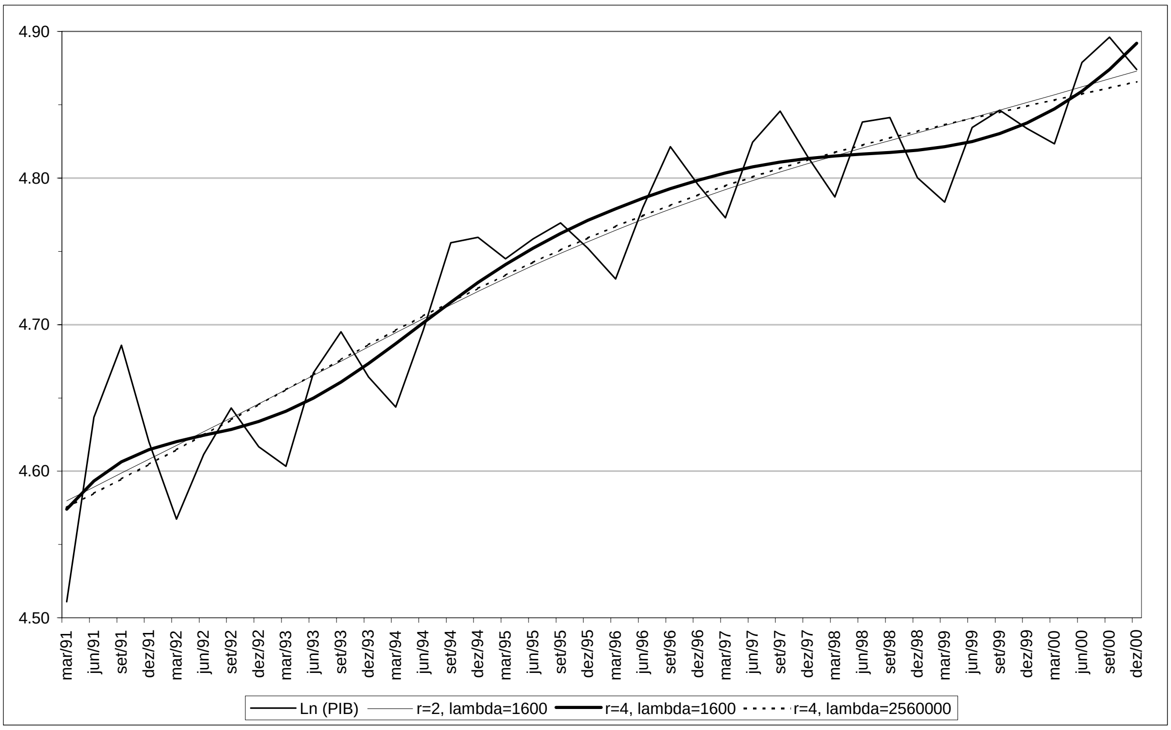
3. An example: the estimation of potential GDP[14]

This section aims at studying the effects of replacing the HP filter by a higher order r-filter. One of the most important HP filtering applications on Economics – the potential output estimation – was chosen to illustrate these effects.

Three series are derived from Brazilian GDP quarterly data. The first two series use the same smoothing factor (λ = 1600) but different kinds of filtering. The HP filter (r = 2) is replaced by a higher order r-filter (r = 4) in the second series. The third series was built to show the equivalence[15] between filters of different orders. All of them could be considered as potential output series.

It is important to highlight, however, that there is no intention to discuss which result can be more easily interpreted in light of the economic facts. It is not on within the scope of this paper to analyze recent Brazilian economic history but only to give some insight into the properties of filters.

[14] Appendix D provides further examples.
[15] As defined on section 3.1



It can be noted from the comparison of the results that oscillations can be easily identified in the second series (r = 4, λ = 1600), while the first series (r=2, λ = 1600) is almost linear. This insight is quite intuitive since the first and second series can be respectively interpreted as deviations of first and third order polynomials.

The third series (r = 4, λ = 2560000) is equivalent to the first one (r = 2, λ = 1600). This fact can be observed from the figure since the third series fits the first one almost perfectly.

4. Conclusion

r-filters are completely characterized by two parameters: *r*, the filter order, and λ, the smoothing factor. In case *r* = 2, it corresponds to the traditional HP filter. The *r*-filters most important characteristics encompass: being a zero-phased filter and DGP independent (analogous to the HP filter), the residuals are stationary even when the DGP presents an $(r\text{-}1)^{st}$ polynomial deterministic trend or is $(r-1)$ order integrated[16] and the flexibility of choosing the filter selectivity.

[16] Appendix D shows evidence of this fact.

The requirement of specifying a second parameter (*r*) increases the complexity of selecting the adequate filter for an application. Despite the fact that it is always possible to find an equivalent HP filter, it is important to bear in mind that equivalence, as defined in section 2.2, does not imply the same spectral properties, but only the same cut-off frequency. The benefit brought by the introduction of this new parameter is illustrated in Figure 1 where many frequency responses for equivalent filters are plotted. On comparing them, it is clear that selectivity increases with the filter order. It is important to highlight that selectivity is a great asset and, for many applications, such as filter composition, can be highly desired.

As discussed in **[7]**, **[9]** and **[11]** the border effect is one of the most important criticisms that have been made with respect to the HP filter. It has already been identified that when a new observation is added to the sample, the output sequence changes mainly with respect to its final elements. This behavior has two explanations: the inclusion of a new observation changes the entire output sequence (two sided impulse response) and the characteristic polynomial of the filter changes for the last two observations. The *r*-filters inherit this problem, worsened by the increase in their order, since the characteristic polynomial of the filter changes for the last *r* observations.

There is a clear trade-off between increasing selectivity and reducing the border effect. Consequently, the optimal choice of these parameters, *r* and λ, will depend on the application. For descriptive studies with extensive[17] data set available, it is worth while using high values of *r*. Reciprocally, when the data set is small or when the main objective of the study is inference, the value of *r* should be lowered.

A possible extension of this work would comprise verifying if the characteristic polynomial of a filter on the last *r* observation could be modified to minimize the border effect. The interest in this approach rises from the fact that it does not require any hypothesis about the data generation process.

[17] The use of the expressions “extensive”, “high” and “small” is informal and their values depend on the application and on the data set itself.

Reference

[1] – Hodrick, R.J. and Prescott, E.C. – "Postwar U.S. Business Cycles: An Empirical Investigation", *Journal of Money, Credit, and Banking*, 29 (1997): 1-16.[*]

[2] – Oppenheim, Alan V. and Schafer, Ronald W. – *Discrete-Time Signal Processing*, Prentice Hall Signal Processing Series (1989)

[3] – Cogley, Timothy and Nason, James M. – "Effects of the Hodrick-Prescott Filter on Trend and Difference Stationary Time Series. Implications for Business Cycle Research", *Journal of Economic Dynamics and Control*, 19 (1995): 253-278

[4] – King, Robert G. and Rebelo, Sergio T. – "Low Frequency Filtering and Real Business Cycles", *Journal of Economic Dynamics and Control*, 17 (1993): 207-231

[5] – Reeves, J.J., Blyth, C.M. and Small, J.P. – "The Hodrick-Prescott Filter, a Generalization, and a New Procedure for Extracting an Empirical Cycle from Series", Working Paper No. 160, Department of Economics, The University of Auckland, Auckland, New Zealand

[6] – Ehlgen, Jürgen – "Distortionary effects of the optimal Hodrick-Prescott filter", *Economic Letters*, 61 (1998): 345-349

[7] – Razzak, W. – "The Hodrick-Prescott technique: A smoother versus a filter. An application to New Zealand GDP", *Economic Letters*, 57 (1997): 163-168

[8] – Baxter, M. and King, Robert G. – "Measure Business Cycles Approximate Band-Pass Filters for Economic Time Series", NBER Working Paper No. 5022

[9] – Apel, Mikael; Hansen, Jan and Lindberg, Hans – "Potential output and output gap", *Quarterly Review* - Sveriges Riksbank (1996 III)

[10] – Giorno, Claude; Richardson, Pete; Roseveare, Deborah and Noord, Paul van den – "Estimating potential output, output gaps and structural budget balance", Economics Department Working Paper nº 152 - OCDE/GD(95)1

[*] This paper is substantially the same as the authors' 1981 working paper.

[11] – Röger, Werner; Ongena, Hedwig – “The Commission Services Cyclical Adjustment Method –Comparison of the Hodrick-Prescott filter with other trend estimation methods”, *Banca d’Italia*

Appendix A – How to Construct the Matrix $B(r)$

In this appendix two methods to build matrix *B(r)*[18] will be presented.

<u>1st method: Calculating an element of the matrix *B(r)*</u>

Initially, it should be recalled that the relation $y = T_{\lambda,r}(x) = (I + \lambda.B(r))^{-1} x$ (Equation 7) was derived from the evaluation of the First Order Condition of the problem defined on equation A.1 (already expressed in Equation 1). In particular, *B(r)* comes from the parcel F_2.

$$\underset{\{y_j\}_{j=1}^N}{\text{Min}} \quad F_{\lambda,r}$$

$$F_{\lambda,r} = \underbrace{\sum_{i=1}^{N}(y_i - x_i)^2}_{F_1} + \lambda \underbrace{\sum_{i=\frac{r}{2}+1}^{N-\frac{r}{2}} \left(\Delta^r y_i\right)^2}_{F_2} \tag{A.1}$$

Thus, before calculating *B(r)*, it is important to observe that the r[th] difference centered in y_k and its derivative are given by:

$$\Delta^r y_k = \sum_{l=-\frac{r}{2}}^{\frac{r}{2}} (-1)^{l+\frac{r}{2}} C_r^{l+\frac{r}{2}} y_{k+l} \quad \Rightarrow \quad \frac{\partial \Delta^r y_k}{\partial y_i} = \begin{cases} (-1)^{\frac{r}{2}+k-i} C_r^{\frac{r}{2}+k-i}, & \text{if} \quad |\text{k-i}| \le \frac{r}{2} \\ 0, & \text{otherwise} \end{cases} \tag{A.2}$$

The results shown in the previous equation are now used to calculate *B(r)*:

$$\begin{aligned} &\frac{\partial}{\partial y_i}\left(\sum_{k=\frac{r}{2}+1}^{N-\frac{r}{2}} \left(\Delta^r y_k\right)^2\right) = 2\sum_{k=1+\frac{r}{2}}^{N-\frac{r}{2}} \left(\Delta^r y_k\right) \frac{\partial}{\partial y_i}\left(\Delta^r y_k\right) = \\ &= 2 \sum_{k=\max\left(1+\frac{r}{2},i-\frac{r}{2}\right)}^{\min\left(N-\frac{r}{2},i+\frac{r}{2}\right)} \left(\Delta^r y_k\right)(-1)^{\frac{r}{2}+k-i} C_r^{\frac{r}{2}+k-i} = 2.\sum_{(k,l)\in\Lambda} (-1)^{k-i+l} C_r^{\frac{r}{2}+k-i} C_r^{\frac{r}{2}+l}.y_{k+l} \end{aligned} \tag{A.3}$$

$$\text{where} \quad \Lambda = \left\{ (k,l) \,\middle|\, \begin{matrix} \max\left(1+\frac{r}{2}, i-\frac{r}{2}\right) \le k \le \min\left(N-\frac{r}{2}, i+\frac{r}{2}\right) \\ |l| \le \frac{r}{2} \end{matrix} \right\}$$

[18] A third method uses the convolution of sequences to calculate a whole line of the matrix *B(r)*. This third method is not presented in this paper because it does not shed new light on the problem.

Let $j = k + l$, so each term of the matrix *B(r)* is given by[19]:

$$b_{i,j}^{(r)} = \sum_{k\in\Gamma} (-1)^{i-j} C_r^{i-k+\frac{r}{2}} C_r^{j-k+\frac{r}{2}} \qquad \text{(A.4)}$$

where $\Gamma = \left\{ k \middle| \max\left(1+\tfrac{r}{2}, i-\tfrac{r}{2}, j-\tfrac{r}{2}\right) \le k \le \min\left(N-\tfrac{r}{2}, i+\tfrac{r}{2}, j+\tfrac{r}{2}\right) \right\}$

Now two important properties are stated and proved.

Property A.1

B(r) is symmetric, that is, $b_{i,j}^{(r)} = b_{j,i}^{(r)}$, $\forall\, i, j = 1,\ldots,N$.

The proof is obvious, since i and j are easily interchangeable.

Property A.2

$b_{i,j}^{(r)} = b_{N-i+1,N-j+1}^{(r)}$, $\forall N, r, i, j$.

Proof:

In order to prove this result, it is necessary to observe that:

$b_{N-i+1,N-j+1}^{(r)} = \sum_{k\in\Gamma'} (-1)^{i-j} C_r^{N-i+1-k+\frac{r}{2}} C_r^{N-j+1-k+\frac{r}{2}}$, where:

$\Gamma' = \left\{ k \middle| \max\left(1+\tfrac{r}{2}, N-i+1-\tfrac{r}{2}, N-j+1-\tfrac{r}{2}\right) \le k \le \min\left(N-\tfrac{r}{2}, N-i+1+\tfrac{r}{2}, N-j+1+\tfrac{r}{2}\right) \right\}$.

Depending on the limits of the summation, $b_{N-i+1,N-j+1}^{(r)}$ will include different parcels. Therefore, 9 cases were identified:

1) $1+\tfrac{r}{2} \le k \le N-\tfrac{r}{2}$
2) $1+\tfrac{r}{2} \le k \le N-i+1+\tfrac{r}{2}$
3) $1+\tfrac{r}{2} \le k \le N-j+1+\tfrac{r}{2}$
4) $N-i+1-\tfrac{r}{2} \le k \le N-\tfrac{r}{2}$
5) $N-i+1-\tfrac{r}{2} \le k \le N-i+1+\tfrac{r}{2}$
6) $N-i+1-\tfrac{r}{2} \le k \le N-j+1+\tfrac{r}{2}$
7) $N-j+1-\tfrac{r}{2} \le k \le N-\tfrac{r}{2}$
8) $N-j+1-\tfrac{r}{2} \le k \le N-i+1+\tfrac{r}{2}$
9) $N-j+1-\tfrac{r}{2} \le k \le N-j+1+\tfrac{r}{2}$

However, just 6 of these cases should be analyzed, since cases 1, 5 and 9 cannot occur.

For case 2, the expression $b_{N-i+1,N-j+1}^{(r)}$ can be written as:

[19] The factor 2 that appears in Equation A.3 is simplified regarding the First Order Condition gives $2(y-x)+2\lambda B(r).y = 0$.

$$b^{(r)}_{N-i+1,N-j+1} = \sum_{k=1+\frac{r}{2}}^{N-i+1+\frac{r}{2}} (-1)^{i-j} C_r^{N-i+1-k+\frac{r}{2}} C_r^{N-j+1-k+\frac{r}{2}} = \sum_{k=1+\frac{r}{2}}^{N-i+1+\frac{r}{2}} (-1)^{i-j} C_r^{r-(N-i+1-k+\frac{r}{2})} C_r^{r-(N-j+1-k+\frac{r}{2})}$$

$$= (-1)^{i-j}\left[C_r^{-N+i+r} C_r^{-N+j+r} + \ldots + C_r^{r} C_r^{j-i+r}\right] = \sum_{k=i-\frac{r}{2}}^{N-\frac{r}{2}} (-1)^{i-j} C_r^{i-k+\frac{r}{2}} C_r^{j-k+\frac{r}{2}} = b^{(r)}_{i,j}$$

The same kind of argument can be used to prove all other cases, thus providing conclusive proof to Property A.2.

2^o method: Decomposing matrix *B(r)*

Matrix $B(r)$ can be decomposed as $B(r) = A(r).D(r).A^t(r)$, where:

$$\underset{NxN}{A(r)} = \begin{bmatrix} C_r^0 & & & & 0 \\ \vdots & \ddots & & & \\ (-1)^{i-1} C_r^{i-1} & & \ddots & & \\ \vdots & & & \ddots & \\ C_r^r & & & & \ddots \\ 0 & C_r^r & \cdots & (-1)^{i-1} C_r^{i-1} & \cdots \quad C_r^0 \end{bmatrix}$$

and

$$\underset{NxN}{D(r)} = \left[\begin{array}{c|c} I_{(N-r)x(N-r)} & 0_{(N-r)xr} \\ \hline 0_{rx(N-r)} & 0_{rxr} \end{array}\right]$$

$$a_{i,j} = \begin{cases} (-1)^{i-j} C_r^{i-j} & 0 \le i-j \le r \iff j \le i \le j+r \le N \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad d_{i,j} = \begin{cases} 1, & i = j \le N-r \\ 0, & \text{otherwise} \end{cases}$$

Let $\tilde{A}(r) = A(r).D(r)$

$$\tilde{a}_{i,l} = \begin{cases} (-1)^{i-l} C_r^{i-l} & \begin{cases} l \le N-r \\ 0 \le i-l \le r \Leftrightarrow i-r \le l \le i \end{cases} \\ 0 & \text{otherwise} \end{cases}$$

$$a^t_{l,j} = \begin{cases} (-1)^{l-j} C_r^{j-l} & \begin{cases} 1 \le l \\ 0 \le j-l \le r \Leftrightarrow j-r \le l \le j \end{cases} \\ 0 & \text{otherwise} \end{cases}$$

So, using these definitions:

$$\left(A.D.A^t\right)_{i,j} = \sum_{l=1}^{N} \tilde{a}_{i,l} a^t_{l,j} = \sum_{l=\max(i-r,\,j-r,1)}^{\min(i,\,j,N-r)} (-1)^{i-j} C_r^{i-l} C_r^{j-l} = \sum_{k=\max\left(i-\frac{r}{2},\,j-\frac{r}{2},1+\frac{r}{2}\right)}^{\min\left(i+\frac{r}{2},\,j+\frac{r}{2},N-\frac{r}{2}\right)} (-1)^{i-j} C_r^{i-k+\frac{r}{2}} C_r^{j-k+\frac{r}{2}} = b^{(r)}_{i,j}$$

where the last step uses a change of coordinates given by $k = l + \frac{r}{2} \Leftrightarrow l = k - \frac{r}{2}$

Property A.3

$rank\left(B(r)\right) = N - r$

Proof:

This property can be easily proven considering decomposition $B(r) = A(r).D(r).A^t(r)$.

Appendix B – $F_{\lambda,r}$ is a Convex Function

The main purpose of this appendix is to show that $F_{\lambda,r}$ (Equation 1) is a convex function of *y* for any fixed *x*, which is a sufficient condition for the zero of the derivative to be a global minimum. In order to do so, it is enough to prove that its Hessian matrix is positive definite.

$$F_{\lambda,r} = \underbrace{\sum_{i=1}^{N}(y_i - x_i)^2}_{F_1} + \lambda \underbrace{\sum_{i=\frac{r}{2}+1}^{N-\frac{r}{2}}\left(\Delta^r y_i\right)^2}_{F_2}$$

It can be seen that the Hessian matrix of the loss function[20] $F_{\lambda,r}$ can be decomposed as the sum of Hessian matrices of F_1 and F_2. It will be shown that the Hessian matrix of F_1 and F_2 are respectively positive and positive semidefinite. Consequently, the $F_{\lambda,r}$ Hessian matrix is positive definite.

It is easy to verify that the Hessian of F_1 is positive definite and it is given by:

$$Hess(F_1) = \begin{bmatrix} \frac{\partial F_1}{\partial y_1 \partial y_1} & \frac{\partial F_1}{\partial y_1 \partial y_2} & \cdots & \frac{\partial F_1}{\partial y_1 \partial y_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_1}{\partial y_N \partial y_1} & \frac{\partial F_1}{\partial y_N \partial y_2} & \cdots & \frac{\partial F_1}{\partial y_N \partial y_N} \end{bmatrix} = 2.\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix}$$

It is now necessary to calculate the Hessian matrix of F_2. As *B(r)* is the Jacobean matrix of F_2, the Hessian matrix will just be its Jacobean matrix. However, *B(r)* represents a linear transformation and its Jacobean matrix will also be the matrix itself. This fact leads to the conclusion that it is necessary to prove that *B(r)* is a positive semidefinite.

In order to verify this, the decomposition exhibited in Appendix A, $B(r) = A(r).D(r).A^t(r)$, should be used. Firstly, it is important to highlight that $A(r): \Re^N \to \Re^N$ is an isomorphism. This is easily shown by the fact that, as A is a lower

[20] According to Equation 8, $F_{\lambda,r}$ is defined in $\Re^N \times \Re^N$. The function whose Hessian matrix is being calculated is a restriction of this function. Thus, its domain is $\Re^N$. By abuse of notation they are both denoted $F_{\lambda,r}$.

triangular matrix, its determinant is given by product of the elements that lie on the main diagonal. Thus, $\det\left(A(r)\right)=1$.

Theorem B.1

$B(r)$ is positive semidefinite.

Proof:

Let $x \in \Re^N$. Since $A(r):\Re^N \to \Re^N$ is isomorphism then there is a unique $y \in \Re^N$ such that $y = A(r).x$. Therefore, $x^t.B(r).x = x^t.A^t(r).D(r).A(r).x = y^t.D(r).y = \sum_{k=1}^{N-r} y_k^2 \geq 0$,

Appendix C – Proof of the Convergence

In section 2.4, it was stated that applying an *r*-filter with $\lambda = +\infty$ is equivalent to obtaining the Least Squares Polynomial Adjustment (LSPA). In this appendix, a proof for this statement is presented by constructing $P : \Re^N \to \Re^N$, the orthogonal projection of point *x* over hyper-plane M^r, showing that $T_{\lambda,r} : \Re^N \to \Re^N$ converges to P[21].

<u>Calculating the orthogonal projection of point *x* over M^r</u>

Let $x = (x_1, x_2, \ldots x_N)$ represent an input series. Point x^* will represent the orthogonal projection of x on hyper-plane M^r if $\left\langle (x - x^*), \bar{j} \right\rangle = 0$, $\forall\ j = 0,1,\ldots, r-1$, where $\langle \cdot,\cdot \rangle$ represents the canonical inner product and $\bar{j} \in \Psi$, the M^r base. Or in matrix notation:

$$\underbrace{\begin{bmatrix} \langle \bar{0}, \bar{0} \rangle & \langle \bar{0}, \bar{1} \rangle & \cdots & \langle \bar{0}, \overline{r-1} \rangle \\ \langle \bar{1}, \bar{0} \rangle & \langle \bar{1}, \bar{1} \rangle & \cdots & \langle \bar{1}, \overline{r-1} \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \overline{r-1}, \bar{0} \rangle & \langle \overline{r-1}, \bar{1} \rangle & \cdots & \langle \overline{r-1}, \overline{r-1} \rangle \end{bmatrix}_{rxr}}_{Q} \cdot \underbrace{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_r \end{bmatrix}_{rx1}}_{a} = \begin{bmatrix} \langle x, \bar{0} \rangle \\ \langle x, \bar{1} \rangle \\ \vdots \\ \langle x, \overline{r-1} \rangle \end{bmatrix}_{rx1} = \underbrace{\begin{bmatrix} \bar{0}' \\ \bar{1}' \\ \vdots \\ \overline{r-1}' \end{bmatrix}}_{R^T}{}_{rxN} \cdot x_{Nx1}$$

where $(a_1, \ldots, a_N)$ are the coordinates of x^* in the base Ψ.

It is important to observe that this system has a unique solution since its determinant [22] will equal zero if and only if $N = 0, 1, -1, 2, -2, \cdots, r-1, -(r-1)$. Remembering that *N* represents the sample size, it must be a positive number. Additionally, it is also required that it be greater than 2*r*, given that applying a filter with an order greater than half of the sample size does not make sense.

Expressing $x^* = a_1.\bar{0} + a_2.\bar{1} + \ldots + a_r.\overline{r-1} = R.a$ and $a = Q^{-1}.R^T.x$, then $x^* = \left(R.Q^{-1}.R^T\right)x$.

Let $P = R.Q^{-1}.R^T : \Re^N \to \Re^N$. Since P is the orthogonal projection over M^r then it is easy to observe that:

---

[21] M^r and $T_{\lambda,r}$ were previously defined in sub-section 2.4 and sub-section 2.3.

[22] $\det = k.N^r.\prod_{j=1}^{r-1} (N-j)^{r-j} (N+j)^{r-j}$

Lemma C.1

x is a fixed point of $P:\Re^N \to \Re^N$ if and only if $x \in M^r$.

Lemma C.2

$P = R.Q^{-1}.R^T$ is symmetric and so diagonalizable. Furthermore P has rank r and exactly r eigenvalues one and $N-r$ eigenvalues zero.

Proof:

Firstly, it is important to verify that $P^2 = P$. Let $x \in \Re^N$ and $x^* = P.x$ respectively represent an input sequence with N observations and its orthogonal projection over M^r.

Thus: $P^2.x = P(Px) = P.x^* = x^* = P.x$

Now consider $v \in \Re^N$ and $\xi \geq 0$ such that $P.v = \xi.v$. Then:

$$P^2.v = P.v = \xi.v \;\Rightarrow\; P(Pv) = \xi.v \;\Rightarrow\; P(\xi v) = \xi.v \;\Rightarrow\; \xi.P(v) = \xi.v \;\Rightarrow\; \xi^2.v = \xi.v$$

Since by definition of eigenvector, v is not the zero vector then $\xi = 0$ or $\xi = 1$. If $\xi = 1$ the $P.v = v$. Applying Lemma C.1, its possible to conclude that $v \in M^r$. So, there are exactly r eigenvalues equals to one. On the other hand $P.v = 0$ if and only if v is in the orthogonal complement of M^r, which has dimension $N-r$. So, there are exactly $N-r$ eigenvalues zero. This concludes proof of the lemma.

Characteristics of $T_{\lambda,r}$

Lemma C.3

Point $x \in \Re^N$ is a fixed point of $T_{\lambda,r}(r):\Re^N \to \Re^N$ if and only if $x \in M^r$

Proof:

$$T.x = x \;\Leftrightarrow\; (I + \lambda B)^{-1}.x = x \;\Leftrightarrow\; (I + \lambda B).x = x \;\Leftrightarrow\; B.x = 0 \;\Leftrightarrow\; x \in M^r$$

The last step uses the fact that $\ker(B(r)) = M^r$ [23].

[23] The fact that $M^r \subset \ker(B(r))$ is easily proven considering that *B(r)* is a linear transformation and *B(r).p = 0* for all *p* that belongs to M^r base. Using the fact that *rank(B(r)) = N – r* (Property A.3), it is easy to see

Theorem C.4

Let x be any point in $\Re^N$. Then the isomorphism $T_{\lambda,r} : \Re^N \to \Re^N$ applies x over y, in such a way that, for any $\lambda > 0$, $\langle (x-y), z \rangle = 0, \ \forall z \in M^r$.

Proof:

$\langle (x-y), z \rangle = \langle x, z \rangle - \langle T_\lambda(r).x, z \rangle = \langle x, z \rangle - \langle x, T_\lambda(r).z \rangle = \langle x, z \rangle - \langle x, z \rangle = 0$, where the last step comes from the fact that $T_{\lambda,r}(r) : \Re^N \to \Re^N$ fixes M^r, and the step before the last is because $T_{\lambda,r} : \Re^N \to \Re^N$ is a symmetric operator.

Lemma C.5

$T_{\lambda,r}^{-1} = I + \lambda.B(r)$ is positive definite. Moreover, for every $\lambda > 0$ the matrix $\lambda.B$ has exactly $N-r$ eigenvalues strictly greater than zero and r eigenvalues equal to zero.

Proof:

$T_{\lambda,r}^{-1} = I + \lambda\, B(r)$ is a sum of two symmetric matrices (Property A.1), one which is positive definite and the other, positive semidefinite (Theorem B1). Thus, $T_{\lambda,r}^{-1}$ is positive definite, leading to the conclusion that $T_{\lambda,r} : \Re^N \to \Re^N$ also is.

Using the fact that $B(r)$ is symmetric and all of its coefficients are real numbers, it is possible to conclude that it is diagonalizable and all its eigenvalues are also real numbers. Moreover, as it is positive semidefinite and its rank is equal to $N-r$ (Property A.3), it has exactly r eigenvalues equal to zero and $N-r$ eigenvalues greater than zero. This property also holds for $\lambda.B(r)$, since $\lambda > 0$.

Lemma C.6

The change of base that diagonalizes $T_{\lambda,r}$ and $B(r)$ is the same, that is, there is an inversible and orthogonal matrix N such that $B(r) = N^{-1}.S.N$ and $T = N^{-1}.\Lambda.N$.

Proof:

Let $N^{-1} = N^T$ be such that $B(r) = N^{-1}.S.N$. Then:

$T_{\lambda,r}^{-1} = I + \lambda\, B = I + \lambda\, N^{-1}.S.N = N^{-1}\left(I + \lambda\, S\right)N$.

So $T_{\lambda,r}^{-1} = N^{-1}\Lambda^{-1}N$, where Λ is a diagonal matrix. So $\Lambda^{-1} = I + \lambda.S$

that dim(ker($B(r)$)) = r = dim(M^r). This shows that M^r = ker($B(r)$). The whole proof is available with the

$$S=\begin{bmatrix}\sigma_1 & & & & & \\ & \ddots & & & 0 & \\ & & \sigma_{N-r} & & & \\ & & & 0 & & \\ & 0 & & & \ddots & \\ & & & & & 0\end{bmatrix};\quad \Lambda^{-1}=\begin{bmatrix}1+\lambda\sigma_1 & & & & & \\ & \ddots & & & 0 & \\ & & 1+\lambda\sigma_2 & & & \\ & & & 1 & & \\ & 0 & & & \ddots & \\ & & & & & 1\end{bmatrix}$$

Convergence of $T_{\lambda,r}^{-1}$ on P

It is important to observe that $T_{\lambda,r}^{-1}$ has all eigenvalues equal or greater than one. As $\lambda \to +\infty$, $N-r$ of them goes to infinite while r of them remains equal to one. Considering that $T_{\lambda,r} = N^{-1}.\Lambda.N$, it is clear that, as $\lambda \to +\infty$, exactly $N-r$ eigenvalues goes to zero and r of them are equal to one. Therefore, it can be concluded that $\lim\limits_{\lambda\to\infty}\Lambda$ and P have the same eigenvalues.

To finalize it is enough to show the following result that says that the directions in M^r are those corresponding to eigenvectors of both P and $T_{\lambda,r}$ associated to eigenvalue one and that the orthogonal complement of M^r is the space generated by the eigenvectors of P corresponding to an eigenvalue zero or generated by eigenvectors of $T_{\lambda,r}$ associated to eigenvalues σ such that $0<\sigma<1$.

Theorem C.7

a) v is an eigenvector of P corresponding to an eigenvalue one if and only if it is an eigenvector of $T_{\lambda,r}$ corresponding to an eigenvalue one.

b) Moreover v is an eigenvector of P corresponding to an eigenvalue zero, that is $Pv=0$ or equivalently $v\in\ker(P)$, if and only if v belongs to the orthogonal complement of M^r.

c) If this happens then v is a linear combination of eigenvectors of $T_{\lambda,r}$ associated to eigenvalues σ_i such that $0<\sigma_i<1$.

d) Furthermore $|T_{\lambda,r}.v|<|v|$.

---

authors.

Proof:

a) Let r be fixed. So $P.v = v$ if and only if $v \in M^r$ if and only if $T_{\lambda,r}.v = v, \quad \forall \lambda > 0$. The first part is given by lemma E.1 and the second by lemma E.2.

b) Now, $P.v = 0$ if and only if v belongs to the orthogonal complement of M^r because P is the orthogonal projection over M^r.

The only if part follows easily because P is the orthogonal projection over M^r.

c) Let v be in the orthogonal complement of M^r. Then:

$v = \sum_{i=1}^{N-r} a_i v_i$ and $\langle v_i, v_j \rangle = 0, \forall i \neq j$, $a_i \in \mathbb{R}$, $\forall i$. This coordinates a_i can be taken all non-negative (if not just change the direction of v_i), so that $a_i = |a_i|$.

where $\{v_i\}_{i=1}^{N-r}$ is a set of orthonormal eigenvectors of $T_{\lambda,r}$ because $T_{\lambda,r}$ is a symmetric operator and by the spectral theorem it has an orthonormal basis of eigenvectors.

It was earlier explained that for every eigenvalue σ of $T_{\lambda,r}$ the relation $0 < \sigma \leq 1$ holds.

All the eigenvectors of $T_{\lambda,r}$ associated to eigenvalues one are in M^r, so all the others must be in the orthogonal complement of M^r.

So $T.v = T_{\lambda,r}\left(\sum_{i=1}^{N-r} a_i v_i\right) = \sum_{i=1}^{N-r} a_i T_{\lambda,r}.v_i = \sum_{i=1}^{N-r} a_i \sigma_i v_i$, where $\{\sigma_i\}_{i=1}^{N-r}$ is the set of eigenvalues of $T_{\lambda,r}$ associated to eigenvectors in the orthogonal complement of M^r and as it was previously proved $0 < \sigma_i < 1$.

d) Finally:

$$|T.v| = \left|T_{\lambda,r}\left(\sum_{i=1}^{N-r} a_i v_i\right)\right| = \left|\left(\sum_{i=1}^{N-r} a_i T_{\lambda,r} v_i\right)\right| = \left|\left(\sum_{i=1}^{N-r} a_i \sigma_i v_i\right)\right| = \sqrt{\sum_{i=1}^{N-r} (a_i \sigma_i)^2} < \sqrt{\sum_{i=1}^{N-r} a_i^{\,2}} = \left|\left(\sum_{i=1}^{N-r} a_i v_i\right)\right| = |v|$$

, where the strict inequality holds because $0 < \sigma_i < 1$, $\langle v_i, v_j \rangle = 0, \forall i \neq j$ and not all of the a_i are zero.

This concludes the proof of this theorem.

Appendix D – Other examples

In this appendix, three controlled experiments are presented to evidence problems related to the use of the HP filter on series that are dominated by a high order polynomial trend[24] or that are integrated of order equal or superior to two. It also shows how high order filters can be used to avoid those problems.

Initially, a series that includes only a polynomial trend was generated. Therefore, applying a filter to this series in order to extract its trend should return the same series. However, when the HP filter is applied to a third order polynomial, the filtered series depends on the smoothing factor, λ. Only when λ is small, the resulting trend approximates the correct trend. This fact does not occur when a fourth order filter[25] is used. The trend is perfectly identified for any smoothing factor, λ.

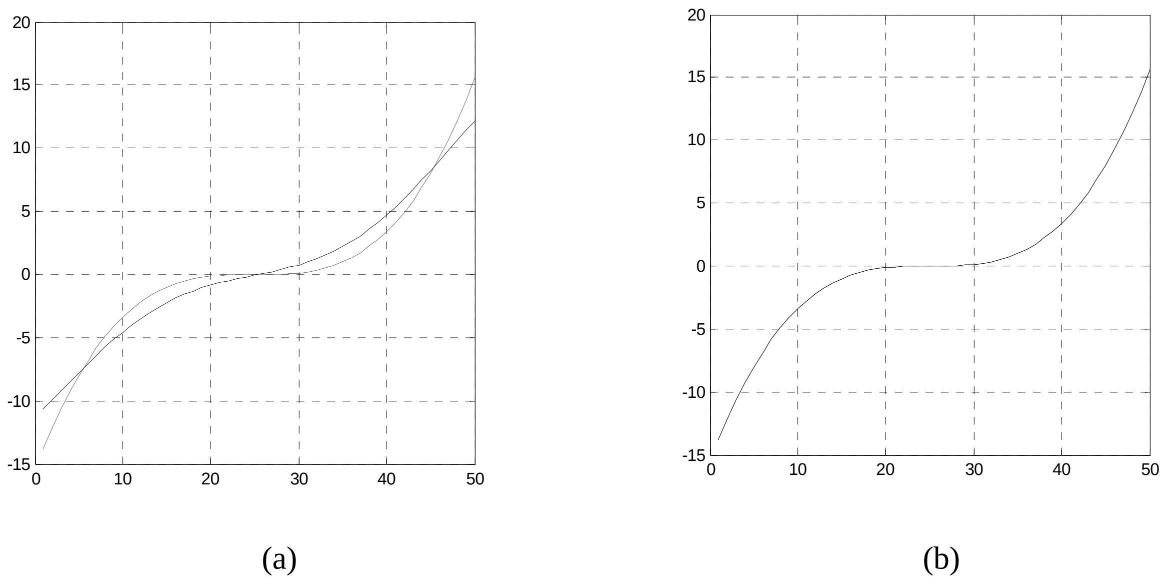


Figure D.1: Gray -> actual trend ; Black -> fitted trend
(a) $r = 2$, $\lambda = 1600$
(b) $r = 4$, $\lambda = 2560000$, fitted and actual trend are coincident. It happens for any λ.

[24] In the present context, high order means second order or superior.
[25] In fact, any filter with order equal or superior to 4 could have been used. The filter order should be superior to the polynomial order.

The same behavior can be observed when a noise[26] is added to the input series. It is important to emphasize that misidentifying the trend can lead to false conclusions. Figure D.2 (a) shows, for instance, a period that the series is continuously below its trend. If it were potential output estimation, the output gap would be negative during this entire period. However, when the trend is correctly identified, it becomes clear that this conclusion does not hold.

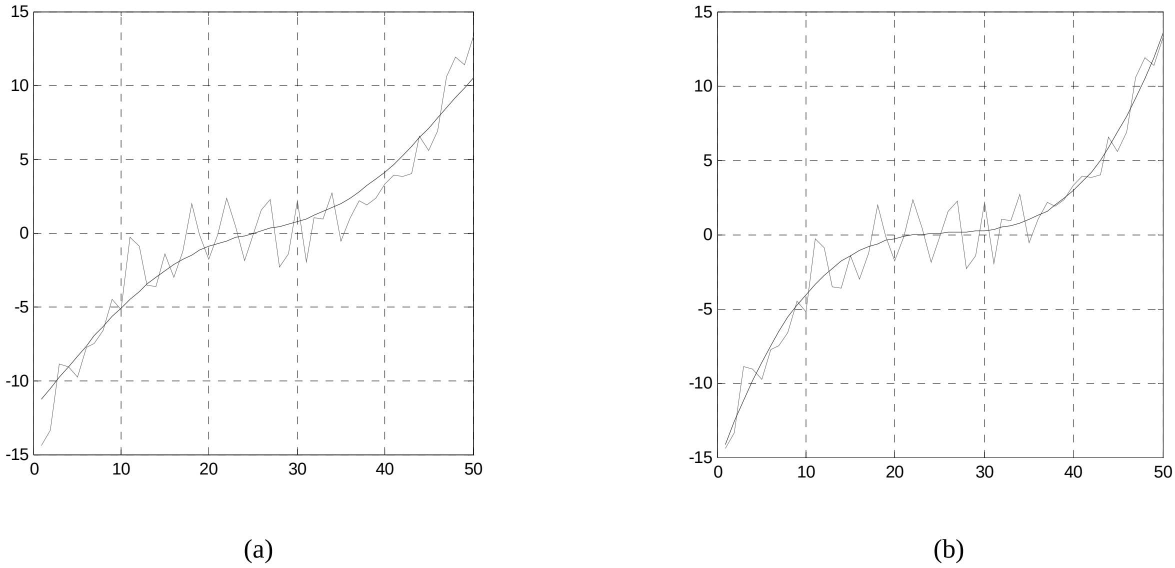
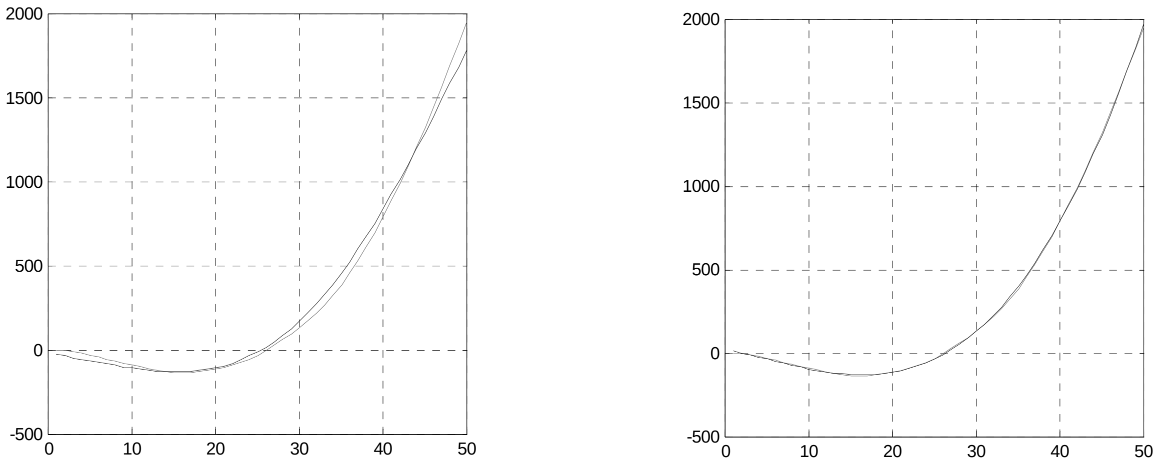


Figure D.2: Gray -> actual data ; Black -> fitted trend
(a) $r = 2$, $\lambda = 1600$
(b) $r = 4$, $\lambda = 2560000$ (same cut-off frequency). As in the previous case the time trend is perfectly fitted .

[26] Uniformly distributed in the interval [-3, 3].

The last experiment considers a series that is integrated of order 3. Figure D.4 gives evidence that using a high order filter can make the residual stationary[27].



(a) (b)

Figure D.3: Gray -> actual data ; Black -> fitted trend
(a) $r = 2$, $\lambda = 1600$
(b) $r = 4$, $\lambda = 2560000$ (same cut-off frequency).

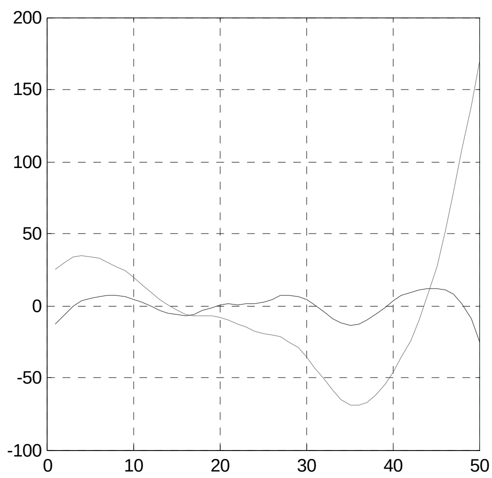


Figure D.4: residues when $r = 2$ (Gray) and $r = 4$ (Black)

[27] Despite the evidence shown, there is no proof that this is a general result.

Banco Central do Brasil

Trabalhos para Discussão

Os Trabalhos para Discussão podem ser acessados na internet, no formato PDF,
no endereço: *http://www.bc.gov.br*

Working Paper Series

Working Papers in PDF format can be downloaded from: http://www.bc.gov.br

1	**Implementing Inflation Targeting in Brazil** *Joel Bogdanski, Alexandre Antonio Tombini and Sérgio Ribeiro da Costa Werlang*	Jul/2000
2	**Política Monetária e Supervisão do Sistema Financeiro Nacional no Banco Central do Brasil** *Eduardo Lundberg*	Jul/2000
	Monetary Policy and Banking Supervision Functions on the Central Bank *Eduardo Lundberg*	Jul/2000
3	**Private Sector Participation: a Theoretical Justification of the Brazilian Position** *Sérgio Ribeiro da Costa Werlang*	Jul/2000
4	**An Information Theory Approach to the Aggregation of Log-Linear Models** *Pedro H. Albuquerque*	Jul/2000
5	**The Pass-Through from Depreciation to Inflation: a Panel Study** *Ilan Goldfajn and Sérgio Ribeiro da Costa Werlang*	Jul/2000
6	**Optimal Interest Rate Rules in Inflation Targeting Frameworks** *José Alvaro Rodrigues Neto, Fabio Araújo and Marta Baltar J. Moreira*	Jul/2000
7	**Leading Indicators of Inflation for Brazil** *Marcelle Chauvet*	Sep/2000
8	**The Correlation Matrix of the Brazilian Central Bank's Standard Model for Interest Rate Market Risk** *José Alvaro Rodrigues Neto*	Sep/2000
9	**Estimating Exchange Market Pressure and Intervention Activity** *Emanuel-Werner Kohlscheen*	Nov/2000
10	**Análise do Financiamento Externo a uma Pequena Economia Aplicação da Teoria do Prêmio Monetário ao Caso Brasileiro: 1991–1998** *Carlos Hamilton Vasconcelos Araújo e Renato Galvão Flôres Júnior*	Mar/2001
11	**A Note on the Efficient Estimation of Inflation in Brazil** *Michael F. Bryan and Stephen G. Cecchetti*	Mar/2001
12	**A Test of Competition in Brazilian Banking** *Márcio I. Nakane*	Mar/2001

13	**Modelos de Previsão de Insolvência Bancária no Brasil** *Marcio Magalhães Janot*	Mar/2001
14	**Evaluating Core Inflation Measures for Brazil** *Francisco Marcos Rodrigues Figueiredo*	Mar/2001
15	**Is It Worth Tracking Dollar/Real Implied Volatility?** *Sandro Canesso de Andrade and Benjamin Miranda Tabak*	Mar/2001
16	**Avaliação das Projeções do Modelo Estrutural do Banco Central do Brasil para a Taxa de Variação do IPCA** *Sergio Afonso Lago Alves*	Mar/2001
	Evaluation of the Central Bank of Brazil Structural Model's Inflation Forecasts in an Inflation Targeting Framework *Sergio Afonso Lago Alves*	Jul/2001
17	**Estimando o Produto Potencial Brasileiro: uma Abordagem de Função de Produção** *Tito Nícias Teixeira da Silva Filho*	Abr/2001
	Estimating Brazilian Potential Output: a Production Function Approach *Tito Nícias Teixeira da Silva Filho*	Aug/2002
18	**A Simple Model for Inflation Targeting in Brazil** *Paulo Springer de Freitas and Marcelo Kfoury Muinhos*	Apr/2001
19	**Uncovered Interest Parity with Fundamentals: a Brazilian Exchange Rate Forecast Model** *Marcelo Kfoury Muinhos, Paulo Springer de Freitas and Fabio Araújo*	May/2001
20	**Credit Channel without the LM Curve** *Victorio Y. T. Chu and Márcio I. Nakane*	May/2001
21	**Os Impactos Econômicos da CPMF: Teoria e Evidência** *Pedro H. Albuquerque*	Jun/2001
22	**Decentralized Portfolio Management** *Paulo Coutinho and Benjamin Miranda Tabak*	Jun/2001
23	**Os Efeitos da CPMF sobre a Intermediação Financeira** *Sérgio Mikio Koyama e Márcio I. Nakane*	Jul/2001
24	**Inflation Targeting in Brazil: Shocks, Backward-Looking Prices, and IMF Conditionality** *Joel Bogdanski, Paulo Springer de Freitas, Ilan Goldfajn and Alexandre Antonio Tombini*	Aug/2001
25	**Inflation Targeting in Brazil: Reviewing Two Years of Monetary Policy 1999/00** *Pedro Fachada*	Aug/2001
26	**Inflation Targeting in an Open Financially Integrated Emerging Economy: the Case of Brazil** *Marcelo Kfoury Muinhos*	Aug/2001

27	**Complementaridade e Fungibilidade dos Fluxos de Capitais Internacionais** *Carlos Hamilton Vasconcelos Araújo e Renato Galvão Flôres Júnior*	Set/2001
28	**Regras Monetárias e Dinâmica Macroeconômica no Brasil: uma Abordagem de Expectativas Racionais** *Marco Antonio Bonomo e Ricardo D. Brito*	Nov/2001
29	**Using a Money Demand Model to Evaluate Monetary Policies in Brazil** *Pedro H. Albuquerque and Solange Gouvêa*	Nov/2001
30	**Testing the Expectations Hypothesis in the Brazilian Term Structure of Interest Rates** *Benjamin Miranda Tabak and Sandro Canesso de Andrade*	Nov/2001
31	**Algumas Considerações sobre a Sazonalidade no IPCA** *Francisco Marcos R. Figueiredo e Roberta Blass Staub*	Nov/2001
32	**Crises Cambiais e Ataques Especulativos no Brasil** *Mauro Costa Miranda*	Nov/2001
33	**Monetary Policy and Inflation in Brazil (1975-2000): a VAR Estimation** *André Minella*	Nov/2001
34	**Constrained Discretion and Collective Action Problems: Reflections on the Resolution of International Financial Crises** *Arminio Fraga and Daniel Luiz Gleizer*	Nov/2001
35	**Uma Definição Operacional de Estabilidade de Preços** *Tito Nícias Teixeira da Silva Filho*	Dez/2001
36	**Can Emerging Markets Float? Should They Inflation Target?** *Barry Eichengreen*	Feb/2002
37	**Monetary Policy in Brazil: Remarks on the Inflation Targeting Regime, Public Debt Management and Open Market Operations** *Luiz Fernando Figueiredo, Pedro Fachada and Sérgio Goldenstein*	Mar/2002
38	**Volatilidade Implícita e Antecipação de Eventos de *Stress*: um Teste para o Mercado Brasileiro** *Frederico Pechir Gomes*	Mar/2002
39	**Opções sobre Dólar Comercial e Expectativas a Respeito do Comportamento da Taxa de Câmbio** *Paulo Castor de Castro*	Mar/2002
40	**Speculative Attacks on Debts, Dollarization and Optimum Currency Areas** *Aloisio Araujo and Márcia Leon*	Apr/2002
41	**Mudanças de Regime no Câmbio Brasileiro** *Carlos Hamilton V. Araújo e Getúlio B. da Silveira Filho*	Jun/2002
42	**Modelo Estrutural com Setor Externo: Endogenização do Prêmio de Risco e do Câmbio** *Marcelo Kfoury Muinhos, Sérgio Afonso Lago Alves e Gil Riella*	Jun/2002

43	**The Effects of the Brazilian ADRs Program on Domestic Market Efficiency** *Benjamin Miranda Tabak and Eduardo José Araújo Lima*	Jun/2002
44	**Estrutura Competitiva, Produtividade Industrial e Liberação Comercial no Brasil** *Pedro Cavalcanti Ferreira e Osmani Teixeira de Carvalho Guillén*	Jun/2002
45	**Optimal Monetary Policy, Gains from Commitment, and Inflation Persistence** *André Minella*	Aug/2002
46	**The Determinants of Bank Interest Spread in Brazil** *Tarsila Segalla Afanasieff, Priscilla Maria Villa Lhacer and Márcio I. Nakane*	Aug/2002
47	**Indicadores Derivados de Agregados Monetários** *Fernando de Aquino Fonseca Neto e José Albuquerque Júnior*	Set/2002
48	**Should Government Smooth Exchange Rate Risk?** *Ilan Goldfajn and Marcos Antonio Silveira*	Sep/2002
49	**Desenvolvimento do Sistema Financeiro e Crescimento Econômico no Brasil: Evidências de Causalidade** *Orlando Carneiro de Matos*	Set/2002
50	**Macroeconomic Coordination and Inflation Targeting in a Two-Country Model** *Eui Jung Chang, Marcelo Kfoury Muinhos and Joanílio Rodolpho Teixeira*	Sep/2002
51	**Credit Channel with Sovereign Credit Risk: an Empirical Test** *Victorio Yi Tson Chu*	Sep/2002
52	**Generalized Hyperbolic Distributions and Brazilian Data** *José Fajardo and Aquiles Farias*	Sep/2002
53	**Inflation Targeting in Brazil: Lessons and Challenges** *André Minella, Paulo Springer de Freitas, Ilan Goldfajn and Marcelo Kfoury Muinhos*	Nov/2002
54	**Stock Returns and Volatility** *Benjamin Miranda Tabak and Solange Maria Guerra*	Nov/2002
55	**Componentes de Curto e Longo Prazo das Taxas de Juros no Brasil** *Carlos Hamilton Vasconcelos Araújo e Osmani Teixeira de Carvalho de Guillén*	Nov/2002
56	**Causality and Cointegration in Stock Markets: the Case of Latin America** *Benjamin Miranda Tabak and Eduardo José Araújo Lima*	Dec/2002
57	**As Leis de Falência: uma Abordagem Econômica** *Aloisio Araujo*	Dez/2002
58	**The Random Walk Hypothesis and the Behavior of Foreign Capital Portfolio Flows: the Brazilian Stock Market Case** *Benjamin Miranda Tabak*	Dec/2002
59	**Os Preços Administrados e a Inflação no Brasil** *Francisco Marcos R. Figueiredo e Thaís Porto Ferreira*	Dez/2002

60	**Delegated Portfolio Management** *Paulo Coutinho and Benjamin Miranda Tabak*	Dec/2002
61	**O Uso de Dados de Alta Freqüência na Estimação da Volatilidade e do Valor em Risco para o Ibovespa** *João Maurício de Souza Moreira e Eduardo Facó Lemgruber*	Dez/2002
62	**Taxa de Juros e Concentração Bancária no Brasil** *Eduardo Kiyoshi Tonooka e Sérgio Mikio Koyama*	Fev/2003
63	**Optimal Monetary Rules: the Case of Brazil** *Charles Lima de Almeida, Marco Aurélio Peres, Geraldo da Silva e Souza and Benjamin Miranda Tabak*	Feb/2003
64	**Medium-Size Macroeconomic Model for the Brazilian Economy** *Marcelo Kfoury Muinhos and Sergio Afonso Lago Alves*	Feb/2003
65	**On the Information Content of Oil Future Prices** *Benjamin Miranda Tabak*	Feb/2003
66	**A Taxa de Juros de Equilíbrio: uma Abordagem Múltipla** *Pedro Calhman de Miranda e Marcelo Kfoury Muinhos*	Fev/2003
67	**Avaliação de Métodos de Cálculo de Exigência de Capital para Risco de Mercado de Carteiras de Ações no Brasil** *Gustavo S. Araújo, João Maurício S. Moreira e Ricardo S. Maia Clemente*	Fev/2003
68	**Real Balances in the Utility Function: Evidence for Brazil** *Leonardo Soriano de Alencar and Márcio I. Nakane*	Feb/2003