

ECONOMIA APLICADA

1998, 2(3)

116
BIBLIOTECA ACADÊMICA
Faculdade de Ciências da Administração e
Contabilidade da Universidade Federal do Rio de Janeiro

116
PCRP

Sumário

ARTIGOS

Efficiency and Equity Aspects of Social Spending in Selected Countries of Latin America and East Asia: A Comparative Approach 445

Maria da Conceição Sampaio de Sousa

Previsão de Séries de Tempo na Presença de Mudança Estrutural: Redes Neurais Artificiais e Modelos Estruturais 487

Wilson R. Corrêa
Marcelo S. Portugal

Pluriatividade e Ruralidade: Aspectos Metodológicos 515

Angela Kageyama

Encargos Sociais e Demanda por Trabalho no Setor Formal da Economia 553

Reynaldo Fernandes

An Econometric Analysis of the Effects of Macroeconomic Variables on Interest Rates in Brazil 579

Fernando Blumenschein

DEBATE

A Crise Asiática e o Papel do FMI 609

Ruben Dario Almonacid

COMO EU PESQUISEI

A Arte do Direito e da Economia: Um Artigo Autobiográfico 627

William M. Landes

Dedicação à Pesquisa 647

Juan Hersztajn Moldau

Revista Economia Aplicada/Departamento de Economia da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo e Fundação Instituto de Pesquisas Econômicas.

-- v.2, n.3 (1998)- -- São Paulo: FEA/USP-FIPE, 1998-

Trimestral

ISSN 1413-8050

I. Economia. I. Universidade de São Paulo. Faculdade de Economia, Administração e Contabilidade. Departamento de Economia. II. Fundação Instituto de Pesquisas Econômicas.

CDD - 330

Efficiency and equity aspects of social spending in selected countries of Latin America and East Asia: a comparative approach

Maria da Conceição Sampaio de Sousa[§]

ABSTRACT

The paper reviews the main features of social expenditures in Latin America and East Asia. The objective is to evaluate the efficiency and equity aspects associated with education, health care, and social security spending in both regions. Emphasis is given to the impact of these expenditures on income distribution and to its effects on the welfare of the low-income population. Furthermore, the relationship between the public and private sectors are examined in order to inquire to what extent public expenditures in social areas, complement, rather than substitute for private activities.

Key words: social expenditures, public spending, Latin America, East Asia.

RESUMO

O artigo examina os principais aspectos dos gastos sociais na América Latina e Leste Asiático com o objetivo de avaliar a eficiência e equidade dos aspectos associados aos gastos com educação, saúde e seguridade social em ambas as regiões. Ênfase é dada ao impacto dos gastos sobre a distribuição de renda e seus efeitos sobre o bem-estar da população de baixa renda. Além disso, é analisada a relação entre os setores público e privado com vistas a investigar em que medida os gastos públicos em áreas sociais complementam, ao invés de substituir, atividades privadas.

Palavras-chave: gasto social, despesa pública, América Latina, Leste Asiático.

§ Do Departamento de Economia Universidade de Brasília (E-mail: mcss@guarany.cpd.unb.br)

The author would like to thank the Conselho Nacional de Desenvolvimento Científico and Tecnológico (CNPq) and Japanese Society for the Promotion of Science (JSPS) for support during the elaboration of this paper.

Recebido em janeiro de 1998. Aceito em abril de 1998.

1 Introduction

Investments in social services, particularly in education and health, are critical elements for economic development. The positive link between the availability of an adequate supply of human capital and development is fairly robust across studies and methodologies.¹ For instance, theoretical models of endogenous growth point to the fact that public social investments produce externalities and may increase the long-term prospects of the economy. Yet, the benefits of public spending are not limited to its possible consequences on growth. They are also supposed to bring about a very much needed equity to the process of economic development. There is now an emerging consensus that access to certain publicly provided goods such as basic education and essential health care may considerably reduce poverty and improve human development.² Experiences in many countries corroborate these results. Public spending policies have played an important role in creating and preserving a more equitable form of economic growth. Yet, this path has not been fully utilized. Recent studies have shown that the potential of public spending to reduce economic inequalities has been underestimated.³ Hence, to promote equality by increasing the availability of public goods seems to be a very promising route.

Once the necessity of appropriate public policies is accepted, the problem becomes how to allocate government spending in such a way as to avoid the crowding out of the typical activities of the private sector. Public expenditures, ideally, should be allocated within and across sectors so as to maximize social welfare including the impact on the poor. The modern economics of the public sector, strongly based on welfare theory, provides a framework within which these issues may be addressed. Accordingly, public expenditures should focus on goods and services not provided (or under-provided) by private markets. Hence, public activities are to take place where clear market failures are detected. In the particular case of education and health, this implies a higher concentration of resources on primary schooling and basic health care. Moreover, it also involves a cutback of spending in higher education as well as a move away from a more expensive and discriminatory curative medicine in favor of greater effort for prevention.

The important question here is to what extent these patterns apply to Latin America and East Asia. Do these regions, when establishing the priorities associated with their public

1 For example, see Romer (1986), Lucas (1988), Barro and Lee (1993), Easterly and Rebelo (1993), among others.

2 See Anand and Ravallion (1993).

3 Harberger (1977), Goode (1984), Sen (1981), Cornes (1982), among others.

expenditures policies, follow those recommendations? Is the rationale for public spending convergent across regions or does their allocation of public funds follow a clearly differentiated pattern? To answer these and other related questions, this paper will investigate the social public spending patterns in both regions, emphasizing the efficiency and equity aspects of such policies. We are particularly interested in evaluating the role played in improving income distribution. We will argue that East Asia was particularly successful in using social expenditures to correct typical market failures whereas in Latin America policies lagged behind in terms of equity and efficiency. Indeed, recent evidence seems to point to a greater efficiency of East Asia in identifying a group of potential reallocations of public spending that would simultaneously improve efficiency and distribution.⁴ By contrast, social public spending in Latin America was trapped by the traditional dilemma between equity and efficiency and seems to have played only a minor role in attenuating inequalities.

This paper is organized as follows. Sections 2 and 3 present and discuss the structure of public spending in education and health in Latin America and East Asia from a comparative perspective. Emphasis will be given to the analysis of equity and efficiency aspects of social spending. Section 4 examines social security spending and Section 5 analyzes the relationship between public expenditure and income distribution. Section 6 reviews the key issues and trends in regional patterns of public spending and investigates the main source of variations across regions. Finally, Section 7 draws some lessons and conclusions from the experience of the two regions.

2 Efficiency and equity aspects of educational services

One can convincingly argue that, whatever objectives are assigned to schooling, it is always desirable that they should be efficiently fulfilled. This is particularly true in developing countries, where the opportunity cost of resources used in education may be very high. An education system is said to be efficient if it produces the desired output at minimum cost or, if for a given availability of resources, it maximizes the desired output. In such a system, resources must be efficiently allocated by prioritizing inputs that effectively enhance learning and are cost-effective, and avoid the use of ineffective and expensive inputs.

4 Birdsall and James (1993).

Still, efficiency is not the only point to be taken into account. Education affects productivity and thus the entry to the labor market. In addition, it has a significant effect on earnings, as the rates of return on education - private as well as social - are high, especially, in developing countries.⁵ Therefore, it may considerably influence poverty as well as the patterns of income distribution.⁶ Hence, if governments want to achieve universal and fair schooling, it is necessary to ensure that all segments of the population have access to the benefits of education and to the economic and social mobility it provides. Education systems should thus be equitable.

This section examines the efficiency and equity of the schooling systems in Latin America and East Asia. A number of aspects deserve to be examined such as what factors are responsible for better results in the learning process, whether the regions are effectively using their limited resources in order to maximize outcomes and whether the education system is accessible to the disadvantaged. These and other related questions will be considered below.

2.1 Cost-efficiency of schooling in Latin America and East Asia

As efficiency in education has been analyzed in a vast literature, the paper does not cover all of its aspects. Instead, it will concentrate on topics considered more relevant for which data are available. Below we will briefly discuss the main outcomes of the educational system in both regions as well as the inputs used to reach such outcomes.

2.1.1 Output efficiency: school completion, repetition and learning achievements

Table 1 presents selected indicators of output efficiency of the education systems for Latin America and East Asia. From the outset, it should be noticed that access to primary schooling is universal in both regions. However, when we consider the patterns of cohort survival, the differences, across and within regions, are impressive. In East Asia, the dropout rate is much lower than in Latin America. Most of the East Asian students complete primary school whereas in Latin America, except for Uruguay, rates of survival are

5 Psacharopoulos (1993b).

6 Recent studies confirm the crucial role played by education levels on the patterns of income distribution in Latin America. See, for instance, Psacharopoulos (1993), Barros (1992), and Cardoso and Helwege (1992).

considerably lower and repetition rates much higher.⁷ Such a result offer clues about the quality of the education system. If students fail to progress, the chances of dropout probably increase thus reducing the survival rates.⁸ Resources spent on these students are partly wasted because they acquire only part of the required cognitive skills.

Notice that low cohort survival rates are not inevitable outcomes in poor countries. The Philippines, Indonesia and Bolivia do much better than Brazil, a country with a higher per capita income. This suggests that policy choices within primary education also play a critical role in the effectiveness of the school system. Just as poor achievement leads to repetition and dropout, high cognitive achievement is associated with promotion to higher levels of schooling. Not surprisingly, enrollment in secondary schooling is also much higher in East Asia than in Latin America.

Table 1
Selected Indicators of Efficiency: Patterns of Cohort Survival in
Primary and Secondary Schools, Repeaters - 1990

Countries	Primary gross enrollment ratio 1990	Percent of population entering grade 1	Percent of entrants surviving to the end of primary	Repeaters in primary - 1990	Secondary gross enrollment	Enrolled in primary going to secondary	Repeaters in the secondary
Latin America							
- Argentina	111.0	n.a	n.a	n.a	71.0	n.a	n.a
- Bolivia	85.0	100.0	44.0	3.0	34.0	n.a	5.0
- Brazil	108.0	100.0	20.0	19.0	39.0	n.a	n.a
- Chile	87.0	n.a	77.0	n.a	72.0	n.a	n.a
- Colombia	111.0	100.0	56.0	11.0	55.0	34.0	20.0
- Costa Rica	101.0	100.0	79.0	11.0	42.0	48.0	12.0
- Mexico	115.0	100.0	72.0	9.0	55.0	58.0	2.0
- Paraguay	108.0	100.0	59.0	9.0	30.0	n.a	n.a
- Peru	126.0	100.0	n.a	n.a	67.0	n.a	10.0
- Uruguay	108.0	99.0	93.0	9.0	81.0	n.a	n.a
- Venezuela	97.0	100.0	48.0	11.0	34.0	42.0	6.0
East Asia							
- Hong Kong	108.0	100.0	97.0	n.a	75.0	n.a	n.a
- Indonesia	117.0	100.0	77.0	5.0	45.0	n.a	1.0
- Korea	107.0	100.0	99.0	n.a	87.0	98.0	0.0
- Malaysia	93.0	88.0	96.0	n.a	56.0	n.a	n.a
- Philippines	112.0	100.0	70.0	2.0	73.0	65.0	2.0
- Singapore	108.0	n.a	100.0	n.a	70.0	n.a	n.a
Thailand	99.0	85.0	87.0	4.0	33.0	n.a	n.a

Source: World Bank (1994, 1996).

7 The UNESCO special survey on repetition shows that in Latin America 29% of all primary students are repeating their grade each year. For the first year this rate is 42%. As a result, Latin American countries spent US\$ 2,5 billion each year to teach these 20 million repeaters.

8 Wolff, Schiefelbein and Valenzuela (1994) found a strong negative correlation between repetition rates and school completion in Latin America. As children become older, their opportunity costs of remaining in school increases. Further repetitions discourage students and reduce the possibility of passing.

There is a striking difference between Latin America and East Asia in learning achievements in primary schools. In various standardized achievement tests East Asia scores are systematically higher than those for Latin America (Table 2). Moreover, East Asian countries do not only perform only better than their Latin American counterparts. They have also overtaken the developed nations. Results of the last TIMSS (Third International Math and Science Study) show that the region's richest countries - Singapore, Hong Kong, and Korea - took three and two, out of the top four places, in the mathematics and science tests respectively. Even Thailand, a relatively poor country, scored above the average and outperformed traditional rich industrial countries like Germany, Switzerland, England, Denmark and the United States. In the same test, Colombia, the only Latin American country present in the sample, left behind only South Africa, the country which got the worse score in both mathematics and science. In the 1992 IEA (International Education Assessment) study of reading literacy, Venezuela scored the lowest of the twenty seven participants with nearly one third of nine years-old students taking the test scored at chance levels or below. In the 1992 IAEP (International Assessment of Educational Progress), thirteen year-old students of the two Brazilian cities included - Fortaleza and São Paulo - scored slightly higher than Mozambique and far below the Asian countries included in the sample (Korea, Taiwan, and China).

Table 2
Education Achievements in Latin America and East Asia:
Results of Selected Standardized Tests

Countries	Reading Literacy	Mathematics		Science	
	Studies 1992/1993 ^b	1992/93 ^c	1994/95 ^a	1992/93 ^c	1994/95 ^a
Latin America					
- Argentina	361	38		38	
Bolivia	284				
Brazil		34		50	
- Chile	366				
- Colombia		40	385	37	411-
- Costa Rica	381	55		56	
-Dominican Rep.	303	40		37	
Ecuador	308			--	
Venezuela	383	40		41	
East Asia					
- Indonesia	394				
- Hong Kong	517		588		522
- Korea		73	607	78	565
- Singapore	515		643		607
Taiwan		73		76	
Thailand		50	522	55	525

Source: Wolff, Schiefelbein and Valenzuela (1994), Burki and Edwards (1996), The Economist, March 29, 1997. a: Third International Maths and Science Study (TIMSS); b: IEA Study of Reading Literacy; c: International Assessment of Educational Progress Test of Mathematics and Science.

Summarizing, all indicators used here seem to demonstrate the higher efficiency of the educational systems of East Asia. Such achievements become all the more remarkable when we compare levels of government spending across regions (Table 3). Spending on education, as a percentage of GDP, is roughly the same in the two regions. Yet East Asia outperforms Latin America in most of the efficiency measures considered. This remarkable result arises because regions and countries differ in the way their education systems are organized and financed. They differ in such aspects as private and public financing, class size, teaching motivations and qualifications. Because some arrangements are more efficient than others, the same expenditure can achieve very different levels of educational development. For this reason, appropriate government policies, as a determinant of education development, are at least as important as the size of public spending. Below, we will investigate the elements behind the large discrepancy between Latin American and East Asian educational patterns.

2.1.2 Cost-effectiveness of educational inputs: a comparative analysis

A cross country comparison of the spending in education reveals a striking difference between regions. In East Asia nearly all countries show a most pronounced emphasis on primary education. This is particularly true for Indonesia, Korea, Philippines and Thailand, where the share of the primary education ranges from 57 to 64 percent, well above the regional average. Compared to Latin America, East Asian governments clearly concentrate resources on the lower levels of education. This pattern of allocation was made feasible through a significant self-financing private sector in higher education. In Latin America, by sharp contrast, continuing a long-term tendency, the share of education budget devoted to primary education declined from 49% to 43% between 1980-89 while the share of higher education was steadily rising.⁹

9 Wolff, Schiefelbein and Valenzuela (1994).

Table 3
Level and Distribution of Public Spending on Education in
Selected Countries of Latin America and East Asia

Countries	Overall Government Spending on Education – Percentage			Distribution of Public Spending by Level of Education		
	GDP		Total government spending	Primary 1985	Secondary 1990	Higher 1990
	1960	1990				
Latin America						
- Argentina			-	37.7 ^a		
- Bolivia	1.5	3.0	20.1	66.0 ^b		23.0
- Brazil	1.9	4.6		45.9	10.1	26.0
- Chile	2.7	3.7		51.0	21.0	22.0
- Colombia	1.7	2.9	21.4	39.2	20.8	21.0
- Costa Rica	4.1	4.6	20.8	35.1	12.9	36.0
- Mexico	1.2	4.1		24.3	n.a	n.a
- Uruguay	3.7	3.1	15.9	37.7	30.3	23.0
- Venezuela	3.7	4.1	18.8	24.5	n.a	n.a
- Regional Average	2.6	3.8				
East Asia						
- Hong Kong		3.0	17.4	71.0 ^b		29.0
- Indonesia ^a	2.5	3.7 ^a		62.0	27.0	9.0
- Korea	2.0	3.6	22.4	57.0	34.0	7.0
- Malaysia	2.9	6.9	18.8	36.0	34.0	15.0
- Philippines	2.3	2.9	10.1	64.0	16.0	15.0
- Singapore	2.8	3.4		65.0 ^b		31.0
- Thailand	2.3	3.8	20.0	58.0	24.0	15.0
- Regional Average	2.5	3.9				

Note: (a) 1984 figures (b) 1990 figures and includes secondary education.

Source: Tan and Mingat (1992); Wolff, Schiefelbein and Valenzuela (1994); Human Development Report (1994).

So far, the discussion has been limited to the distribution of the education budget among levels. This section investigates how such funds are spent on inputs that particularly influence the effectiveness of the school system. The focus will be mainly on factors affecting the efficiency in the physical organization of school, class size, availability of teaching material, teaching motivations, and qualifications. Tables 4, 5 and 6 will be used in the discussion of those factors.

The student to school ratio in primary school is substantially higher in East Asia than in Latin America. Korea and Singapore present the highest ratios whereas Costa Rica and Colombia have the smaller schools. Due to the existence of economies of scale in the production of educational services, with average operating costs falling sharply with enrollments, this could be a significant advantage of the East Asian systems.¹⁰ Furthermore,

¹⁰ Evidence for China and Philippines confirm the existence of decreasing operating costs respectively in higher and secondary education. (Tan and Mingat, 1992)

East Asia seems to be making better use of equipment. Latin America schools fail to exploit the economies of scale and, clearly, are very far from an optimal utilization of its physical facilities. Of course, one should take into account the fact that in large countries, with scattered rural populations, small schools are more adequate. However, this fact does not explain the under-utilization of schools in Latin America.¹¹

As for the student to teacher ratio, both Latin America and East Asia follow the worldwide tendency to reduce class size in primary education. This result has been more remarkable in East Asia where these ratios have declined substantially over the last years. The causes behind this common trend are different, though. In East Asia, ratios are declining due to reduced demographic pressure that has decreased enrollment growth. In Latin America, the declining student to teacher ratio probably reflects political pressures for hiring (unskilled) teachers rather than an explicit policy aiming at increasing efficiency, evidenced by the fact that the decline in the ratio has not improved learning. Besides, the argument that links reduced class size and better learning has been strongly challenged by recent research. Several studies have found that, when other learning factors are controlled, an increase in class size does not seem to have a clear negative effect on achievement. Only 8 out of 30 studies reviewed by Harbison and Hanushek (1992) find support for smaller classes in primary schools. Korea illustrates the case. Its relatively crowded classes did not prevent it from building an excellent educational system, one of the best in the world. In fact, good management theory suggests that, until funds for instructional material are ensured, teacher student ratios should be kept constant or slightly increasing. Increasing the ratio could significantly reduce educational expenses.¹² A reduction of class size to fewer than forty is, therefore, very expensive and thus constitutes a blind alley to improve the quality of the school system in developing countries.

According to recent research, the availability of teaching materials is critical for the learning process. The importance of an adequate provision of textbooks and writing materials is widely endorsed.¹³ Furthermore, educational material is also highly cost effective. Harbison and Hanushek (1992) have shown that for every dollar invested in

11 Even in relatively large countries, it is possible to benefit from economies of scales, if appropriate measures are taken. For instance, the Thai government offered free bicycles for students living far from schools.(Lokeed and Vespoor , 1991) Besides, Costa Rica, a quite small country, has also the smallest schools, thus making clear that in Latin America the under-utilization of school space could not be explained only by the size of the country.

12 Wolff, Schiefelbein and Valenzuela (1994) argue that increasing student/teacher ratios from 27 to 30 would save 11% on the recurrent budget for primary education.

13 See Lokeed and Vespoor (1991) for an excellent survey of world research on the determinants of learning.

textbooks and other educational materials, the primary education system saved on average four dollars from reduced output and repetition. Here, Latin America fits the typical pattern of developing countries where primary students either lack textbooks and other instructional materials entirely or share them extensively with other students. The few data available (Table 5) and previous studies confirm that the financing of learning materials, including textbooks, is highly inadequate in Latin America. Indeed, in this region public expenditures on instructional materials is typically less than one percent of GDP per capita. This constitutes a serious shortcoming of the educational system and may partly explain its poor quality. In addition, it has important implication for equity as lack of public provision of such items puts enormous pressures on poor parents thus hindering equity.

Table 4
Student to School and Student to Teacher Ratios in Selected
Countries of Latin America and East Asia

Countries	Student School Ratio		Student Teacher Ratio		
	Primary		Primary		Secondary
	1965	1985	1965	1990	1990
Latin America					
Argentina	164.0	222.0	20.0	19.0	7.0
- Bolivia	72.0	148.0	28.0	25.0	18.0
- Brazil	80.0	132.0	28.0	23.0	14.0
- Colombia	96.0	121.0	36.0	30.0	20.0
- Chile	212.0	240.0	37.0	29.0	17.0
- Costa Rica	132.0	117.0	27.0	32.0	19.0
- Mexico	201.0	197.0	47.0	31.0	17.0
Paraguay	141.0	146.0	30.0	25.0	n.a
Peru	127.0	153.0	36.0	28.0	21.0
- Uruguay	146.0	148.0	31.0	22.0	n.a
Venezuela	133.0	210.0	34.0	23.0	9.0
Regional Average	136.7	166.7	32.2	26.1	15.8
East Asia					
Hong Kong	351.0	670.0	29.0	27.0	23.0
Indonesia	227.0	177.0	41.0	23.0	23.0
Korea	969.0	747.0	62.0	34.0	25.0
Malaysia	263.0	328.0	29.0	20.0	19.0
Philippines	166.0	270.0	31.0	33.0	33.0
- Singapore	714.0	962.0	29.0	26.0	n.a
Thailand	178.0	221.0	35.0	18.0	18.0
Regional Average	409.7	482.1	36.6	25.9	23.5

Source: Lockheed and Vespoor (1991), tables A-8 and A-9; Human Development Report (1995).

There are also important differences in teacher qualifications and working conditions across regions. The quality of teaching, a key determinant of student achievement, is basically determined by teacher training and motivation. Yet, the typical pattern of

developing countries encompasses a teaching force that is neither motivated nor trained. Do Latin America and East Asia share this pattern or they deviate from this doomed rule? Are there enough incentives in these regions to motivate teachers to perform well? These questions will be addressed below.

A critical point concerns teacher salaries. Here again, Latin American and East Asian patterns clearly diverge. Salaries are considerably higher in the latter region (Table 6) and there is no sign that the differential is decreasing. Corvalán (1990), analyzing five Latin American countries (Argentina, Bolivia, Colombia, Costa Rica and, Guatemala), for the period 1980-1987, finds that, except for Colombia, salaries had declined for all teaching levels. Similar evidence was found for Brazil. (Harbison and Hanushek, 1992) Latin America presents also large disparities across sectors. Psacharopoulos, Valenzuela and Arends (1993) reveal that, except for Costa Rica, urban teachers earn less than their rural counterparts. Since teachers need to supplement their income with other jobs, absenteeism is thus likely to be higher in Latin America than in East Asia.

Table 5
Expenditures on Teaching Materials School Teachers - Primary School - 1985

Countries	Expenditures on Teaching Material per Student			Average Teacher Salary		
	Year	1985 dollars	Percent of GDP per capita	Year	1985 dollars	Percent of GDP per capita
Latin America						
- Argentina				1984	3839.5	1.8
Bolivia				1982	1525.4	2.6
Brazil (1990 US \$)				1989	4560.0	1.6
Colombia				1981	2630.0	2.2
- Chile	1981	0.7	0.05	1981	2845.7	2.0
- Costa Rica	1983	1.7	0.13	1986	4760.3	3.2
Ecuador	1986	0.1	0.01	1986	2241.2	2.0
Mexico	1986	1.8	0.09	1986	1733.3	0.9
Peru	1985	0.2	0.02	1985	1145.4	1.3
Uruguay	1980	10.9	0.52	1986	1955.3	1.0
Regional Average		2.6	0.14		2723.6	1.9
East Asia						
Korea	1986	5.6	0.24	1979	9841.6	4.2
Malaysia	1984	13.6	0.71	1984	6191.5	3.2
Hog Kong	1984	1.7	0.03	1984	12843.4	2.1
Singapore	1982			1982	11063.6	1.8
Thailand	1986	3.1	0.39	1986	1861.6	2.3
- Regional Average		6.0	0.34		8360.0	2.7

Source: Lockheed and Vespoor (1991), and Wolff, Schiefelbein and Valenzuela (1994).

However, increasing wages will not necessarily improve the quality of teaching in Latin America as in this region the training of teachers is, most certainly, deficient. The proportion of untrained teachers is quite high and could reach as much as 40% in Peru and Nicaragua. Teachers have rarely completed a secondary education. Psacharopoulos *et alii* (1993) calculated that the mean years of schooling is insufficient. Consequently, without a serious program to improve teacher abilities, higher wages will only imply higher costs and, very probably, will not bring about any significant improvement in the quality of teaching.

Low wages, inadequate training, poor working conditions, and uncertain career paths are the obvious signs of the low status attributed to the teaching profession in Latin America. Thus, it is not surprising that the teaching profession fails to attract the most able students. Students entering the profession have an academic profile which is lower than that of other students in higher education. (Gysling, 1991) In East Asia, particularly in Korea and Singapore, the overall situation of teachers is undoubtedly superior. Salaries are competitive and the training much better than in Latin America. The situation will tend to further improve in the near future as declining enrollment growth in East Asia will allow educational resources to be redirected toward training and motivating teachers as well as toward improving managerial and administrative support. Furthermore, in East Asia the prestige associated with the profession is relatively high. For instance, in Korea teachers used to receive the same honors as the king and the parents (Lockeed and Vespoor, 1991) and in other countries they also enjoy a relatively good status. Hence, the profession attracts capable people and has low attrition rates.

Last but not least, the impact of inputs ultimately depends on how schools effectively use available resources. Good teaching needs an orderly environment, clear goals, and strong instructional leadership, which are only provided by an efficient institutional framework. Managerial capability and institutional strength are crucial to implement and sustain high-quality education systems. Here again, East Asian countries are well ahead of their Latin American counterparts as they are actively trying to improve school management by training principals and other school staff. Malaysia, Thailand, the Philippines and Korea are leading the way in developing coherent policies to tackle these issues.¹⁴ Korea, in particular, has created institutes that assess and provide information about the educational system and disseminate examples of good practices among schools. These institutions

14 Den Hartog Georgiades and Jones (1989).

effectively contribute to the remarkable performance of Korean students in international achievement tests. It is worth noticing that examples of radical education management reforms are more probably to be found in Latin America. While East Asian countries follow a more predictable pattern, Latin American reforms often introduce bold changes. To illustrate this point, let us consider the examples of Venezuela and Colombia as regards the degree of centralization. These countries constitute two extreme cases. The excessive centralization of the education system in Venezuela has precluded regional autonomy, and has resulted in rigid standardization and inefficiencies, while in Colombia the degree of decentralization has been so high that schools have frequently ignored the directions given by the central administration.

2.2 Equity aspects

Education policies are particularly relevant for equity because they affect access to schooling and, subsequently, to jobs and income. The provision of universal primary school and the wide access to secondary and higher education substantially increase the opportunities for upward mobility. Recent studies have consistently found that additional years or levels of schooling are translated into a large labor market premium.¹⁵ Also, there is by now a growing consensus that education is a critical determinant of income distribution. Indeed, current research points to education as the single most important determinant of inequality.¹⁶

In such a context, it is very important to evaluate the equity aspects of the school system in the two regions. The relevant questions here are: “How to reach to specific levels of education?” and “How are the benefits of education are distributed across income classes?” Answers to these questions involve analyzing three characteristics of the education system that are believed to have significant implications for equity. They are the existence of a bias toward higher education, the equity in the distribution of benefits, and the balance between the public and private sector as school providers. This section discusses the bias toward higher education while the two other aspects will be examined in Section 5.

15 Barros (1992), Pscharapoulos (1993).

16 See, for instance, Fizbein and Psacharapoulos (1992), World Bank (1993), Barros (1992).

2.2.1 The bias towards higher education

The bias toward higher education relates to the emphasis given in some countries to tertiary education, not infrequently in detriment of other levels of education. Such bias is reflected in the priorities given to this sector with respect to the extent of coverage, resource availability and government subsidization. Table 6 shows, for all the three levels of education, the actual school enrollment ratios and the ratios that would be expected on the basis of per capita income levels alone. The latter are derived from cross-section regressions of educational indicators on per capita income and its square.

Table 6
Actual and Predicted Enrollment Ratios in Latin America and East Asia –1992

Countries	GNP per capita- US\$	Primary		Secondary		Higher	
		Actual	Predicted	Actual	Predicted	Actual	Predicted
Latin America							
Argentina	6170	100.0	92.8	78.6	77.5	51.6	27.4
Bolivia	750	89.0	79.9	43.3	53.2	22.0	16.0
Brazil	2810	79.2	85.9	68.8	63.4	27.7	21.3
Colombia	1350	73.2	81.8	66.6	56.4	23.9	17.7
- Chile	2780	83.9	85.8	84.4	63.4	29.6	21.2
- Costa Rica	2010	86.6	83.7	50.6	59.7	23.4	19.4
Ecuador	1100	92.2	81.0	73.6	55.1	37.4	17.0
Mexico	3510	100.0	87.7	60.1	66.6	19.1	22.8
Paraguay	1410	86.5	82.0	46.0	56.7	11.6	17.8
Peru	1350	100.0	81.8	74.6	56.4	34.6	17.7
Uruguay	3470	90.3	87.6	82.3	66.4	44.2	22.8
Venezuela	2920	93.8	85.9	61.6	64.0	25.6	21.6
East Asia							
Hong Kong	15710	99.5	97.8	84.0	84.6	23.4	28.4
Indonesia	680	94.1	79.7	60.2	52.8	18.8	15.8
Korea	7220	100.0	95.0	83.0	79.1	34.6	28.8
Malaysia	2830	93.1	86.0	64.8	63.6	7.0	21.3
Philippines	790	83.0	80.0	75.2	53.4	26	16.0
Singapore	16970	100.0	96.6	87.4	82.7	21.5	26.7
Thailand	1840	81.4	83.3	37.0	58.9	17.5	19.0

Source: Social Indicators of Development, 1994; Predicted enrollment ratios are derived from regression shown in Appendix 1.

Latin America and East Asia fit the typical pattern of developing countries in the sense that both regions have enormously increased enrollment in universities. For the majority of

the countries reviewed, enrollment in tertiary institutions is greater than predicted from their income levels, pointing to the existence of a bias toward higher education. However, in Latin America, the bias is clearly stronger than in East Asia. Except, for Mexico and Paraguay, actual enrollments are considerably superior to the predicted values. Consider, for example, the case of Brazil. Its gross enrollment ratio in primary education is 79.2 percent compared with a predicted ratio of 85.9 percent whereas in higher education the actual figure is 27.7 percent, 6.4 percentage points above its predicted value. This pattern clearly indicates a strong favoring of higher education. In East Asia, Indonesia, the Philippines and Korea also presented enrollment ratios superior to predicted levels. However, in these countries the bias was not against the intermediate levels. Primary as well as secondary education have an actual coverage which is higher than their predicted values. It should also be noted that for countries in this region, the public sector share of enrollment in higher education is inferior to 50 %. To meet excess demand for higher education, Indonesia and Korea have relied on private institutions and open universities whereas in the Philippines, the demand for higher education has been satisfied almost entirely by the private sector.

A second source of bias toward higher education is the concentration of resources by student and by level of education. To assess the implicit priorities attributed to the various levels of education, we used data on public expenditures by student, expressed as a percentage of per capita GNP. Table 7 presents those figures for Latin America and East Asia as well as their deviations from the regional average. For purpose of comparison it is convenient to have a condensed indicator that combines data on deviations into a single number. This rate was calculated by establishing the mid-points of deviations between primary and secondary education, and between secondary and higher education than taking the difference between these mid-points, which shows the pattern of deviation for each country. So, for Brazil the overall pattern is of rising deviations with higher levels of education, whereas for Korea it declines.

Table 7
Expenditures per Student and Deviation from Regional Average in Selected Countries of Latin America and East Asia –1992

Countries/Regions	Expenditures by student in percentage of per capita GDP			Deviation of expenditures from regional average (percent)			Rate of spending bias toward higher education (percent)
	Primary	Secondary	Higher	Primary	Secondary	Higher	
Latin America							
- Argentina	7.8	17.8	18.4	-26.4	43.0	-60.3	-16.9
- Bolivia	14.4	10.4	31.0	36.5	-16.7	-33.1	-34.8
- Brazil	11.0	19.6	111.1	4.33	57.4	139.7	67.7
- Colombia	9.3	8.3	39.2	-12.4	-33.5	-15.4	-1.5
- Chile	11.4	14.1	32.7	7.4	13.1	-29.5	-18.5
- Costa Rica	10.1	11.5	60.2	-4.5	-7.4	30.0	17.3
- Dominican Republic	5.3	3.4	27.6	-50.1	-72.8	-40.4	4.9
- Ecuador	7.2	7.5	31.8	-32.3	-40.2	-31.3	0.5
- Guatemala							
- Mexico	7.1	15.7	102.8	-33.7	26.0	121.8	77.8
- Paraguay	5.7	18.2	64.1	-45.9	46.2	38.2	42.1
- Peru	6.4	10.4	26.9	-39.4	-16.4	-42.0	-1.3
- Uruguay	9.8	9.0	46.2	-7.7	-27.9	-0.3	3.7
- Venezuela	6.1	16.8	55.8	-42.0	34.8	20.5	31.3
East Asia							
- Hong Kong	17.2	7.9	59.0	62.2	-36.8	27.2	-17.5
- Indonesia	8.4	10.0	19.1	-20.2	-19.6	-58.8	-19.3
- Korea	22.9	14.0	7.0	116.6	12.0	-84.9	-100.8
- Malaysia	15.4	23.5	117.7	45.9	88.6	153.9	54.0
- Philippines	11.2	6.7	22.4	5.51	-45.9	-51.8	-28.6
- Singapore	11.6	11.5	34.8	9.8	-7.7	-24.9	-17.4
Thailand	19.7	22.2	31.0	86.4	77.9	-33.0	-59.7
-Regional Average	10.6	12.5	46.3				

Source: Author's calculation based on UNESCO data.

Here again, the bias toward higher education is much more pronounced in Latin America than in East Asia. Indeed, in this region, except for Malaysia, the reduction of deviations with the level of education is manifest as most of the indices of spending bias toward higher education are negative. The cases of Korea and Thailand are particularly remarkable. Their respective rates are -100.8% and -59.7% indicating a clear emphasis on basic education. Notice also that the positive bias for Malaysia is explained by the very low enrollment in the sample year. In Latin America, many countries present a positive bias, denoting the obvious priority of tertiary education. Mexico presents the highest bias toward university education with an rate as high as 77,8%. In this regard, Brazil also has one of the strongest biases against primary education in that it is second only to Mexico's.

Even if in poor countries the favoring of tertiary education over lower levels seems to constitute a general pattern, the emphasis on higher education, as measured by resource intensity, should not be considered an inevitable consequence of under-development. Educational systems in poor countries do not necessarily favor higher education as could be seen by looking at the figures for Indonesia and Philippines. In fact, the intensity of the bias depends as much on policies adopted as on the specific conditions of the countries.

The level of the bias is probably more linked to income distribution than with the level of income. Simple calculations seem to corroborate this hypothesis. The correlation between the bias toward higher education and the top quintile's share of income to bottom's quintile is positive and equal to 0.64 (the correspondent rank correlation is as high as 0.56) meaning that the bias tends to be associated with inequalities. So, the huge income disparities in Latin America may partly explain its strong favoring of higher education.

Finally, the extent of government subsidies of higher education also determines the size of the bias. The extent of subsidies depends on the availability of low cost alternatives (such as distance education), private financing, and user charges. East Asian countries use all these options to reduce public spending on tertiary education. Thus, Korea exploits these three options well while distance education is largely used in Thailand, and the Philippines rely heavily on a self-financing private sector. As for user charges, even in public institutions, cost recovery may be quite substantial. On average, fees finance a larger proportion of the costs of public education in East Asia than in any other developing region. (Tan and Mingat, 1986) Here two patterns are observed: in Korea, the Philippines and Malaysia, the share of cost recovery through fees consistently rises with the level of education. In Indonesia and Thailand, the rate of cost recovery rises steeply from primary to secondary education and then drops for higher education.

Public subsidies of higher education is clearly more common in Latin America than in East Asia. Latin American countries spend a substantial share of public funds to finance universities. In this region, the public sector controls and totally finances the best universities and offers the majority of good quality graduate programs in higher education.¹⁷ Cost recovery in public institutions is practically nonexistent although an important private sector may exist in some countries.

Why has university education with its bias expanded so enormously? Has the performance of the universities justified the amount of resources spent on higher education, considering that such resources could certainly have had alternative uses?

The extension of the middle class in Latin America, and the formation of a new professional class in East Asia that gave priority to educating their children explain most of the enlargement of the university system.¹⁸ In Latin America, this huge enrollment in higher

17 Winkler (1994). For the Brazilian case see Negri (1996).

18 Ranson, Khoo and Selvaratnam (1993).

education has certainly helped to lower the returns on education. Indeed, in this region, the favoring of higher education is particularly harmful as the rates of returns of this level of education are among the lowest in the world and have been declining since 1980.¹⁹ Although its quality has been slowly improving over the past two decades, specially with respect to the training of personnel, Latin American universities suffer from the chronic problems that affect tertiary education in the third world. Such problems include overcrowding, deteriorating physical facilities, poor library resources, and insufficient scientific equipment. As a result, these institutions are characterized by low internal efficiency, graduate unemployment and limited scientific output. Furthermore, most of the increase in university recruitment has gone to careers without a direct impact on the development process because they are less expensive and the deficiencies in the training less visible. This goes against the efficiency prescriptions according to which only capital market failures may justify public funding for expensive scientific equipment and low income students. Moreover, in Latin America, the political functions of the universities have predominated over the educational and economic tasks they could have performed.

In contrast, in East Asia, Korea and Taiwan, and to a lesser extent, Malaysia and Singapore, have managed to break up the pattern by means of a deliberate policy to invest in technical higher education.²⁰ This policy involved recruiting foreign faculty and attracting highly trained nationals back to the country as well as improving the working conditions in institutions. Korea, for instance, adopted science and technological policies designed to create effective incentives to increase expertise in the country. Institutions were built and equipped with the most advanced equipment and qualified personnel. As a by-product, Korea has become a leader in the micro-technology industry, which greatly helped to increase its external competitiveness. In Singapore, the Government's commitment has moved the higher education system away from the hegemony of the traditional elitist British university towards a well-structured, multifunctional and labor-market driven university, well adapted to the increasing needs of a fast growing economy.²¹ As for Malaysia, because of its well-known ethnic problems, a small higher education system tried to reconcile an elitist selection policy, adopted to ensure a high degree of completion, with the need for social equity. National selection policies based on quotas complement the selection by merit criteria to reduce ethnic inequalities.

19 Winkler (1994).

20 See Ranson, Khoo and Selvaratnam (1993), and Selvaratnam (1994).

21 Selvaratnam (1994).

3 Public health care in Latin America and East Asia

As was the case with education, in the health sector, private markets display a wide range of market failures which provide a *rationale* for government intervention in the production of public health services. Safe water, sanitation treatment and the control of infectious and parasitic diseases are typical examples of pure public goods in the sense that they are characterized by non-excludability in consumption. In such cases, it is well-known that the private sector will tend to under supply, or not supply this kind of goods at all. Also, the existence of information asymmetries and moral hazard problems imply that private health care costs will tend to be excessively high. There are also strong efficiency arguments to support public financing of basic health care as they have large positive externalities. Finally, public health programs address the imperfect information of consumers as they tend to underestimate the benefits of preventive care. Hence, public health services tends to be under supplied by private markets making a clear case for government intervention. This section discusses to what extent such principles are followed in both regions. The pertinent questions concern the extent of public health programs, the balance between preventive and curative programs, as well as the equity and efficiency aspects of public provision of health services.

3.1 Public health conditions: comparative statistics

Developing countries have achieved remarkable reductions in morbidity and mortality over the past thirty years. A natural question to ask is, how much of this improvement is due to government policy, and how much simply reflect income growth over the period? Establishing a relationship between public programs and health indicators is difficult as health outcomes are influenced by factors other than government spending. Nevertheless, when controlling for such other factors (e.g. the growth of income), attempts to measure the impact of public expenditure on health conditions suggested that public health programs - such as immunization programs, access to safe water and better sanitary conditions - have a very substantial effect on reducing maternal and infant mortality as well as morbidity rates.²²

In such a context, Latin America and East Asia have not been exceptions. Health indicators have improved steadily in both regions as can be seen from Table 10. However, they have improved at a slower pace in Latin America. This made the region slip in international comparisons of health achievements. Before the eighties, it presented the

22 Hammer, Nabi and Cercone (1995), World Bank (1992), Van den Walle (1995), among others.

highest levels of health in the developing world. By now, the provision of health services in Latin America is worse than in regions with comparable levels of development.²³ In spite of recent improvements, sanitary conditions in the region are still very far from adequate. This was clearly reflected in the epidemics of cholera and other diseases in the late 1980s and early 1990s.

Table 8
Selected Indicators of Public Health : 1975-80, 1987-1992

	Access to Safe Water		Infant Mortality (per thousand live birth)		Immunization - percent of age group				Maternal mortality rate - per 100000 live births		Life Expectancy - total years	
	80/85	87/92	80/85	87/92	Measles		DPT		80/85	87/92	80/85	87/92
					80/85	87/92	80/85	87/92				
Latin America												
- Argentina	66	n.a	36	29	90	99	66	84	85	140	70	71
- Bolivia	53	66	103	103	17	73	24	58	450	48	54	60
- Brazil	84	87	71	57	80	83	67	75	150	140	63	69
- Colombia	89	100	40	21	52	75	60	84	130	110	67	66
- Chile	87	86	24	24	77	77	84	84	55	40	71	71
- Costa Rica	90	94	19	14	83	90	82	95	26	18	74	76
- Ecuador	57	70	62	45	40	54	36	89	220	170	64	67
- Mexico	70	78	49	35	30	78	26	64	92	110	67	70
- Peru	55	58	74	52	32	82	26	81	165	280	59	65
- Uruguay	98	98	33	20	17	82	57	88	56	36	71	72
-Venezuela	80	89	39	33	25	54	27	54	65	200	69	70
- Regional Average	75,4	75,1	50	39	49,4	77	50,4	78	136	117	66,3	69
East Asia												
- Hong Kong	89	99	8,8	6	n.a	42	87	83	4,0	4,0	76	78
- Indonesia	38	42	95	66	15	80	15	86	450	390	56	60
- Korea	73	93	30	13	89	93	76	74	34	30	68	71
- Malaysia	84	79	28	14	20	79	59	90	59	20	68	71
- Philippines	52	81	51	40	49	85	52	88	80	74	62	65
- Singapore	100	100	9,3	5	75	92	78	91	11	10	72	75
- Thailand	56	72	44	26	22	60	47	69	270	37	64	69
- Regional Average	70,3	80,9	38,0	24	45	76	59,1	83	130	85,8	66,6	70

Source: World Bank (1994).

In East Asia, the considerable emphasis on primary health care has undoubtedly translated into better health conditions. Even in a poor country like Indonesia, where the health conditions are amongst the lowest in East Asia, large investments in primary health care have been clearly associated with a reduction in both the rate and duration of reported morbidity.(Deolalikar, 1995)

23 Londoño (1996), calculated also the health-related losses in productivity.

This divergent behavior is not explained by the lack of financial resources devoted to health. Latin American governments spend approximately 3 percent of GDP annually on health, a high figure when compared with its East Asian counterparts. Such a result also holds when measured in terms of expenditure per capita. For instance, in Argentina, in 1992, the government spent US\$ 192 per capita whereas the equivalent figure for Korea was only US\$183. However, almost all Korean indicators on health care were better than those for Argentina. Brazil's inefficiency is again quite remarkable. In spite of the fact that its spending is within the range of the regional average, its rate of infant mortality is the second highest in the region.

As was the case in education, the answer is to be found on the cost-effectiveness associated with input use. This can be seen at Table 9, which provides information on selected health inputs for Latin America and East Asia. The proportion of doctors to nurses is much higher in Latin America.²⁴ This could be a sign of inefficiency as, compared to doctors, the costs of training nurses as well as their salaries are much lower. If properly trained, they can be extremely useful in providing the basic care for a needy population at a much lower cost, thus permitting a better use of public resources. Furthermore, the relatively high number of doctors in Latin America is very probably concentrated in hospitals and urban areas. For instance, in Peru, two thirds of all doctors live in Lima, which has only 29 percent of the country's population. In most rural areas there is only about one doctor per ten thousand inhabitants.²⁵ Such a situation is confirmed by the figures on hospital beds shown in Table 10. In Latin America there are more beds per inhabitant than in East Asia. However, hospitals absorb a large share of resources assigned to the health sector but provide services only for a relatively few people and hence, with important distributional consequences. In particular, a major effect of reliance on hospital care on income distribution is the increase of rural-urban imbalances. The benefits of hospitals services are not distributed equally among the population but are instead received disproportionately by residents of urban areas. In Colombia, Indonesia, and Malaysia, the average health subsidy received by urban households is up to five times larger than that enjoyed by rural residents. (World Bank, 1987) This particularly hurts Latin America as, compared to East Asia, this region has a larger rural population as well as more pronounced income

24 Although there is no optimal ratio between nurses and doctors, a ratio of fully qualified nurses to physicians of 2 and 4 to 1 and 0.1 to 0.2 physicians per 1000 population is considered to be adequate. (World Bank, 1993) In Latin America, these proportions are reversed as doctors outnumber nurses.

25 World Bank (1987).

disparities. To summarize, all the information above suggests the existence, in Latin America, of a bias toward curative medicine in detriment of basic health care, the typical public good.

Table 9
Selected Health Indicators in Latin America and East Asia

Countries	Public Health Spending (percent of GNP) -1990	Public Health Spending Per Capita (US\$ 1992)	Health Inputs per 1000 population					
			Physicians		Nurses		Hospitals beds	
			1980-85	1987-1992	1980-85	1987-1992	1980-1985	1987-1992
Latin America								
- Argentina	5.9	190.9	2.70	3.03	1.02	0.56	5.59	4.63
- Bolivia	2.4	16.3	0.65	0.39	0.41	0.13	2.00	1.21
- Brazil	2.8	77.6	1.46	1.18	0.83	0.29	5.00	3.32
- Colombia	3.0	38.7	0.81	0.94	1.54	0.38	1.68	1.42
- Chile	3.4	80.6	0.87	0.46	2.69	2.98	3.40	3.12
- Costa Rica	7.6	152.0	0.99	0.97	2.10	0.45	3.31	n.a
- Ecuador	2.6	27.9	1.22	1.02	1.63	1.62	1.87	1.60
- Mexico	1.6	55.5	0.84	1.61	1.19	n.a	1.32	1.25
- Paraguay	1.2	16.1	0.66	0.80	0.82	0.14	1.67	0.92
- Peru	1.9	18.0	0.93	1.06	n.a	n.a	1.70	1.41
- Uruguay	2.6	86.8	1.95	n.a	n.a	n.a	3.25	4.52
- Venezuela	2.0	58.0	1.20	1.58	1.58	3.03	2.70	2.60
- Regional Average	2.9	75.7	1.00	0.89	1.09	0.34	2.29	2.36
East Asia								
- Hong Kong	1.1	169.2	0.83	n.a	1.26	n.a	4.90	4.27
- Indonesia	0.7	4.7	0.11	0.14	0.80	0.35	0.55	0.66
- Korea	2.7	183.3	0.72	1.05	0.84	2.20	1.68	3.33
- Malaysia	1.3	36.3	0.32	0.41	0.72	2.66	1.28	2.32
- Philippines	1.0	7.7	0.15	0.12	0.37	n.a	1.74	1.28
- Singapore	1.1	173.2	0.96	1.22	3.11	n.a	4.26	3.64
- Thailand	1.1	20.0	0.17	0.22	0.54	1.08	1.54	1.61
- Regional Average	1.3	85.0	0.24	0.25	0.71	0.87	1.45	2.44

Source: World Bank (1994) and calculations of the author.

3.2 The balance between curative and preventive care

In the balance between preventive and curative care, we find a clearly divergent pattern between the two regions. In Latin America, the predominance of a costly curative medicine in detriment of preventive care has greatly helped to increase health costs and subsidies. The emphasis in hospital services, very intensive in capital and technological inputs, turned attention away from the more equitable ambulatory assistance. For instance, Brazil, in 1982, assigned 70% of public health funds to cover expenses for doctors and hospital care, including expensive procedures. (Birdsall and James, 1993) McGreevey (1984) reports

also that in 1979 the Brazilian budget assigned more funds to these high technologies services utilized by only 10,000 individuals from the Southern and Southeastern regions than to the 41 million people living in the poor areas of the North and Northeast.²⁶ Peru is another typical example of such a situation, where 87 percent of public expenditure in health care in 1980-1981 was spent on curative care. The exception in Latin America is Chile, where increased reliance on private hospitals was followed by a shift of public resources toward basic health care.

This concentration of public resources on curative care cannot be justified on economic or social grounds. Government provision of what is essentially a private good merely substitutes for market supply, thus not being an efficient use of public funds. Recent empirical studies confirm the prescriptions of the modern analysis of public expenditure. According to these studies, government spending on public health services is more likely to improve health outcomes than with expensive clinical services.²⁷ The implication is that public spending should be restructured toward public health programs and away from clinical services that could be otherwise provided by the market.

Clearly, in Latin America, the utilization of sophisticated and expensive medical procedures for a few people makes the extension of basic care unreasonably difficult for the great majority. The cost of extending the Bismarckian model of social security to the whole population would be prohibitive for most Latin America countries. Mesa-Lago (1989) calculated such costs and found that they would represent a large proportion of the GDP. Exceptions are Mexico, Brazil and Venezuela; these countries could extend their coverage at moderate costs. In the poorest countries, however, costs associated with this extension are virtually unbearable. If Ecuador, Nicaragua, El Salvador, Colombia and Peru wished to cover all the population, it would cost 59,7%, 25,3%, 21,0% 18,4% and 15,7% of GDP respectively. Under current conditions such costs are clearly incompatible with the financial means of these countries.

26 A typical example of the ineptitude of hospital care is cesarean deliveries. These procedures have great implication for the misuse of hospital resources. In selected countries where medical training and the financial reimbursement system provide adverse incentives, the proportion of total deliveries by cesarean increased the potential risk of giving birth and unnecessarily burdened the health system. Brazil again is the typical example of such misuse of funds. The cost of unwarranted cesareans in Brazil is calculated to be around US\$53 million per year. (Barnum and Kutzin, 1993)

27 Hammer, Nabi and Cercone (1995), Van den Walle (1995), World Bank (1993), among others.

What about East Asia? Compared to Latin America, the region seems to rely less on expensive public hospitals and costly medical personnel. The share of public funds going to hospitals and clinics is lower than in Latin America (Appendix, Table A-1). The greater number of nurses and midwives, compared to doctors, indicates a higher use of essential, more cost effective, outpatient services. Although doctors are needed to supervise essential clinical care and handle complicated cases, most of the services in the minimum clinical package can be provided by nurses and midwives.²⁸ In Indonesia sub-centers (units of the primary health care system) are headed by a nurse or midwife, have few staff and offer curative and maternal and child care. (Van den Walle, 1995) Malaysia has also invested heavily in the infrastructure needed to improve peripheral health units providing essential clinical care for the poor (World Bank, 1993), and in Thailand, public-sector nurse-midwives successfully perform female sterilization. Furthermore, as shown below, the substantial participation of the private sector in the provision of health care suggests that public resources are concentrated where they should be, namely, the provision of basic health care, the truly public good. Last but not least, such a concentration of public funds on essential health services may have strong positive distributive consequences as they deal with problems that mainly affect the poor.

3.3 The financing of public health care

Although public funds are an important source for financing health services in both regions, they are by no means the only source available. Private spending, particularly in East Asia, accounts for a significant share of health care spending (Table 10). Except for the Philippines and Singapore, the ratio between private and public expenditures is higher than one. In Hong Kong and Thailand, expenditures by families account for more than 70 percent of total health expenditures, and in Indonesia this proportion could reach 65 percent. In sharp contrast, in Latin America, the participation of private sources, even if not negligible, is much lower. Except for Peru, all the ratios between private and public spending are less than unity. Clearly, in this region, the bulk of health care financing is falling on the public sector, a pattern similar to that observed in the rich welfare states. This is hardly surprising since the model of social protection adopted by Latin America was inspired by the Bismarckian model. In Latin America, government-run health care systems

28 The essential package includes prenatal and delivery packages, family planning services, management of the sick child, treatment of tuberculosis, and case management of sexually transmitted disease. For details see World Bank (1993).

are overextended and are very probably crowding out the private sector. Hence, their sizes needed to be reassessed.

How do Latin American governments finance this huge spending? Basically, the Latin American health care system is characterized by a dual nature: a ministry of health that completely supports services for the destitute and a social security system which provides services for workers of the formal sector. Spending by ministries of health is financed by general tax revenues whereas social security institutions, in most of Latin America, are financed by payroll and earmarked taxes (see Section 4). User charges in government facilities usually cover only a small share of expenses. As expected, with the public provision of goods which are private in nature, there is a frequent tendency to the over-consumption health care services, particularly, expensive curative procedures. Consumers have little incentive to restrain the use of high level facilities when they are either free of charge or command the same fees as simpler facilities. This significant over-consumption takes various forms such as the unjustified use of cesareans as well as the excessive consumption of antibiotics.²⁹ Additionally, an increasing coverage, together with a high rate of evasion and lack of contribution by the state made the contributions for the social security system insufficient to cover the escalating expenses.³⁰ The predictable result is the existence of chronic budget deficits in the health care system. Such deficits have been financed by cross-subsidizing the health sector with surpluses from the pensions' accounts as well as via reductions of funds available for preventive health care programs. Some Latin American countries began adopting policies of cost recovery by creating or increasing fees. Up to now, such attempts have clearly been insufficient to mitigate the huge deficits. Moreover, those policies are very much criticized on equity grounds as Latin Americans tend to see any cost recovery policy as an attempt to dismantle their system of social protection and strongly react to any changes, even though they may be extremely necessary.³¹

29 Zschock (1986) and Mesa-Lago (1989).

30 For details on how evasion occurs, see Mesa-Lago (1988).

31 See Soria (1997).

Table 10
Public and Non- Government Health Expenditures - (US dollars per capita) – 1991

Countries	Total	Public	Non- Government	Non-Govern/Public
Latin America				
Argentina	334.31	204.56	73.45	0.36
Bolivia	28.27	19.94	11.77	0.59
Brazil	114.87	76.58	38.29	0.50
Chile	109.88	79.45	32.77	0.41
Colombia	50.87	22.88	27.99	1.22
Costa Rica	165.09	134.51	29.14	0.22
Mexico	108.54	54.27	54.28	1.00
Paraguay	40.38	17.06	23.32	1.37
Peru	70.76	41.96	28.80	0.69
Uruguay	141.29	76.44	64.85	0.85
Venezuela	97.17	53.98	43.18	0.80
East Asia				
Hong Kong	663.90	128.12	535.78	4.18
Indonesia	12.85	4.50	8.36	1.86
Korea	431.37	176.45	254.89	1.44
Malaysia	77.45	33.56	43.89	1.31
Philippines	14.30	7.14	7.16	1.00
Singapore	271.42	157.08	114.34	0.73
Thailand	81.58	17.94	63.66	3.55

Source: World Bank (1993) and authors calculations.

Note: Non-government expenditures is defined as the sum of private expenditures on health care, expenditures by missions, and expenditures by Non-Governmental Organizations.

In East Asia, Korea has a social insurance system, financed by mandatory contributions from employees and employers, which is very similar to that prevalent in Latin America. Health care for the poor is paid out of general revenues. All Koreans are covered, and the government subsidy to extend coverage to the elderly and indigent is highly progressive. Korea's health providers are predominantly private and are paid on a fee-for-services basis which has encouraged the excessive use of expensive procedures. As in Latin America, this over-consumption of sophisticated medical services has led to a dramatic increase in health spending. The share of GDP devoted to health care rose from 3.7 % in 1980 to 6.6% in 1990. In the future, this may jeopardize the prospects of the Korean economy as it may crowd out other sectors and raise labor costs, thus undermining the country's international competitiveness.

4 Social security spending

A major difference between public spending in the two regions concerns social security. In Latin America, the adoption of the Bismarckian model created a pattern of social protection similar to that prevalent in the rich European countries. All Latin American countries have old-age pension schemes, sickness-maternity health plans, and disability benefits (Table 10). A few countries have also unemployment insurance plans. Coverage of these different social security programs varies greatly across countries. By 1980, the coverage was as high as 95.6 percent of the economically active population in Brazil but the corresponding figure was only 11.6 percent for Guatemala.³² However, on average, a substantial part of the Latin American population benefits from some form of social protection.³³

Table 11
Selected Characteristics of Social Security in Latin America - 1983

Countries	Expenditure - (percent of GDP)	Revenues (percent of GDP)	Coverage Pensions	Dependency Ratio ¹	Administration Costs (percent of GDP)
Argentina	8.6	6.4	69.1	26.3	4.3
Bolivia	2.1	2.2	18.5	13.2	19.3
Brazil	5.7	5.3	95.6	13.8	9.9
Colombia	2.2	2.3	22.4	16.7	12.4
Costa Rica	6.3	8.8	68.3	12.9	6.9
Chile	14.4	8.7	61.7	12.6	6.2
Ecuador	4.2	5.4	23.2	13.4	23.7
Mexico	2.8	1.3	42.0	11.9	13.3
Uruguay	11.0	6.7	81.2	32.9	6.4
Venezuela	1.5	1.5	49.8	11.4	13.8

Source: Mackenzie (1988).

1: Percentage of the population aged 60 or more in the population aged between 20-59.

Regarding the form of financing social protection, the combined contributions of workers, employer and governments represent more than 30 percent of wage bill in many countries (Ahmad, 1991). This imposes a serious constraint on the financing of social

³² See Mackenzie (1988) and Mesa-Lago (1989). Of course, these figures apply mainly to the formal sector. However, in some countries informal employment may be very significant and those workers are by and large outside the social protection system.

³³ For a detailed analysis of Latin American Social Security System see Mackenzie (1988).

security as further increases in already high payroll and social security taxes are difficult to implement. Excessively high contribution rates lead to evasion and encourage the development of informal labor markets. Moreover, most of the pension plans are financed on a pay-as-you-go basis in which current contributions pay for current benefits.³⁴ Any difference between these values is to be met by reserves or transfers from other governmental sources.³⁵

The premature adoption of the Bismarckian model of social comprehensive protection turned out to be a precarious choice as the Latin American countries lacked the productivity foundations on which was based the European model. As a result, the system is a source of fiscal disequilibrium in Latin America, especially, in countries where the coverage is extensive and the dependency ratio high as in Argentina and Uruguay. Not surprisingly, reform of social security programs is currently on the agenda of most Latin American governments.

In East Asia, as a rule, the social insurance principle, at least as it is financed through taxation is almost absent. When it does exist, it takes the form of forced savings, in sharp contrast to the Latin American experience. Singapore, Malaysia and the Philippines extensively use the funded principle for financing social security expenses. Instead of constituting a burden to government finances, in East Asia social security schemes have encouraged savings and helped to stabilize the long-term development of financial markets. As Shome (1989) pointed out, in these countries, social security funds have been major sources of government debt financing.

It is important to understand the reasons behind such different behavior. Besides economic elements, socio-political factors also play a role. The strong populist tradition in Latin America very likely helped to shape social protection. Indeed, one of the main characteristics of populism in Latin America was an early development of a relatively strong labor movement. The social demands of labor unions included the implementation of a better social security system along the lines followed by more developed countries.³⁶

34 At first, most Latin American countries offered full-funding of social insurance through trust funds. Subsequently, the increase of the number of retirees, the practice of cross-subsidizing the health care system with surpluses from retirement accounts, together with an inappropriate investment policy, depleted the trusts.

35 In 1994, Colombia replaced its state-run, pay-as-you go pension system with a privately-run, fully funded scheme.

36 Deyo (1990).

Such a phenomenon was more marked among the prosperous countries of the Southern cone. Uruguay, with its highly sophisticated welfare state, became the typical example of such a tendency. Furthermore, the political instability of the continent created the need to form populist alliances to guarantee support for the country's political leaders. In this process, social benefits were eventually extended to more and more labor categories.

In East Asia, the situation was quite different. In this region, with the probable exception of Korea, social unrest was much lower and the workforce much less demand-oriented. Due a number of reasons, including the absence of widespread populism and the adoption of a successful labor intensive development strategy resulting in higher wages, labor movements were less strong and had less to say in the East Asian nations.³⁷ These arguments can also be put differently. Weak labor movements help to explain the absence of leftist and populist coalitions that would have supported nationalistic economic policies. Hence, the state was able to impose relatively free labor markets, keeping wage pressures down and thus contributing to raising profits and increasing managerial flexibility and employment. In such a context, there was no pressing need for government intervention in the labor markets to guarantee worker welfare as was the case in Latin America, and no need to provide social benefits for a discontented population. Finally, external political conditions may also have played a significant role in keeping down labor demand for social protection. As argued by Haggard (1990), "*because of their adverse external political situation divided countries facing communist adversaries - little ideological or organizational space was allowed for socialist, leftist or populist forces, nor was labor allowed an independent voice.*" Labor markets were relatively free from the interference of trade unions. Therefore, governments could restrain labor claims, including those linked to social security benefits.

5 Public spending and income distribution

There is a marked difference between Latin America and East Asia in the role played by fiscal policies in income distribution. In Latin America, attempts to redistribute income have tended to rely more on tax policy while in East Asia the tax system has been directed at generating the revenues required by the expansion of government expenditure policies. East Asian public spending policies have been quite efficient and have played an important

37 Even when labor movements were apparently strong, they played a minor role in the process of setting the relationship between trade unions, government and the firm's management. Good discussions of labor relations in East Asia can be found in Freeman (1994) and Chen and Taira (1995).

role in creating and preserving a more equitable form of economic growth. In particular, public spending on basic education and essential health care have helped to increase the supply of skilled labor.³⁸ Consequently, productivity gains associated with a very efficient absorption of modern technology, increased wages and employment thus reducing poverty and attenuating income disparities.

In the light of the new evidence, East Asia seems to have made a better choice. Indeed, public spending policies appear to be more efficient than tax policies in attenuating income disparities and reducing poverty levels. The progressive nature of social spending, highlighted in various studies,³⁹ is shown in Tables 13 and 14, which present the rate of benefit of social spending in selected countries of both regions.

In Latin America, except for Bolivia, expenditures on health care are the most progressive, with Gini's coefficient ranging from -0.12 to -0.32. Yet, as previously stated, recent studies consistently emphasize the fact that the provision of the essential package of health services, dispensed at primary health care centers, is more equitable than hospital services. In particular, hospital services benefit the better-off urban population and hence may have a negative impact on equity. Indeed, a recent review of various countries studies show that the bottom 40 percent of the population received 50 percent of the subsidies of public health and only 29 percent of the subsidies accruing to hospitals. (Barnum and Kutzim, 1993) This occurs because hospitals tend to be located in cities and more than proportionately serve urban middle-class households. Demographics and disease patterns also determine this income bias. Higher income groups tend to be older and to suffer from diseases requiring hospitalization whereas the poorest households are more often young families with children for whom many health problems are those included in basic health care package that do not require higher-level hospital care. The extent of this bias varies across regions and countries but is likely to be more pronounced in Latin America because the public health care system tends to emphasize curative medicine, which is traditionally more hospital-oriented. Of course, this does not imply that hospitals should be eliminated as they are an essential part of any health system, even one whose main emphasis is on primary health care. The real issue is that public health spending is excessively hospital-oriented and benefits the middle and upper classes more.

38 Barro and Lee (1993).

39 Petrei (1987), Cominetti (1994), Van de Walle and Nead (1995), among others.

The distributive impact of education in Latin American is second only to health care, with Gini's coefficients oscillating between -0.01 and -0.18. Note that progressivity results from the substantial distributive impact associated with primary education. As poor families are larger and as the rich could rely on the private sector to educate their children, public spending in basic education functions as a strong income equalizer. For tertiary education, the situation is radically different. As previously mentioned, the strong bias toward higher education that characterizes Latin America discriminates against poor households in favor of the middle and upper classes. In this region, while poor students in the rural areas often lack basic learning materials, middle and upper class students obtain very large government subsidies thus making public spending in higher education highly regressive.

Table 13
Benefit Incidence of Various Expenditure Programs- Percentage

	Education				Health	Social Security	Housing	Total	Total Excluding Social Security
	All	Primary	Secondary	Higher					
<u>Argentina</u>									
Lower 40%	48	65	47	17	62	23	97	36	54
Upper 40%	34	20	32	65	18	57	0	46	29
Gini Coefficient	-0.10	-0.30	-0.11	0.31	-0.32	0.23	-0.68	0.06	-0.18
<u>Bolivia</u>									
Lower 40%	46	56		22	26	27		41	
Upper 40%	33	23		57	58	54		39	
Gini Coefficient	-0.09	-0.22		0.22	0.18	0.18		-0.03	
<u>Chile</u>									
Lower 40%	50	62	50	21	58	13	38	32	53
Upper 40%	31	20	28	62	21	72	42	51	32
Gini Coefficient	-0.12	-0.28	-0.14	0.28	-0.25	0.41	0.02	0.13	-0.17
<u>Colombia</u>									
Lower 40%	51	67	48	16	54			52	
Upper 40%	27	14	28	63	28			27	
Gini Coefficient	-0.17	-0.36	-0.13	0.25	-0.18			-0.17	
<u>Costa Rica</u>									
Lower 40%	42	52	36	17	48	19	17	40	43
Upper 40%	41	19	34	72	32	65	65	43	39
Gini Coefficient	-0.01	-0.22	-0.07	0.37	-0.12	0.28	0.36	0.02	-0.04
<u>Uruguay</u>									
Lower 40%	54	74	59	12	55	36	34	38	53
Upper 40%	29	14	23	66	24	51	56	42	28
Gini Coefficient	-0.18	-0.44	-0.23	0.36	-0.22	0.17	0.14	0.03	-0.18

Source: Cominetti (1994, table 16, p. 67-68).

Notice also that social security expenses, a significant component of public spending in Latin America, is strongly regressive. This is probably due to incomplete coverage and to the existence of social insurance which tends to favor the more expensive and regressive hospital based system. This effect is substantial enough to reverse the progressive nature of

social spending. The positive distributional impact of public spending becomes negative when social security expenses are included.

In East Asia, results obtained were similar in the sense that in this region too, social public spending, particularly that related to basic health care and primary education, are progressive. Here too, primary and secondary education are more pro-poor than tertiary education. In Malaysia, except for higher education, social service expenditures have been progressive and became more so after 1974 as the poor began to receive the largest share of social expenditure benefits. (Hammer *et alii*, 1995) Results for Indonesia also show that government spending on basic primary health care provides the best option for reaching the poor. (Van den Walle, 1995) The progressivity of public spending in this region is enhanced by the fact that, as shown above, East Asia concentrated public funds precisely on primary education and basic health care. Notice also that in the educational sector, the level of public spending appears to be more equitably distributed in countries with significant levels of private financing in higher education such as Indonesia, Korea and the Philippines (Tan and Mingat, 1992), suggesting that private financing in higher education, instead of reducing, may increase equity.

Table 14

Benefit Incidence of Health and Education Subsidies In East Asia in Selected Years

Countries/Spending	Year of Survey	Percentage of Government Subsidy received by Income Group		
		lower 40 percent	Middle 40 percent	Upper 20 percent
All Education				
Indonesia	1978	46	25	29
Malaysia	1974	41	41	18
Higher Education:				
Indonesia	1978	7	10	83
Malaysia	1974	10	38	52
Public Health				
Indonesia	1974	19	36	45
Malaysia	1974	47	37	17
Philippines	1975	27	33	40
Hospitals				
Indonesia	1974	23	53	23
Malaysia	1974	36	34	20

Source: Pradhan (1996).

6 Regional divergences: tentative explanations

This section tries to summarize the divergent pattern across regions. Why do public expenditure policies seem to be systematically better implemented in East Asia whereas in Latin America they so frequently fail to achieve their objectives? Why did Latin America waste the possibilities of redistribution through an increased supply of public goods whereas East Asia systematically used social spending to promote a more egalitarian style of development? These are not easy questions to answer. Although there are complex and numerous forces at play, three, non-excluding factors, had paramount importance, i.e., the role of government, intellectual beliefs, and the existence in East Asia of a highly motivated bureaucracy that assured the implementation of expenditure policies in a remarkably efficient way. These points will be briefly discussed below.

The role played by the government was essential to determine the way social policies were implemented. In East Asia, a clear-sighted political leadership consistently emphasized economic and social policies as complementary (rather than conflicting) priorities. This certainly played a key role in the region's outstanding performance. Furthermore, the absence of large inequities meant that the government had no need to struggle against powerful industrial or landed groups.⁴⁰ Consequently, it was much easier to design and implement the required expenditure policies. For instance, it is widely believed that one of the main reasons why Korea and Singapore have developed so quickly is that their governments have made determined and successful efforts to raise educational standards. In Latin America, the state, under the influence of competing and income-seeking groups, has been unable to accomplish the required social reforms. Even when well intentioned, these governments, still under the influence of the populist state, tried to implement a combination of mutually inconsistent policies which led to macroeconomic and political instability and failed to fulfill the much needed social demands. In such a context, it was clearly more difficult to prioritize and efficiently implement a consistent set of public policies. Finally, the very high levels of income disparity that characterize Latin America were too often used to discard more efficiency-oriented public spending policies. Recurrently, poor income distribution served as an excuse to adopt "quick fixes" instead of more consistent long-term policies.

40 In Taiwan and Korea, land reform and foreign occupation helped to weaken those groups. (Rodrik, 1994)

Secondly, the role played by intellectual beliefs should not be understated. The predominance in Latin America of the ECLA economic doctrine which downplayed the market element and clearly pushed the state beyond its traditional functions may, unintentionally, have contributed to the neglect of the social area. The typical Latin American state, involved in multiple activities, was unable to concentrate its actions on areas where the presence of market failures was indisputable as in most of the social services. Instead, it chose to replace private initiative in many other sectors. Such a dispersion of objectives resulted in inefficiency, corruption and inequity. In East Asia, there was also a strong government intervention in a quite distinct way. As Tanzi and Shome (1992) pointed out, in this region, *“the government role was limited to its traditional functions, namely, the provision of economic infrastructure, the maintenance of a stable macroeconomic framework and the promotion of growth.”* Government actions were not conflicting with market forces, but corrected and reinforced these same forces.

Finally, in East Asia, the implementation of expenditure policies was greatly helped by the existence of a competent, honest and efficient bureaucracy capable to administer the multiple tasks inherent to reforms. This is not easy to accomplish as in many countries of both regions such a bureaucracy is not available. In East Asia, particularly in Singapore, Taiwan, and Korea, the higher concern for education helped to create a relatively skilled labor force that permitted the formation of a qualified, well-paid and socially respected bureaucracy. This is reflected in the ability of the public sector to attract capable people. Indeed, in these countries, civil servants' salaries are quite competitive compared to those of the private sector. (Appendix, Table A-3) This bureaucracy has certainly played a crucial role in the region's remarkable performance in the social areas and has certainly contributed to the successful implementation of expenditure policies. In Latin America, the situation has been different. In many countries, the administrative organization still reflects an era where centralization and planning were dominant. A poorly paid bureaucracy, with little social prestige, and full of populist ideas has been prone to inefficiency and corruption. As a result, the provision of social services, particularly health and education, has been undermined, and this has resulted in a strong negative impact on the region's welfare.

7 Conclusions

This paper has reviewed the main features of social expenditure in Latin America and East Asia. The objective has been to evaluate the efficiency and equity aspects associated with education, health care, and social security spending in both regions. Emphasis was given to the impact of these expenditures on income distribution and to their effects on the

welfare of the low-income population. Furthermore, the relationship between the public and private sectors was examined in order to inquire to what extent public expenditures in social areas, complemented rather than substituted private activities. The main findings are summarized below.

First, the results suggest a greater overall efficiency of East Asian countries concerning the use of public funds in the social sectors. Government activities complemented private activities instead of substituting them. A competent, well-paid and socially respected bureaucracy greatly contributed to the efficient implementation of public policies, thus making social spending a key element to create and preserve a more equitable form of economic growth. In sharp contrast, Latin America failed to assure to its citizens of an efficient provision of social services. The neglect of these areas has resulted in poverty and led to large income inequalities. This has not been due to a lack of financial resources but to the poor utilization of available funds. The absence of administrative capacity and the lack of appropriate incentives to civil servants has created in Latin America a situation of serious mismanagement that has jeopardized the efficiency and accountability of the public sector.

In terms of educational expenditures, East Asia has performed better. Outcomes have, unquestionably been superior to those ones obtained by Latin America. Indeed, for some countries - Korea, Taiwan and Hong Kong - educational results, as measured by standardized international tests, put them on the top of the world performers. Remarkably, these results have been obtained in spite of relatively low levels of public spending, indicating a very effective utilization of inputs. As for equity, the concentration of public resources on primary education, a reduced bias toward higher education, a high degree of cost recovery, and a strong participation of the private sector in higher educational levels, introduced a higher degree of equity to the East Asian schooling system.

Despite the fact that primary education shows the highest rates of social profitability, Latin America has under invested in this sector. Universities have received a larger share of the education budget, which has created a strong bias towards higher education and helped to increase the inequality of the educational system. Additionally, Latin America, compared to East Asia, has managed its resources less efficiently. There are plenty of indications of misuse of inputs, leading to a needless increase of educational costs. As for outcomes, high dropout and repetition rates together with a poor quality of schooling testify the existence of many shortcomings in Latin American school and call for a serious reformulation of its educational system.

As was the case in education, the problems of health sector in Latin America are more related to cost effectiveness than to lack of resources. In many Latin American countries, the emphasis on curative medicine instead of basic preventive care is not only highly regressive but also inefficient. Public funding of basic health care activities provides larger benefits for the poor, thus contributing to reduce inequalities. Those programs also effectively improve health outcomes. As for East Asia, the substantial participation of the private sector in the provision of health care suggests that public resources are concentrated where they should be, namely, the provision of basic health care. The concentration of public funds on essential health services may have strong positive distributive consequences as they deal with problems that mainly affect the poor. On the other hand, hospitals, which in many countries of both regions absorb a large part of resources, mainly serve the urban areas, thus helping to increase rural-urban imbalances. Even though hospitals have a valid and important role, it is likely that in terms of cost effectiveness, government resources are excessively concentrated on them.

Another marked difference between Latin America and East Asia concerns social security spending. Such a divergent behavior cannot be attributed uniquely to economic variables. Socio-political factors - including a strong populist tradition - helped to explain why social protection is widespread in Latin America. In East Asia, except for Korea, the absence of populist coalitions, together with a labor-intensive development strategy that has increased wages, have kept down labor demands and restrained the claims for social security benefits.

Finally, the results obtained indicate that appropriate public expenditure policies are powerful instruments to improve income distribution. In both regions, a substantial part of government subsidies for primary education and basic health care have gone to low-income households. Security social and higher education expenditures were found to be highly regressive. As Latin America spends relatively more on these last two components, we conclude that, overall, Latin America has failed to seize the significant distributive potential associated with social spending, thus missing an opportunity to effectively reduce its tremendous income inequalities.

References

Anand, S. and Ravallion, M. Human development in poor countries: on the role of private incomes and public services. *Journal of Economic Perspectives* 7, p. 133-50, 1993.

- Ahmad, E. Social security and the poor: choices for developing countries. *The World Bank Research Observer* 6, p. 105-127. 1991.
- Barnum, H. and Kutzim, J. *Public hospitals in developing countries: resource use, cost, financing*. Baltimore: John Hopkins University Press, 1993.
- Barro, R. and Lee, J. International comparison of educational attainment. *Journal of Monetary Economics* 32, p. 363-394, 1993.
- Barros, R. P. Welfare, inequality and social conditions in Brazil in the last three decades. Paper presented at the *Conference on Poverty and Inequality in Latin America*, Washington, D.C.. Brookings Institution, 1992.
- Birdsall, N. and James, E. Efficiency and equity in social spending: how and why governments behave. In Lipton, M. and Van der Haag, J. (eds.), *Including the poor*. Washington, D.C: The World Bank, 1993.
- Burki, S. J. and Edwards, S. Dismantling the populist state. *World Bank Latin American and Caribbean Studies*, Washington, D.C., 1996
- Cardoso, E. and Helwege, A. Below the line: poverty in Latin America. *World Development* 20, p. 19-37, 1992.
- Chen, S. J and Taira, K. Industrial democracy, economic growth and income distribution in Taiwan. *American Asian Review* 13, p. 49-77, 1995.
- Cominetti, R. El gasto social en América Latina: un examen cuantitativo y cualitativo. *Cuadernos de la CEPAL*, Diciembre de 1994:
- Corvalán, A. M. *El financiamiento de la educación en periodos de austeridad presupuestaria*. Santiago, Chile: Comilación UNESCO-OREALC, 1990.
- De Geyndt, W. Managing the quality of health care in developing countries. *World Bank Technical Paper N° 259*, Washington, DC.: World Bank, 1995.
- Den Hartog Georgiads, W. and Jones, H. *A review of research on headmaster and school principalship in developing countries*. Washington, D.C.: World Bank, 1989.
- Hammer, J., Nabi, I. and Cercone, J. Distributional effects of social sector expenditures in Malaysia, 1974-1989. In: Van de Walle, D. and Nead, K. (eds.), *Public spending and the poor: theory and practice*. Baltimore: John Hopkins University Press, 1995.

- Hammer, J. The economics of malaria control. *The World Bank Research Observer* 8, p. 1-22, 1993.
- Harbison, R. W. and Hanushek, E. A. *Educational performance of the poor*. Oxford University Press, 1992.
- Lockweed, M. E. and Vespoor, A. M. *Improving primary education in developing countries*. Oxford University Press, 1991.
- Londoño, J. L. *Is there a health gap in Latin America?* Technical Department for the Latin American and Caribbean Region, Washington, D.C.: The World Bank, 1996.
- McGreevey, W. *et alii. Política e financiamento do sistema de saúde brasileiro: uma perspectiva internacional*. Brasília, DF: IPEA/IPLAN, 1984.
- Meerman, J. *Public expenditure in Malaysia: who benefits and why*. New York: Oxford University Press, 1979.
- Mesa-Lago, C. Financiamento de la atención a la salud en America Latina y el Caribe, con focalización en el seguro social. *Documento de la Serie de Estudios de Seminarios del IDE, N° 42*. Washington, D.C, 1989.
- Negri, B. *O financiamento da educação no Brasil*. Brasília: Ministério da Educação e Cultura, 1996.(mimeo).
- Petrey, A. H. *El gasto publico social y sus efectos distributivos*. Rio de Janeiro: Estudos Conjuntos de Integração Econômica da América Latina, 1987.
- Pradhan, S. Evaluating public spending. *World Bank Discussion Paper 323*, Washington, D.C., 1996.
- Psacharopoulos, G. Returns to investment in education: a global update. *Working Paper 1067*. Washington, D.C.: World Bank, 1993a.
- _____. *Poverty and income distribution in Latin America: the story of the eighties*. Washington, D.C.: World Bank, 1993b.
- Psacharopoulos, G., Valenzuela, J. and Arends, M. Teachers' salaries in Latin America: assessing priorities for schooling investments. *Working Paper 1056*, Washington, D.C.: World Bank, 1993.
- Purkey, S. L. and Smith, M. S. Effective schools: a review. *The Elementary School Journal* 83, p. 427-452, 1983.

- Ransom, A. Khoo, S. M. and Selvaratnam, V. *Improving higher education in developing countries*. Washington, D.C.: World Bank, 1993.
- Romer, P. Increasing returns and long run growth. *Journal of Political Economy* 94, p. 1002-37. 1986.
- Sampaio de Sousa, M. C. Patterns of taxation in Latin America and East Asia. *Revista Brasileira de Economia*. Forthcoming.
- Selvaratnam, V. *Innovations in higher education: Singapore at the competitive edge*. Washington, D.C.: World Bank, 1994.
- Soria, V. The Nafta, the unique thought, and the transformation of the Mexican protection system. Paper presented at the *9th International Conference on Socio Economics, SASE*, Montreal, Canadá, July 5-7, 1997
- Tan, J. P. and Mingat, A. Public funding of education: a comparison. *Comparative Education Review*: 30, p. 260-270, 1986.
- _____. *Education in Asia: a comparative study of cost and financing*. World Bank Regional and Sectoral Studies, The World Bank, 1992.
- Van de Walle, D. and Nead, K. *Public spending and the poor: theory and practice*. London: Oxford University Press, 1995a.
- _____. The distribution of subsidies through public health services in Indonesia -1978-87. In: Van de Walle, D. and Nead, K. (eds.), *Public spending and the poor: theory and practice*. London: Oxford University Press, 1995b.
- Winkler, D. *La educación superior en América Latina: cuestiones sobre eficiencia y equidad*. Documentos para Discusión del Banco Mundial, 77S, Washington, D.C., 1994.
- World Bank *Financing health services in developing countries: an agenda for reform*. A World Bank Policy Study, Washington, D.C., 1987.
- _____. *Malaysia: fiscal reform for stable growth*. Report N° 105512-UG. East Asia and Pacific Regions, Washington, D.C.: World Bank, 1992
- Zchock, D. K. Medical care under social insurance in Latin America. *Latin American Research Review* 21, p. 99-122, 1986.

Appendix I

Table A-1

Means of Government Expenditures in Developing Countries, Average of 1985-1989

Government Expenditures	Latin America		East Asia	
	Percentages of GDP	Percentages of total Expenditures	Percentages of GDP	Percentages of total Expenditures
General Public Services	5,30	30,63	6,34	28,22
- Defense	1,62	9,74	2,07	10,87
Social Services	5,64	31,34	7,96	33,51
- Education	2,77	14,85	4,8	20,45
- Primary and Secondary	1,73	9,69	3,26	15,34
Tertiary	0,17	1,12	0,63	2,64
- Health	1,22	7,21	1,75	7,03
- Hospitals	0,77	4,77	0,86	3,89
- Clinic	0,07	0,88	0,14	0,78
Social Security and Welfare	0,88	5,48	0,66	3,09
Economic Services	3,65	18,56	6,12	25,17
Others	4,05	18,63	4,21	18,26

Source: Pradhan (1996).

Table A-2

Ordinary Least Square Regression Between Gross Enrollment Ratios and per-capita GNP, by Education Cycle in East Asia and Latin America – 1992

	Latin America and East Asia			World		
	primary	secondary	higher	primary	secondary	higher
Per capita GNP	0,00328	0,00581	0,00303	0,00057	0,00119	0,00040
	(2,19)	(4,28)	(3,24)	(3,,20)	(8,3)	(5,3)
(Per capita GNP) ²	-1,2729e-007	-2,2500e-007	-1,3390e-007	-3,2e-007	-0,53e-007	-0,16e-007
	(-1,62)	(-3,16)	(-2,72)	(2,5)	(5,1)	(3,0)
Intercept	77,38	48,95	13,78	82,74	24,31	5,06
Number of observations	49	49	49	91	91	91
R ²	0,14	0,38	0,21	0,12	0,58	0,44

Source: Data on enrollment ratios and GNP are from World Bank (1994).

Table A-3
ICP Estimates of Per Capita GDP and Ratio of Public Sector to Private Sector
Salaries, 1992

Economy/Region	GDP Per Capita US \$	Senior Level		Mid -Level	
		A	B	A	B
Latin America`					
- Argentina	4680	24,1 ^b	24,1	28,6	28,6
- Chile	6190	70,4	63,2		
- Uruguay	6000			37,1 ^b	37,1
- Venezuela	6740	29,5	42,4	53,4	53,4
East Asia					
- Korea	7190	98,8 ^b	98,8	81,3	82,2
- Malaysia	5900	40,0	33,3	34,3	50,0
- Philippines	2320	27,7	24,3	25,0	32,5
- Singapore	14920	114,0 ^b	114,0	115,0 ^b	115
- Taiwan	7954	65,2	60,3	63,5	65,8
- Thailand	4610	47,6	27,5	33,6	43,0

Source: World Bank (1993).

Previsão de séries de tempo na presença de mudança estrutural: redes neurais artificiais e modelos estruturais*

Wilson R. Corrêa§
Marcelo S. Portugal†

RESUMO

As políticas de estabilização de preços e o processo de liberalização comercial realizados nas últimas décadas aumentaram muito a dificuldade de geração de previsões precisas de séries econômicas, em função das mudanças estruturais que foram ocasionadas no processo gerador dos dados. Este artigo apresenta uma avaliação empírica da capacidade preditiva de modelos de Redes Neurais Artificiais (RNA) e modelos estruturais de série de tempo (MEST) quando existe uma mudança estrutural. Nossa preocupação básica é avaliar a capacidade relativa destes dois modelos de previsão em identificar a ocorrência da mudança estrutural e a rapidez no ajuste das previsões um passo à frente após a mudança, com e sem intervenção no modelo. Nestes exercícios empíricos usamos séries reais, tais como a taxa de inflação e o valor das importações brasileiras, e criadas por simulação. As séries geradas por simulação seguem diferentes processos ARIMA com imposição de mudanças estruturais na média e tendência. Os resultados mostram uma capacidade levemente superior das RNAs em identificar a ocorrência de quebras estruturais e realizar previsões um passo à frente nos períodos imediatamente seguintes a esta.

Palavras-chave: mudança estrutural, redes neurais, séries de tempo.

ABSTRACT

The Brazilian price stabilisation policies and trade liberalisation measures of this decade have considerably increased the difficulty in generating accurate time series forecasts due to structural changes in the data generation processes. In this paper we provide an empirical evaluation of the forecasting performance of Artificial Neural Networks (ANN) and Structural Time Series models (STS) in the presence of structural change. We are basically interested in evaluating the capability of ANN and STS models in terms of both identifying that a structural change has happened and the speed of adjustment of the one step ahead forecasts after the change. We use both real and simulated time series in these exercises. The simulated series are generated from ARIMA processes with imposed structural changes in the mean and trend. On the other hand, we also use real time series data for the Brazilian inflation rate and total imports. The results for the one step ahead forecasts show that the ANN models present a marginally better performance than the STS in the periods just after the structural change.

Key words: structural change, neural networks, time series.

* Os autores gostariam de agradecer a colaboração do bolsista de iniciação científica do CNPq Gregório Silva Caetano. Uma versão anterior deste artigo foi apresentada no XVIII International Symposium on Forecasting, em Edinburgh, Escócia.

§ Mestrando em economia na UFRGS e professor substituto na UFRGS e no Centro Regional de Economia e Administração da Universidade de Passo Fundo (rotatori@vortex.ufrgs.br).

† Professor do Curso de Pós-Graduação em Economia da UFRGS e pesquisador do CNPq (msp@vortex.ufrgs.br).

1 Introdução

A partir da década de setenta, tem sido grande a proliferação de modelos univariados de previsão com aplicação em economia. O objetivo básico destes modelos é a realização de previsões de curto prazo, utilizando apenas os dados referentes à própria série que se deseja prever. Em algumas extensões, outras séries podem ser utilizadas, mas, de forma geral, não é necessário a especificação de um modelo econométrico completo. A experiência prática na utilização de modelos univariados para a previsão de curto prazo tem se mostrado bastante satisfatória.

A renovação e o aperfeiçoamento dos modelos têm sido constante. Os modelos ARIMA, que foram largamente utilizados nas décadas de setenta e oitenta, foram rapidamente substituídos pelos chamados modelos estruturais de séries de tempo, de enfoque clássico ou bayesiano. Mais recentemente, mesmo estes modelos estruturais já começam a ser questionados quanto à sua eficiência em gerar previsões de curto prazo pelas redes neurais, desenvolvidas originalmente no campo da inteligência artificial.

O objetivo central deste artigo é avançar no estudo da comparação entre modelos de previsão univariados, de forma a avaliar a capacidade de cada modelo em lidar com a existência de quebras estruturais na série que é objeto da previsão. Nossa análise se restringirá aos modelos estruturais de séries de tempo e a modelos de redes neurais. Estes modelos foram escolhidos pela sua já demonstrada capacidade de reconhecer padrões e regularidades em séries de tempo. Nossa análise tentará avaliar a capacidade relativa de cada modelo em identificar o ponto de quebra estrutural e a rapidez do ajustamento das previsões um passo à frente após a quebra, com e sem a utilização de intervenção sobre o modelo.¹

As seções seguintes apresentam, brevemente, a metodologia a ser usada e os resultados empíricos.

2 Modelos de previsão

Nesta seção introduziremos os modelos de previsão a serem utilizados. Como mencionado acima, trataremos basicamente de dois tipos de metodologias de previsão: modelos estruturais de séries de tempo e modelos de redes neurais.

1 Comparações do desempenho preditivo destes modelos para séries de tempo sem quebra estrutural podem ser encontradas em Portugal (1995) e Fernandes, Navaux e Portugal (1996).

2.1 Modelos estruturais de séries de tempo

Nos últimos anos, especialmente após a introdução do filtro de Kalman em econometria, a literatura sobre modelos estruturais tem avançado substancialmente. O filtro de Kalman permitiu a operacionalização dos modelos estruturais, ao fornecer uma forma fácil de estimar os componentes estruturais do modelo que são variantes no tempo.²

Seguindo-se Harvey (1989), podemos dizer que a idéia básica é ver uma série de tempo como o resultado do somatório de vários componentes não diretamente observados. Em geral, assume-se que a série tenha um componente de tendência (μ_t), cíclico (ψ_t), sazonal (γ_t) e irregular (ε_t). Em termos analíticos

$$y_t = \mu_t + \gamma_t + \psi_t + \varepsilon_t.$$

Esta forma de modelagem de séries de tempo tem o grande atrativo de oferecer uma descrição das principais características da série. Ela nos possibilita a análise de mudança de tendência e alterações no padrão sazonal ou no comportamento cíclico da série. Por permitir uma direta interpretação dos componentes da série este tipo de modelo pode ser considerado mais interessante que os tradicionais modelos ARIMA. Na verdade, para cada tipo de modelo de decomposição em componentes não observados existe um modelo ARIMA correspondente.

O modelo completo é obtido após a especificação de cada um dos componentes isoladamente. Por razões de espaço vamos utilizar uma versão reduzida do modelo apresentado acima, chamado na literatura de modelo estrutural básico, que exclui o componente do ciclo. Esta simplificação não prejudica a exposição, pois o componente cíclico é modelado de forma semelhante ao componente sazonal, usando-se funções seno e cosseno. O modelo estrutural básico pode então ser especificado como

$$y_t = \mu_t + \gamma_t + \varepsilon_t \tag{1}$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \tag{2}$$

2 Na verdade, os modelos estruturais de séries de tempo fazem parte de uma classe mais genérica de modelos de decomposição de uma série de tempo em seus componentes não diretamente observados. Neste artigo trataremos apenas dos modelos estruturais de inspiração clássica. Modelos muito semelhantes, de inspiração bayesiana, chamados de modelos lineares dinâmicos, podem ser encontrados em West e Harrison (1997).

$$\beta_t = \beta_{t-1} + \zeta_t \quad (3)$$

$$\gamma_t = \sum_{j=1}^{s/2} \gamma'_t$$

onde γ'_{jt} é formado por

$$\begin{bmatrix} \gamma_{jt} \\ \gamma_{jt}^* \end{bmatrix} = \begin{bmatrix} \cos \lambda_j & \text{sen } \lambda_j \\ -\text{sen } \lambda_j & \cos \lambda_j \end{bmatrix} \begin{bmatrix} \gamma_{j,t-1} \\ \gamma_{j,t-1}^* \end{bmatrix} + \begin{bmatrix} \omega_{jt} \\ \omega_{jt}^* \end{bmatrix}$$

e ε_t , η_t e ζ_t são erros aleatórios com distribuição normal e variâncias σ_ε^2 , σ_η^2 e σ_ζ^2 , ω_t e ω_t^* são também erros aleatórios normais com variância igual a σ_ω^2 , λ_j é a frequência e γ'_{jt} aparece por construção.

Para simplificar a exposição vamos supor que estamos lidando com dados trimestrais, de forma que $s = 4$. Desta forma, a frequência $\lambda_j = 2\pi j/4$ para $j = 1$ e 2 , o que produz $\lambda_j = \pi, \pi/2$. Usando relações trigonométricas básicas, podemos rescrever o sistema de equações como

$$\begin{aligned} \gamma_{1t} &= \gamma_{1,t-1}^* + \omega_{1t} \\ \gamma_{1t}^* &= -\gamma_{1,t-1} + \omega_{1t}^* \\ \gamma_{2t} &= -\gamma_{2,t-1} + \omega_{2t} \\ \gamma_{2t}^* &= -\gamma_{2,t-1}^* + \omega_{2t}^* \end{aligned} \quad (4)$$

A estimação do modelo com parâmetros variáveis pode ser feita utilizando-se o filtro de Kalman. O modelo formado pelas equações (1) a (4) pode ser facilmente colocado no formato do estado de espaço, como mostrado abaixo.

$$y_t = [1 \ 0 \ 1 \ 0 \ 1] \alpha_t + \varepsilon_t \quad (5)$$

$$\alpha_t = \begin{bmatrix} \mu_t \\ \beta_t \\ \gamma_{1t} \\ \gamma_{1t}^* \\ \gamma_{2t} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \beta_{t-1} \\ \gamma_{1,t-1} \\ \gamma_{1,t-1}^* \\ \gamma_{2,t-1} \end{bmatrix} + \begin{bmatrix} \eta_t \\ \zeta_t \\ \omega_{1t} \\ \omega_{1t}^* \\ \omega_{2t} \end{bmatrix} \quad (6)$$

Nesta especificação temos uma variável para representar o nível da série (μ_t) e outra para a tendência propriamente dita (β_t). Nas séries que não apresentam uma tendência clara o valor de β_t será aproximadamente zero, enquanto que se houver uma taxa de crescimento constante teremos $\beta_t = \beta$.

Como mencionamos acima, a estimação dos modelos de decomposição em componentes não observáveis é feita por meio do Filtro de Kalman. O filtro é um algoritmo desenvolvido no âmbito da engenharia para resolver o problema de extração de sinal.³ Utilizando a metodologia apresentada por Atkinson, Koopman e Shephard (1997) temos os modelos representados na forma de espaço de estados, seguindo a formulação Gaussiana, dados pela equação de medida (7) e transição (8).

$$y_t = Z_t \alpha_t + \varepsilon_t \quad (7)$$

$$\alpha_t = T_t \alpha_{t-1} + \eta_t \quad (8)$$

$$\alpha_0 / Y_0 \sim N(a_0, P_0), \varepsilon_t \sim N(0, H_t), \eta_t \sim N(0, Q_t)$$

$$E(\varepsilon_t, \eta_t) = 0$$

$$E(\varepsilon_t, \alpha_0) = 0$$

$$E(\eta_t, \alpha_0) = 0$$

$$E(\varepsilon_i, \varepsilon_j) = \delta_{ij} \sigma^2 H_t$$

$$E(\eta_i, \eta_j) = \delta_{ij} \sigma^2 Q_t$$

onde δ_{ij} é o delta de Kronecker, y_t representa o vetor de medida, α_t o vetor de estado e Z_t e T_t são matrizes conhecidas.

3 Ver Kalman (1960). Diferentes interpretações do Filtro de Kalman na área da econometria e estatística clássica e Bayesiana podem ser encontradas em Duncan and Horn (1972), Meinhold e Singpurwalla (1983) Harvey (1987), Diderrich (1985).

O que o Filtro de Kalman faz é gerar a melhor previsão para o vetor de estado e sua matriz de covariância (P_t) a cada período de tempo. Uma vez que uma nova observação se torne disponível, o filtro atualiza esta previsão levando em conta o erro de previsão cometido. Neste sentido, o filtro aprende com os seus erros. A estimação das variâncias (σ^2), às vezes denominadas hiperparâmetros, é feita por máxima verossimilhança.⁴ O filtro consiste ainda no seguinte conjunto de equações

$$a_{t+1/t} = T_{t+1}a_{t/t-1} + K_t v_t \quad (9)$$

$$P_{t+1/t} = T_{t+1}P_{t/t-1}L_t' + Q_{t+1} \quad (10)$$

$$v_t = y_t - Z_t a_{t/t-1} \quad (11)$$

$$K_t = T_{t+1}P_{t/t-1}Z_t'F_t^{-1}$$

$$F_t = Z_t P_{t/t-1} Z_t' + H_t$$

$$L_t = T_{t+1} - K_t Z_t \quad t = 1, \dots, n$$

O Conjunto de equações (9) a (11) deixa claro o processo de aprendizagem que é realizado pelo filtro de Kalman. A cada momento no tempo, a atualização do vetor de estados é feita com base no erro de previsão um passo à frente descontado pelo ganho do filtro (K_t). O ganho do filtro embute os pesos relativos que o modelo atribui ao passado da série e a mais recente observação. Estes pesos, por sua vez, são função das variâncias estimadas da perturbação aleatória presentes nas equações de medida e transição.

A partir da implementação do filtro, utilizamos o método de *disturbance smoother*, o qual é passado em toda a amostra iniciando-se com $r_n = N_n = 0$ e realizando-se as recurssões dadas por (12) e (13).

$$e_t = F_t^{-1} v_t - K_t' r_t \quad (12)$$

$$r_{t-1} = Z_t' F_t^{-1} v_t + L_t r_t \quad (13)$$

$$D_t = F_t^{-1} + K_t' N_t K_t$$

$$N_{t-1} = Z_t' F_t^{-1} Z_t + L_t' N_t L_t \quad t = n, \dots, 1$$

4 Uma explicação mais detalhada do Filtro de Kalman e da forma de estimação dos hiperparâmetros, assim como a derivação das equações de previsão e atualização, pode ser encontrada em Portugal (1993).

Em particular, podemos observar que $H_t e_t$ é a estimativa de ε_t e $Q_t r_{t-1}$ a de η_t com os erros quadráticos médios dados respectivamente por $H_t - H_t D_t H_t'$ e $Q_t - Q_t N_{t-1} Q_t'$. Essas estimativas são também denominadas *smoothed estimates* ou resíduos auxiliares (*auxiliary residuals*). De acordo com Harvey e Koopman (1992) podemos utilizar a simples inspeção gráfica destes resíduos, em conjunto com a análise dos testes de normalidade, como indicativo da existência de quebra estrutural, bem como da presença de *outliers* na série a ser analisada. Na existência de quebras estruturais e *outliers* são realizadas intervenções no modelo com a especificação de variáveis *dummy* para captar o comportamento irregular da série. No presente trabalho utilizamos esta metodologia para a especificação dos períodos de ocorrência da quebra e modelagem das irregularidades apresentadas.⁵ A inclusão das variáveis *dummy* como variáveis explicativas é relativamente simples e requer a utilização das recursões apresentadas por De Jong (1989).⁶

Deve-se destacar, ainda, que mesmo quando o modelo é corretamente especificado os resíduos são serialmente correlacionados e, de acordo com Harvey e Koopman (1992), seguem um processo ARIMA(2,0,0), influenciando, assim, os testes de normalidade apresentados. Contudo, no presente trabalho tais testes são ajustados para a presença autocorrelação serial.

2.2 Redes neurais

As Redes Neurais Artificiais (RNA) surgiram como uma área da informática cujo objetivo básico era o de criar modelos artificiais do cérebro humano, de forma a permitir que computadores "pensem". Em outras palavras, uma RNA é, basicamente, um modelo não paramétrico do cérebro humano.⁷ Uma RNA funciona de forma similar ao cérebro humano, tentando reconhecer padrões e regularidades nos dados que lhe são apresentados. Elas podem aprender por experiência e fazer generalizações com base em seu conhecimento passado.

5 Uma metodologia diferenciada de determinação do período de ocorrência de quebra estruturais pode ser encontrada em Atkinson, Koopman e Shephard (1997).

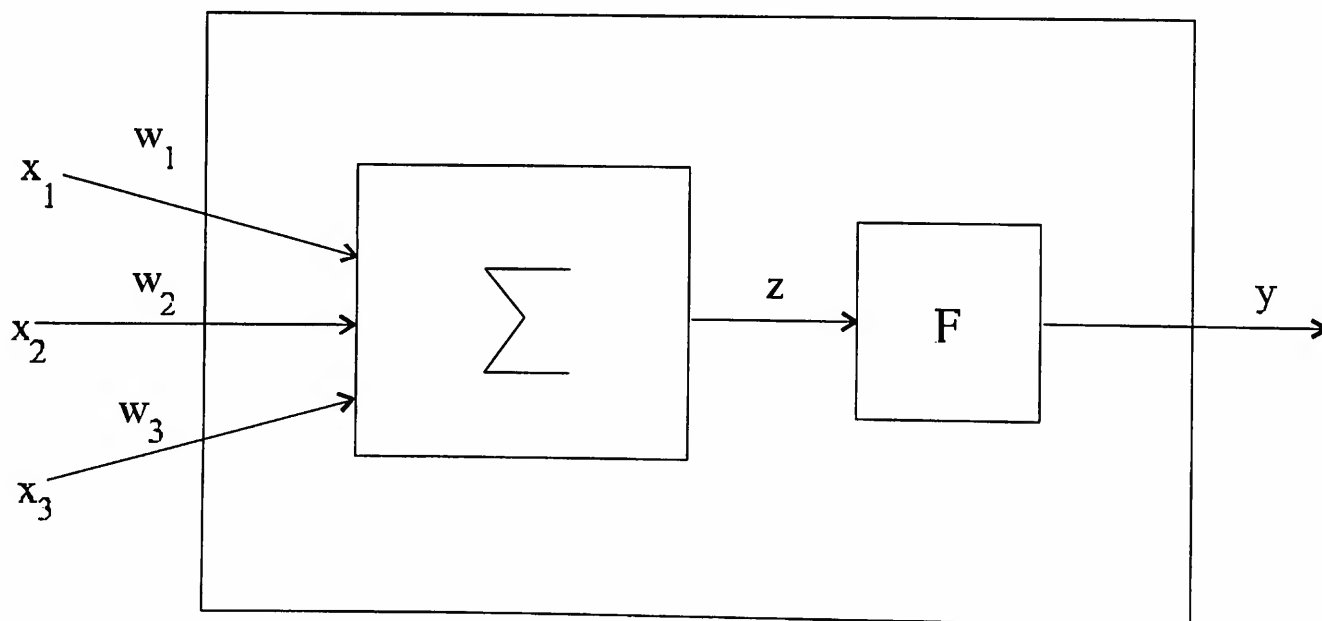
6 Por motivo de espaço omitiremos as mesmas, mas uma análise detalhada pode ser encontrada em Koopman, Harvey, Doornick e Shephard (1995) e Atkinson, Koopman e Shephard (1997).

7 Para uma introdução geral às redes neurais, ver Haykin (1994).

O cérebro humano, com seus bilhões de neurônios interligados em rede, é bem distinto de um computador. As experiências e sensações, que são observadas por meio dos sentidos humanos, geram sinais elétricos que passam pelos neurônios criando elos de ligação entre eles. Estas ligações entre neurônios são novamente sensibilizadas sempre que uma outra experiência ou sensação similar ocorre, permitindo, então, o seu reconhecimento. O aprendizado de algo novo é sempre feito em relação àquilo que já era conhecido, isto é, ele confirma ou modifica as ligações preexistentes entre os neurônios. Na verdade, diferentes experiências ou sensações são armazenadas no cérebro com diferentes ligações, ou diferentes intensidades de ligações, entre neurônios.

Cada neurônio, individualmente, funciona de maneira bastante simples. O neurônio recebe sinais elétricos das outras células do corpo humano por meio de pontos de contato (sinapses), processa estes sinais e, se o resultado deste processamento em um dado período de tempo excede um certo valor, produz outros sinais que são, então, transmitidos a outras células. Aparentemente, não é o funcionamento do neurônio, mas sim o alto grau de ligação entre eles, que é responsável pelo grande poder "computacional" do cérebro humano.⁸

Figura 1



⁸ Ver, por exemplo, Wasserman (1989, p. 194). O processamento em paralelo efetuado pelo cérebro humano permite que sejam realizadas diversas tarefas de forma simultânea. Neste sentido, é possível, por exemplo, falar e caminhar ao mesmo tempo.

Um neurônio artificial atua exatamente como um neurônio comum, como pode ser visto na Figura 1. O neurônio artificial recebe vários sinais entrada (x), calcula uma média ponderada destes sinais (z), e quando esta média é aplicada a uma função de ativação (F) produz então um sinal de saída (y). Neste sentido, usando o exemplo da Figura 1, temos que

$$z = \sum_{i=1}^3 x_i w_i.$$

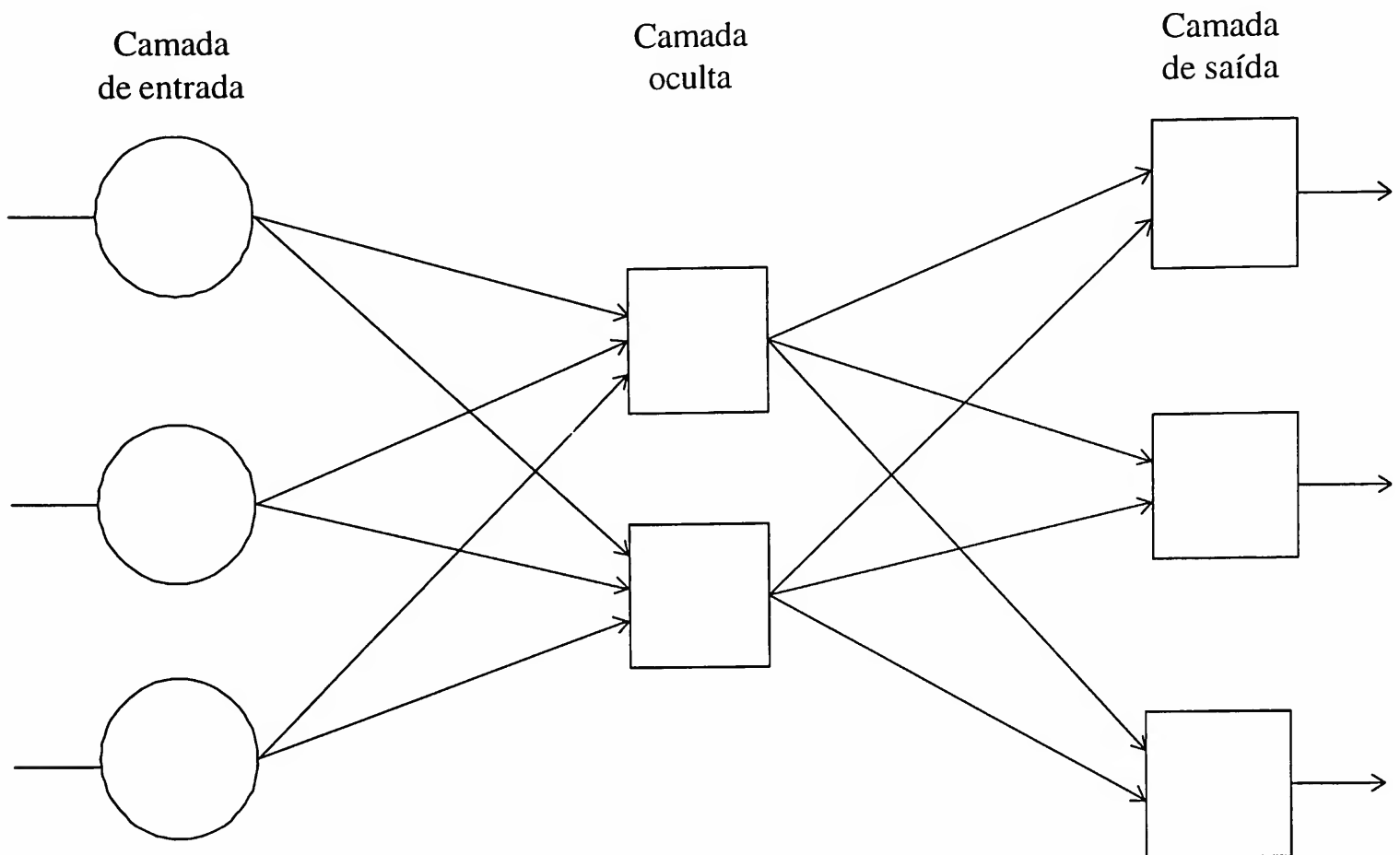
A função de ativação pode, em princípio, tomar diversas formas. Ela pode ser uma função logística, degrau, ou mesmo uma simples função linear. A função mais usada em aplicações práticas é uma função logística do tipo

$$F(z) = \frac{1}{1 + e^{-z}}$$

Uma RNA é composta por um certo número de neurônios artificiais organizados em diferentes camadas, como pode ser visualizado pela Figura 2. A primeira camada, chamada usualmente de camada de entrada, serve apenas para propiciar a entrada dos dados, não executando qualquer computação. É nas camadas intermediárias ou ocultas e na camada de saída que ocorre o efetivo processamento dos dados. A rede mostrada na Figura 2 é usualmente chamada de *feed forward network*, pois apresenta todas as ligações em um só sentido. É possível, contudo, conceber redes em que as informações movimentam-se em várias direções.

A questão principal para operacionalizar uma RNA é como obter os vetores de pesos (w), que correspondem às sinapses que se estabelecem entre os neurônios biológicos. No cérebro humano as sinapses são estabelecidas, e fortalecidas ou não, por meio de um processo de aprendizado. Para uma RNA precisamos também constituir um processo de aprendizado para a determinação dos pesos que ligam os neurônios artificiais. Para a realização de previsões, nós vamos usar apenas algoritmos de aprendizado supervisionado. A idéia básica é treinar a RNA até que esta seja capaz de reconhecer padrões e regularidades nos dados, para então extrapolar quanto ao comportamento futuro.

Figura 2



Para treinar a RNA nós podemos usar um algoritmo de aprendizado supervisionado onde, dado um conjunto de dados de entrada, os pesos são ajustados continuamente de forma a gerar uma saída desejada. O treinamento de uma RNA consiste, portanto, no ajuste dos pesos de forma a produzir um resultado desejado. O algoritmo de treinamento mais usado na literatura é conhecido como *backpropagation*.⁹ Por meio dele os pesos são ajustados de forma a minimizar os erros de previsão dentro do período amostral. Em cada interação os pesos são revisados de acordo com a inclinação da superfície de erros, tentando atingir um mínimo global. As equações (14) a (16) apresentam o algoritmo de *backpropagation*.

$$\Delta w_{ij,k}^n = \eta \delta_{j,k} y_{i,k-1} \quad (14)$$

9 Na verdade, o algoritmo de *backpropagation* é apenas uma das várias alternativas disponíveis para a realização do aprendizado da rede. O maior problema na sua utilização diz respeito à possibilidade de obtenção de mínimos locais. Na prática, contudo, existem diferentes formas de tentar evitar este tipo de ocorrência. Ver, a este respeito, Portugal e Fernandes (1996, p. 68).

$$w_{ij,k}^{n+1} = w_{ij,k}^n + \Delta w_{ij,k}^n \quad (15)$$

$$\delta_{j,k} = \left(\frac{dy}{dz} \right)_{i,k} (y_{i,k} - y_{i,k}^T); \quad (16)$$

para $i = 1, 2, \dots, I$; $j = 1, 2, \dots, J$ e $k = K$

onde; η é o coeficiente de aprendizado (usualmente entre 0.01 e 1.0);

$y_{i,k}$ é a saída do neurônio i na camada k ;

$y_{i,k}^T$ é o valor desejado para $y_{i,k}$;

$w_{ij,k}^n$ é o peso ligando o neurônio i na camada $k-1$ ao neurônio j na camada k no passo n .

Note que o algoritmo acima só pode ser aplicado para a última camada da rede, pois ele pressupõe o conhecimento do valor desejado de saída, informação esta que só está disponível para a última camada ($k = K$). Para as camadas ocultas o valor de $\delta_{i,k}$ na equação (16) tem de ser obtido de outra forma. O algoritmo de *backpropagation* deriva seu nome do fato de propagar para trás o valor de $\delta_{i,k}$ usando os pesos em cada camada, como é indicado pela equação (16a).

$$\delta_{j,k} = \left(\frac{dy}{dz} \right)_{i,k} \left(\sum_j \delta_{j,k} w_{ij,k} \right); \quad (16a)$$

para $k = 1, 2, \dots, K-1$.

Para atingir a convergência para o mínimo de forma mais rápida, pode ser utilizada uma transformação auto-regressiva das equações (14) e (15).¹⁰

$$\Delta w_{ij,k}^{n+1} = \eta (\delta_{j,k} y_{i,k-1}) + \alpha (\Delta w_{ij,k}^n) \quad (17)$$

10 Ver Rumelhart, Hinton e Williams (1986).

$$w_{ij,k}^{n+1} = w_{ij,k}^n + \Delta w_{ij,k}^{n+1} \quad (18)$$

onde α é o *momentum coefficient* (normalmente em torno de 0.9).

3 Dados utilizados e metodologia

A partir destas considerações iniciais sobre a aplicação e uso do filtro de Kalman e das RNAs, vamos iniciar a análise das séries com uma breve apresentação das suas características básicas. Num segundo momento faremos uma discussão a respeito da metodologia de análise empregada para a realização de previsões considerando a ocorrência das quebras estruturais e sua influência na realização destas últimas.

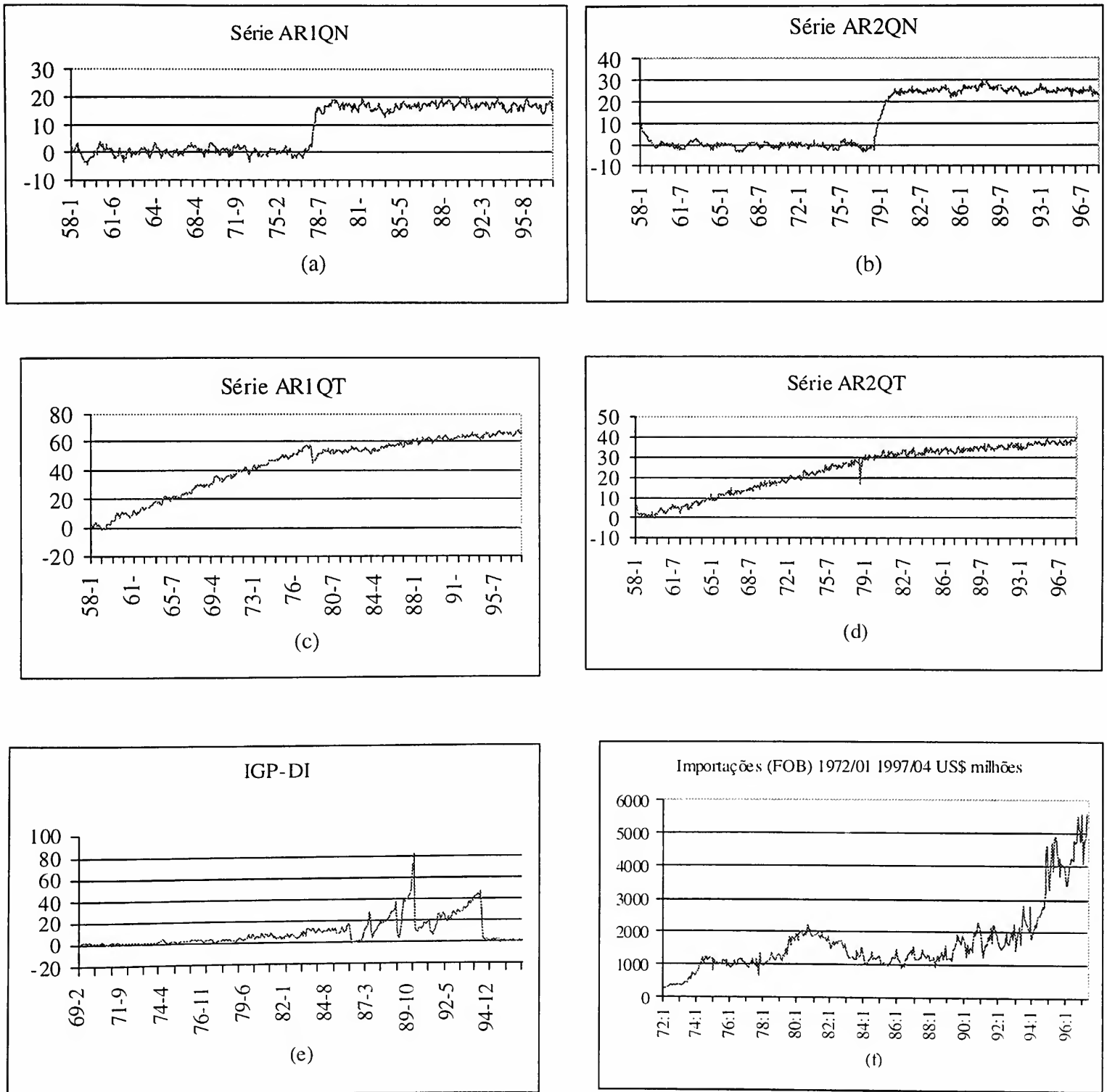
As séries passíveis de análise neste trabalho podem ser subdivididas em dois grupos distintos. O primeiro constitui-se de séries construídas artificialmente a partir de um processo estocástico ARIMA, e o segundo pelas séries denominadas “reais”, as quais correspondem ao total das importações brasileiras, no período compreendido entre janeiro de 1972 e abril de 1997, e a taxa de inflação medida pelo IGP-DI para o período de fevereiro de 1969 até junho de 1987

As séries geradas pelo processo ARIMA procuram emular dois comportamentos distintos, o de quebra estrutural no nível da série e o de quebra na inclinação da tendência. Desta forma temos 4 séries para dois processos ARIMA distintos AR(1) e AR(2). As seis séries estudadas são apresentadas nos Gráficos 1(a) a 1(f).

No que concerne às séries reais, cabe aqui apenas um comentário a respeito da restrição do período de análise do IGP-DI até junho de 1987. Este inclui a primeira tentativa de estabilização econômica com o Plano Cruzado, quando as taxas de inflação apresentaram uma queda substancial nos seus valores de março a dezembro de 1986, com a ocorrência inclusive de uma deflação em abril de 1986. Posteriormente, durante o primeiro semestre de 1987, a inflação retoma a sua trajetória de crescimento anterior, ultrapassando inclusive os valores observados pré-implementação do Plano. Devemos ressaltar aqui, que apesar do Plano Real também representar uma quebra estrutural significativa, como podemos visualizar no Gráfico 1(e), esta se assemelha bastante à ocorrência da quebra estrutural no nível das séries geradas pelos processos ARIMA, onde existe uma abrupta alteração no nível da série e a permanência destes valores em torno da nova média. Neste sentido, a análise comparativa das previsões resultantes da utilização dos modelos estruturais

(MEST) e das RNAs para a inflação pós-Plano Real implicaria apenas uma duplicação dos resultados já obtidos.¹¹

Gráfico 1



¹¹ Além disso, as frequentes quebras estruturais resultantes dos diversos planos de estabilização de preços dificultam a obtenção das estimativas de máxima verossimilhança confiáveis para os hiperparâmetros.

Considerando que o objetivo central deste trabalho consiste em realizar uma avaliação empírica da performance dos MEST e das RNAs na presença de quebras estruturais no DGP (*Data Generation Process*), concentraram-se esforços no sentido determinar o período de ocorrência da quebra na série. A partir desta determinação, realizaram-se previsões procurando investigar a capacidade preditiva e, fundamentalmente, de correção desta previsão com a introdução de novas informações. Nosso estudo baseia-se então na avaliação da capacidade relativa de ambos os modelos em reconhecer alterações no DGP mediante o aprendizado com os erros de previsão passados. No caso dos MEST, este aprendizado ocorre via decomposição dos erros de previsão, e para as RNAs, pelo processo de treinamento da rede com a minimização dos erros de previsão dentro do período amostral em questão.

Em particular, para a determinação do período de quebra foram utilizados os resíduos auxiliares gerados pelo Filtro de Kalman. Valores padronizados destes resíduos maiores do que 2 em módulo são considerados como um indicativo da presença de quebras estruturais, como discutido anteriormente.¹² A validade desta metodologia de determinação do período de ocorrência da quebra estrutural no DGP pode ser avaliada considerando-se as quatro séries geradas pelos processos ARIMA, onde evidentemente os períodos de quebra são conhecidos. Na Tabela 1 são apresentados os resultados obtidos para a determinação dos pontos de quebra.¹³

Tabela 1

Série	Modelo Especificado	Ponto de Quebra Estimado ¹	Ponto de Ocorrência da Quebra
AR2QN	Tendência Fixa	1978/07	1978/09
AR1QT	Tendência Fixa	1978/06	1978/06
AR1QN	Tendência Fixa	1977/12	1977/12
AR2QT	Tendência Fixa	1978/06	1978/06

Notas: (1) Para a determinação do ponto de quebra utilizamos o maior valor absoluto dos resíduos auxiliares estimados.

12 Nas estimações deste artigo foram utilizados os pacotes econométricos STAMP 5.0, para a estimação dos MEST, e RATS 4.0, para estimação das RNAs.

13 Uma apresentação mais detalhada dos modelos utilizados e dos testes de especificação é apresentada na seção 4.

Fica claro, portanto, a eficácia empírica da metodologia aplicada na determinação dos períodos de ocorrência de quebra. Contudo, existe a necessidade de considerarmos a simplicidade dos modelos estimados e o padrão relativamente regular no comportamento dos dados, à exceção da ocorrência da quebra. Dada a simplicidade dos modelos, fica relativamente facilitada a determinação do período de quebra pela decomposição dos resíduos auxiliares em irregulares e de nível, uma vez que a tendência é fixa. Não obstante a maior complexidade do modelo estrutural básico (BSM) estimado para a série de importações, os resultados se mostraram compatíveis com as informações externas disponíveis sobre a data de quebra, que neste caso não são conhecidas com certeza. Na verdade, quanto mais sofisticado for o modelo maior a dificuldade em se diferenciar entre os vários tipos de quebra estrutural. Os valores dos resíduos auxiliares para as importações indicaram uma quebra no nível em novembro de 1994 e outra em julho de 1995, além de uma quebra na inclinação da tendência em abril de 1992, resultados coerentes com o comportamento das importações brasileiras diante das alterações ocorridas em seu volume e valores a partir da abertura comercial iniciada no começo da década de 90.

Por fim, antes de nos determos na discussão dos resultados obtidos com as previsões, algumas observações se tornam importantes com relação à forma de realização das mesmas. Em particular, para os modelos estruturais a ocorrência da quebra é considerada em uma análise de previsão efetuada dentro da amostra, isto é, a estimação do vetor de estados é feita para todo o período amostral. Neste caso podemos efetuar as previsões de duas maneiras distintas. A primeira consiste em produzir estimativas para os valores da série a partir do ponto de ocorrência da quebra sem utilizarmos o processo de intervenção no modelo, que consiste basicamente na introdução de variáveis *dummy* na tentativa de captar a alteração no DGP, a segunda considerando a referida intervenção. Podemos definir três tipos de variáveis *dummy* procurando captar, respectivamente, alterações causadas pela presença de *outliers*, quebra no nível e quebra na inclinação da tendência.

A título de validar a necessidade de intervenção nos modelos estimados, com vistas à realização das previsões, investigou-se o comportamento do Erro Absoluto Médio (EAM) para ambas as situações. Embora, como veremos adiante, a inclusão das variáveis *dummy* melhore substancialmente os resultados dos testes de especificação dos modelos estimados, principalmente o teste para normalidade dos resíduos, elas têm pouco efeito em termos de redução do EAM. O EAM foi calculado a partir das previsões um passo à frente geradas por cada modelo estimado.

$$\text{impulse dummy } D = \begin{cases} 1 & t = t^* \\ 0 & t \neq t^* \end{cases}$$

$$\text{step dummy } D = \begin{cases} 1 & t \geq t^* \\ 0 & t < t^* \end{cases}$$

$$\text{staircase dummy } D = \begin{cases} (T - t^*) & t \geq t^* \\ 0 & t < t^* \end{cases}$$

De maneira semelhante, consideramos o EAM obtido com as previsões realizadas pelas RNAs como fator de comparação entre a capacidade preditiva de ambos os modelos (RNA e MEST). Em particular, para as RNAs realizamos as previsões um passo à frente treinando inicialmente com toda a amostra, de tal forma que a rede pudesse reconhecer o padrão de mudança no DGP. Deste modo podemos estabelecer uma comparação entre o desempenho de ambos os modelos na capacidade de reconhecer as alterações ocorridas.

4 Resultados empíricos

Dando prosseguimento à análise das séries “reais” e “construídas” vamos considerar inicialmente os resultados obtidos com as previsões para as séries geradas pelos processos ARIMA. Para facilitar a compreensão denominamos as séries em questão com o DGP utilizado na construção de cada uma. Neste caso temos AR1QN para a série gerada por um processo auto-regressivo de ordem 1 com uma quebra no nível, AR1QT para um processo auto-regressivo de ordem 1 com quebra na inclinação da tendência, e assim por diante. Apresentamos na Tabela 2 os resultados obtidos para a estimação dos diferentes modelos estruturais para as respectivas séries.

Como mencionado anteriormente, podemos observar que a utilização de intervenções nos MEST não melhora significativamente a sua capacidade de previsão, alterando, porém, a especificação do mesmo. A título de exemplificação apresentamos no Gráfico 2 as previsões um passo à frente com intervenção (2 a) e sem intervenção (2b), intervenção para a variável AR1QN, onde claramente não se observam diferenças, demonstrando que a ordem de grandeza desta é mínima.

Tabela 2

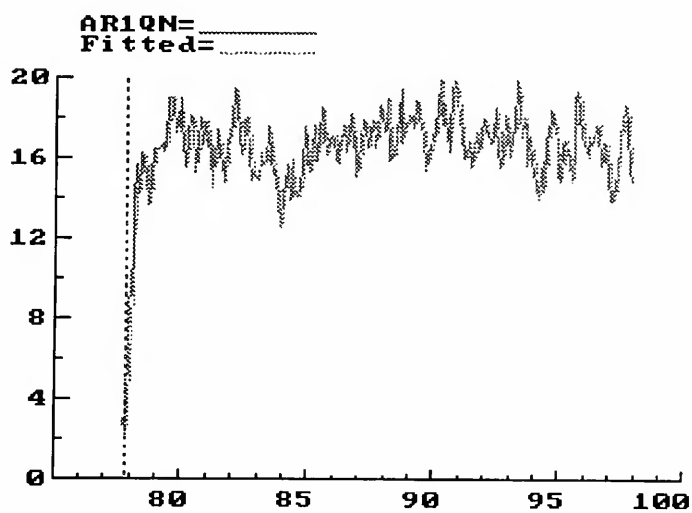
Variável	Testes de Especificação ¹	Modelo Escolhido	Critério de Ajuste	Erros Absoluto Médio	Coefficiente da Variável de intervenção ²
AR1QN Sem intervenção	Normalidade=13.48 H(159) = 0.96338 Q(20,19) = 32.55	Tendência Fixa $\sigma_{\zeta} = 0$	$R^2 = 0.9817$ $R_d^2 = 0.0141$ $R_s^2 = -0.013$	EAM=0.962414 t = t* em 77/11	
Com intervenção	Normalidade=2.735 H(159) = 0.96604 Q(20,19) = 39.56	Tendência Fixa $\sigma_{\zeta} = 0$	$R^2 = 0.9825$ $R_d^2 = 0.0737$ $R_s^2 = 0.0473$	EAM=0.933858 t = t* em 77/11	Nível 1977/12 6.1171 (5.5916)
AR2QN Sem intervenção	Normalidade=32.49 H(159) = 1.353 Q(20,19) = 107.4	Tendência Fixa $\sigma_{\zeta} = 0$	$R^2 = 0.9907$ $R_d^2 = 0.2179$ $R_s^2 = 0.2070$	EAM=0.99791 t = t* em 78/10	
Com intervenção	Normalidade=9.710 H(155) = 1.373 Q(20,18) = 61.70	Tendência Fixa $\sigma_{\zeta} = 0$	$R^2 = 0.9916$ $R_d^2 = 0.2702$ $R_s^2 = 0.2618$	EAM=0.991097 t = t* em 78/10	Nível 1978/09 4.950 (5.0934)
AR1QT Sem intervenção	Normalidade=3873 H(159) = 0.9668 Q(20,19) = 34.84	Tendência Fixa $\sigma_{\zeta} = 0$	$R^2 = 0.9963$ $R_d^2 = 0.0133$ $R_s^2 = -0.0108$	EAM=0.9167584 t = t* em 78/7	
Com intervenção	Normalidade=3.885 H(159) = 0.9663 Q(20,19) = 45.31	Tendência Fixa $\sigma_{\zeta} = 0$	$R^2 = 0.9969$ $R_d^2 = 0.1934$ $R_s^2 = 0.1737$	EAM=0.9267144 t = t* em 78/7	Nível 1978/6 -11.033 (-10.23) Inclinação 1979/9 -0.1939 (-2.3945)
AR2QT Sem intervenção	Normalidade=4890 H(159) = 0.8733 Q(20,19) = 53.00	Tendência Fixa $\sigma_{\zeta} = 0$	$R^2 = 0.9882$ $R_d^2 = 0.4183$ $R_s^2 = 0.3980$	EAM=0.950283 t = t* em 78/7	
Com intervenção	Normalidade=1.366 H(159) = 0.8780 Q(20,19) = 54.96	Tendência Fixa $\sigma_{\zeta} = 0$	$R^2 = 0.9906$ $R_d^2 = 0.5363$ $R_s^2 = 0.5201$	EAM=0.928755 t = t* em 78/7	Irregular 1978/6 -11.789 (-11.057)

Notas: (1) Os testes aqui apresentados correspondem aos testes de Bowman-Shenton para normalidade, que tem distribuição χ^2_2 ; H(h) é o teste para heterocedasticidade, com distribuição aproximadamente F(h,h); Q(p,d) é o teste Q de Box-Ljung baseado nas primeiras p autocorrelações dos resíduos com distribuição aproximadamente χ^2_{d-1} .
(2) Os valores entre parênteses indicam a estatística t dos coeficientes.

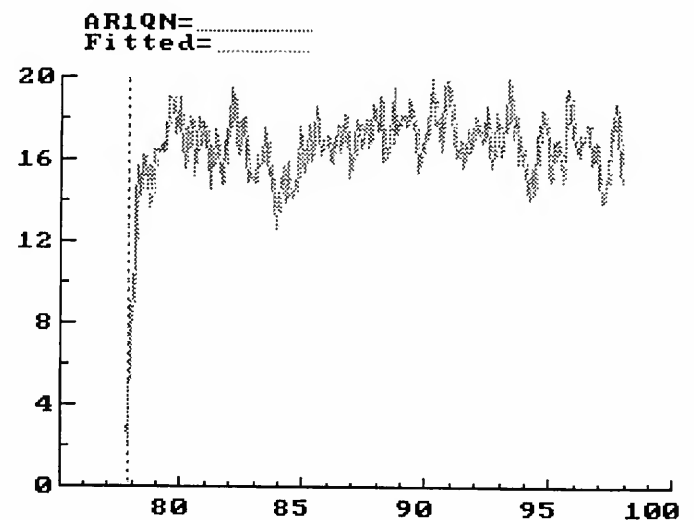
Para as séries geradas a partir de processos ARIMA observamos que o EAM chega inclusive a ser ligeiramente inferior quando não utilizamos intervenção, apesar de uma deterioração sensível dos testes de especificação do modelo, principalmente no que diz

respeito aos testes de normalidade dos resíduos. Os critérios de ajuste R_d^2 e R_s^2 ¹⁴ apresentam ganhos substanciais nos modelos com intervenção, em relação aos modelos sem intervenção. Contudo, devemos levar em conta que estes critérios representam resultados calculados dentro da amostra, enquanto que o EAM constitui um critério de análise baseado em extrapolações realizadas pelas previsões um passo à frente. De acordo com Harvey (1989), não obstante os resultados obtidos com o EAM, os modelos sem intervenção devem ser preteridos em relação aos modelos com intervenção, com base nos resultados obtidos nos testes de especificação do modelo. Em particular, os critérios de ajuste R_d^2 e R_s^2 também devem ser considerados. Neste caso modelos que apresentem ambos os critérios de ajuste com valores negativos devem ser rejeitados em favor da especificação mais simples apresentada na Tabela 2 seguindo critérios de parcimônia.

Gráfico 2



(a)



(b)

A boa performance dos modelos estruturais sem a utilização de intervenção demonstra sua poderosa capacidade preditiva, a qual está intimamente relacionada às equações de atualização e previsão bem como com o ganho do filtro, uma vez que estas envolvem parâmetros que são variáveis no tempo. Neste sentido, os modelos estruturais se mostraram capazes de prever e detectar corretamente os pontos de mudanças estruturais sem a necessidade de uma intervenção subjetiva no modelo.

14 Os critérios de ajuste R_d^2 e R_s^2 são definidos, respectivamente, por $(1 - SSE) / \sum_{t=2}^T (\Delta y_t - \hat{\Delta y}_t)^2$ e $1 - SSE/SSDSM$, onde SSE é definido como a soma dos quadrados dos resíduos e SSDSM é a soma dos quadrados dos resíduos da primeira diferença em torno das médias sazonais ($\Delta y_t = \sum_{j=1}^4 \gamma_j^* Z_{tj} + \eta_t, t = 2, \dots, T$).

Com relação às séries de importações e de inflação, o resultado se mostrou bastante diverso do anterior, como podemos observar na Tabela 3. Considerando-se inicialmente a série de importações, podemos observar uma evidente melhora nos testes de especificação do modelo. Novamente esta melhora está relacionada com o teste para normalidade dos resíduos e com critérios de ajuste para a especificação do modelo BSM com tendência fixa e a realização de uma intervenção em 94/09 no nível e em 92/04 na inclinação da tendência. Os resultados apontaram ainda para um ganho marginal em termos de EAM com a especificação das intervenções. Esta conclusão fica evidenciada quando consideramos as previsões um passo à frente para ambos os modelos que são apresentadas no Gráfico 3, onde temos o Gráfico 3(b) representando o modelo com intervenção e novamente reforçando a hipótese de independência destas com a capacidade preditiva do modelo.

Tabela 3

Variável	Testes de Especificação ¹	Modelo Escolhido	Critério de Ajuste	Erros Absoluto Médio	Coefficiente da Variável de intervenção ²
Importações Sem intervenção	Normalidade=283.2 H(97) = 7.621 Q (16,13) = 40.36 Chi ² (11) = 81.5152	Estrutural Básico (BSM)	$R^2 = 0.9500$ $R_{dl}^2 = 0.3442$ $R_s^2 = 0.1917$	EAM=459.062 t = t* em 94/11	
Com intervenção	Normalidade=175.7 H(97) = 7.305 Q (16,14) = 34.18 Chi ² (11) = 82.6465	Estrutural Básico com tendência fixa (BSM)	$R^2 = 0.9528$ $R_{dl}^2 = 0.3815$ $R_s^2 = 0.2377$	EAM=457.8173 t = t* em 94/11	Inclinação 47.055 em 92/4 (3.0982) Nível 572.00 em 94/9 (2.9565)
Com intervenção	Normalidade=202.3 H(97) = 7.160 Q (16,14) = 33.852 Chi ² (11) = 86.3604	Estrutural Básico com tendência fixa (BSM)	$R^2 = 0.9532$ $R_{dl}^2 = 0.3866$ $R_s^2 = 0.2440$	EAM=353.807 t = t* em 95/08	Inclinação 53.955 em 92/4 (3.3444) Nível 533.97 em 94/9 (2.7338) Nível -346.50 em 95/4 (-1.770)
IGP-DI Sem intervenção	Normalidade=294.5 H(73) = 18.15 Q (13,12) = 25.62	Nível Local <i>Random Walk plus Noise</i>	$R^2 = 0.8277$ $R_{dl}^2 = 0.0036$ $R_s^2 = -0.09178$	EAM=2.598021 t = t* em 86/04	
Com intervenção	Normalidade=345.9 H(73) = 14.22 Q (13,12) = 13.76	Nível Local <i>Random Walk plus Noise</i>	$R^2 = 0.8593$ $R_{dl}^2 = 0.1864$ $R_s^2 = 0.1085$	EAM=2.705026 t = t* em 86/04	<i>Dummy</i> para o período 86/03 - 86/12 -9.3261 (-8.3614)

Notas: (1) Chi²(s-1) é o teste de sazonalidade com distribuição aproximadamente χ_{s-1}^2 .

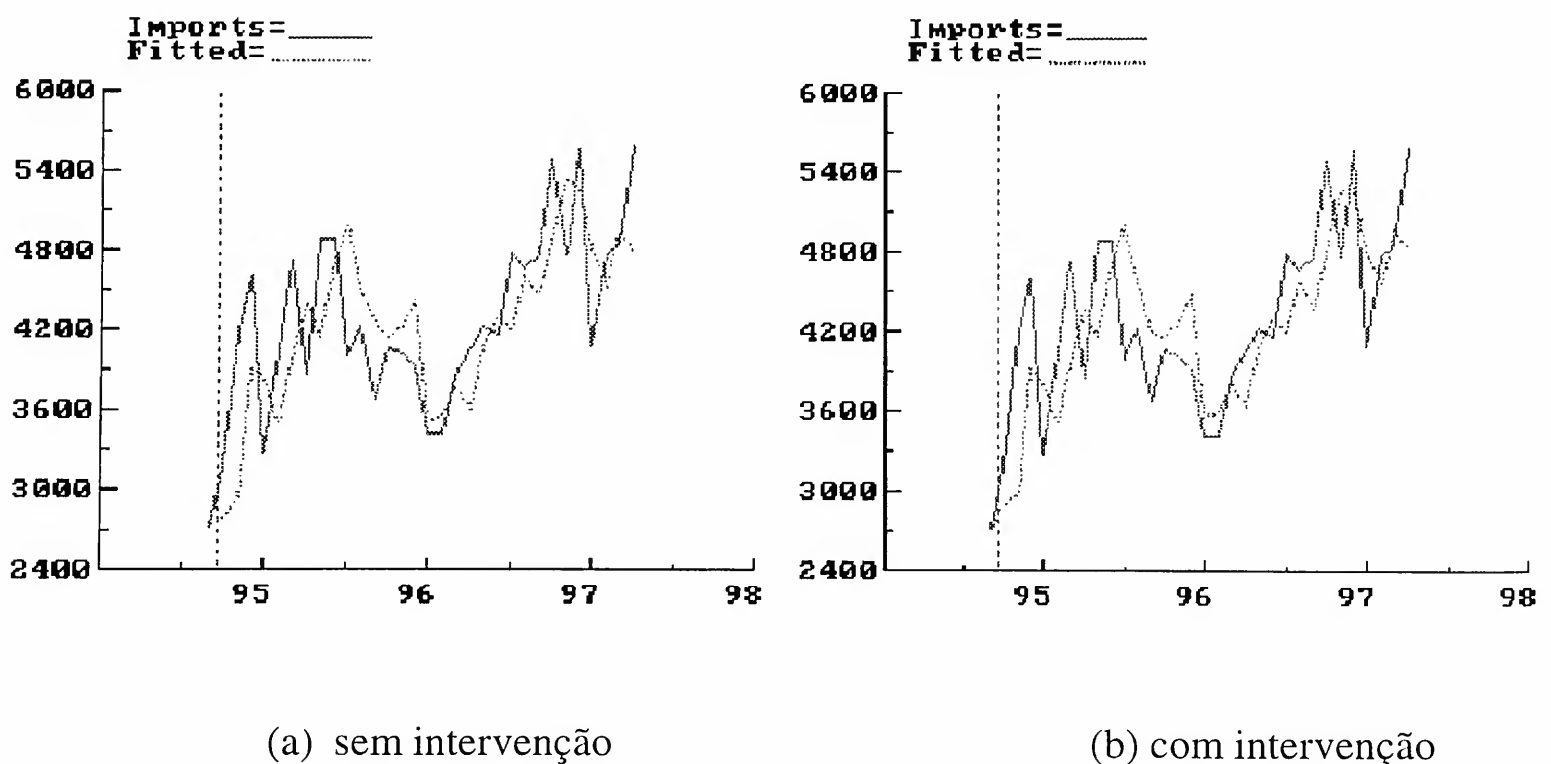
(2) Os valores entre parênteses indicam a estatística t dos coeficientes.

Não obstante a melhora no teste de normalidade dos resíduos, como ocorrera com as séries geradas por processos ARIMA, esta não se deu na mesma dimensão daquelas, no sentido em que o modelo especificado não é capaz de satisfazer este pressuposto básico.

Em particular, isto se deve ao comportamento bastante irregular da série de importações, observado a partir do início da década de 90. Nos Gráficos 4 apresentamos os componentes sazonal 4(b), irregular 4(a) e de tendência 4(c) para a série de importações.

Claramente a decomposição apresentada pelos modelos estruturais nos permite visualizar as mudanças ocorridas na série de importações, com destaque para o forte aumento da sazonalidade da série a partir da década de noventa, atingindo um patamar de 500 milhões de dólares, e a abrupta mudança no nível da série ocorrida em meados da década de 90 para algo em torno de 4 bilhões de dólares. Estas mudanças na série de importações estão intimamente ligadas à liberalização comercial que foi deflagrada a partir dos anos 90. O controle quantitativo das importação implicava não apenas em uma redução artificial do nível, mas também das características sazonais da série.

Gráfico 3



O reflexo destas alterações encontra-se na distribuição de freqüência dos resíduos do modelo especificado, onde observamos a ocorrência de valores em módulo muito distantes da média esperada destes com destacada freqüência, violando o pressuposto de normalidade dos mesmos. A título de exemplo apresentamos no Gráfico 4 as distribuições de freqüência dos resíduos para o modelo em questão 4(e), e para a série AR1QN 4(d), a qual apresentou resíduos seguindo, aproximadamente, uma distribuição normal.

Tabela 4

Variável	EAM / EPAM ¹ Redes Neurais	EAM/EPAM Modelos Estruturais	Período Observado	Configuração Básica da Rede ²
AR2QN	1,0111 / 4.32%	0.9911 / 4.28%	78/10 - 97/12	y_{t-1} y_{t-2} , <i>dummy</i> , 4 neurônios na camada oculta
AR1QT	0.9910 / 1.70%	0.9313 / 1.57%	77/12 - 97/12	y_{t-1} , <i>dummy</i> , 4 neurônios na camada oculta
AR1QN	0.8685 / 5.48%	0.9354 / 5.79%	77/12 - 97/12	y_{t-1} , <i>step dummy</i> , 5 neurônios na camada oculta
AR2QT	1.1003 / 3.19%	0.9287 / 2.87%	78/07 - 97/12	y_{t-1} , tendência, <i>staircase dummy</i> , 3 neurônios na camada oculta
Importações	544.236 / 11.87%	457.8173 / 10.29%	74/10 - 97/04	y_{t-1} , tendência, <i>step dummy</i> , <i>dummy</i> Sazonais (11), 3 neurônios na camada oculta
Inflação	2.0289 / 60.85%	2.5246 / 110.97%	86/04 - 87/04	y_{t-1} y_{t-2} , <i>dummy</i> para o período 86/3-86/12
Média dos EPAM/EQMP	14.57%	22.63%		
Média dos EPAM/EQMP sem o IGP-DI	5.31%	4,96%		

Notas: (1) As abreviações correspondem, respectivamente, ao Erro Absoluto Médio e ao Erro Porcentual Absoluto Médio.

(2) As variáveis *dummy* inseridas no processo de treinamento e geração de resultados na rede seguem os resultados obtidos com os modelos estruturais. A função de ativação utilizada foi sempre uma logística.

Gráfico 4

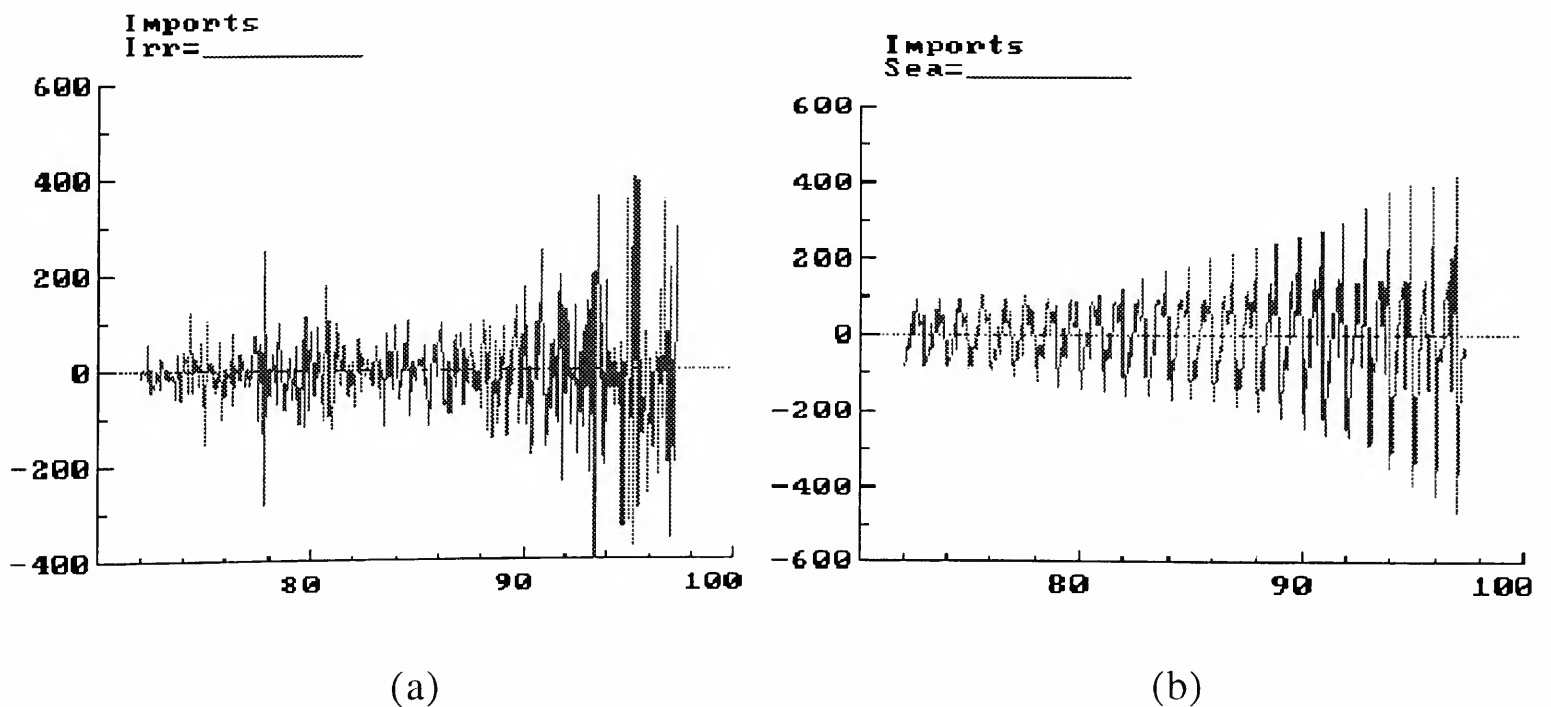
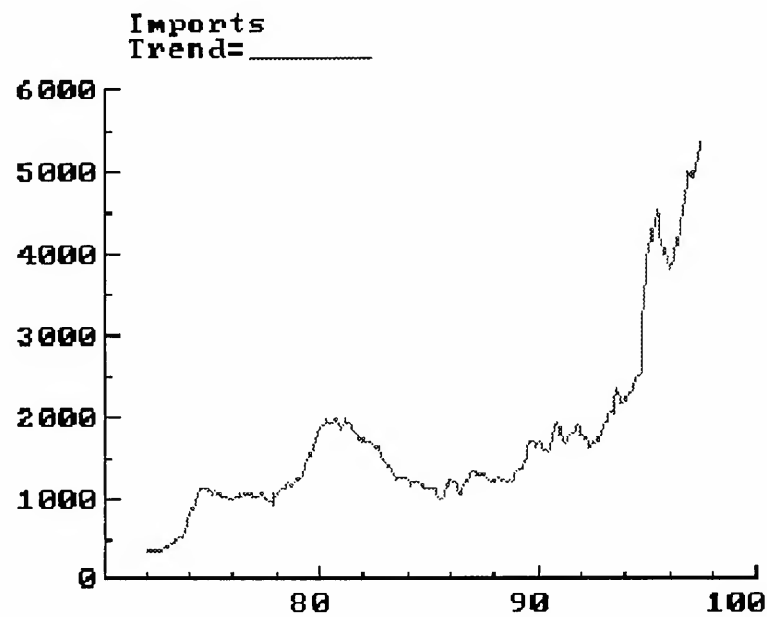
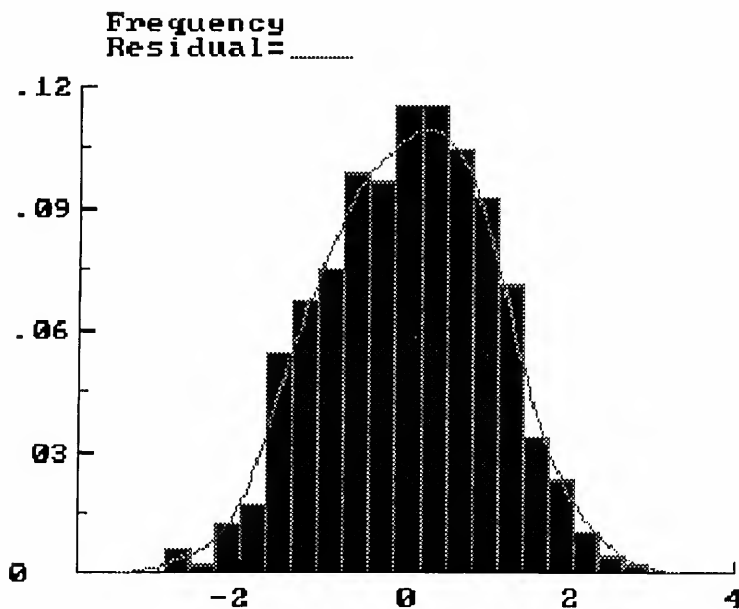


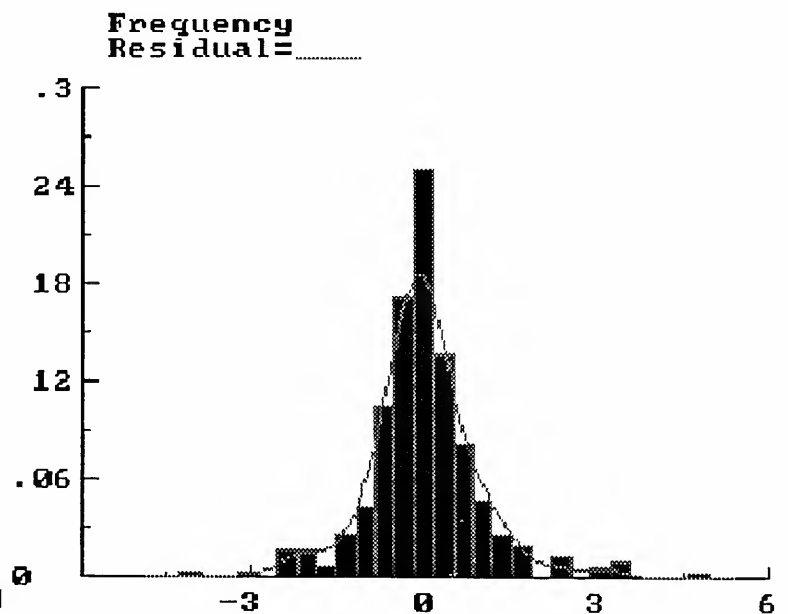
Gráfico 4 (Continuação)



(c)



(d)



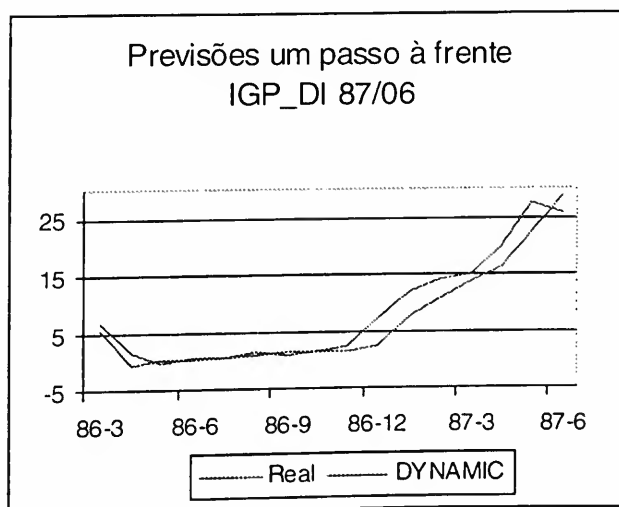
(e)

Apesar das irregularidades presentes na série, principalmente no que diz respeito ao nível da mesma, que apresenta um significativo aumento a partir de meados de 1995, a especificação de uma nova intervenção no nível em 1995/04 não resultou em melhoras nos testes de especificação, sendo que as melhoras obtidas nos critérios de ajuste se mostraram marginais. Desta forma podemos dizer que a esta nova especificação do modelo estrutural básico com intervenção mostrou-se inócua.

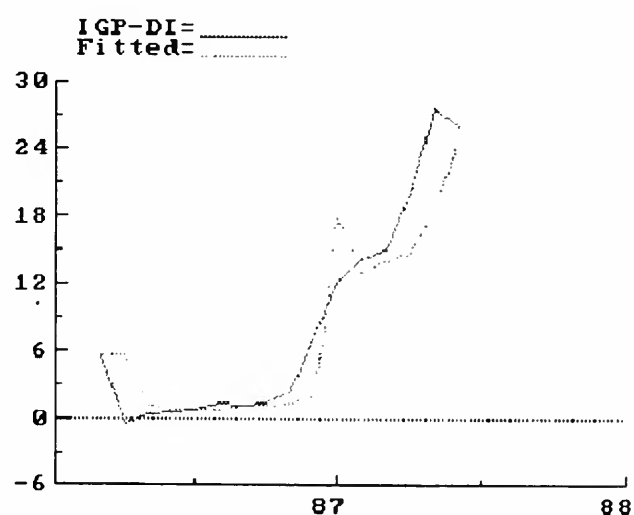
Para a série de inflação observamos um movimento distinto com a especificação do modelo com intervenção, representado pela variável *dummy* no período de março a dezembro de 1986. Neste caso observou-se uma piora na especificação dos modelos em termos dos testes de normalidade, contrapondo-se à sensível melhora observada no critério de ajuste, o qual, por sua vez, indica a rejeição do modelo inicial sem intervenção em favor do modelo com intervenção, seguindo os critérios de parcimônia, como discutido anteriormente. Novamente observamos uma perda em termos de EAM para o modelo especificado com intervenção; contudo, esta se mostrou justificável devido aos resultados obtidos com o critério de ajuste.

No que concerne à avaliação dos resultados obtidos pelas RNAs, que são modelos não-lineares com estimadores não paramétricos, e aqueles gerados pelo MEST efetuamos uma comparação com base no EPAM. Os resultados desta comparação são apresentados na Tabela 4, juntamente com a arquitetura ótima obtida para a rede neural. Levou-se em consideração as previsões um passo à frente realizadas pelas diversas configurações das redes para as mesmas variáveis e para as mesmas datas de início da geração de resultados. Devemos ressaltar que as comparações entre as RNAs e os MEST devem se limitar à capacidade preditiva de cada modelo, uma vez que para as RNAs não existe um conjunto de testes de especificação a ser utilizado para se obter a parametrização ótima do modelo. Neste sentido, para obtermos a arquitetura ótima da rede podemos utilizar características observadas a partir da estimação dos modelos estruturais, bem como do DGP utilizado para a geração das séries seguindo processos ARIMA. Outra característica que deve ser ressaltada é a de que as RNAs são modelos não-lineares, em contraposição à linearidade imposta pelo filtro de Kalman, o que não permite, portanto, maiores comparações entre as especificações dos modelos presentes em um ou outro.

Gráfico 5



(a) RNA



(b) MEST

É importante considerar ainda que para as RNAs a previsão um passo à frente é caracterizada pelo treinamento da rede na totalidade do período amostral, o que permite que esta reconheça o padrão de mudanças no DGP. Também para os modelos estruturais, as previsões são baseadas na totalidade do período amostral para o algoritmo, com a estimação do vetor de estados realizada por meio do processo de alisamento (*smoothing*). Para os MEST as previsões um passo à frente são obtidas a partir do ponto determinado de quebra com a incorporação de novas observações a cada ponto no tempo. Este procedimento possui semelhanças com o processo de treinamento da rede, o qual é realizado para todo o período amostral, e a partir da fixação dos pesos atribuídos para cada sinapse procede-se a geração do vetor de resultados (camada de saída), com a incorporação da nova informação ocorrendo com a inclusão de uma observação a mais no vetor de entradas da rede (camada de entrada).

Tabela 5
Erros Absoluto e Porcentual Médio Posteriores à Quebra (3 períodos)

Variável	EAM / EPAM	EAM / EPAM	Períodos	
	Redes Neurais	Modelos Estruturais		
AR2QN	1.8265 / 16.47%	2.8846 / 26.68%	78/10	78/12
AR1QT	0.7781 / 1.64%	0.8520 / 1.80%	78/07	78/09
AR1QN	2.1907 / 25.36%	0.8186 / 14.28%	77/12	78/02
AR2QT	1.3980 / 5.17%	1.6555 / 5.84%	78/07	78/09
Importações	596.2881 / 14.66%	735.9333 / 17.82%	94/10	94/12
Inflação	0.8853 / 173.51%	2.3763 / 448.98%	86/04	86/06
Média dos EPAM	39.47%	85.90%		
Média dos EPAM sem o IGP-DI	12.66%	13.28%		

Os resultados apresentados na Tabela 4 indicam que a diferença medida em termos de erros percentuais entre as RNAs e os MEST não se mostrou significativa, com exceção da série de inflação, o que nos encorajou a detalhar mais a capacidade de previsão de ambos os modelos a partir da ocorrência da quebra. Para isto apresentamos nas Tabelas 5 e 6 o EAM e o Erro Porcentual Absoluto Médio (EPAM) para, respectivamente, as três e seis primeiras observações posteriores à data de ocorrência da quebra. Levando-se em consideração as três observações posteriores à quebra, observamos um desempenho superior das RNAs em cinco das seis séries estudadas. A única cujo resultado dos MEST se mostrou superior foi a série AR1QN, com uma diferença pequena, porém não desprezível, em termos percentuais. Considerando-se os resultados obtidos para as seis observações posteriores à quebra, apresentados na Tabela 6, não é mantido o comportamento anterior, com as RNAs mostrando-se superiores em apenas 3 das seis séries apresentadas. Ainda assim, para as três séries onde os MEST são superiores, AR2QT,

importações e AR1QN, os ganhos em termos percentuais se mostram marginais. Esses resultados demonstram uma melhor capacidade das RNAs em perceber alterações no DGP nos períodos imediatamente posteriores à quebra, utilizando-se da sua reconhecida capacidade de aprender por experiência e fazer generalizações com base no seu conhecimento passado.

Tabela 6
Erros Absoluto e Porcentual Médio Posteriores à Quebra (6 períodos)

Variável	EAM / EPAM Redes Neurais	EAM / EPAM Modelos Estruturais	Período
AR2QN	2.2931 / 17.72%	2.5411 / 20.95%	78/10 - 79/03
AR1QT	0.9513 / 1.67%	1.069 / 2.18%	78/07 78/12
AR1QN	1.6344 / 16.33%	1.5100 / 15.14%	77/12 - 78/05
AR2QT	1.3905 / 4.86%	1.3504 / 4.76%	78/07 78/12
Importações	702.3402 / 17.80%	673.0333 / 16.15%	94/10 - 95/03
Inflação	0.6207 / 102.01%	1.3169 / 234.72%	86/04 - 86/09
Média dos EPAM	26.73%	48.98%	
Média dos EPAM sem o IGP-DI	11.67%	11.83%	

Em particular, para a série de inflação observamos um EPAM extremamente alto, tanto para as previsões dos três primeiros períodos quanto para aquelas geradas para os seis períodos após a quebra. Estes valores altos se deveram à ocorrência de uma deflação em abril de 1986, a qual representou uma queda muito abrupta no valor da série, visto que em fevereiro de 1986 a inflação alcançava os 14%. Os resultados obtidos demonstram uma evidente incapacidade de ambos os modelos em conseguir prever este tipo de alteração no DGP. Tal fato é representado pela existência de um erro de previsão extremamente alto para abril de 1986, como pode ser visualizado no Gráfico 5, onde são apresentadas as previsões um passo à frente para a série de inflação para as RNAs e para os MEST.

Quando observamos apenas a média dos resultados obtidos pelos modelos em consideração nas Tabelas 5 e 6, notamos uma sensível superioridade das RNAs em relação aos MEST, independente do período analisado. Contudo, este resultado está enviesado pela série de inflação, onde a ocorrência da deflação mencionada acima reduz severamente a capacidade de previsão de ambos os modelos. Ao excluirmos a série de inflação os resultados mostram uma equivalência entre as RNAs e os MEST para a previsão na totalidade do período, bem como para os períodos imediatamente posteriores à ocorrência da quebra, como exposto nas Tabelas 4, 5 e 6. Para o período de previsão como um todo os MEST mostram-se marginalmente superiores, enquanto que para as previsões próximas do momento da quebra são as RNAs que se mostram marginalmente superiores.

5 Conclusão

Neste trabalho procurou-se investigar o comportamento dos modelos estruturais e das RNAs com relação à sua capacidade em reconhecer mudanças estruturais no DGP

Os MEST conseguiram realizar previsões de maneira eficiente, ainda que desconsiderando o processo de intervenção, resultando em uma independência entre capacidade de previsão e intervenções nos modelos especificados. Em outras palavras, o processo de atualização do vetor de estados a cada período de tempo, que é realizado pelo filtro de Kalman, foi capaz de ajustar muito rapidamente o modelo à quebra estrutural sem a necessidade de uma intervenção subjetiva. Contudo, esta independência presente nos MEST tem como resultado, para todas as variáveis analisadas, uma sensível perda em termos de má especificação do modelo, o que desencoraja a utilização dos MEST sem intervenção em um estudo comparativo com outros modelos, ainda que não-paramétricos.

Considerando-se as séries de maneira isolada, nossos resultados indicam ainda uma leve superioridade na capacidade de previsão e, portanto, reconhecimento da alteração no DGP, por parte das RNAs nos momentos imediatamente posteriores à ocorrência da quebra estrutural. No momento da quebra as RNAs geram melhores previsões para 5 das 6 séries utilizadas. Contudo, este resultado não se mantém quando estendemos as previsões para a totalidade do período, onde os ganhos obtidos pelos MEST em relação às redes é marginal. Este resultado é válido para todas as séries analisadas neste trabalho, à exceção da série de inflação. Nesta última a ocorrência de uma deflação em abril de 1986 produz uma quebra estrutural no nível da série de tal ordem que ambos os modelos se mostraram incapazes de percebê-la, resultando em uma perda não desprezível na capacidade de previsão. Apesar deste fenômeno, as RNAs se mostraram capazes de se ajustar à ocorrência da quebra estrutural na série de inflação, de maneira muito superior aos MEST. Neste sentido, podemos concluir que as RNAs conseguem identificar mais rapidamente as quebras estruturais.

Além dos MEST e das redes neurais, a questão da mudança estrutural poderia ser também modelada por meio de mudanças de regime por cadeias de Markov (Markov Switching). Uma apresentação desta metodologia com exemplos ilustrativos pode ser encontrada em Hamilton (1994, cap. 22).

Por fim, é preciso considerar uma limitação da análise presente neste trabalho com relação ao número de séries modeladas. A extensão das conclusões obtidas exigiria a realização de um experimento de Monte Carlo. Contudo, a realização deste tipo de

experimento foge ao escopo inicial deste trabalho, que procura incluir em sua análise, não somente séries construídas, mas também séries observadas. A realização de um experimento de Monte Carlo fica, portanto, como uma sugestão para um trabalho posterior.

Referências bibliográficas

- Alexander, I. and Morton, H. *An introduction to neural computing*. London: Chapman and Hall, 1990.
- Atkinson, A. C., Koopman, S. J. and Shephard, N. Detecting shocks: outliers and breaks in time series. *Journal of Econometrics*, 80, p. 387-422, 1997
- Box, G. E. P. & Jenkins, G. M. *Time series analysis, forecasting and control*. San Francisco: Holden-Day, 1976.
- De Jong, P. Smoothing and interpolation with the state space model. *Journal of the American Statistical Association*, 84, 1085-1088, 1989.
- Diderrich, G. T. The Kalman filter from the perspective of Goldberg-Theil estimators. *The American Statistician*, 39, 193-198, 1985.
- Duncan, D. B. and Horn, S. D. Linear dynamic estimation from the viewpoint of regression analysis. *Journal of the American Statistical Association*, 67, 815-821, 1972.
- Fernandes, L. G. L., Navaux, P. O. A. & Portugal, M. S. Previsão de séries tempo: redes neurais e modelos estruturais. *Pesquisa e Planejamento Econômico*, v. 26, n. 2, p. 253-276, 1996.
- Hamilton, J. D. *Time series analysis*. Princeton: Princeton University Press, 1994.
- Harvey, A. C. Applications of the Kalman filter in econometrics. In: Ewley, T. F. (ed.), *Advances in econometrics Fifth World Congress*, v. I, Cambridge: Cambridge University Press, 1987
- _____. *Forecasting structural time series models and the Kalman filter*. Cambridge: Cambridge University Press, 1989.
- Harvey, A. C. and Koopman, S. J. Diagnostic checking of unobserved componentes time series models. *Journal of Business and Economic Statistics*, 10, 377-389, 1992.
- Haykin, S. *Neural networks: a comprehensive foundation*. Nova York: Macmillan, 1994.

- Kalman, R. E. A new approach to linear filtering and prediction problems. Transactions of the ASME, series D. *Journal of Basic Engineering*, 35-45, 1960.
- Koopman, S. J, Harvey, A. C., Doornick, J. A. and Shephard, N. *STAMP 5.0: structural time series analyser, modeller and predictor*. Londres: Chapman & Hall, 1995.
- Meinhold, R. J. and Singpurwalla, N. D. Understanding the Kalman filter. *The American Statistician*, 37, 123-127, 1983.
- Portugal, M. S. Modelos de parâmetros variáveis: uma resenha crítica. *Pesquisa e Planejamento Econômico*, v. 23, n. 1, p. 99-134, 1993.
- _____ Neural networks versus time series methods: a forecasting exercise. *Revista Brasileira de Economia*, v. 49, n. 4, p. 611-629, 1995.
- Portugal M. S. & Fernandes L. G. L. Redes neurais artificiais e previsão de séries econômicas: uma introdução. *Nova Economia*, v. 6, n. 1, p. 51 74, 1996.
- Refenes, A. N. Constructive learning and its application to currency exchange rate Forecasting. In: Turban, E. and Trippi, R. (eds.), *Neural networks applications in investment and finance services*. New York: Probus Publishing, 1991.
- Rumelhart, D. E., Hinton, G. E. and Williams, R. J. Learning internal representations by error propagation. *Parallel Distributed Processing*, Cambridge: MIT Press, v. I, p. 318-362, 1986.
- Souza, R. C. and Zandonade, E. Forecasting via neural networks: a comparative study. PUC-RJ: Departamento de Engenharia Elétrica, 1993, *mimeo*.
- Wasserman, P. D. *Neural computing: theory and practice*. New York: Van Nostrand Reinhold, 1989.
- West, M. and Harrison, J. *Bayesian forecasting and dynamic models*. 2ª ed. New York: Springer Verlag, 1997.

Pluriatividade e ruralidade: aspectos metodológicos[§]

Angela Kageyama[†]

RESUMO

O artigo resume algumas visões recentes sobre os conceitos de pluriatividade e ruralidade a partir de uma revisão bibliográfica sobre o tema, enfocando basicamente seus aspectos metodológicos. Dois conceitos chaves são destacados: o de economia local e o de inserção das unidades econômicas (famílias ou empresas) nessas economias. Ao final são feitas sugestões para aplicação dos conceitos em pesquisa empírica, ressaltando as limitações e dificuldades envolvidas.

Palavras-chave: agricultura, pluriatividade, economias locais.

ABSTRACT

The paper summarizes some recent visions about the concepts of pluriactivity and rurality found in the literature, focusing mainly their methodological aspects. Two key concepts are emphasized: local economy and insertion of the economic units (families or households) in those economies. Some suggestions are made aiming at the use of the concepts in empirical research, taking into consideration the limitations and difficulties involved.

Key words: agriculture, pluriactivity, local economies.

[§] Trabalho desenvolvido com Bolsa de Pesquisa do CNPq para o Projeto Rurbano, representando uma parte da pesquisa bibliográfica desse projeto. Em continuidade estão sendo realizadas pesquisas empíricas em diversos estados do Brasil, seguindo os passos metodológicos gerais sugeridos pela revisão da literatura. Para maiores detalhes acessar <http://www.eco.unicamp.br/projetos/rurbano>.

A autora agradece os comentários e sugestões de José Graziano da Silva e Rodolfo Hoffmann a uma versão preliminar do trabalho. Agradece também as sugestões de um *referee* da revista.

[†] Professora do Instituto de Economia da Universidade Estadual de Campinas

1 Dificuldades preliminares: definições

Os dois conceitos que abrem o título deste artigo são largamente utilizados. Servem de base não apenas a classificações e descrições empíricas, mas também à formulação e implementação de políticas, orientam institutos nacionais de estatísticas em sua coleta de dados, dão origem a múltiplas teses de mestrado e doutorado e, no entanto, sua definição nem sempre é precisa ou uniforme.

Há porém um certo consenso de que, na maioria dos países desenvolvidos e em alguns Estados do Brasil (São Paulo, Paraná, Santa Catarina), há uma tendência de crescimento da importância da pluriatividade para as famílias e regiões antes centradas na agricultura, ao lado de uma redução dos territórios antes classificados como eminentemente ou exclusivamente rurais, ainda que não se utilize uma definição universal desses termos.

Nossa primeira tarefa será recolher da bibliografia selecionada as principais definições daqueles dois termos, para posteriormente refletirmos melhor sobre seu significado analítico.

1.1 Tempo parcial ou pluriatividade?

O fenômeno da redução do peso das atividades agrícolas no emprego e na renda das pessoas, famílias e regiões “rurais”, dando lugar aos “empregos múltiplos” e fontes de renda diversificadas, tem sido referido na literatura, às vezes sem muita distinção, como pluriatividade ou como agricultura (e agricultores) de tempo parcial.

No capítulo sobre emprego do Relatório de 1992 da Comissão das Comunidades Europeias sobre a situação da agricultura na Comunidade as estatísticas distinguem claramente os trabalhadores (empresários, familiares ou empregados) a **tempo parcial** (isto é, pessoas que trabalham apenas uma fração do tempo de trabalho anual de uma pessoa empregada “a tempo inteiro”) e em outra classificação os **pluriativos** (isto é, pessoas que possuem “outra atividade lucrativa” além da atividade principal). Em 1987, por exemplo, para o conjunto da Europa dos 12 (exceto Dinamarca) havia 8,4% de pluriativos entre os empresários agrícolas¹ ocupados em tempo integral, 33% de pluriativos entre os

1 Foram considerados apenas os empresários que são simultaneamente os próprios chefes das explorações agrícolas, responsáveis pela sua gestão corrente e cotidiana.

de tempo parcial tipo 1 (de 50% a menos de 100% do tempo de trabalho completo) e 43% de pluriativos entre os de tempo parcial de tipo 2 (menos de 50% do tempo completo).

Por esses dados fica evidente que os dois fenômenos não podem ser totalmente superpostos: o tempo parcial numa atividade (agricultura, por exemplo) é condição necessária para poder dedicar-se a outras atividades, mas não é suficiente e nem é sinônimo de pluriatividade. Mais ainda: é preciso esclarecer se o tempo parcial se refere a uma atividade (“agricultura de tempo parcial”) ou ao número total de horas trabalhadas pela pessoa. O primeiro aspecto é o que parece estar mais diretamente relacionado com a pluriatividade, enquanto o segundo (jornada de trabalho) seria mais pertinente a análises do subemprego ou subocupação. Exemplificando, pode-se ter um indivíduo pluriativo que trabalha apenas metade da jornada normal (ele seria pluriativo e tempo parcial), assim como uma pessoa ocupada em tempo integral que pratica agricultura em tempo parcial porque se dedica a outras atividades (urbanas, por exemplo) no resto do tempo (ele seria um pluriativo, com ocupação em tempo integral e praticando agricultura de tempo parcial). Esses casos, se aparentemente esdrúxulos quando precisamente formulados, não são incomuns nas áreas rurais mais dinâmicas. Mas, ao contrário das estatísticas, nem sempre os autores que analisam a pluriatividade distinguem claramente os aspectos acima indicados.

Existe, de fato, um longo debate histórico acerca desses termos. Na ótima revisão empreendida por Fuller (1984) podemos acompanhar as diferentes conotações e conceituações que a expressão “agricultura de tempo parcial” (*part-time farming*, termo cunhado por Rozman em 1930) foi adquirindo desde os anos 30 até meados dos anos 80, em função da importância econômica e política atribuída ao fenômeno nos diferentes países da América, Europa e Japão. Em outro trabalho (Fuller, 1990) pode-se acompanhar a passagem do termo “agricultura de tempo parcial” para “empregos múltiplos” (*multiple job-holding*, termo usado inicialmente por Fuguitt nos anos 50-60) e, finalmente, para “pluriatividade” (*pluriactivity*, expressão típica do fim dos anos 80 na Europa, no contexto das reformas da PAC).

Resumindo, até os anos 50, tanto nos Estados Unidos como na Europa a agricultura de tempo parcial tinha como foco de análise o produtor - geralmente à frente de um pequeno estabelecimento -, as fontes de renda e com uma abordagem geralmente descritiva e regional. No final dessa década, porém, dois elementos decisivos foram atribuídos ao conceito: primeiro, a idéia de que a unidade de análise pertinente no caso é a **família**, e não o “operador” (*farm operator*) porque a família é a instância no âmbito da qual é tomada a decisão de praticar agricultura no todo ou em parte do tempo, uma vez que essa decisão

depende do conjunto dos recursos disponíveis e de escolhas e necessidades do grupo familiar; o segundo elemento novo surgido no período, principalmente no nível das estatísticas europeias, foi a incorporação do **tempo de trabalho** (em número de dias ou de equivalentes de força de trabalho por ano) como medida do grau de atividade fora da fazenda que serviria para identificar uma família como de tempo parcial ou não. Anteriormente, apenas o critério de renda servia para esse fim.

Na década seguinte os estudos empíricos sobre diferentes tipologias da agricultura de tempo parcial dominaram o debate, sendo dessa época algumas importantes contribuições envolvendo o novo termo “*multiple job-holding*”, numa tentativa de dirigir o foco prioritariamente para fora da fazenda, para as atividades não-agrícolas, embora ainda o objetivo central fosse o seu efeito sobre o desempenho da propriedade agrícola.

De meados da década de 70 até os primeiros anos da década de 80 uma série de seminários internacionais,² além de diversos livros e números especiais de periódicos (ver Fuller, 1984) ensejaram um verdadeiro renascimento da pesquisa e do debate sobre a agricultura de tempo parcial, ficando bem estabelecidos dois aspectos: que a unidade de análise relevante no caso é a família ou, para alguns autores, o estabelecimento familiar (*household*) e que a agricultura de tempo parcial não é um fenômeno temporário nem de transição no desenvolvimento agrícola, constituindo, ao contrário, uma forma bem definida e persistente de relacionamento intersetorial em muitos países.

Fuller (1984) propõe, a partir de seu estudo histórico sobre o conceito de agricultura de tempo parcial, as seguintes distinções:

- * **agricultor de tempo parcial** : pessoa que trabalha parte do seu tempo na agricultura;
- * **unidade agrícola (estabelecimento) de tempo parcial**: estabelecimento que, com o nível corrente de recursos, não consegue gerar uma demanda de trabalho correspondente ao trabalho completo de um ano (para seus membros);
- * **unidade agrícola de tempo integral**: aquela em que nenhum membro da família tem emprego remunerado fora;

2 Guelph Symposium, Canadá, 1975; Wye College Workshop, 1976 e Seminar, 1977, Inglaterra; Mixed Households, Yugoslavia, 1981.

* **estabelecimentos/famílias com empregos múltiplos (*multiple job-holding households*)**: em que existem rendas agrícolas e não-agrícolas, podendo a renda externa superar ou não a renda agrícola obtida no estabelecimento. Este último conceito enfatiza a natureza multissetorial, a pluralidade de atividades econômicas das famílias agrícolas, que irá desembocar na idéia de pluriatividade.

O conceito de MJHFH (*multiple job-holding farm households*), embora não novo na época, foi a base do programa de pesquisas do Arkleton Trust³ nos anos 80 e procurava, de um lado, remover a ênfase no setor agrícola e, de outro, isentar o termo de conotações pejorativas ligadas à idéia de “tempo parcial”.(Fuller, 1990) A pesquisa colocou em relevo a importância de outros fatores além da lógica interna da unidade familiar, tais como: composição e ciclo vital da família, presença de mercados de trabalho não-agrícolas e diversas características do contexto regional e local, inclusive políticas, na conformação e evolução dos empregos múltiplos (MJH). Estes, por definição, referiam-se especificamente a atividades remuneradas em base contratual convencional (*gainful employment*).

No final dos anos 80, a ampliação do conceito visando incluir atividades ou trabalhos não necessariamente remunerados em dinheiro levou a pesquisa do Arkleton Trust à adoção do termo **pluriatividade**, que pode incluir: emprego em outros estabelecimentos agrícolas, p. ex. trabalho assalariado; atividades “paraagrícolas” como alimentos e bebidas processados; atividades não-agrícolas no estabelecimento, p. ex. turismo e alojamento; atividades externas não-agrícolas, p. ex. assalariamento.(Fuller, 1990, p. 367)

Nas palavras do autor:

“Pluriatividade descreve, assim, uma unidade produtiva multidimensional, em que se empreendem atividades agrícolas e não-agrícolas dentro e fora do estabelecimento, e pelas quais diferentes tipos de remuneração são recebidos (rendimentos, rendas em espécie e transferências).”(Fuller, 1990, p. 367)⁴

3 Esse programa de pesquisa foi um marco no estudo da pluriatividade, cobrindo 24 regiões da Europa, com *surveys* em 1987 e 1991 em uma amostra estratificada de 300 estabelecimentos e painéis com 70 estabelecimentos nos anos intermediários. No *survey* levado a efeito em 1987 o projeto adotou como definição operacional de estabelecimentos pluriativos aqueles em que algum membro adulto tivesse uma atividade remunerada sazonal ou regular além da atividade primária agrícola.

4 Para tornar a leitura mais fluente, esta e outras citações foram traduzidas de sua língua original pela autora.

Na visão de Fuller, portanto, há uma evolução que vai do conceito de agricultura de tempo parcial para empregos múltiplos (MJH) e, finalmente, para pluriatividade, no sentido de ampliar o peso dado às atividades e condições externas à agricultura, bem como às relações que se estabelecem entre as unidades produtivas e os contextos regionais, especialmente os mercados de trabalho. É no bojo dessas discussões que há uma “redescoberta”, segundo Fuller, do desenvolvimento local e das comunidades.

Do fim dos anos oitenta até os primeiros anos da década de 90 seguiram-se vários trabalhos conceituais e metodológicos sobre tempo parcial e pluriatividade na agricultura. Uma discussão específica sobre definições encontra-se no *Journal of Agricultural Economics* (Inglaterra) do período. Lund (1991), por exemplo, criticando o uso do critério de “**outro trabalho remunerado**” na definição de “agricultura em tempo parcial” “agricultor em tempo parcial” e “família agrícola em tempo parcial” (tal como em Gasson, 1988, cit. por Lund), propõe que inicialmente se classifiquem os **indivíduos em tempo parcial** em oposição aos de tempo integral segundo os seguintes critérios:

- (a) trabalham menos do que uma jornada integral estipulada, no estabelecimento agrícola;
- (b) possuem outra(s) ocupação(ões) remunerada(s);
- (c) possuem alguma outra ocupação remunerada e a agricultura não é a principal atividade, seja em termos de tempo de trabalho ou de renda.

Para Lund o termo “agricultor em tempo parcial” deve ser reservado para os que preenchem apenas o primeiro critério (a), enquanto “pluriativo” “múltiplo emprego” ou “mais de uma ocupação remunerada” devem ser usados para identificar os que preenchem os critérios (b) e (c). Estes dois últimos critérios, desde logo, podem ser usados para unidades como família e estabelecimento, enquanto a idéia de *part-time* só se aplica em referência às práticas laborais individuais.

Gasson (1991), em réplica à proposta de Lund, concorda que

“Parece de fato racional descrever agricultores que têm outras ocupações como ‘dual-job ou multiple-job holders’ e reservar o termo ‘part-time’ para os agricultores que trabalham menos do que a jornada integral estipulada na agricultura.”(Gasson, 1991, p. 200)

Admite que deve concordar que *part-time* não é a melhor forma de descrever uma **atividade** (p. 201), mas mesmo assim prefere manter o conceito de *part-time farming* sob

o argumento de que para as políticas agrícolas de sustentação das unidades agrícolas nas áreas rurais é importante dispor de uma definição que contenha a idéia de combinação de diferentes fontes de renda e sua contribuição para um “fundo comum” de sustento da família, sejam essas fontes provenientes de pessoas da família, de outros membros do domicílio ou mesmo de parceiros de negócios.

Nos trabalhos ligados à pesquisa do grupo Arkleton muitas vezes pluriatividade e tempo parcial aparecem indistintamente como sinônimos. Por exemplo, Newby (1987, p. 157), assim se expressa:

“(...) a exploração agrária pluriativa (isto é, a exploração agrária a tempo parcial) ...”

Na mesma coletânea de pesquisas vemos a opinião de Barthez (1987) indo em direção oposta: para essa autora a pluriatividade pode ser uma consequência da atividade (especificamente a agricultura) em tempo parcial, mas não são a mesma coisa. A população agrícola ativa a tempo parcial define-se pela não utilização de todo o seu tempo de trabalho na atividade agrária (tomou-se como referência a jornada semanal de 39 horas). Coloca-se então o problema da utilização do tempo restante, que pode ser a **inatividade** ou sua alocação numa **atividade distinta**, geralmente externa. Só neste último caso configura-se a pluriatividade. Mas nem isto é assim tão simples porque, com seu desenvolvimento, a “atividade” passa a ser um atributo pessoal, e na família pluriativa cada indivíduo tende a exercer uma única atividade. Isto parece ser mais evidente no caso das esposas de agricultores, as quais, encontrando uma nova atividade fora da exploração agrícola, tendem a abandonar a “dupla atividade” e a dedicar-se em tempo integral à atividade externa.(Barthez, 1987. p. 172)

No Brasil, veja-se, por exemplo, o artigo de Marcondes (1962), que estudando as relações entre a industrialização e o surgimento da agricultura em tempo parcial em São Paulo define seu objeto em termos de duas variáveis: tempo aplicado na atividade e existência de outras fontes de renda:

“Agricultor em tempo parcial é todo aquele que sistematicamente emprega parte do seu tempo e do de sua família nas lides agrícolas, auferindo alguma renda, quase sempre inferior àquela que regularmente retira de uma outra atividade profissional ou econômica.”(Marcondes, 1962, p. 31)

A principal diferença entre esse trabalho e os outros aqui citados é que ele enfoca centralmente o indivíduo urbano que, com a tecnificação industrial ou a aposentadoria, passa a dispor de tempo livre para se ocupar da agricultura como *hobby* ou com finalidade econômica, mas de qualquer modo como “segunda atividade”. Nos demais trabalhos a agricultura é vista como a atividade principal e as outras fontes de renda como complementares.

Outro trabalho mais recente (dos Anjos, 1995), embora traga em seu título a expressão “agricultura em tempo parcial” e não utilize o termo pluriatividade sequer uma vez, trata basicamente deste último fenômeno. Já em Graziano da Silva (1997a) a distinção aparece claramente: tempo parcial relaciona-se com tempo de trabalho, pluriatividade relaciona-se com diversificação de atividades, geralmente em mercados não-agrícolas:

“Pode-se dizer que as atividades agropecuárias já constituem uma atividade de tempo parcial para um contingente expressivo de trabalhadores no Brasil, não apenas porque as atividades agrícolas não demandam todo o tempo de trabalho disponível das famílias rurais, mas também porque as atividades agrícolas não geram renda suficiente para todas as pessoas ocupadas em tempo integral, que também deveriam ser consideradas parte dos subocupados.” (Graziano da Silva, 1997a, p. 53)

“A pluriatividade (...) é conseqüência desse esforço de diversificação dos pequenos produtores para se inserirem nos novos mercados locais que se abrem.” (p. 62)

Mas em outro trabalho do mesmo autor - Graziano da Silva (1997b, p. 77-78) - confundem-se os dois conceitos:

*“No mundo rural dos países desenvolvidos esse novo paradigma ‘pós-industrial’ tem um ator social já consolidado: o **part-time farmer**, que podemos traduzir por agricultores em tempo parcial. A sua característica fundamental é que ele não é mais somente um agricultor ou pecuarista: ele combina atividades agropecuárias com outras atividades não-agrícolas, dentro ou fora de seu estabelecimento (...). Em resumo, o **part-time** não é mais um fazendeiro especializado, mas um trabalhador autônomo que combina diversas formas de ocupação (assalariadas ou não). Essa é a sua característica nova: uma pluriatividade que combina atividades agrícolas e não-agrícolas.”*

Parece claro que a idéia de pluriatividade diz respeito ao desempenho simultâneo de atividades econômicas diversificadas por um indivíduo ou uma família, como apontam Le Heron *et alii* (1994):

*“No fim dos anos 80 o debate refletiu a visão de que os desenvolvimentos contemporâneos na produção agrícola e nas interações entre os agronegócios, os estabelecimentos e as famílias ligados à agricultura deveriam ser teorizados, primeiramente, como parte da evolução mais ampla do capitalismo e, segundo, como parte de uma rede multissetorial de atividades que iam muito além da fazenda. Desta perspectiva, o fenômeno da **pluriatividade** ou múltiplos empregos e múltiplas fontes de renda é visto como parte integral da geografia da acumulação e da contínua mercantilização das relações sociais de produção e reprodução.”*

Para esses autores a pluriatividade é potencialmente um resultado de diversas “negociações” ou estratégias entre indivíduos que possuem diferentes posições na família e no negócio agrícola, e que ajustam suas atividades levando em conta o ciclo familiar, a organização do estabelecimento e do trabalho doméstico, os riscos envolvidos na menor dedicação à atividade agrícola, visando assegurar a continuidade da família no negócio agrícola de uma geração à seguinte. Assim, a pluriatividade deve ser entendida em contextos específicos; no caso em questão (Nova Zelândia), por exemplo, a inserção das unidades familiares em sistemas de produção de *commodities* altamente capitalizados, as reformas políticas dos anos 80, o aumento do desemprego e da participação das mulheres na força de trabalho, a redução do tamanho médio das famílias e a tendência a aumentar os gastos em ativos não-produtivos foram elementos cruciais na conformação de um contexto em que a pluriatividade estendeu-se e ganhou características específicas (grande participação de mulheres, sobretudo em ocupações profissionais qualificadas, como ensino e enfermagem; manutenção de alta dependência da renda proveniente da atividade agrícola na cadeia de *commodity* a que se alinha o estabelecimento; pluriatividade derivada mais de fatores externos do que de retornos inadequados da própria atividade agrícola).

Também Carneiro (1995, p. 50) aponta a importância do contexto na compreensão do que vem a ser a pluriatividade:

“As atividades complementares ou suplementares à produção agrícola exercidas por um ou vários membros de um grupo doméstico são reconhecidas tanto por aqueles que estudam a questão agrária (...)

quanto pelos agentes sociais nela implicados (...) através da noção de pluriatividade. (...) este termo homogeneiza fenômenos sociais distintos, com significados que variam segundo a dinâmica de reprodução das pequenas unidades produtivas.”

A autora analisa três situações históricas (numa aldeia dos Alpes franceses) para mostrar como a prática da pluriatividade preenche funções distintas, como reforço para a reprodução ou para a modernização da agricultura familiar ou como expressão do abandono da atividade agrícola.

Um outro estudo (Cabral, 1987), em duas regiões de Portugal, aponta as diferenças no papel da pluriatividade nas estratégias de abandono parcial das atividades agrícolas pelos camponeses. Num caso - o dos parceiros pobres do Noroeste - a sobrevivência depende mais do assalariamento de membros da família fora da agricultura do que da atividade agrícola, e estes se caracterizam, na verdade, como assalariados que exercem a agricultura em tempo parcial. A pluriatividade representa, para esses operários-camponeses, uma solução que evita a proletarização completa. No outro caso - o dos camponeses abastados do Oeste - o abandono parcial da agricultura se dá “pelo alto”; a combinação de atividades agrícolas com outras, sobretudo o comércio nas feiras e mercados rurais, fortalece a economia camponesa:

“A pluriatividade nos parece, assim, menos como uma etapa em direção ao abandono da agricultura do que como uma orquestração, em torno da agricultura, de uma multiplicidade de atividades - produtivas e não-produtivas para uma população rural cujos membros continuam simultaneamente estimuladores e instrumentos de estratégias de crescimento da renda global indivisa dos grupos domésticos camponeses.”(Cabral, 1987, p. 202)

A partir dessa revisão parece-nos adequado adotar uma posição próxima à de Lund (1991), estabelecendo que a pluriatividade refere-se à combinação de atividades - por indivíduos ou famílias - em diferentes setores, portanto diferentes mercados, da economia; para o que nos interessa, no entanto, restringiremos o “conceito” para o caso em que um desses setores seja a agricultura. “Conceito” está entre aspas para chamar a atenção para o fato de que nem todos os autores concordariam com essa classificação. Veja-se, por exemplo, a posição de um autor expressa num dos mais importantes programas de pesquisa sobre pluriatividade, o Arkleton Research:

“O termo ‘pluriatividade’ não é um conceito, mas apenas uma noção empírica que corresponde às situações em que a família agrícola não obtém a totalidade de suas rendas da produção agrária primária nem tampouco dedica a esta a totalidade de seus recursos.” (Blanc, 1987, p. 89)

Os trabalhos de Shucksmith (Shucksmith *et alii*, 1989 e Shucksmith & Smith, 1991) vêm a pluriatividade como uma idéia provavelmente menos profícua analiticamente (se tomada de forma isolada) do que o foco nas **estratégias familiares**, dentro das quais a diversificação (ou pluriatividade) pode desempenhar algum papel. O foco desses trabalhos é a compreensão da persistência das propriedades agrícolas familiares em certas regiões da Escócia e no Reino Unido em geral, concluindo que três fatores principais foram responsáveis por esse fenômeno: a) a multiplicação de propriedades relativamente grandes com uma queda simultânea do uso de trabalho assalariado, ensejando uma substituição de trabalho contratado por trabalho familiar; b) a alta flexibilidade das relações familiares, que possibilitou a adoção de diferentes estratégias diante das mudanças nas condições dos mercados; c) a presença de fontes de renda adicionais à agricultura, especialmente turismo e recreação (pluriatividade). Nesse contexto, o conceito relevante mais geral é o de estratégias das unidades familiares (*farm business strategies*), sendo a pluriatividade apenas um de seus possíveis componentes, uma forma potencial de incrementar as rendas. Essa proposta, a nosso ver, abre uma linha de investigação interessante porque contextualiza a pluriatividade no conjunto de decisões das empresas familiares, visando à acumulação ou, na pior das hipóteses, à sobrevivência. O que os autores pretendem é tirar o caráter “inclusivo” da categoria pluriatividade (*catch-all category*), colocando-a como uma entre outras estratégias, tais como a busca de novos mercados ou as tentativas de redução de riscos.

1.2 A definição de rural

“Rural: Uma sociedade ou comunidade pode ser classificada como rural, e não urbana, pelos seguintes critérios, que indicam áreas de um continuum: menor densidade demográfica, menor diferenciação social, menor mobilidade social e espacial; índices mais baixos de mudança social; agricultura como ocupação principal; posse da terra como o centro convergente do sistema político-econômico.” (Dicionário de Ciências Sociais, FGV, p. 1090)

“Na situação URBANA consideram-se as pessoas e os domicílios recenseados nas áreas urbanizadas ou não correspondentes às cidades (sedes municipais), às vilas (sedes distritais) ou às áreas urbanas isoladas. A situação RURAL abrangeu toda a área situada fora desses limites, inclusive os aglomerados rurais de extensão urbana, os povoados e os núcleos.”(Censo Demográfico de 1991, IBGE, p. 9)

Ressalta nessas definições, que de uma forma ou de outra reflete a maioria das definições encontradas nas estatísticas nacionais ou nos trabalhos analíticos, o caráter **relativo** do rural, como oposição, complemento ou resíduo de sua contraparte, o urbano. Mais do que isso, os elementos que definem o rural são indicadores de “desenvolvimento” (especificamente, de “menor” desenvolvimento) e, portanto, relativos também no tempo. Está-se tentando definir, por conseguinte, um objeto cambiante no espaço - porque o avanço da urbanização redefine seus limites - e no tempo - porque os elementos que o definem estão em permanente processo de mudança. Saraceno (1994a p. 468) resume bem o argumento:

“O paradoxo da ruralidade é que ela é definida em termos negativos e só pode permanecer rural se não sofrer mudança ou se declinar. É impossível para uma área rural desenvolver-se sem que automaticamente se torne não-rural.”

Esse problema tem sido contornado, de forma recorrente, principalmente pelas instituições responsáveis pelas estatísticas e políticas territoriais dos diversos países, por meio de “categorias intermediárias” como periurbano, semi-rural, “campo urbanizado” (Bagnasco, 1997), “novo rural” (Graziano da Silva, 1997b), que tentam preencher o espaço do suposto *continuum* entre aqueles dois extremos.

A OCDE, por exemplo, no âmbito de seu Programa de Desenvolvimento Rural, vem utilizando um esquema territorial com três níveis, combinando critérios demográficos. Para os 26 países membros, cerca de 50.000 comunidades locais são agrupadas de acordo com a densidade demográfica⁵ em 2.000 regiões, classificadas como: **predominantemente rurais** (mais de 50% da população vivendo em comunidades rurais); **significativamente rurais** (entre 15% e 50%) e **predominantemente urbanizadas** (abaixo de 15%).(OECD, 1996)

5 Se tiver menos de 150 habitantes por km² a comunidade é definida como rural. Para o Japão o limite é de 500 hab/km²

O INSEE, na França, criou a categoria “periurbano” para designar comunidades rurais localizadas em volta de cidades. Seu crescimento seria interpretado apenas como alargamento das atividades urbanas, dando continuidade ao padrão anterior das migrações. Nessa categoria estão incluídas todas as comunidades industriais com no mínimo 100 empregos industriais e uma empresa com no mínimo 20 assalariados. O rural é restrito a comunidades com menos de 2.000 habitantes em áreas contíguas e sem qualquer indústria relevante. Nessa perspectiva as áreas periurbanas podem crescer não apenas pelo deslocamento dos rurais para essas áreas mas também pelo deslocamento dos urbanos que preferem morar longe do centro. E as áreas rurais, pela própria definição, só podem ir perdendo população e emprego porque para ser classificado como rural não pode haver indústria.(Saraceno, 1994b)

Uma outra estratégia foi seguida pelos italianos. O ISTAT dividiu a população em 4 categorias: o rural, o semi-rural, o semi-urbano e o urbano, mas essas categorias nunca foram de fato usadas em pesquisas relevantes, embora o Instituto tenha estatísticas detalhadas sobre elas. O problema é que a dicotomia rural-urbano não atende mais às necessidades teóricas de explicar as diferenças regionais na Itália, aparecendo o tamanho das comunidades como uma variável muito mais relevante. Assim, por exemplo, a reversão das migrações tendeu a ser explicada, nesse país, mais em função das desigualdades regionais do desenvolvimento (uma explicação nacional, portanto) e não pelo corte rural-urbano.(Saraceno, 1994b)

Finalmente, outro exemplo conspícuo é o da nossa PNAD, que permite atualmente desagregar a situação do domicílio em 5 tipos de áreas: urbana urbanizada, urbana não urbanizada ou isolada, rural extensão do urbano, aglomerado rural e rural exclusive.⁶

O problema, no entanto, não se resolve criando categorias intermediárias, porque o conceito de rural tem limitações - se se pretende utilizá-lo como categoria analítica, por exemplo para explicar as desigualdades territoriais do desenvolvimento - na sua própria concepção, o que requer uma revisão mais profunda de seu significado.

Carneiro (1997), por exemplo, tem uma visão bastante crítica tanto da idéia de dicotomia (extremos) como da de *continuum* entre urbano e rural. Recorrendo a autores que sustentam a necessidade de analisar o rural centrado-se nas relações sociais, argumenta:

6 Comentários sobre as definições podem ser encontrados em Graziano da Silva (1997c).

“Orientar o foco de análise para os agentes sociais deste processo [de decomposição e recomposição de sistemas sociais] e não mais para um espaço geográfico reificado possibilita, por exemplo, que a distinção entre ‘cidade’ e ‘aldeia’ ou ‘urbano’ e ‘rural’ desapareça ou torne-se inútil como questão sociológica. Isso porque cada espaço contém em si contradições e conflitos resultantes da relação entre sistemas de valores e interesses distintos, quer sejam eles tidos como de origem ‘urbana’ ou ‘rural’ ”(Carneiro, 1997. p. 155)

Considerando que a ruralidade não é uma realidade “empiricamente observável” mas uma “representação social” definida culturalmente por atores sociais (p.181), propõe a noção de **localidade** como expressão das múltiplas formas de interação desses atores (que desempenham atividades heterogêneas, agrícolas ou não) com a sociedade e a economia global:

“(...) a noção de localidade (...) denota apenas a referência espacial como qualificadora de um universo de relações sociais específico. Em outras palavras, a noção de localidade não define, de forma alguma, a natureza rural ou urbana do grupo ou das práticas e relações sociais que ele desenvolve.”(Carneiro, 1997 p. 162)

Apesar do enfoque antropológico, que não é o que nos interessa mais de perto, o trabalho citado toca no ponto essencial: a necessidade de abolir a noção de rural como conceito básico para os fins analíticos da economia, sociologia e desenvolvimento regional, mas preservar, em algum outro conceito que venha a substituí-lo, o componente espacial ou territorial. A noção de “localidade” é desse tipo; a de “economia local” também, como veremos adiante.

Encontramos uma noção de **localidade** utilizada como instrumento empírico, bem como analítico, nos estudos sobre desenvolvimento rural coordenados por Murdoch e Marsden (1994), também visando à superação da classificação rural-urbano. Para esses autores a compreensão dos vários padrões de diferenciação que levam os espaços a se distinguirem uns dos outros requer analisar os complexos processos que compõem e conformam cada localidade ou região. No caso da recente diferenciação do chamado “rural” (na Inglaterra), esses processos dizem respeito ao que os autores denominam “*land development*”· primeiro, e mais importante, as mudanças no uso da terra, que geraram “mercados segmentados de terra”, orientados a diferentes setores de produção e consumo (agricultura, preservação ambiental, habitação, lazer, turismo); segundo, novas atitudes sociais que

criaram os “novos residentes rurais” pessoas de classe média em busca de um lugar para morar no campo, embora seus vínculos econômicos estejam nas cidades, além de novas atividades de recreação nas áreas rurais; terceiro, a pressão das indústrias manufatureiras e serviços em busca de terra e de construções fora dos grandes centros urbanos. As “localidades rurais” são então caracterizadas como pontos de encontro de diversos conjuntos de relações sociais que interagem em forma de redes, que conformarão o seu caráter material e social. Nessa abordagem combinam-se níveis micro e macro (atores individuais x contexto), como ilustra a seguinte passagem:

“Sugerimos que o processo de desenvolvimento territorial [‘land-development’] pode ser estudado por essa forma e, com isso, a localidade rural pode ser considerada como uma série de resultados físicos e sociais surgidos à medida que os atores perseguem seus objetivos no interior dessas redes. A localidade é assim um ponto de encontro onde redes se interceptam e onde alguns atores podem impor seus interesses sobre outros. Com o tempo, esses resultados vão se compondo para dar lugar a uma formação sócio-espacial.”(Murdoch e Marsden (1994, p. x-xi)

Apenas para não deixar passar em branco, lembramos que a identificação do corte rural-urbano (territorial) com o corte agricultura-indústria (setorial), ou seja, a identificação do rural com o *locus* das atividades agrícolas, obviamente já foi superada, razão pela qual não nos estenderemos sobre o assunto. Como aponta Carneiro (1997), dois fenômenos recentes têm contribuído ainda mais para que o espaço rural não se defina exclusivamente pela atividade agrícola: o crescimento da pluriatividade entre a população rural e a procura crescente de formas de lazer ou opções de moradia no campo por pessoas vindas das cidades, que o transformam em lugar de vida, em vez de espaço de produção. Para todos os efeitos, parece clara a idéia de que o rural é uma categoria **espacial** que independe, **para sua definição**, da exclusividade ou predominância da atividade agrícola, embora quase sempre esta ainda represente a maior parte de seu produto e de seus empregos.

2 O problema dos níveis de análise

A passagem dos conceitos teóricos ao campo de análise empírico quase nunca se dá sem dificuldades (nem nas ciências sociais e nem sequer nas ciências ditas exatas), e o caso presente não constitui exceção.

No caso da pluriatividade, para muitos autores, a unidade de análise - seja para construir tipologias de cunho quantitativo seja para estudos sociológicos e antropológicos mais aprofundados - deve ser a **família** (*family*) ou a **unidade de exploração familiar** (*farm household*), porque é nesse âmbito que são tomadas as decisões relativas às estratégias de sobrevivência do grupo doméstico e de reprodução dos ativos aplicados na exploração agrícola. (Ver dos Anjos, 1995, p. 50-52)

Marsden (1990), por exemplo, compara dois sistemas pluriativos: um sistema de reprodução semi-autônomo, no qual os fatores e *inputs* são produzidos internamente nas propriedades familiares e só há relações de mercado para a venda dos produtos, e um sistema de reprodução dependente do mercado, em que os fatores e produtos são comprados e vendidos. Neste último, a unidade familiar integra-se em diversos circuitos mercantis, envolvendo o emprego de diferentes fatores de produção, como pequenas unidades industriais, lojas, novas máquinas e trabalho, além dos associados à própria agricultura. Ela se integra em diferentes mercados, com diferentes níveis de apoio estatal e sujeitos a diferentes políticas. Nesse contexto,

“uma importante batalha se desenvolve, assim, pelo controle das habilidades e trabalho entre a unidade familiar pluriativa e os agentes externos que procuram influenciar o processo de produção diretamente.”(Marsden, 1990, p. 378)

Nesse modelo, níveis subseqüentes de mercantilização se expressam em termos de divisão e redivisão do trabalho na família, segundo gênero e gerações. O sistema semi-autônomo é sustentado pela reconversão de ativos agrícolas em novos investimentos, por exemplo, transformação da propriedade em ofertadora de serviços (em vez de produtos), como *bed-and-breakfast*, alojamento para cavalos, pousadas para turismo etc. Neste modelo, é a família que resiste à mercantilização, aumentando as rendas não-agrícolas da propriedade rural.

Na Nova Zelândia, Le Heron *et alii* (1994) destacam que a pluriatividade deve ser conceituada como um *continuum*, no sentido de que uma multiplicidade de situações sociais emerge do fato de as famílias agrícolas terem várias atividades. Por exemplo, a pluriatividade pode representar tanto uma estratégia de sobrevivência (permitindo postergar a perda da propriedade ou a diferenciação para baixo) como uma estratégia de expansão de capital (o aumento da renda permite a compra de mais terra ou outros ativos produtivos, favorecendo a diferenciação para cima). Destacam também que a adoção da unidade empresa familiar (*farm-household*) na análise permite tratar a fazenda como um

campo relacional em que interagem o negócio, a família e o domicílio (como conjunto dos habitantes) e que é necessário explorar a interação entre a organização familiar, o mercado e as relações de produção, sem pressupor separação ou dominância entre esses níveis. A pluriatividade é uma das respostas das famílias a mudanças nas circunstâncias econômicas e sociais, devendo portanto sua análise explorar tanto o contexto como a unidade produtiva familiar, ou seja, o nível micro.

O trabalho de De Benedictis (1995) retoma as linhas metodológicas da pesquisa internacional feita para a CE sob o patrocínio da fundação inglesa Arkleton Trust. Sob uma hipótese “espacial” de trabalho de que são múltiplas as formas de pluriatividade, determinaram áreas de estudo ou sistemas territoriais representando diferentes contextos em que as empresas familiares estavam se diferenciando e se desenvolvendo. O tema implícito na metodologia adotada é o “processo de diversificação da realidade territorial” e o papel da empresa familiar agrícola nesse processo, sendo inseparáveis, portanto, as unidades familiares e o contexto em que operam. Constatando que a pluriatividade não era, como se chegou a pensar, o prenúncio da saída da agricultura, mas uma solução funcional no nível micro (ao elevar a renda familiar e viabilizar a permanência no campo) e meso (viabilizando indústrias locais); que as características da família (dimensão, composição, ciclo de vida, escolaridade) eram extremamente importantes na determinação das estratégias da empresa familiar; e que o contexto (econômico, social, institucional e cultural) era uma importante variável explicativa dos processos de diferenciação, adotou-se um esquema analítico que tinha uma área relativa à unidade familiar e outra relativa ao contexto. Assim, a pesquisa empírica deveria considerar as características do contexto (meio físico, mercados de trabalho e produtos, políticas) e os mecanismos de interação entre contexto e unidade familiar, a fim de tornar compreensíveis as decisões de curto e longo prazo que formam as estratégias familiares e seus impactos sobre a diferenciação socioeconômica da estrutura agrária.

A referência espacial relevante (“região”) é a economia local (a cidade e seu entorno rural) que forma o “ambiente produtivo” ou contexto. A análise da capacidade de desenvolvimento das áreas rurais - dada pela sua diversificação interna e suas formas de integração com o exterior - deve ter portanto dois pólos, as empresas (ou as famílias) e o contexto ou economia local em que se inserem.

No mesmo grupo de pesquisa encontra-se o trabalho de Mantino (1995), que propõe dois pontos metodológicos importantes: 1) que a unidade de análise para a pesquisa deve ser a família pluriativa (pelo menos um membro com atividade não ocasional fora da propriedade); 2) que a pluriatividade é função das características da família e da

segmentação do mercado de trabalho. Para o autor, no estudo da pluriatividade e mercados de trabalho é importante que a unidade analisada seja a família, por três razões principais: a família é a unidade de decisão de alocação dos recursos familiares e da empresa; há forte interdependência entre as decisões da família e o desenvolvimento da economia local (contexto), notadamente nas áreas de industrialização difusa; a unidade familiar tem sido objeto de importantes políticas agrícolas e sociais. A família representa pois

“uma estrutura social organizada, com uma divisão interna de trabalho, com papéis diferenciados, estrutura hierárquica e processos internos de socialização e controle coerentes.”(Mantino, 1995, p. 267, citando Bagnasco, 1988, p. 89)

A classe socioprofissional, o ciclo vital, a instrução, a capacidade empresarial, o sexo, a idade e os papéis intrafamiliares são as variáveis explicativas da alocação do trabalho entre a produção própria, o mercado de trabalho e as atividades domésticas, e as pesquisas sobre industrialização difusa na Itália ressaltaram que a família é uma unidade flexível capaz de absorver as crises e se aproveitar das oportunidades do mercado, constituindo um fator específico de desenvolvimento de algumas economias locais.

No modelo adotado a alocação de trabalho é fruto da influência recíproca entre família e contexto, e a segmentação do mercado de trabalho é a chave para a sua compreensão. A família não se defronta com um mercado de trabalho indistinto, mas com vários mercados, em que a dotação geral de recursos e o capital humano da família darão diferentes condições para cada um de seus membros. A família seria a primeira “agência” da oferta de trabalho, capaz de alocar trabalho em ocupações diversas de diferentes segmentos ou setores. Assim, a pluriatividade não é um fato totalmente individual e nem uma escolha puramente da lógica familiar. A família representa o núcleo que toma as decisões sobre a produção do estabelecimento, e a presença de um de seus membros que trabalhe fora influencia tais decisões. A alocação de trabalho entre o estabelecimento familiar e o mercado é fruto de decisões interdependentes no interior da família, e a intensidade da pluriatividade dentro da família não depende só da demanda externa de trabalho, mas sobretudo das variáveis endógenas “clássicas” da família. A pluriatividade, nesse esquema, não é uma escolha ocupacional qualquer, mas sim uma opção precisa, de certo tipo de trabalho, com duração, retribuição e perspectiva de carreira bem determinadas e sua difusão contribui fortemente para incentivar processos de homogeneização social entre famílias agrícolas e não-agrícolas.

Como vimos anteriormente, Shucksmith *et alii* (1989) defendem que a pluriatividade - vista basicamente como diversificação de fontes de renda - é um resultado de uma ou

várias estratégias das famílias agrícolas e, portanto, só tem sentido referi-la ao âmbito familiar. Entre essas estratégias os autores incluem: a busca de novos mercados em resposta a alterações ou expectativas de alterações nos preços relativos; as tentativas de redução de riscos; as mudanças no ciclo familiar (saída ou entrada de filhos no mercado de trabalho, por exemplo); as mudanças na dinâmica interna da família; e as respostas à percepção de novas oportunidades de mercado (no mercado de trabalho, novos produtos e novos usos da terra, por exemplo).

Em outro trabalho, Shucksmith & Smith (1991) mostram que a pluriatividade é uma das explicações para a persistência da produção agrícola familiar na Europa e, nesse contexto, a unidade de análise relevante é, de novo, a família. O trabalho desses autores indica que há, no entanto, limites claros à pluriatividade como solução para os problemas enfrentados pela agricultura familiar: pelo lado da mobilidade do trabalho há limites porque os agricultores resistem a aceitar trabalhos manuais de baixa qualificação e baixo *status* fora da agricultura (e em muitas regiões esses são os tipos de trabalho disponíveis para eles); também a diversificação de atividades dentro da propriedade, reorientando o uso da terra (como turismo, recreação etc.), tem limites porque sua implantação geralmente exige recursos de capital de que a família não dispõe.

Embora no nível da mensuração se possa facilmente definir indivíduo pluriativo (seja em termos de tempo de trabalho ou de fontes de renda) ou família pluriativa (a que tem um ou mais membros pluriativos ou a que tem membros que trabalham em diferentes setores de atividade ou a que tem diversas fontes de renda), a discussão no plano mais conceitual é quase interminável. No âmbito do projeto Arkleton, Newby (1987) argumenta que se trata de uma questão metodológica mais geral, de saber até que ponto o estabelecimento familiar agrícola é um mero agrupamento de estratégias individuais ou desenvolve seus próprios objetivos coletivos. O trabalho de Blanc (1987) parte da hipótese de que a pluriatividade decorre da existência de **mobilidade do trabalho** (intersetorial, entre atividades organizadas com base em diferentes relações de produção, entre postos de trabalho que exigem diferentes qualificações e aprendizados) e esta por sua vez pode observar-se tanto no nível dos indivíduos (mudança de atividade no curso da vida ativa) como no plano familiar (mudança de atividade entre uma geração e outra). Após analisar detalhadamente as características da pluriatividade desde o segundo pós-guerra, conclui que o papel da família é importante num plano macrossocial (a socialização dos jovens, por exemplo), porém

“No nível microeconômico isto leva a considerar que a unidade de tomada de decisão é o indivíduo e não a família e, em conseqüência, a interessar-se muito particularmente pelas trajetórias profissionais das

peçoas, colocando em relevo em primeiro lugar sua atividade principal, que muitas vezes não será agrária.”(Blanc, 1987, p. 109)

Para Barthez (1987), à medida que o tempo parcial e a pluriatividade passaram a ser persistentes e estatisticamente relevantes seu significado também se alterou. No início, quando esses fenômenos estavam referidos à economia camponesa, a unidade de observação era a **exploração agrícola**, em que cada membro (chefe, mulher, filhos) desempenhava tarefas específicas e em tempos determinados. Como essas tarefas eram internas à exploração, provavelmente quase todas as famílias podiam ser consideradas “pluriativas” porque nesse contexto a noção de pluriatividade adotada pela autora é bastante abrangente.⁷ Veja-se, por exemplo, a seguinte afirmação:

“Por conseguinte, a família agrícola é pluriativa se se consideram os critérios utilizados em princípio para expressar o trabalho na agricultura como atividade profissional: varia individualmente a duração do trabalho e a natureza das tarefas. Se se considera a exploração agrária como lugar de atividade profissional de uma família, esta é pluriativa na medida em que a unidade de produção oferece possibilidades de divisão do trabalho em várias ‘atividades’ especializadas. Quando um se dedica à produção de cereais e outro se especializa na criação de animais, produz-se uma situação de pluriatividade.”(Barthez, 1987, p. 165)

Com o desenvolvimento da economia impulsionando a modernização agrícola e o surgimento dos tempos “inativos” (tempo parcial) alguns membros das famílias passam a ter atividades externas, inclusive em outros setores, de modo que a exploração ou estabelecimento agrícola passa a constituir “o centro de uma atividade dotada de satélites” (p.168), os quais dependerão das qualificações individuais e das estratégias de adaptação das famílias à nova situação. Nesse momento a unidade de exploração deixa de ser o único modo de compreender a pluriatividade:

“A família, universo doméstico, aparece então como o elemento essencial na interpretação desse fenômeno de pluriatividade. De tal ponto de vista, através da análise das relações familiares se deve poder compreender como se organizam as diferentes atribuições profissionais de uns e outros, assim como a co-gestão das rendas correspondentes.

⁷ Como apontam Shucksmith *et alii* (1989, p. 346), embora o fenômeno da pluriatividade seja bastante antigo, o que mudou foi a **natureza** das outras fontes de renda e emprego, em direção a atividades e negócios não-agrícolas.

Enquanto que antes a exploração agrária formava o centro unificador da família, com o desenvolvimento da pluriatividade a família se convertia, por sua vez, no centro unificador das diversas atividades exercidas por seus membros.”(p. 169)

Paradoxalmente, na evolução subsequente a própria diversidade de ocupações e as inserções individuais em mercados de trabalho organizados acaba por destruir o postulado da unidade familiar como centro da definição da pluriatividade. Cada vez mais os membros da família agrícola estão inseridos no universo profissional externo, num conjunto de relações mercantis que não se confundem mais com o universo familiar. Agora, como observa Barthez (1987, p. 176-177),

“Sem dúvida alguma, a profissionalização individual dos membros das família modifica consideravelmente os termos de troca familiares. Mas para estudar suas modalidades o ponto de partida da análise não pode ser a família como unidade em si, mas preferentemente os indivíduos que nela coabitam, tentando-se compreender como, a partir de sua dependência pessoal, conseguem integrar-se num projeto comum que os mantêm numa vida de grupo.”

(...) “A exploração agrícola como lugar único da atividade se quebra e dá lugar a uma pluralidade de adscrições profissionais. Simultaneamente, desfaz-se a família agrícola baseada na unidade profissional pela exploração agrária. O grupo de coabitantes que resulta desse duplo movimento não pode identificar-se como uma comunidade em si, fundada em regras específicas; aparece, mais bem, como o desenlace de trajetórias individuais e variadas que, num dado momento, convergem em torno de um interesse comum, formando o grupo familiar.”

Essa consideração da família a partir da independência profissional de seus membros abre, segundo a autora, a possibilidade de analisar a pluriatividade em termos de mobilidade do trabalho e, portanto, no plano individual.

Concluindo, podemos dizer que o estudo da pluriatividade pode ser feito em diversos níveis analíticos, e isto não se define abstratamente nem *a priori*, mas está relacionado com o fato de que esse fenômeno, mesmo que pensado como mera noção descritiva empírica, tem significados distintos conforme o nível (ou estágio) de desenvolvimento da economia agrícola familiar e seu contexto.

No nível micro, a unidade pode ser a **família** - seja no sentido estritamente demográfico do termo, ou enquanto unidade familiar de produção - porque a decisão de um de seus membros de combinar atividades externas (não-agrícolas) com a atividade agrícola implica realocar os recursos produtivos de toda a família. O cálculo do balanço entre tempos de trabalho e rendas pertence ao âmbito familiar. Mas essa unidade também pode ser o **indivíduo**, se a família passou a ser o agregado de qualificações e estratégias profissionais de seus membros, com a atividade agrícola já restrita a um papel secundário.

Mas a “segunda atividade” não se exerce em abstrato ou descolada do resto da economia, e sim num **mercado de trabalho**, que deve constituir, portanto, um outro nível analítico. Este não é propriamente macro, mas talvez “**mesoeconômico**” (contexto), pois a manutenção do vínculo com a propriedade rural (inclusive como moradia) implica atividades em mercados de trabalho locais não muito distantes.⁸ Essas atividades podem estar localizadas indiferentemente nas áreas ditas rurais (serviços ligados ao turismo rural, por exemplo) ou em áreas urbanas próximas (indústrias, comércio, serviços de todo tipo). O que importa é a existência de mercados de trabalho com necessidades que possam ser satisfeitas pelos membros das famílias rurais pluriativas, o que por sua vez não depende da delimitação urbano-rural, e sim de características econômicas, sociais e culturais das **economias locais**, um nível analítico relativamente novo, ao qual dedicaremos a próxima seção.

3 Um nível mesoanalítico: as economias locais

A noção de pluriatividade tem geralmente seu ponto de origem na atividade agrícola - quer dizer, está referenciada a um **setor econômico básico** e outras atividades complementares - mas pode também ser aplicada sem a restrição de que uma das atividades seja a agricultura.⁹ Qualquer que seja o caso, ela não pode ser concebida nem analisada no

8 Há que pensar o caso dos pequenos agricultores e trabalhadores agrícolas que se deslocam para outros estados ou países (caso do sul da Europa) durante certas épocas do ano para exercer uma “segunda atividade”, mas é de se admitir que esses tipos de deslocamento dificilmente poderiam ser generalizados para as famílias agrícolas de todo o país.

9 Não estamos tratando aqui, por exemplo, das famílias urbanas em que cada membro exerce atividades em diferentes setores da indústria, comércio ou serviços, embora também nesse caso se possa dizer que são famílias pluriativas; mas podemos, por outro lado, incluir as famílias que moram na zona rural e têm todos os seus membros trabalhando numa indústria local e no comércio, se o nosso corte privilegiar o territorial e não o setorial. Este último seria o caso de um estudo do “novo rural”, em que se buscassem as características do território, com seus novos tipos de famílias e atividades. Poderíamos observar, por exemplo, a transformação de famílias monoativas agrícolas em famílias pluriativas que abandonaram a agricultura mas continuaram a morar na propriedade rural.

âmbito exclusivo da unidade produtiva, porque implica, necessariamente, relações mercantis com outros setores de atividade, sejam seus mercados de trabalho (se a “segunda atividade” for assalariada), sejam seus mercados de produtos e insumos (se for autônoma, um estabelecimento comercial ou um serviço, por exemplo). Uma família rural agrícola que tenha membros em idade ativa disponíveis e qualificados para certos trabalhos não-agrícolas não basta, obviamente, para desencadear um processo de pluriatividade; é preciso que no seu entorno, ou na sua localidade, ou na cidade próxima haja um mercado de trabalho apto a absorver sua força de trabalho. Em outros termos, é preciso que, de um lado, a modernização agrícola (ou no outro extremo a sua inviabilidade) tenha gerado uma oferta de trabalho interna às unidades produtivas familiares e, de outro, o “contexto” territorial próximo tenha dinamismo suficiente para sustentar mercados de trabalho (ou para novos produtos, serviços autônomos etc.), ou seja, é preciso haver **demanda** adequada para o trabalho pluriativo.

Tanto para a oferta como para a demanda, nessa situação, o corte territorial é imprescindível, porque a mobilidade espacial do trabalho tem limites, pois pressupomos que: a) a família continua vinculada de alguma forma à atividade agrícola, portanto a uma área de terra especificamente localizada; b) continua a existir uma co-gestão familiar da renda e dos ativos aplicados nas várias atividades. Mesmo que a unidade de observação seja o indivíduo pluriativo, se admitirmos que uma das atividades seja agrícola coloca-se também algum tipo de limite ao deslocamento.

A noção de rural, por sua vez, como foi visto, é (ou deveria ser) estritamente **territorial ou espacial**, não podendo ser identificada com setor de atividade nem, acrescentamos, com mercado enquanto *locus* de transações mercantis específicas. Mas sua importância, mesmo enquanto critério de classificação puramente espacial, tem sido reduzida (e questionada) à medida que se mesclam em seu interior atividades econômicas, consumos, hábitos e culturas que antes eram típicas das áreas urbanas.

Já vimos algumas idéias que tentam superar o problema, e em todas elas aparece como elemento chave a **relação** do que era chamado de rural com a cidade, vila ou área urbanizada próxima, isto é, surge uma idéia de **entorno** ou **contexto** - uma espécie de nível meso de análise,¹⁰ porque ultrapassa a unidade produtiva, a família e os atores individuais

10 Trata-se, mais propriamente, de um nível meso de **agregação** espacial que leva em conta características econômicas e sociais, e que não pode evidentemente substituir os níveis de análise micro e macroeconômico e suas unidades analíticas (firmas, mercados etc.). Para uma discussão a respeito do assunto, aplicada à noção de complexos industriais, ver Possas (1992).

(micro) e fica aquém do nível geral de desenvolvimento do país ou da grande região (macro) -, no qual a economia agrícola familiar, incluindo a pluriatividade, está inserida. Esse tipo de enfoque, ainda que com denominações diversas, é sobejamente utilizado na literatura internacional, prestando-se em especial às análises do caso italiano, dada a situação peculiar daquele país (a chamada “industrialização difusa”). Nesta seção incluímos algumas ilustrações desse enfoque, com o objetivo de defender a necessidade de rever os cortes territoriais tradicionais presentes nas nossas estatísticas e de suscitar esforços no sentido de propor novos agrupamentos descritivos ou analíticos mais adequados.

No seminário sobre desenvolvimento descentralizado patrocinado pela FAO no México em 1997 vários trabalhos adotam ou propõem esse tipo de abordagem, como ilustra a seleção a seguir.

Gordillo de Anda (1997), por exemplo, centra a argumentação no surgimento das “cidades médias” e na descentralização (industrial, fiscal, institucional etc.) como elementos que forçaram uma redefinição do rural e a adoção de novos recortes analíticos. A nova definição de **espaço rural** seria

“a zona fora da área metropolitana que integra os povoados rurais e as cidades médias que evoluíram a partir desses povoados.”(Gordillo de Anda, 1997. p. 3)

O autor analisa a necessidade de levar em conta as múltiplas dimensões dos espaços rurais e a heterogeneidade de seus habitantes para a reconstrução das **instituições**, que facilite a interação entre os atores (mercado, Estado, sociedade civil) nesse novo contexto.

A tese central de outro trabalho (Paniagua, 1997) é que o desenvolvimento da agricultura é um processo recíproco do desenvolvimento urbano, contra a percepção generalizada que vê uma precedência histórica do desenvolvimento rural sobre o urbano. Argumenta que o potencial de desenvolvimento depende da inserção de cada espaço nos mercados, da existência de vantagens comparativas naturais e da dotação de externalidades, sobretudo a proximidade a centros urbanos de maior dinamismo.(Paniagua, 1997. p. 8-9)

Considerando que na maioria dos países latino-americanos o fenômeno da transição demográfica já está resultando numa redução da migração aos grandes centros urbanos e numa multiplicação de cidades de porte médio, torna-se importante perceber que estas cidades podem dar um novo alento ao desenvolvimento regional, à descentralização econômica interna e ao surgimento de mercados locais, “*criando um conjunto inédito de*

externalidades para a economia agrícola local e regional. Não só se incrementa a demanda de alimentos derivada do crescimento da população urbana, mas também o desenvolvimento de pequenas e microempresas gera consumos intermediários que exigem encadeamentos cada vez mais estreitos com a produção agrícola do entorno. O artesanato, o turismo, os serviços, a construção etc. fazem parte da nova realidade urbana regional, distanciando-se da imagem das megalópolis subdesenvolvidas.” (Paniagua, 1997, p. 10-11) Os espaços rurais que ficam no interior desses sistemas urbanos interconectados, liderados pelas cidades médias, podem ser concebidos como “**clusters agroindustriais**”, e são esses espaços que têm potencial para o desenvolvimento econômico.

Também para Figueroa (1997) a economia rural está sempre integrada a uma cidade, e a economia regional, combinação de uma cidade e seu entorno rural, é o contexto adequado para compreender o desenvolvimento rural. Seu trabalho representa uma tentativa de inserir teoricamente o tema da vinculação rural-urbana no marco da Nova Economia das Instituições (a partir do trabalho de Douglas North). A proposição teórica da escola institucionalista é de que as formas das trocas observadas dependem dos **custos de transação**, prevalecendo aquelas com menores custos de transação. Na economia rural esses custos são elevados porque os mercados são pouco desenvolvidos, de forma que parte das transações da economia rural tende a dar-se fora do mercado. Em países da América Latina com alto peso de economias camponesas alguns mercados não existem (por exemplo, o mercado de seguros) e outros são incipientes (mercados financeiros), tornando mais difícil a modernização da agricultura. Em cada economia regional existem então transações que são de mercado e outras de “não-mercado” (baseadas em relações de parentesco, reciprocidade e normas tradicionais); os mercados são **regionais** (e não rurais) e quanto maior o tamanho da cidade em relação a seu entorno rural, maior será o peso das transações efetuadas em mercados no total das transações da região. Quanto maior a cidade, mais a economia se baseará nas regras de mercado, mais impessoais serão as relações e mais isso se transmitirá ao meio rural:¹¹

“Um país com áreas urbanas compostas de poucas megalópolis, ou de muitas cidades pequenas, não permitiria o desenvolvimento generalizado dos mercados regionais. Ao contrário, um país com áreas urbanas compostas de muitas cidades de tamanho médio geraria condições favoráveis para o desenvolvimento dos mercados e, portanto, para o desenvolvimento rural. Certamente a urbanização ajudaria o

11 Para o autor, isso tornaria as instituições rurais mais eficientes para a modernização.

desenvolvimento rural, mas não qualquer urbanização e sim aquela geograficamente descentralizada.”(Figuerola, 1997, p. 13)

No caso brasileiro, segundo Faria (1976), as características do sistema urbano, ao contrário da maioria das economias de origem primário-exportadora, afastaram-se do modelo urbano *primate city*, dando lugar a uma assimetria entre a distribuição da população (dispersa) e a distribuição da industrialização (concentrada). O sistema urbano resultante moveu-se em direção a uma distribuição **lognormal** da população por tamanho da cidade, isto é, uma população mais bem distribuída entre os diversos estratos de tamanho de cidades, formando uma rede urbana mais diferenciada, dispersa e integrada. Isto seria, em termos do tema tratado nesta seção, uma vantagem para o desenvolvimento das áreas rurais e dos mercados de trabalho locais.

Garofoli (1997), referindo-se ao caso italiano, argumenta que desde meados dos anos 70 é possível observar um novo modelo de desenvolvimento regional baseado na criação de “territórios” e na proliferação de pequenas empresas. O espaço assume características de um “território”, isto é, um *clustering* de relações sociais e um fator estratégico de oportunidades de desenvolvimento. Nessas áreas as relações entre a economia e o ambiente (contexto) e as redes locais entre firmas tornam-se fundamentais, constituindo a base para **economias externas à firma mas internas à área**. É o modelo conhecido como “industrialização difusa”, que depende de **fatores endógenos**, como a presença de trabalho autônomo e de *know-how* local, coesão cultural, habilidades profissionais, organização familiar dos tempos de trabalho e rendas etc., e de **condições externas** favoráveis, como as novas tecnologias e a crise dos mercados de produção massiva. O “território” tem papel central no processo de desenvolvimento econômico, cobrindo aspectos como informação, produção, mercado de trabalho e governança sócio-institucional e os fatores de sucesso são eminentemente locais. Para que se forme um mercado de trabalho, para que os rurais possam ter múltiplas atividades, é preciso um certo desenvolvimento industrial descentralizado que crie áreas dinâmicas, de forma que os estudos devem ter um nível de agregação espacial o menor possível.

No mesmo sentido vai o trabalho de Bagnasco (1997). Discutindo os efeitos da industrialização difusa, presente no Centro e Nordeste da Itália, mostra como as cidades médias com grande número de pequenas empresas mostram-se fundamentais para gerar o desenvolvimento. Mas as regiões só puderam aproveitar essas novas possibilidades da industrialização difusa por contarem com certas características **locais**, entre as quais destaca-se a sua inserção em redes compactas de cidades grandes e pequenas que conferem ao território funções urbanas (comércio, comunicações, estradas, bancos) que rebaixam os custos de transação. Além disso, essas cidades contavam com um tipo especial de articulação com o campo - caracterizado por uma estrutura social específica, a da família

camponesa autônoma - que viabilizou um mercado de trabalho flexível. Essas famílias camponesas puderam ir combinando, assim, rendas agrícolas e de trabalho industrial, acumulando às vezes um capital que permitia a um filho abrir seu próprio negócio artesanal ou industrial no futuro, formando o que autor chama de “**campo urbanizado**” Hoje, no entanto, a família camponesa já não é mais tão importante, e quase todos os trabalhadores vivem em centros urbanos. (Entre 1951 e 1991 a população agrícola na Itália passou de 42,2% para 8,5% do total).

Entre as mais importantes contribuições para a discussão da ruralidade e da industrialização difusa na Itália encontram-se os trabalhos de Saraceno (1994a, 1994b, 1996, 1997).

Segundo essa autora, até os anos 70 a industrialização e terciarização da economia coincidiam espacialmente com o processo de urbanização e as diferenças espaciais eram interpretadas como desequilíbrios que poderiam ser corrigidos pelo mercado ou pelas políticas econômicas.

Recentemente, verificou-se uma ruptura, na economia industrial, no sentido de mostrar que pequenas e médias empresas podem, em certos contextos, operar de forma eficiente e exibir vantagens competitivas em relação às grandes. Constatou-se em várias regiões que uma multiplicidade de sistemas tecnológicos e organizacionais pode coexistir e se reproduzir, sem se mover em direção a um padrão de eficiência predeterminado, ou seja, há múltiplas trajetórias para o desenvolvimento econômico e as diferenças espaciais persistem e podem ser relevantes para a manutenção da competitividade. Na França, as áreas periurbanas foram as que tiveram o maior crescimento do emprego total (agrícola e não-agrícola) e da população, enquanto no rural o emprego declinou; a redistribuição da população foi mais importante que a do emprego. Na Itália, ao contrário, a redistribuição do emprego foi maior que a da população, favorecendo as pequenas e médias cidades. Mas as áreas rurais intermediárias, como quer que sejam classificadas, foram as que mais diversificaram espontaneamente as atividades. Conclui então que

“A referência espacial relevante parece ser a região ou, mais apropriadamente, a economia local.”(Saraceno, 1994b, p. 326)

Em vista dessas mudanças, o aparato conceitual deve evoluir, incorporando de forma crucial o conceito de “ambiente produtivo” uma unidade de referência em que uma específica combinação de fatores (sociais, econômicos, tecnológicos, institucionais) é organizada de uma forma particular e determina sua competitividade, isto é, esse contexto local é que vai determinar a capacidade de desenvolvimento das áreas rurais, na presença ou não de centros urbanos importantes.

“É a economia regional ou local, em que operam as pequenas e médias empresas, que deve ser descrita como um todo: seus aspectos e inter-relações multissetoriais, seus vínculos sociais, sua organização institucional e capacidade de reproduzir-se no tempo e sua integração com outras economias locais, nacionais ou internacionais.”(idem, p. 327)

Conclui que o enfoque rural-urbano da diferenciação espacial era relevante enquanto os processos de urbanização e industrialização operavam da maneira clássica, típica da primeira geração de países desenvolvidos. Com o descolamento entre espaço e setor (industrialização difusa, novas funções de lazer das áreas rurais, descentralização dos serviços públicos) surge um novo tipo de área, mais dinâmica, que era uma área rural, mas tornou-se periurbana ou de industrialização difusa. O melhor termo para descrevê-las é **economia local** e sua análise deve conter pelo menos duas dimensões: diversificação interna e integração externa.

Esta idéia é aplicada, por exemplo, para explicar a persistência da pequena exploração familiar agrícola no período recente de industrialização (anos 70 e 80) na Itália. Os fatores que contribuíram para essa persistência foram: a criação de “**economias de diversificação**”, pela divisão do trabalho; a **pluriatividade**, que resolveu o problema das baixas rendas (os pequenos estabelecimentos tornaram-se estruturalmente dependentes da pluriatividade e do mercado de trabalho diversificado) e, finalmente, as **economias externas** geradas pelas cadeias agroindustriais, que mantiveram rentável a própria produção agrícola.

Os casos estudados na Itália e Europa em geral indicam que o crescimento da competitividade está associado com dois elementos novos: as economias de escopo¹² ou

12 Economias de escopo ou de produção conjunta são obtidas quando a produção conjunta de diferentes produtos por uma firma é maior do que se fosse realizada por duas (ou mais) firmas, cada uma produzindo um único produto, usando a mesma quantidade de insumos. Produzindo dois ou mais produtos a firma pode usufruir vantagens de custo resultantes da utilização conjunta de insumos ou de meios de produção, programas conjuntos de *marketing*, economias decorrentes de uma administração comum; recursos gerenciais, maquinaria e trabalho qualificado podem ser compartilhados, gerando economias de custo em comparação com a produção separada dos bens em questão. Tipicamente, a presença de economias de escopo significa que as curvas de transformação dos produtos (que descrevem as diferentes combinações de dois produtos que podem ser obtidas com a mesma quantidade fixa de insumos) são negativamente inclinadas e **côncavas** para baixo. (Pindick e Rubinfeld, 1995, p. 216 ss) De acordo com Chandler (1990) pode-se usar o termo economias de escopo também na distribuição: por exemplo, usando a mesma rede de refrigeração um atacadista pode distribuir laticínios, carnes, frutas e vegetais frescos com vantagens de custo que poderão beneficiar inclusive as empresas compradoras. As economias de escopo são, na verdade, o estímulo mais comum à diversificação tanto na produção como na distribuição e *marketing* e inclusive nas atividades de P&D (quando não sejam pesquisas produto-específicas).

diversificação e a demanda segmentada. No passado o princípio que guiou tanto as áreas rurais (agrícolas) quanto as áreas urbanas (indústrias e serviços) era o das economias de escala; hoje esse princípio deu lugar a um melhor aproveitamento dos recursos disponíveis, ampliando a gama de oportunidades de trabalho e a capacidade de resposta em períodos de crise e favorecendo a integração entre setores de diferentes níveis tecnológicos. Este novo princípio é o das **economias de escopo**, que tem conduzido a uma diversificação interna e maior integração externa das regiões, como é o caso das “economias locais” e da industrialização difusa. O segundo elemento é a crescente **segmentação da demanda**¹³ de produtos, sendo as pequenas e médias empresas a resposta da economia às novas condições do mercado, com padrões de consumo crescentemente diferenciados e para os quais essas empresas, ligadas entre si por redes mais ou menos formais, se mostram mais eficientes que as grandes. (Saraceno, 1997)

Nesses novos contextos, a capacidade explicativa dos antigos critérios de diferenciação espacial, entre os quais destaca-se o binômio urbano-rural, foi-se exaurindo e, segundo a autora, as “economias locais” são mais satisfatórias para captar as dinâmicas em curso:

“No curso dos anos 80 verificou-se em muitas regiões, não apenas européias, uma inversão ou desaceleração dos fluxos migratórios tradicionais entre as zonas rurais e urbanas, de um lado, e uma difusão das atividades econômicas que privilegiava não mais apenas as zonas urbanas, mas também os centros menores e as zonas rurais. Na Itália esse fenômeno foi particularmente significativo e foi a origem da individualização da “Terceira Itália” e dos fenômenos de industrialização difusa. Muitas explicações assinalam a importância dos processos de internacionalização da economia na origem dessas inversões de tendência. Essa evolução inesperada deve ser ainda analisada e interpretada sem pré-julgamentos, mas é claro que as suas implicações para o desenvolvimento rural (e agrícola) são extremamente importantes desde que colocam em jogo a competitividade das áreas rurais e o significado da diversificação em curso.

13 Refere-se ao peso crescente da procura por produtos diferenciados e específicos (vinho de tal região ou safra, leite em pó para diferentes faixas etárias, produtos com baixo teor de calorías, produtos sem agrotóxicos, novas fontes vegetais de proteínas etc.) ou nichos de mercado, em contraposição ao consumo massivo de *commodities* e seus derivados.

Parece portanto que nos processos reais há, de um lado, uma variação dos termos de troca entre zonas urbanas e zonas rurais que nos obriga a rever as nossas categorias de análise e, de outro, há também um deslocamento espacial dos 'centros' de desenvolvimento entre regiões e países: ambas estas tendências colocam em crise a capacidade explicativa tradicionalmente atribuída às dinâmicas urbano-rurais e assim também ao conceito de ruralidade. Emerge também a possibilidade, freqüentemente tomada de forma otimista, de que o destino das zonas rurais localizadas no interior de regiões dinâmicas tenha muito mais oportunidades que no passado.”(Saraceno, 1996, p. 5)

Sua tese central é que a “lógica regional” explica muito mais, em cada período do desenvolvimento (e não só no período recente), do que a lógica “rural-urbano” mesmo que se agreguem à definição, além da densidade demográfica ou tamanho populacional, outros indicadores mais sofisticados (atividades agrícolas, mudanças sociais e econômicas etc.). A redistribuição do emprego industrial em favor das cidades menores, com diferentes padrões regionais, um significativo processo de difusão territorial de serviços, inclusive a descentralização dos serviços públicos e o descolamento entre agrícola e rural levaram, portanto, à necessidade de introduzir um novo critério espacial, definido como **economia local**:

“O enfoque da economia local é baseado num espaço mais precisamente definido, não necessariamente coincidente com uma região administrativa. Idealmente, ele seria delimitado pela maior área contígua em que tenha lugar o fenômeno da comutação¹⁴ e/ou empresas troquem quantidades relevantes de bens e serviços segundo um padrão consistente.”(Saraceno, 1994a, p. 456)

(...) “Em contraste com a categoria rural/urbano, baseada na homogeneidade e não-contigüidade de suas características espaciais, o enfoque da economia local/regional é baseado nos efeitos de heterogeneidade e contigüidade, procurando descrever as várias formas de integração e trocas que se desenvolvem entre seus segmentos espaciais e setores de atividade. A presença ou ausência de centros

14 *Commuting* no original. Refere-se a viagens ou deslocamentos diários para o trabalho.

urbanos de diferentes tamanhos em termos demográficos numa região ou numa economia local é um fato relevante, mas não um critério a priori.” (idem, p. 468-469)

A idéia de substituir os critérios tradicionais de “áreas homogêneas” e descontinuidade urbano-rural por agrupamentos espaciais diversificados e contíguos, nos quais as informações e transações fluem em redes mais ou menos complexas, dinâmicas ou eficientes - de forma quase independente da localização urbana ou rural das atividades¹⁵ - parece constituir uma linha metodológica profícua para realizar agrupamentos espaciais em níveis subnacional e subestadual nos estudos de pluriatividade.

4 Sugestões para pesquisa

A discussão anterior sugere evitar a aplicação mecânica de termos que, uma vez disseminados, vão perdendo seus limites conceituais mais precisos tornando-se, com isto, meras referências sem muito poder analítico. Tal pode vir a ser o caso da pluriatividade que, embora há muito disseminada na realidade e na literatura econômica e sociológica dos países desenvolvidos, só mais recentemente vem sendo objeto de análise em nosso País.

Os estudos citados deixam claro que a unidade relevante para a definição e análise empírica da pluriatividade não pode ser estabelecida *a priori*, mas depende das características específicas da região ou do grupo econômico que se quer analisar.

Haverá casos - por exemplo, regiões de agricultura familiar consolidada - em que o estabelecimento rural ou a família é a unidade de decisão relevante, devendo-se pensar então em **famílias pluriativas**, definidas a partir das diferentes combinações de atividades e rendas em seu interior.¹⁶ Uma idéia importante, neste caso, é a de “recursos partilhados”, isto é, a família mantém suas características de unidade de renda e de gastos, mesmo que parte desses recursos seja gerida de forma mais ou menos independente por membros individuais.

15 Não se pode dizer, obviamente, que a localização das atividades industriais, financeiras e de serviços seja totalmente independente das concentrações urbanas. Está-se querendo chamar a atenção para o fato de que a manutenção de limites rígidos entre urbano e rural não ajuda a esclarecer as dinâmicas regionais e, particularmente, a pluriatividade.

16 Deve-se lembrar que a identidade entre família e estabelecimento agrícola faz sentido na Europa e EUA, mas no nosso caso é fundamental distinguir domicílios rurais, famílias e explorações familiares, dada a diversidade de tipos de “famílias expandidas” presentes no meio rural.

Mas em outros casos - em que os objetivos do estudo são o emprego e os mercados de trabalho - o **indivíduo pluriativo** é a unidade de observação adequada, porque cabe a cada um dos membros da família individualmente, dadas as suas qualificações e disponibilidades, procurar as melhores oportunidades e a melhor forma de inserção nos diferentes mercados de trabalho e setores de atividade.

Quanto à categoria espacial, os textos indicam a necessidade de ultrapassar os cortes formais das fontes de dados (urbano-rural e derivados) e buscar uma unidade territorial que tenha algum significado econômico e social mais adequado para melhorar nossa capacidade analítica ante as atuais transformações do território. Como se viu antes, a idéia de “economia local” é uma forma interessante de fazer essa abordagem.

Isto coloca um problema empírico de difícil solução, mas para o qual vale a pena dirigir esforços, qual seja, o de gerar agregados territoriais com significado relevante a partir de variáveis mensuráveis e disponíveis. Nos levantamentos diretos em áreas restritas (estudos de caso) parte do problema deixa de existir porque podemos criar nossas próprias variáveis e seus indicadores e delimitar espaços livres de limitações administrativas, censitárias e estatísticas. Já nas análises mais abrangentes, a partir de estatísticas censitárias ou amostrais, é preciso realizar um esforço adicional no sentido de gerar tipologias ou agrupamentos (*clusters*) a partir de variáveis e níveis de agregação previamente estabelecidos (por exemplo, municípios, distritos censitários etc.).

Um aspecto chave na discussão, que é a forma de associar os dois temas propostos neste trabalho - atividade e território - do ponto de vista econômico e social, é o de “inserção na economia local” isto é, como se dá a interação e integração das famílias, domicílios, indivíduos ou empresas nas redes econômicas e sociais locais, e nos mercados locais e nacionais. A pluriatividade é uma dessas formas de inserção. As empresas familiares, por exemplo, podem se inserir numa economia local via *filière*, isto é, conectadas a um sistema local com predomínio de uma cadeia (produto), como nos casos de integração agroindustrial; outras formas de inserção podem ser os mercados diversificados de produtos de massa (*commodities*) ou de especialidades, ou ainda os mercados de trabalho (agrícolas ou urbanos) ou as inserções especializadas (por exemplo, nas áreas de usinas açúcar ou nos distritos industriais).

Os temas para pesquisa sugeridos pela revisão bibliográfica indicam que é ampla a área de conhecimento a ser preenchida por estudos gerais ou de casos específicos, regionais e locais. Alguns desses temas são a seguir arrolados:

- * importância relativa dos fatores internos à família (composição, ciclo vital, qualificações etc.) e das condições externas (características das economias locais: mercados de trabalho, produtos e serviços, presença de concentração urbana) no surgimento e na evolução da pluriatividade;
- * importância relativa das diferentes fontes de renda da família pluriativa, em termos de atividades agrícolas e não-agrícolas e de outras fontes (pensões, aluguéis, rendas, remessas), segundo o tipo de família e de economia local;¹⁷
- * formas de inserção da pluriatividade nas estratégias econômicas da família: diversificação da aplicação de capital visando obter economias de escopo e reduzir custos de transação (ou modificar sua estrutura) vs. estratégia de sobrevivência para evitar a “diferenciação para baixo”;
- * formas de pluriatividade e sua relação com as estratégias familiares: entrada em mercados de trabalho (agrícolas e não-agrícolas), inserção em novas cadeias produtivas fora da propriedade (novos produtos agrícolas, processados, não-agrícolas) e destinação dos recursos da propriedade a novos usos (turismo, alojamentos etc.);
- * possibilidades de gerar tipologias de economias locais e de famílias pluriativas que permitam obter uma compreensão mais abrangente da problemática;¹⁸
- * relações entre os diversos tipos de sistemas urbanos (mais concentrados ou mais espalhados no interior) e as características da pluriatividade na agricultura;
- * análise das implicações dos pontos anteriores para políticas setoriais e regionais voltadas para as populações agrícolas (ou rurais);¹⁹

17 No caso descrito por Shucksmith & Smith (1991), por exemplo, constatou-se que a maior parte das famílias agrícolas para as quais as rendas não-agrícolas superavam as rendas agrícolas eram na verdade de “*new entrants*” isto é, famílias que provinham de atividades não-agrícolas e só recentemente haviam adquirido ou herdado uma propriedade agrícola. Não se tratava, portanto, de um movimento de diversificação de atividades a partir da agricultura, como nos casos usualmente estudados.

18 A maioria dos estudos revisados consiste de análises em profundidade em localidades bastante restritas.

19 A maioria dos trabalhos europeus sobre pluriatividade e economias locais, incluindo o do Arkleton Trust, insere-se no contexto de discussão das políticas, notadamente das reformas da PAC no fim dos anos 80.

- * análise das possibilidades, vantagens e desvantagens das diferentes fontes de dados empíricos, como PNAD e Censo Demográfico, para a caracterização da pluriatividade na agricultura. O Censo Agropecuário, por ter como unidade de observação o estabelecimento agropecuário, não se presta, obviamente, para apreender o fenômeno em questão.

A implementação de uma proposta como essa requer, como trabalho empírico, primeiramente a delimitação e identificação da “região” ou economia local, a partir do(s) centro(s) urbano(s) mais próximo(s) da área objeto de estudo e, posteriormente, a identificação das formas de inserção do objeto (famílias, empresas, áreas rurais etc.) nessa economia local, envolvendo as relações mercantis não só dos produtos, mas também do trabalho. As dificuldades de um tal empreendimento são óbvias, mas esforços em sua direção poderão tornar mais produtivos os resultados das pesquisas sobre os temas aqui abordados.

Referências bibliográficas

- Bagnasco, A. La función de las ciudades en el desarrollo rural: la experiencia italiana. *SEMINARIO INTERNACIONAL INTERRELACIÓN RURAL-URBANA Y DESARROLLO DESCENTRALIZADO*. Taxco, México: FAO/ONU, abril de 1997
- Barthez, Alice. Familia, actividad y pluriactividad en la agricultura. Arkleton Research. *Cambio Rural en Europa*. Colóquio de Montpellier, 1987, p. 161-177
- Blanc, Michel. Pluriactividad y movilidad del trabajo: un enfoque macroeconomico. Arkleton Research. *Cambio Rural en Europa*. Colóquio de Montpellier, 1987, p. 89-109.
- Cabral, Manuel Villaverde. Pluri-activité et stratégies paysannes d’abandon de l’agriculture- deux exemples. *Information sur les sciences sociales*, v. 26, n. 2, p. 319-344, 1987 Reproduzido em Godelier, M., *Transitions et subordinations au capitalisme*, Paris: Ed. de la Maison des Sciences de l’Homme.
- Carneiro, Maria José. Pluriatividade: uma resposta à crise da exploração familiar? In: XVII Encontro Nacional do PIPSA, 1994, Porto Alegre. *Anais...* Porto Alegre: UFRGS. Programa de Pós-graduação em Sociologia, 1995, p. 50-58.
- _____. Ruralidade: novas identidades em construção. *Anais do XXXV Congresso Brasileiro de Economia e Sociologia Rural*, Brasília, 1997. p. 147-185.

Chandler, Jr. Alfred D. *Scale and scope dynamics of industrial capitalism*. Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1990.

Comissão das Comunidades Europeias. A situação da agricultura na Comunidade, Relatório 1992. Bruxelas, Luxemburgo, 1993.

De Benedictis, Michele. Agricultura familiare in transizione: nodi teorici ed evidenze empiriche. *In: De Benedictis, M. Agricultura familiare in transizione*. Studi e Ricerche, Itália: INEA, 1995, p. 11-42.

dos Anjos, Flávio S. A agricultura em tempo parcial: elementos de discussão para um enfoque sociológico. *Cadernos de Ciência e Tecnologia*, Brasília, v. 12, n. 1/3, p. 39-54, 1995.

Faria, Vilmar. O sistema urbano brasileiro- Um resumo das características e tendências recentes. *Estudos CEBRAP* 18, S.Paulo: Ed.Cebrap, p. 91-115, 1976.

Figueroa, Adolfo. Hacia un nuevo enfoque sobre el desarrollo rural en América Latina. *SEMINARIO INTERNACIONAL INTERRELACIÓN RURAL-URBANA Y DESARROLLO DESCENTRALIZADO*, Taxco, México: FAO/ONU, abril de 1997.

Fuller, Anthony M. Part-time farming: the enigmas and the realities. *In: Schwartzweller, H. (ed.), Research in Rural Sociology and Development*. JAI Press Inc. v. 1, p. 187-219, 1984.

_____. From part-time farming to pluriactivity: a decade of change in rural Europe. *Journal of Rural Studies*, v. 6, n. 4, p. 361-373, 1990.

Fundação Getúlio Vargas, Instituto de Documentação. *Dicionário de Ciências Sociais*. 2.^a ed. Rio de Janeiro, 1987.

Garofoli, Gioacchino. Rural development and diffuse industrialisation: learning from Italian experience. *SEMINARIO INTERNACIONAL INTERRELACIÓN RURAL-URBANA Y DESARROLLO DESCENTRALIZADO*. Taxco, México: FAO/ONU, abril de 1997

Gasson, Ruth. Part-time farming: a note on definitions - comment. *Journal of Agricultural Economics*, Ashford, UK., v. 42, n. 2, p. 200-201, 1991.

Gordillo de Anda, Gustavo. Restructuración institucional y revalorización de los vínculos rural-urbano. *SEMINARIO INTERNACIONAL INTERRELACIÓN RURAL-URBANA Y DESARROLLO DESCENTRALIZADO*. Taxco, México: FAO/ONU, abril de 1997.

Graziano da Silva, José *et alii*. O emprego rural e a mercantilização do espaço agrário. *São Paulo em Perspectiva*, S. Paulo, v. 11, n. 2, p.50-64, 1997a.

Graziano da Silva, José. O novo rural brasileiro. In: Shiki, S. *et alii* (org.), *Agricultura, meio ambiente e sustentabilidade do cerrado brasileiro*, Uberlândia, 1997b, p. 75-100.

_____. Sobre a delimitação do rural e do urbano no Brasil: testando as aberturas geográficas nas novas PNADs. *Anais do XXXV Congresso Brasileiro de Economia e Sociologia Rural*, Brasília, 1997c, p. 114-146.

IBGE. *Censo Demográfico de 1991 Brasil*. Rio de Janeiro, 1994.

Le Heron, Richard; Roche, Mike; Johnston, Tom. Pluriactivity in New Zealand Agriculture in the 1980s - a benchmark study of livestock and fruit agro-commodity systems. *Geoforum*, v. 25, n. 2, p. 155-172, 1994.

Lund, Philip J. Part-time farming: a note on definitions. *Journal of Agricultural Economics*, Ashford, UK, v. 42, n. 2, p. 196-199, 1991.

Mantino, Francesco. Famiglia, pluriattività e mercati del lavoro: il quadro teorico e i risultati della ricerca. In: De Benedictis, M. *Agricoltura familiare in transizione*. Studi e Ricerche INEA, Itália, 1995. p. 255-283.

Marcondes, J. V. Freitas. A agricultura em tempo parcial no estado de São Paulo e a industrialização. S. Paulo, *Sociologia*, v. 24, n. 1, p. 29-40, 1962.

Marsden, Terry. Towards the political economy of pluriactivity. *Journal of Rural Studies*, v. 6, n. 4, p. 375-382, 1990.

Murdoch, Jonathan & Marsden, Terry. *Reconstituting rurality - class, community and power in the development process*. London: UCL Press, 1994.

Newby, Howard. La familia y la explotación agraria. Arkleton Research. *Cambio rural en Europa*. Colóquio de Montpellier, 1987, p. 155-159.

OECD. *Territorial indicators of employment focusing on rural development*. Paris, 1996.

Paniagua V., Alberto. Criterios y procedimientos para identificar espacios rurales capaces de articularse positivamente a ciudades intermedias. *SEMINARIO INTERNACIONAL INTERRELACIÓN RURAL-URBANA Y DESARROLLO DESCENTRALIZADO*, Taxco, México: FAO/ONU, abril de 1997.

Pindick, Robert S. & Rubinfeld, Daniel L. *Microeconomics*. 3rd edition. Prentice-Hall International Inc., 1995.

Possas, Mario L. Concorrência, inovação e complexos industriais: algumas questões conceituais. *Texto para Discussão* n. 9. Campinas: Instituto de Economia, UNICAMP, junho de 1992.

Saraceno, Elena. Alternative readings of spatial differentiation: the rural versus the local economy approach in Italy. *European Review of Agricultural Economics*, Berlin, v. 21, p. 451-474, 1994a.

_____. Recent trends in rural development and their conceptualisation. *Journal of Rural Studies*, v. 10, n. 4, p. 321-330, 1994b.

_____. Il concetto di ruralità: problemi di definizione su scala europea. Programa de Seminários INEA sobre DESENVOLVIMENTO NAS ÁREAS RURAIS, Métodos de análise e políticas de intervenção. Roma, 30 de outubro de 1996.

_____. Urban-rural linkages, internal diversification and external integration: a European experience. *SEMINARIO INTERNACIONAL INTERRELACIÓN RURAL-URBANA Y DESARROLLO DESCENTRALIZADO*. Taxco, México: FAO/ONU, abril de 1997.

Shucksmith, Mark *et alii*. Pluriactivity, farm structures and rural change. *Journal of Agricultural Economics*, Ashford, UK., v. 40, n. 3, p. 345-360, 1989.

Shucksmith, Mark e Smith, R. Farm household strategies and pluriactivity in Upland Scotland. *Journal of Agricultural Economics*, Ashford, UK., v. 42, n. 3, p. 340-353, 1991.

Encargos sociais e demanda por trabalho no setor formal da economia

Reynaldo Fernandes[§]

RESUMO

Este artigo analisa teoricamente o debate do peso dos encargos sociais no Brasil. Ele utiliza como critério de avaliação o bem-estar dos trabalhadores. A medida proposta para o peso dos encargos sociais é a diferença entre os custos do trabalho impostos pela legislação e o quanto os trabalhadores avaliam os benefícios dela decorrente. Assim, o peso dos encargos sociais possui um significado que, para efeitos do funcionamento do mercado de trabalho, é equivalente a um imposto sobre o trabalho e, portanto, fornece uma medida dos incentivos que empregados e empregadores possuem em transformar suas relações de trabalho de formal para informal. O artigo também avalia a proposição de substituir a fonte de financiamento dos programas sociais, da folha de pagamentos para o valor adicionado, concluindo que ela não é uma proposição tão simples de ser defendida, indicando a necessidade de novos estudos sobre a questão.

Palavras-chave: encargos sociais, demanda por trabalho, imposto sobre folha de salários.

ABSTRACT

This paper analyses theoretically the Brazilian debate about employees' pay roll social contribution. To evaluate the social contribution burden on pay roll it uses employees welfare as criterion of the analysis. It proposes as measure of this burden the difference between costs and benefits both imposed by law, where the benefits are a worker's subjective dimension. Therefore, the social contribution burden has a precise meaning. It is similar to pay roll tax and it provides a measure of the incentive to employers and employees to move from formal to informal labor market. This paper deals also with the matter of the changing the collection source, from pay roll tax to value added tax. It concludes it isn't a proposition to be defended easily.

Key words: social contribution, demand for labor, pay roll tax.

§ Do departamento de economia da FEA/USP, campus de Ribeirão Preto.

O autor agradece a Samuel Abreu Pessôa, Ricardo Paes de Barros, Paulo Picchetti, Antônio Gledson de Carvalho e a dois pareceristas anônimos desta revista pelos comentários e à FIPE – Fundação Instituto de Pesquisas Econômicas pelo suporte financeiro concedido a esta pesquisa.

Recebido em outubro de 1997. Aceito em fevereiro de 1998.

I Introdução

A idéia de que os encargos sociais no Brasil são demasiadamente elevados, com impactos negativos sobre o nível de empregos e salários do setor formal da economia, é hoje bastante difundida. Alguns estudos têm procurado avaliar o peso desses encargos sobre a folha de salários, mostrando o quanto ele é elevado, seja quando comparado ao início da década, seja na comparação com outros países. Por exemplo, Almeida (1992) procura demonstrar que *“há, atualmente, excessiva carga impositiva sobre a folha de salários, o que tem repercutido de modo desfavorável, em especial sobre a estrutura do mercado de trabalho brasileiro, bem como sobre os próprios níveis da arrecadação previdenciária.”* A autora, ao analisar a evolução recente da estrutura de contribuições das empresas e dos empregados para a previdência social, ressalta que as *“soluções para a superação das consecutivas dificuldades financeiras por que passou a instituição foram, via de regra, buscadas na criação de novos tributos ou na elevação da carga impositiva.”* As alíquotas do INSS passaram de 8% para os trabalhadores e 8% para as empresas (até o limite de 20 SM), no início dos anos 80, para 8 a 10% para os trabalhadores (até o limite de 20 SM de referência) e 20 a 22,5% para as empresas.

Em outro estudo, Pastore (1994) calcula em 102% o peso dos encargos sociais sobre a folha de salários. Comparando o Brasil com 9 países europeus (Alemanha, Bélgica, Dinamarca, França, Holanda, Inglaterra, Irlanda, Itália e Luxemburgo) o autor conclui que o Brasil possui *“a mais alta carga na comparação efetuada.”*

Do ponto de vista da política econômica, estes trabalhos apontam para a existência de um grande espaço para se reduzir encargos trabalhistas, com impactos significativos sobre o nível de empregos e salários do setor formal da economia.

Esta idéia, apesar de bastante difundida, não é consensual. Alguns autores argumentam que estes cálculos incluem nos encargos sociais itens que, a rigor, não deveriam ser considerados como tal. Em especial, argumenta-se que muitos destes itens vão direto para o “bolso” do trabalhador e, portanto, deveriam ser considerados como salários. Redefinindo o conceito de encargos sociais, autores como Amadeo (1994a e 1994b) e Santos (1996) calculam que o peso dos encargos sociais sobre a folha de salários está em torno de 25%, o que é compatível com os padrões internacionais, existindo, pois, muito pouco espaço para se reduzir encargos. Como argumenta Amadeo (1994b), retirando-se os itens que vão diretamente para o trabalhador, *“estamos falando de redução dos 24,5% que vão para o Estado e dos 1,6% que vai para as empresas. É pouco o que há para cortar e há que se pensar em como financiar os gastos financiados por estes encargos.”*

A princípio, não deveria causar estranheza que diferentes definições de encargos sociais causem diferenças no peso destes encargos na folha de salários. O fato é que não existe um conceito de aceitação geral para encargos sociais ou trabalhistas e, portanto, a definição apropriada fica dependendo dos objetivos do analista.

Este trabalho procura reavaliar tal debate tendo como perspectiva o bem-estar dos trabalhadores. Assim, ele explora a possibilidade de melhorar as condições de vida dos trabalhadores via uma alteração na legislação trabalhista, no que tange aos “encargos” sobre a folha de salários.

O artigo foi desenvolvido em três seções, além desta introdução: a seção 2 discute alguns aspectos metodológicos sobre o cálculo dos encargos sociais como proporção da folha de salários; a seção 3 analisa o impacto de um imposto sobre a folha de salários, comparativamente a um imposto sobre o valor adicionado, na demanda por trabalho; e, por fim, na seção 4 tecem-se as considerações finais.

II O peso dos encargos sociais sobre a folha de salários: alguns aspectos metodológicos

Os cálculos do peso dos encargos sociais sobre a folha de salários seguem, grosso modo, o seguinte procedimento:

- i) Definem-se os itens que, além do salário básico, compõem o custo do trabalho;
- ii) Divide-se o custo do trabalho entre salários e encargos sociais;
- iii) Calcula-se a razão entre encargos sociais e salários.

Desta forma, as diferenças encontradas no peso dos encargos sociais sobre a folha de salários derivam de diferenças nos itens que devem ser incluídos no custo do trabalho e nos itens, incluídos no custo do trabalho, que devem ser considerados como encargos sociais ou como salários. Por exemplo, Pastore (1994) define encargos sociais como sendo todos os itens do custo do trabalho, à exceção do salário-base, e inclui no custo do trabalho itens como: 13º salário, férias, adicional de férias, INSS, FGTS, salário-educação, INCRA, SEBRAE, SESI, SENAI, repouso semanal, feriados, auxílio enfermidade, aviso prévio e despesas de rescisão contratual. Por outro lado, autores como Santos (1996) ponderam que itens como 13º salário, adicional de férias, FGTS, aviso prévio e outras despesas com rescisão contratual deveriam ser considerados como salários, o que reduziria o numerador

e aumentaria o denominador da razão encargos-salários e, portanto, reduziria o peso dos encargos sociais sobre a folha de pagamentos. Além disso, Santos não inclui no custo do trabalho itens como férias, feriados e repouso remunerado.

O ponto central deste artigo é o bem-estar dos trabalhadores. Assim, questão a ser destacada refere-se à parcela dos custos do trabalho que não reverte em benefícios para os trabalhadores. É evidente que o FGTS, por exemplo, proporciona um benefício direto para os trabalhadores, e pela forma que a questão é colocada em estudos como o de Pastore (1994) tem-se a impressão que o recolhimento para este fundo deveria ser tratado como um simples imposto sobre a folha de salários. Entretanto, colocar o FGTS como parte dos salários significa assumir que os trabalhadores são indiferentes entre ter este benefício e receber em dinheiro o que as empresas arcam para mantê-lo. Se, por hipótese, os trabalhadores estivessem dispostos a trocar este benefício por um aumento correspondente a 4% do salário bruto (metade do que custa para as empresas), um montante de 4% dos salários deveria ser considerado um imposto sobre a folha de pagamentos, ou mesmo um “peso morto”

Antes de nos atermos a esta questão de forma um pouco mais detalhada, seria interessante destacar dois aspectos metodológicos que são comuns à grande maioria dos estudos que calculam o impacto dos encargos sociais sobre a folha de pagamentos.

O primeiro aspecto refere-se ao fato de que em tais estudos consideram-se como encargos as despesas tidas como de responsabilidade das empresas (aquelas que não fazem parte do salário contratual, ou salário bruto dos trabalhadores). Isto é marcadamente importante no caso das contribuições para o INSS.

Hoje, cerca de 30% do salário bruto de um trabalhador é “taxado” para financiar a previdência (8% a 10% referente à parcela do trabalhador e 20% a 22,5% referente à parcela do empregador). Vamos admitir uma situação onde o trabalhador possua um salário bruto de 100, sendo a contribuição para a previdência de 8% para o trabalhador e 20% para a empresa. Restringindo-nos apenas ao INSS, temos, neste caso, o salário líquido do trabalhador igual a 92, o custo do trabalho igual a 120 e a arrecadação da previdência igual a 28. Utilizando-se da metodologia de cálculo desses estudos, o peso dos encargos sociais sobre a folha de salários seria de 20%.

Vamos imaginar agora que o salário bruto do trabalhador seja elevado para 120 e que toda a arrecadação do INSS seja feita sobre o trabalhador – portanto, uma alíquota de 23,33%. Isso em nada altera a situação do trabalhador, da previdência ou da empresa.

Entretanto, o peso dos encargos sociais sobre a folha de salários seria igual a zero. O mesmo raciocínio se aplica caso o salário contratual fosse reduzido para 92 e a empresa tivesse que arcar com 28 para a previdência. Novamente, nada mudaria para as partes envolvidas, mas o peso dos encargos seria de 30,44%.

Assim, três situações economicamente indistinguíveis apresentam diferenças significativas no peso dos encargos sociais sobre a folha de pagamentos.

O segundo aspecto a ser destacado é que nem todos os itens de custo do trabalho são incluídos na discussão do peso dos encargos trabalhistas sobre a folha de salários. A discussão centra-se nos chamados encargos compulsórios ou legais. E mesmo entre estes apenas uma parcela é considerada, aquela mais facilmente generalizável para o conjunto dos trabalhadores.

Uma parte importante dos chamados “custos não-salariais do trabalho” são acordados voluntariamente entre empregados e empregadores. Tais custos variam entre empresas e entre trabalhadores e, por este motivo, são de difícil mensuração. De qualquer modo, por serem voluntários, não são vistos como um atrito no mercado de trabalho.¹

Os estudos do peso dos encargos trabalhistas sobre a folha de salários estão inseridos na discussão sobre o impacto da legislação trabalhista no funcionamento do mercado de trabalho, mas que, evidentemente, não a esgota. Um contrato de trabalho pode ser visto como um arranjo bastante complexo, envolvendo diversos aspectos ou dimensões. Ao estabelecerem um contrato de trabalho empregados e empregadores acertam uma série de cláusulas, tais como: salário; condições e ambiente de trabalho; descanso semanal; férias; jornada de trabalho; esforço e comprometimento no trabalho; planos de seguros etc. Muitas dessas cláusulas são explícitas, outras apenas implícitas.

A legislação trabalhista pode afetar a eficiência dos contratos de trabalho - e, conseqüentemente, o bem-estar dos trabalhadores - de diferentes formas. Ela restringe o conjunto de contratos possíveis e, deste modo, pode impedir que arranjos mais proveitosos para as duas partes sejam estipulados² ou pode “permitir” a implementação de cláusulas

1 Entretanto, uma das possíveis explicações para a existência e o crescimento desses custos reside na própria legislação. Tais benefícios podem ser uma maneira de empresas e trabalhadores “fugirem” dos encargos legais e, assim, podem refletir uma ineficiência alocativa no mercado de trabalho. Para uma discussão desses benefícios ver, por exemplo, Hart (1984).

2 Na realidade, como a legislação pode ser burlada, e isto envolve custos, a legislação trabalhista pode ser vista como “encarecendo” certos tipos de contratos.

que, apesar de vantajosas, pelo menos para uma das partes, não seriam estipuladas livremente.³ A legislação pode, também, impor uma diferença entre o que o trabalhador recebe e o que a empresa gasta para mantê-lo empregado, assemelhando-se à implementação de um imposto, ou subsídio, sobre o trabalho.⁴ É exatamente sobre este último aspecto que os cálculos do peso dos encargos sociais sobre a folha de salários poderiam fornecer uma estimativa.

Como o objetivo é avaliar o impacto da legislação trabalhista sobre o bem-estar dos trabalhadores, temos que comparar o trabalhador representativo, ou mediano, em duas situações distintas: uma que a legislação não está presente (que chamaremos de estado 0) e outra que a legislação está em vigor (que chamaremos de estado 1). Assim, teríamos que avaliar EU_1 e EU_0 , sendo $EU_1 = \pi_1 u(W_1, B_1) + (1 - \pi_1) u(R_1)$ e $EU_0 = \pi_0 u(W_0, B_0) + (1 - \pi_0) u(R_0)$, onde:

EU_i = utilidade esperada do trabalhador representativo no estado i

W_i = salário líquido no estado i

B_i = vetor de benefícios contratados no estado i

R_i = vetor de benefícios do desemprego no estado i (salário-desemprego; renda do setor informal etc.)

$u(.)$ = função de utilidade instantânea

π_i = probabilidade de estar empregado no estado i

Admitindo que o mercado de trabalho no setor formal seja competitivo ou, alternativamente, caso estivéssemos interessados em comparar o bem-estar apenas dos trabalhadores efetivamente empregados no setor formal poderíamos nos restringir a comparar $u(W_1, B_1)$ e $u(W_0, B_0)$. Neste caso, poderíamos definir um salário líquido W^* , tal que $u(W^*, B_0) = u(W_1, B_1)$, sendo que $(W_0 - W^*)$ nos daria uma medida da perda (ou ganho) de bem-estar do trabalhador em virtude da legislação trabalhista.

3 A presença de problemas de coordenação pode impedir cláusulas contratuais que seriam viáveis apenas se existisse uma adesão generalizada. Certos programas assistenciais poderiam ser vantajosos a um determinado custo, o qual seria factível apenas se realizado em uma escala relativamente grande e, por esse motivo, impediria que uma empresa isoladamente o oferecesse.

4 Uma parte da diferença entre o que o trabalhador recebe e o custo de mantê-lo empregado refere-se aos custos de transação (obtenção de informações; elaboração e monitoria de contratos etc.). Tais custos não estão sendo considerados aqui, mesmo porque uma grande parte deles são pagamentos feitos a outros trabalhadores.

Vale notar que tanto W_0 como W^* são variáveis não observáveis no estado 1. Assim, numa economia que já opera com base numa legislação trabalhista, tal medida não poderia ser obtida diretamente e quaisquer estimativas de W_0 e W^* deveriam ser realizadas apenas com informações disponíveis no estado 1. Isto tornaria muito difícil realizar uma estimativa para W_0 . Entretanto, se fossemos capazes de fixar B_0 - ou seja, os benefícios que vigorariam na economia caso nenhuma legislação trabalhista estivesse presente - as informações disponíveis no estado 1 poderiam nos proporcionar uma estimativa para W^* pelo menos para os empregados no setor formal da economia.

Uma estimação para B_0 poderia ser realizada, por exemplo, investigando-se os benefícios que vigoram no setor informal da economia. Com base na estimativa de B_0 passaríamos a investigar a elevação mínima no salário líquido capaz de induzir o trabalhador representativo do setor formal a “reduzir” os benefícios de B_1 para B_0 , o que poderia ser feito perguntando-se diretamente aos trabalhadores.⁵ Ou seja, com base nas informações disponíveis no estado 1 seria possível estimar W^* e, portanto, o quanto os trabalhadores avaliam os benefícios existentes em função da legislação trabalhista,⁶ o qual seria dado por $(W^* - W_1)$.

Com base nas informações do estado 1, também seria possível estimar em quanto os custos do trabalho se elevam quando as empresas são obrigadas a alterar o vetor de benefícios de B_0 para B_1 , a um dado salário líquido. Isto nos daria uma estimativa do hiato, entre o que a empresa gasta e o que o trabalhador recebe, imposto pela legislação trabalhista.⁷ É com relação a este ponto que os estudos do peso dos encargos trabalhistas sobre a folha de salários podem ser úteis.

III O impacto de um imposto sobre a folha de salários

Como visto, a legislação trabalhista pode impor um hiato entre o que os trabalhadores “recebem” e o que as empresas arcam para manter os contratos de trabalho. Cabe ressaltar

5 Para isto poderíamos utilizar técnicas de precificação de bens contingentes.

6 Aqui cabe fazer uma distinção entre benefícios existentes em função da legislação e benefícios garantidos pela legislação. Um benefício garantido pela legislação poderia existir independentemente da mesma. Neste caso, a legislação não importaria nenhuma alteração no funcionamento do mercado de trabalho.

7 Para uma estimativa dos custos devemos incluir, além dos pagamentos de benefícios, os itens que são tipicamente impostos, ou seja, aqueles que não possuem, como contrapartida, nenhum benefício direto aos trabalhadores, como, por exemplo, as contribuições para o SEBRAE.

que tal hiato não é exatamente uma medida para a variação de bem-estar dos trabalhadores, dada a imposição da legislação trabalhista, que seria $(W_0 - W^*)$. Entretanto, ele deve guardar uma relação inversa com essa variação caso todo custo associado à legislação providenciasse um benefício direto aos trabalhadores. O problema é que parte dos chamados encargos sociais, como, por exemplo, o salário-educação, não providenciam nenhum benefício direto aos trabalhadores, possuindo, assim, um caso típico de imposto sobre a folha de salários.

Para se avaliar o bem-estar dos trabalhadores, numa situação que o imposto está presente e numa situação onde o imposto está ausente, seria necessário considerar o uso dos recursos arrecadados. É possível que - se tais recursos forem bem aplicados em fatores que, por exemplo, promovem o crescimento (educação, infra-estrutura etc.) - no equilíbrio de longo prazo trabalhadores estejam melhores na situação em que o imposto sobre o trabalho está presente do que na situação onde este imposto, e os recursos por ele arrecadados, está ausente. Neste caso, não seria possível falar numa relação inversa entre imposto sobre folha de salários e bem-estar dos trabalhadores.

Mesmo levando em conta este aspecto, a questão a ser considerada é que tais gastos poderiam ser financiados por outras fontes que não um imposto sobre o trabalho. Portanto, as situações a comparar referem-se àquelas em que tais gastos são financiados por um imposto sobre a folha de salários ou por um imposto sobre outra base. Aqui, tem-se, freqüentemente, argumentado que a relação negativa entre bem-estar dos trabalhadores e imposto sobre o trabalho deve ser restabelecida. Isto porque, do ponto de vista dos trabalhadores, um imposto sobre o trabalho seria a pior forma de tributação.

A base tributária usualmente apontada para substituir a folha de salários é a do valor adicionado. Acreditamos que a escolha do valor adicionado tem levado em consideração aspectos como abrangência e facilidade de arrecadação e fiscalização, quando comparado a outras bases (o Imposto sobre a Renda, por exemplo), bem como a facilidade de adaptação, uma vez que o mesmo já é utilizado na estrutura tributária vigente. Deste modo, substituir os impostos sobre folha de salários por um imposto sobre o valor adicionado pode não ser, do ponto de vista dos trabalhadores, a escolha ótima, mas levaria a uma situação melhor do que a atual. O objetivo desta seção é exatamente avaliar esta proposição.

III.1 Um modelo de curto prazo

Para iniciar a discussão da proposição acima, utilizaremos um modelo bastante simplificado, no qual o governo, para obter uma dada arrecadação tributária, pode utilizar-se de dois tributos: um sobre a folha de salários e outro sobre o valor adicionado.

Vamos admitir a existência de uma única empresa atuando no setor formal da economia, a qual opera sob regras competitivas. A função de produção dessa firma é dada por:

$$(1) \quad q = f(k, n) \quad \text{onde,}$$

q = quantidade produzida do único produto da firma
 k = estoque de capital (o qual vamos considerar fixo)
 n = quantidade de trabalho

Supondo que o produto seja o numerário, a equação de custo total é representada por:

$$(2) \quad CT = a.q + (1+g).w.n + p.k \quad \text{onde,}$$

a = alíquota do imposto sobre o valor adicionado (neste caso igual ao faturamento)
 g = alíquota do imposto sobre a folha de salários
 w = salário
 p = preço dos serviços do capital (que é dado para firma).

A equação de lucro total é:

$$(3) \quad LT = f(k, n) - a.f(k, n) - (1 + g).w.n - p.k,$$

onde a condição de primeira ordem para a maximização de lucro é igual a:

$$(4) \quad f_n(k, n) = F.w,$$

$$(5) \quad F = \frac{(1+g)}{(1-a)}$$

Vamos denominar F de cunha fiscal.⁸ Assim, tanto um aumento na alíquota do imposto sobre a folha de salários como do imposto sobre o valor adicionado aumentam a cunha fiscal e, portanto, reduzem a demanda por trabalho.

Neste modelo, a arrecadação do governo seria dada por:

$$(6) \quad A = g.w.n + a.f(k, n)$$

A um dado salário de mercado e a um dado nível de emprego podemos determinar as combinações de “a” e “g” que geram a mesma arrecadação. Tal equação, que poderíamos chamar de isoarrecadação, fica representada por:

$$(7) \quad a = \frac{A - g.w.n}{q}$$

Uma questão a ser investigada é: a um dado nível de emprego e salários, qual é a combinação de “a” e “g” que mantém uma dada arrecadação e minimiza F ?

Substituindo (7) em (5), a equação da cunha fiscal torna-se,

$$(8) \quad F = \frac{(1 + g).q}{q - A + g.w.n} \quad e,$$

$$\frac{dF}{dg} = \frac{q.(q - A - w.n)}{(q - A + g.w.n)^2}$$

O sinal de $\frac{dF}{dg}$ fica dependendo do sinal de $(q - A - w.n)$. Note que a equação de lucro total pode ser reescrita como:

$$(3') \quad LT = q - A - w.n - p.k \quad \text{ou,}$$

$$LT + p.k = q - A - w.n$$

8 De modo geral, o termo cunha fiscal é utilizado para representar a diferença entre o que o comprador paga e o vendedor recebe. Aqui, o termo é utilizado para representar a diferença entre o salário real e a produtividade marginal do trabalho.

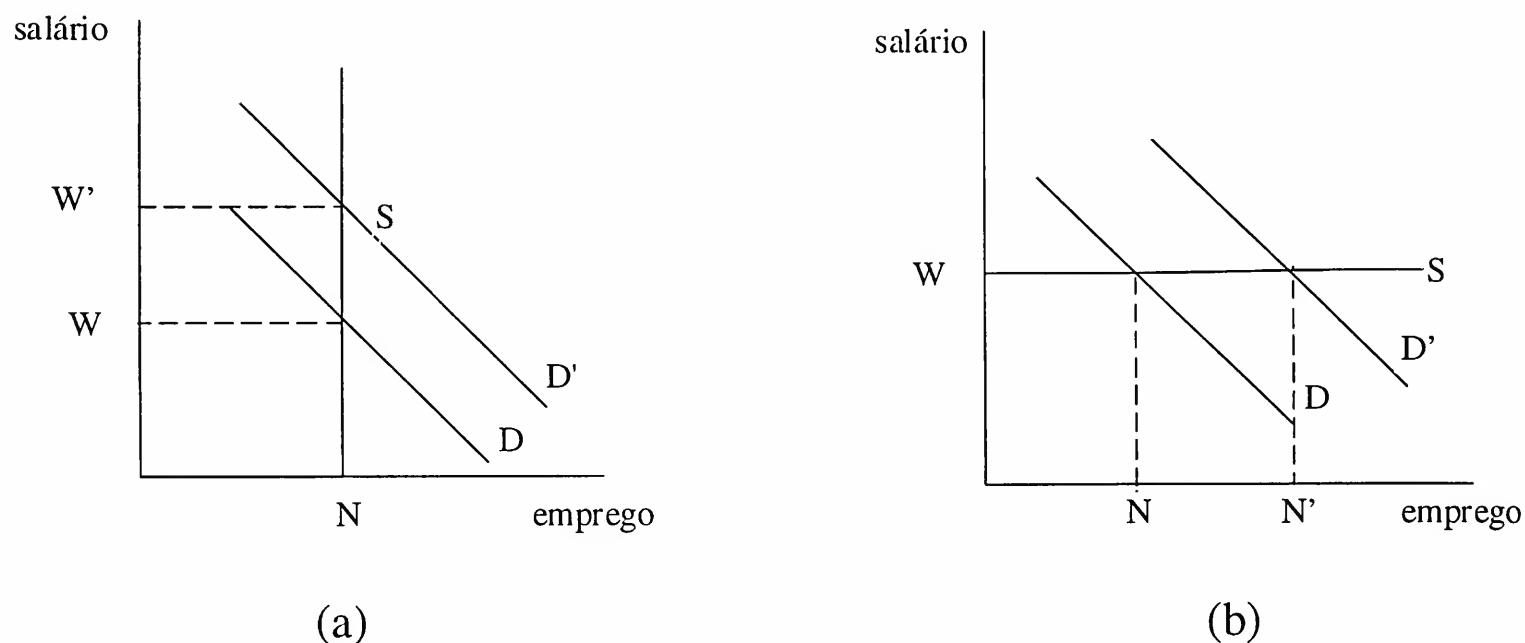
Assim, considerando que a empresa opera de modo a, pelo menos, remunerar o capital, então $\frac{dF}{dg} > 0$. Ou seja, não existiria um ponto de mínimo e quanto menor “g” compensado por um aumento em “a”, menor seria a cunha fiscal. Entretanto, se limitarmos “a” e “g” para serem não negativos, a combinação ótima seria dada por $g = 0$ e $a = \frac{A}{q}$.

Pelo modelo acima o governo, na existência de um imposto sobre a folha de salários, sempre pode reduzir a cunha fiscal e manter sua arrecadação. Isto se dá pela eliminação do imposto sobre a folha de salários e é compensado com o imposto sobre o valor adicionado. Por outro lado, uma redução da cunha fiscal tende a elevar a demanda por trabalho, com impactos positivos sobre o nível de emprego e/ou salários do setor formal da economia, o que, inclusive, tende, num segundo momento, a elevar a própria arrecadação tributária. A forma como esses impactos ocorrem depende da elasticidade da oferta de trabalho. A seguir, passaremos a analisar dois casos polares: quando a oferta de trabalho é perfeitamente inelástica e quando ela é perfeitamente elástica.

A Figura 1.a mostra o caso da oferta de trabalho perfeitamente inelástica. A curva de demanda D representa a situação onde existe um imposto sobre a folha de salários e D’ a situação onde este imposto foi substituído por um aumento na alíquota do imposto sobre valor adicionado. Neste caso, o salário dos trabalhadores se eleva para w' e o nível de emprego se mantém estável. Como o nível de emprego é o mesmo da situação inicial, o nível de produto e de arrecadação tributária também será o mesmo e o aumento de salário dos trabalhadores será proporcional à redução do lucro da firma.

A Figura 1.b representa o caso em que a oferta de trabalho é perfeitamente elástica. Tal situação pode ser pensada, como no modelo de Lewis, para contemplar a existência de abundância de mão-de-obra num setor de subsistência, o qual gera uma renda w . Aqui, o deslocamento da demanda por trabalho no setor formal (de D para D’) não afeta o rendimento dos trabalhadores. Entretanto, como o emprego e o produto aumentam, tanto o lucro como a arrecadação do governo se elevam. Mesmo neste caso poderíamos pensar que a substituição do imposto sobre folha de salários para um imposto sobre valor adicionado é benéfica para os trabalhadores, desde que uma parte do aumento de arrecadação reverta para eles.

Figura 1



Na situação intermediária, onde a elasticidade da oferta de trabalho é maior que zero e menor que infinito, a substituição de um imposto sobre a folha de salários por um imposto sobre o valor adicionado teria como resultado um aumento dos salários, da arrecadação tributária e da parcela de empregados no setor formal da economia. Note que a elevação da demanda por trabalho no setor formal da economia deve ter um impacto positivo nos rendimentos do setor informal e, assim, a situação de todos os trabalhadores seria melhor ao final.

O modelo acima mostra que a substituição do imposto sobre folha de salários por um imposto sobre o valor adicionado melhora a situação dos trabalhadores. Entretanto, é necessário reconhecer que tal modelo é bastante simplificado, contendo apenas uma firma no setor formal com estoque de capital fixo e o mercado de trabalho operando sob regras competitivas. Portanto, seria interessante especular em que medida seria razoável fazer generalizações do resultado obtido, ou se o mesmo está fortemente restrito às hipóteses particulares do modelo.

Uma primeira ressalva que poderia ser feita em relação à análise acima refere-se ao uso da hipótese de competição no mercado de trabalho, ao invés da hipótese de segmentação. Aqui, entretanto, não é possível ver como a hipótese de mercado de trabalho segmentado (formal-informal) possa alterar a conclusão. Admitindo, à semelhança dos modelos de salário eficiência, que a segmentação não altera a condição de equilíbrio da firma, onde a produtividade marginal do trabalho é igual ao custo do trabalho, e considerando que a

análise do impacto da mudança de impostos sobre a cunha fiscal foi realizada no ponto de produto em emprego inicial, a conclusão de que a substituição do imposto sobre folha de salários por um imposto sobre valor adicionado eleva a demanda por trabalho no setor sujeito a tributação não deve ser alterada. Por outro lado, a elevação da demanda no setor formal, que é o setor de salários mais elevados, é benéfica para todos os trabalhadores. Neste sentido, Fernandes (1996) elabora um modelo com mercado de trabalho segmentado entre os setores formal e informal, e conclui que uma redução da cunha fiscal eleva o retorno esperado dos trabalhadores, seja por aumentar os salários nos dois mercados, seja por aumentar a probabilidade de um determinado trabalhador estar empregado no setor formal da economia.

Outra ressalva que poderia ser feita é com relação à hipótese da existência de uma única firma, que é equivalente a supor que as firmas do setor formal são fixas e idênticas. A questão que se coloca é se a existência de firmas com diferentes razões capital-trabalho não pode afetar o resultado.

Para um conjunto de firmas fixas deve existir, para a economia como um todo, uma elevação na alíquota do imposto sobre o valor adicionado capaz de compensar a perda de arrecadação provocada pela eliminação do imposto sobre a folha de salários. Se todas as firmas possuem a mesma razão capital-trabalho, o efeito de tal medida é o descrito no modelo acima. Entretanto, se as firmas possuem diferentes razões capital-trabalho, esta medida deve impor uma elevação na arrecadação tributária das empresas capital intensivo em relação às empresas trabalho intensivo. Assim, para as empresas com baixa razão capital-trabalho a demanda por trabalho deve elevar-se tanto pelo motivo acima exposto como pela redução do volume de impostos pagos. Já para as empresas com elevada razão capital-trabalho o resultado é ambíguo, pois, para estas empresas, o montante de impostos pagos deve se elevar. De qualquer modo, parece mais razoável admitir que a elevação da demanda por trabalho, para a economia como um todo, deve prevalecer, ao invés de diminuir.⁹

Uma outra questão que pode ser levantada diz respeito à evasão de impostos. No modelo acima admitiu-se que as empresas do setor formal pagam corretamente seus impostos, enquanto as empresas do setor informal nada pagam. Entretanto, pode-se argumentar que impostos sobre o valor adicionado são mais fáceis para sonegar, o que pode ter implicações sobre a análise efetuada. Mesmo admitindo a veracidade desta proposição, o resultado da

9 Isto será mais verdadeiro quanto mais inelástica for a demanda por trabalho das empresas capital intensivo (em geral, as grandes empresas) em relação à demanda por trabalho das empresas trabalho intensivo.

troca de base tributária sobre a sonegação seria incerto, pois também é razoável supor que as grandes empresas possuem maior dificuldade de sonegar impostos que as pequenas e, como foi visto, a substituição de um imposto sobre a folha de salários para um imposto sobre o valor adicionado tende a aumentar a participação das grandes empresas no volume de impostos arrecadados.

Por fim, a última ressalva a ser feita na análise acima refere-se à hipótese de curto prazo de que tanto o número de firmas como o estoque de capital da economia são fixos. Note, por exemplo, que na análise do Figura 1.a o ganho dos trabalhadores é proporcional à redução de lucro das firmas. O problema aqui é que uma redução da lucratividade tende a reduzir o estoque de capital de equilíbrio - seja por reduzir o estoque ótimo de cada firma, seja por reduzir o número de firmas -, e a redução do estoque de capital tende a reduzir a demanda por trabalho de longo prazo, não tendo como garantir que a nova curva de demanda se situe acima de D . Ou seja, se o ganho de curto prazo dos trabalhadores ocorrer como uma contrapartida à redução do lucro dos empresários, isto tende a reduzir o estoque de capital no equilíbrio de longo prazo, o que pode acontecer numa magnitude em que a situação dos trabalhadores seria, ao final, pior do que se os impostos continuassem a ser arrecadados sobre a folha de salários.¹⁰ É exatamente esse ponto que passaremos investigar a seguir.

III.2 Um modelo de longo prazo com retornos constantes de escala

Para analisar os efeitos de longo prazo da substituição de impostos adotaremos um modelo que, além de permitir que as empresas ajustem seu estoque de capital e de incluir explicitamente a oferta de trabalho, difere do modelo da subseção anterior em relação à tecnologia e à forma de cobrança dos impostos. Vamos assumir competição perfeita no mercado de trabalho e também que as empresas se defrontam com uma tecnologia que apresenta retornos constantes de escala. Deste modo, não haverá lucro econômico e o estoque de capital ótimo da economia pode se dar com qualquer número de firmas.

10 Minha posição inicial em relação a esse ponto era de que, embora teoricamente possível, as condições necessárias para que a troca de um imposto sobre folha de salários por um imposto sobre valor adicionado provocasse, via redução do estoque de capital, uma piora da situação dos trabalhadores eram muito especiais e que, portanto, isso seria muito pouco provável de acontecer. Eu acreditava que um modelo simples de curto prazo, levando-se em consideração apenas os aspectos da demanda por trabalho, seria suficiente para tratar da questão. Hoje, entretanto, não me parece mais tão simples sustentar que a troca de impostos eleva o bem-estar dos trabalhadores. Para ter uma maior clareza acerca desses problemas foi de fundamental importância os comentários recebidos de Samuel Abreu Pessôa, Ricardo Paes de Barros e de um dos pareceristas anônimos desta revista.

Quanto à cobrança de impostos, o presente modelo possui duas diferenças em relação ao modelo anterior. A primeira é que aqui estamos admitindo que o volume de impostos arrecadado pelo governo é fixo, independentemente do nível de produto e salário.¹¹ A segunda é que o imposto sobre salário será cobrado do trabalhador ao invés da empresa. Essa última diferença é apenas formal, uma vez que ela em nada altera os resultados. As firmas pagam aos trabalhadores um salário bruto (w_b), sobre o qual incide o imposto sobre o trabalho. Os trabalhadores estão interessados no salário líquido (w), onde $w_b (1-g) = w$.

O estoque de capital ótimo se dará no ponto onde a produtividade marginal do capital se iguala ao preço do serviço do capital, dividido por um menos a alíquota do imposto sobre o valor adicionado, enquanto a quantidade de trabalho ótima se dará no ponto onde a produtividade marginal do trabalho é igual ao salário bruto dividido por um menos a alíquota do imposto sobre o valor adicionado. O preço do serviço do capital é considerado constante, refletindo uma taxa de juros de *steady state* determinada, por exemplo, pelas preferências intertemporais dos agentes. Assim, no equilíbrio das firmas temos:

$$(9) \quad (1-a)\phi'(k^*) = p$$

$$(10) \quad (1-g)(1-a)[\phi(k^*) - k^*\phi'(k^*)] = w \quad \text{para } a \text{ e } g < 1$$

onde,

$$f(k,n) = n\phi(k^*);$$

$$k^* = \frac{k}{n} = \phi'^{-1}\left(\frac{p}{1-a}\right) e;$$

$$\frac{dk^*}{da} = \frac{\phi'}{\phi''} \frac{1}{1-a} < 0$$

Vamos admitir, também, que a oferta de trabalho é crescente com o salário líquido e possui uma elasticidade constante, mais especificamente:

¹¹ No modelo anterior a arrecadação foi considerada dada apenas para o nível de produto e salário inicial.

$$(11) \quad n = w^\varepsilon \quad \text{para } \varepsilon > 0 \quad \text{onde;}$$

$\varepsilon =$ elasticidade da oferta de trabalho

Para finalizar a especificação do modelo fica faltando a equação de arrecadação do governo, a qual é dada por:

$$(12) \quad a\phi(k^*)n + g(1-a)[\phi(k^*) - k^*\phi'(k^*)]n = A \quad \text{ou,}$$

$$(12') \quad (1-g) = \frac{(1-a)[\phi(k^*) - k^*\phi'(k^*)]n - A + a\phi(k^*)n}{(1-a)[\phi(k^*) - k^*\phi'(k^*)]n}$$

Uma vez que estamos interessados no bem-estar dos trabalhadores, o problema a ser resolvido é o de encontrar a combinação de alíquotas (a e g) que gera a arrecadação "A" e maximiza o montante de salários da economia. Como "w" e "n" possuem uma relação crescente, maximizar a massa salarial da economia é o mesmo que maximizar o nível de emprego.

Substituindo (12') em (10) e utilizando-se de (11) obtemos:

$$(13) \quad n = \left\{ \frac{(1-a)[\phi(k^*) - k^*\phi'(k^*)]n - A + a\phi(k^*)n}{n} \right\}^\varepsilon \quad \text{ou,}$$

$$(13') \quad n^{\frac{1+\varepsilon}{\varepsilon}} + A = [\phi(k^*) - k^*p]n$$

Diferenciando (13') dos dois lados e com alguma álgebra obtemos:

$$(14) \quad \frac{dn}{da} = \frac{X}{Z} \quad \text{onde,}$$

$$X = n[\phi'(k^*) - p] \frac{dk^*}{da} \quad \text{e;}$$

$$Z = \frac{1+\varepsilon}{\varepsilon} n^{\frac{1}{\varepsilon}} - \phi(k^*) + k^* p$$

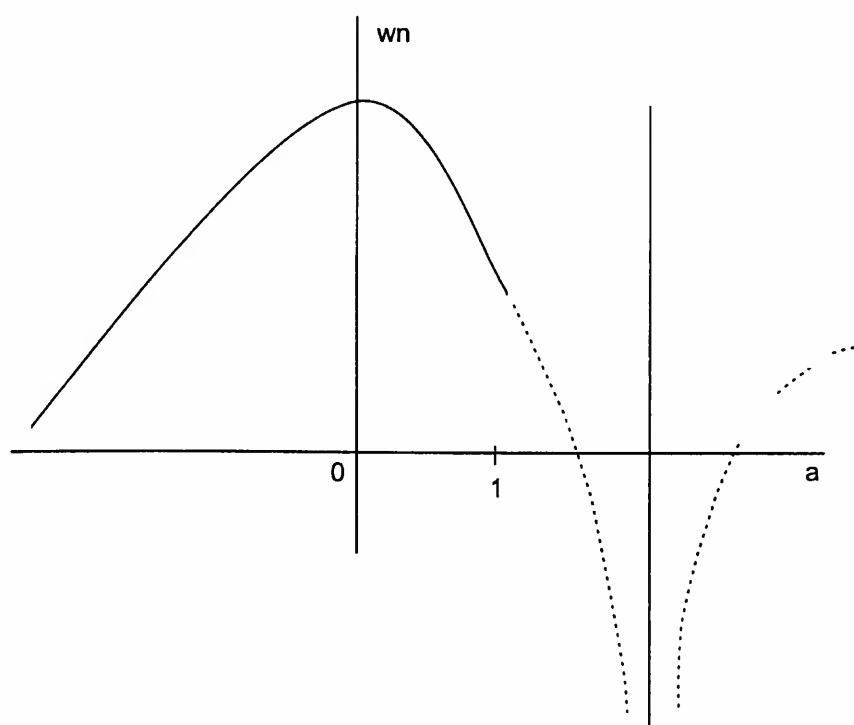
Vale notar que quando $a = 0$, $\phi(K^*) = p$ e, portanto, $X = 0$. Deste modo, $a = 0$ é um valor crítico de a , que gera um ponto de máximo se $Z > 0$ e um ponto de mínimo caso $Z < 0$. Passaremos, então, a estudar o sinal de Z no ponto em que $a = 0$. No ponto em que $a = 0$ temos:

$$(15) \quad n^{\frac{1}{\varepsilon}} = \frac{[\phi(k^*) - k^* p]n - A}{n} \quad \text{então,}$$

$$(16) \quad Z = \frac{[\phi(k^*) - k^* p]n - (1+\varepsilon)A}{\varepsilon n} = \frac{wn - \varepsilon A}{\varepsilon n}$$

Assim, o sinal de Z fica dependendo do sinal de $\left(\frac{wn}{\varepsilon} - A\right)$. No caso em que a massa salarial da economia dividida pela elasticidade da oferta de trabalho for maior que arrecadação tributária do governo, no ponto onde $a = 0$, esse será um ponto de máximo da massa salarial. Caso contrário, esse será um ponto de mínimo.

Figura 2



A Figura 2 apresenta o caso onde a massa salarial assume um ponto de máximo quando $a = 0$. Enquanto "A" e " ϵ " são parâmetros do modelo, " wn " é a variável que se deseja maximizar. À medida que se vai aumentando "a" a massa salarial vai se reduzindo. É possível que a partir de um momento uma elevação de "a" inverta o sinal de "Z" e, portanto, a massa salarial volta a crescer com aumentos de "a". No exemplo da Figura 2 isso ocorre para $a > 1$ e, portanto, $a = 0$ gera um máximo global.

Para uma dada elasticidade da oferta de trabalho, quanto menor for "A" mais provável será que $a = 0$ gere um ponto de máximo. E para uma dada arrecadação do governo, quanto mais elástica for a oferta de trabalho mais provável será que $a = 0$ minimize a massa salarial da economia. Em relação à elasticidade da oferta de trabalho, vale notar que quando $\lim_{\rightarrow \infty} \epsilon < 0$ e, deste modo, $a = 0$ minimiza a massa salarial da economia. Por outro lado, quando $\lim_{\rightarrow 0} \epsilon > 0$, $Z > 0$ e, do ponto de vista dos trabalhadores, é preferível que toda arrecadação seja realizada sobre o trabalho.

A análise acima chama a atenção para a importância da elasticidade da oferta de trabalho na determinação da combinação ótima de alíquotas de imposto, quando se tem em vista o bem-estar dos trabalhadores.¹² Vários estudos internacionais têm procurado estimar a elasticidade da oferta de trabalho.¹³ Pencavel (1986) faz um *survey* extensivo dos trabalhos que estimam a elasticidade da oferta de trabalho para homens, contemplando diversos países e períodos, com diferentes técnicas de estimação. Ele mostra que esses estudos convergem para um valor da elasticidade próximo de zero. Se o fato de se admitir uma oferta de trabalho perfeitamente inelástica constitui uma boa hipótese, o resultado do modelo acima é que, do ponto de vista dos trabalhadores, o melhor seria que toda arrecadação tributária tivesse como base o salário ao invés do valor adicionado. Entretanto, tal resultado está sujeito a uma série de qualificações.

Em primeiro lugar, esses resultados empíricos limitam-se apenas à oferta de trabalho de homens adultos não aposentados e não a oferta de trabalho total. Por exemplo, no caso da elasticidade da oferta de trabalho para mulheres os resultados obtidos são, em média, maiores e bem mais divergentes.¹⁴

12 É interessante observar que, de modo geral, a taxa ótima será tudo sobre o trabalho ou tudo sobre o valor adicionado. Uma combinação de alíquotas com "a" e "g" diferentes de zero dificilmente ocorrerá.

13 Não se encontrou qualquer estudo sobre a elasticidade de oferta de trabalho para o Brasil.

14 Killingsworth e Heckman (1986) reportam valores para estimativas da elasticidade da oferta de trabalho de mulheres que vai de -0.89 a 14.79.

Um segundo aspecto refere-se ao fato que tais estudos sobre a elasticidade da oferta dizem respeito à relação entre salários e horas de trabalho por pagamento no mercado. Como no nosso caso o interesse recai sobre a oferta de trabalho do setor formal da economia, o setor sujeito a tributação, é de se esperar que a oferta de trabalho do setor formal seja mais elástica do que a oferta de horas de trabalho por pagamento no mercado de trabalho.

Por fim, mesmo supondo uma oferta de trabalho fixa, o resultado acima pode estar fortemente dependente de hipóteses muito restritivas do modelo: retornos constante de escala e existência de apenas dois fatores de produção (capital e trabalho) como o preço do serviço do capital fixo no longo prazo. Incluir uma oferta de trabalho fixa implica, nesse modelo, que o produto fica dependendo apenas do estoque de capital da economia. Nesse caso a equação (9) poderia ser reescrita como:

$$(9') \quad (1-a)f'(k) = p \quad \text{onde,}$$

k = estoque de capital da economia

A massa salarial da economia seria $wn = f(k) - kp - A$. Assim, a questão reduz-se a encontrar o estoque de capital que maximiza a massa salarial, sendo que k fica dependendo apenas de "a". A condição de primeira ordem para maximizar a massa salarial é: $f'(k) = p$. Note, por (9'), que essa condição ocorre quando $a = 0$.

O problema surge quando um terceiro fator de produção é incluído na análise. Vamos supor que esse terceiro fator seja também fixo, como, por exemplo, terra ou talento empresarial. Neste caso teríamos $r = f(k) - kp - A$ onde r é o montante de recursos destinado ao pagamento dos fatores fixos. Note que $a = 0$ é, ainda, a condição que maximiza r . Entretanto, como a combinação de alíquotas define não apenas o montante de recursos a ser destinado aos fatores fixos, mas, também, a distribuição desses recursos entre os dois fatores fixos, não é mais possível afirmar que $a = 0$ é a melhor situação para os trabalhadores.

III.3 Um modelo de longo prazo com oferta de trabalho e número de firmas fixos

Para melhor analisar esse ponto, vamos admitir, a exemplo da subseção III.1, que as empresas operam com retornos decrescentes de escala, sendo $q = f(k, n)$ e o preço do

serviços do capital, p , é dado. Cada firma pode alterar o seu estoque de capital, mas o número de firmas no setor formal da economia é fixo. Assim, existiria na economia uma dotação de empresários formais, que pode estar representando talento empresarial, recursos ou outros fatores que estão ausentes nos trabalhadores de modo a impossibilitá-los de tornarem-se um empresário. O lucro econômico das firmas seria a renda desses empresários. Caso os empresários sejam homogêneos, podemos admitir a existência de uma única firma operando no setor formal da economia.

Admitindo que o número de trabalhadores é dado, podemos notar que a economia possui dois tipos de capital humano com oferta fixa: um que torna os indivíduos trabalhadores e outro que os tornam empresários.

As condições de primeira ordem para maximização de lucro da firma são:

$$(17) \quad (1-a)f_n - (1+g)w = 0$$

$$(18) \quad (1-a)f_k - p = 0 \quad \text{onde,}$$

$$f_{nn} < 0 ; f_{kk} < 0 \text{ e } f_{nn} \cdot f_{kk} > f_{nk}^2$$

O jacobiano deste sistema é $|J| = (1-a)^2 [f_{nk}^2 - f_{kk} f_{nn}] < 0$ e as derivadas parciais da demanda por trabalho, n^d , em relação às variáveis exógenas para firma (w , a e g) são:

$$(19) \quad \frac{\partial n^d}{\partial w} = \frac{-(1-a)(1+g)f_{kk}}{|J|} < 0$$

$$(20) \quad \frac{\partial n^d}{\partial a} = \frac{(1-a)[f_{nk} \cdot f_k - f_{kk} \cdot f_n]}{|J|} < 0 \quad \text{para } f_{nk} > 0$$

$$(21) \quad \frac{\partial n^d}{\partial g} = \frac{-(1-a)f_{kk} \cdot w}{|J|} < 0$$

Assim, a demanda por trabalho reduz-se com a elevação de "a", "g" e "w". Como estamos interessados no bem-estar dos trabalhadores, a questão a investigar é o que acontece com os salários quando, para uma dada arrecadação do governo, ocorre uma substituição de impostos. O sistema a ser resolvido é:

$$(22) \quad n^d - \bar{n} = 0$$

$$(23) \quad g \cdot w \cdot \bar{n} + a \cdot f(k, n) - \bar{A} = 0$$

A equação (22) representa o equilíbrio do mercado de trabalho e a equação (23) representa a restrição orçamentária do governo. As variáveis endógenas do sistema são

"w" e "g" e o jacobiano é dado por $|J| = \frac{-f_{kk} \cdot w \cdot \bar{n}}{(1-a)[f_{nk}^2 - f_{kk} \cdot f_{nn}]} < 0$

Para avaliarmos o impacto da substituição do imposto sobre trabalho pelo imposto sobre valor adicionado temos que analisar a derivada de "w" em relação a "a", que é dada por:

$$(24) \quad \frac{dw}{da} = \frac{[f_{kk} \cdot f_n - f_{nk} \cdot f_k] \cdot w \cdot \bar{n} - f(k, n) \cdot f_{kk} \cdot w}{(1-a)[f_{nk}^2 - f_{kk} \cdot f_{nn}] |J|}$$

Caso $\frac{dw}{da} > 0$, para todo "a" e "g", toda substituição de imposto do trabalho para imposto sobre o valor adicionado seria benéfica aos trabalhadores. Se limitarmos "a" e "g" para serem não negativos, a combinação ótima de alíquotas, do ponto de vista dos trabalhadores, seria, novamente, $g = 0$ e $a = \frac{A}{q}$. Para que essa condição seja satisfeita é necessário que o numerador em (24) seja positivo, o que implica:¹⁵

$$(25) \quad \frac{f(k, n)}{n} - f_n > -\frac{f_{nk}}{f_{kk}} \cdot f_k$$

¹⁵ Note que $f_{kk} < 0$

A condição (25) não é de interpretação simples e depende, fundamentalmente, da forma específica da função de produção. Admitindo que a função de produção seja do tipo Cobb-Douglas com retornos decrescentes de escala, ou seja, $q = n^\alpha k^\beta$, com $\alpha + \beta < 1$, essa condição seria:

$$(25') \quad (1 - \alpha) > -\frac{\alpha\beta}{(\beta - 1)}$$

É fácil verificar que (25') está atendida. Assim, admitindo uma função de produção Cobb-Douglas com retorno decrescente de escala, seria melhor para os trabalhadores que toda arrecadação de impostos tivesse como base o valor adicionado.

Por fim, a proposição de que a substituição dos impostos sobre folha de salários por um imposto sobre o valor adicionado melhora a situação dos trabalhadores está fortemente dependente de uma série de hipóteses, em especial acerca da tecnologia das empresas e das elasticidades de oferta dos insumos de produção, as quais não nos parece muito simples para uma tomada de posição *a priori*.

IV Considerações finais

Este artigo tem por objetivo realizar uma reflexão teórica acerca do debate do peso dos encargos sociais no Brasil. Ele surge em virtude de uma insatisfação com o modo pelo qual esse debate vem sendo conduzido no País. Os critérios utilizados para se avaliar a questão não são claramente definidos e a produção de uma disparidade de números tem contribuído para trazer mais confusão do que luz ao problema.

Uma vez que a principal justificativa para a existência da maior parte dos itens freqüentemente classificados como encargos sociais é a “proteção social” dos trabalhadores, utilizou-se como critério de avaliação o bem-estar dos trabalhadores. Com isso acredita-se dar maior clareza e relevância social ao debate, além de evitar desvios da discussão com aspectos secundários e muitas vezes desnecessários.

Toda a análise foi realizada tendo como perspectiva uma economia fechada. De modo geral, um item de encargo que possa ser considerado positivo para os trabalhadores em uma economia fechada será também positivo em uma economia aberta. Por outro lado, distorções no mercado de trabalho, causadas pela legislação trabalhista, que afetem

negativamente os trabalhadores, devem ser combatidas sempre e não apenas em virtude da abertura comercial. A inclusão de aspectos de comércio internacional no debate sobre encargos sociais tem sido, via de regra, pouco cuidadosa, limitando-se a comparar o custo da mão-de-obra entre o Brasil e o resto do mundo. Tal comparação, além de ser de difícil realização, tem levado o debate para um campo “espinhoso” sobre competitividade internacional, num exemplo claro do que considero um desvio do foco da discussão.¹⁶

Para avaliar o impacto da legislação trabalhista sobre o bem-estar dos trabalhadores dois pontos foram destacados. O primeiro diz respeito à distinção entre benefícios existentes em função da legislação e benefícios garantidos pela legislação. Um benefício garantido pela legislação poderia existir extensivamente no mercado independentemente da mesma e, assim, essa legislação teria pouco impacto no funcionamento do mercado de trabalho. Uma proposta para se avaliar, ao menos para os trabalhadores menos qualificados, os benefícios que existiriam independentemente da legislação seria utilizar os benefícios normalmente concedidos no setor informal da economia. O segundo refere-se à necessidade de dividir os encargos entre aqueles que providenciam um benefício direto aos trabalhadores e aqueles que possuem muito mais um caráter de impostos, sendo utilizados para financiar programas mais gerais.

O primeiro grupo inclui itens como SESI/SESC; indenização por demissão sem justa causa; seguro-acidente; e o FGTS. Neste caso a questão seria avaliar se esses benefícios custam mais do que os trabalhadores os avaliam. Em caso afirmativo, uma melhora no bem-estar dos trabalhadores poderia ser alcançada substituindo tais benefícios por um aumento nos salários, correspondente aos seus custos. Uma proposta para avaliar esse hiato é investigar, para os trabalhadores formais, o quanto seria necessário elevar os salários para compensar a perda dos benefícios existentes em função da legislação. Este último ponto refere-se a um problema de disposição a receber e sobre o qual existe uma vasta literatura.¹⁷

Neste primeiro grupo poderíamos incluir, ainda, o vale-transporte; as férias de 30 dias;

16 Pastore (1994), por exemplo, considera que os encargos sociais elevam demasiadamente o custo do trabalho e, assim, reduzem a “competitividade” do país. Ele enxerga a relação entre países como se fosse uma competição entre empresas, não levando em consideração a literatura padrão sobre comércio internacional. A tese da competitividade das nações tem sido severamente combatida por Krugman (1996) no que ele denomina **Interacionalismo Pop**. O interessante é que grande parte dos críticos de Pastore aceitam a tese de competitividade internacional, discordando apenas da importância dos encargos como forma de encarecer demasiadamente o custo do trabalho no Brasil. Para uma discussão que envolve a questão de custos e comércio internacional ver Pessôa (1997).

17 No Brasil, uma interessante discussão e aplicação do Método de Avaliação Contingente é fornecida por Belluzzo (1995).

o 13º salário e o adicional de 1/3 das férias. Entretanto, seria necessário verificar primeiro se tais benefícios não são extensivamente utilizados no setor informal da economia. No caso do 13º e do adicional de férias, como são benefícios em dinheiro, não seria de se esperar um diferença significativa entre custos e benefícios de tais itens, a qual poderia surgir em virtude dos mesmos não estarem homogeneamente distribuídos ao longo do ano e a taxa real de juros ser diferente de zero.

O segundo grupo inclui itens como salário-educação; SEBRAE; SENAI/SENAC; e, em grande medida, o INSS.¹⁸ Aqui teríamos que avaliar se esses programas devem ser mantidos com recursos públicos na forma que estão, ou devem ser reformados ou, mesmo, eliminados. Caso devam ser mantidos, teríamos, ainda, que discutir a forma de arrecadação desses recursos. Uma proposta recorrente é a de transferir a arrecadação desses recursos da folha de pagamentos para o valor adicionado. Esta proposta foi analisada na seção III deste artigo e os resultados não são claros. A possibilidade de melhorar as condições de vida dos trabalhadores via essa substituição é fortemente dependente de uma série de hipóteses, em especial acerca da tecnologia das empresas e das elasticidades de oferta dos insumos de produção.

Existe entre os empresários e entre parte dos analistas econômicos brasileiros um sentimento de que os chamados encargos sociais possuem um peso excessivo, sendo uma das principais explicações para o enorme setor informal que se verifica no País.¹⁹ Este artigo não apresenta uma resposta a esta questão. Ele se limita a realizar uma reflexão crítica sobre os termos nos quais esse debate vem se desenvolvendo e a apontar alguns caminhos que investigações futuras poderiam trilhar.

O primeiro desses caminhos insere-se, ainda, no âmbito da idéia de peso dos encargos sociais. Qualquer avaliação dos custos do trabalho no Brasil mostra que uma parcela significativa deles não faz parte do que os trabalhadores formais levam em dinheiro mensalmente para suas casas. O que os trabalhadores recebem mensalmente em dinheiro é, no máximo, 92% do seu salário contratual, sendo o restante custos para financiar programas sociais ou outros benefícios, incluindo aí o 13º salário, o adicional de férias e as férias de

18 O INSS é um caso híbrido, pois ele fornece um benefício direto para o contribuinte, provavelmente menor que o custo, mas dado o sistema interno de redistribuição e pagamento de benefícios para idosos que pouco ou nunca contribuíram ele possui uma caráter de um programa social mais geral. Assim, uma parte das contribuições para o INSS é tipicamente seguro, outra parte é tipicamente imposto para financiar gastos sociais

19 Hoje menos da metade dos trabalhadores brasileiros possui emprego com carteira de trabalho assinada.

30 dias. Dos benefícios teríamos que investigar quais deles são impostos pela legislação e quais existiriam independentemente dela. Para tanto, como já mencionado, poderiam ser utilizados os benefícios concedidos no setor informal da economia. Os custos dos benefícios impostos pela legislação mais aqueles para financiar os programas sociais formariam os custos do trabalho impostos pela legislação. Tais custos menos o quanto os trabalhadores avaliam os benefícios impostos pela legislação formaria o que poderíamos chamar de peso dos encargos sociais. Aqui, o peso dos encargos sociais possui um significado preciso, ele é, para efeitos do funcionamento do mercado de trabalho, equivalente a um imposto sobre o trabalho e, portanto, fornece uma medida dos incentivos que empregados e empregadores possuem em transformar suas relações de trabalho de formal para informal. Por outro lado, o peso dos encargos sociais, assim definidos, mais o custo dos demais impostos sobre as empresas fornecem uma medida dos incentivos das firmas de caminharem totalmente para a ilegalidade.

O segundo caminho diz respeito às reformas visando desonerar a folha de pagamentos. O artigo aponta para a necessidade de se avaliar separadamente cada um dos itens incluídos nos chamados encargos sociais. No caso dos benefícios diretos é necessário comparar os custos com o quanto os trabalhadores os avaliam. Isto deve ser feito benefício por benefício. Para os programas sociais é necessário analisar se eles justificam o montante de recursos públicos utilizados. Em relação à substituição da fonte de financiamento desses programas, da folha de pagamentos para o valor adicionado, o artigo mostra que, do ponto de vista teórico, não é uma proposição tão simples de ser defendida, indicando a necessidade de novos estudos sobre a questão, principalmente estudos empíricos. Seria importante encontrar formas de comparar emprego e salários entre firmas com diferentes composições da carga tributária. Como no Brasil não existem bases de dados com informações sobre arrecadação e emprego por firmas, estudar melhor as experiências internacionais pode ser uma alternativa, em especial para os países onde a legislação tributária possui uma maior diversidade entre os estados.

Na subseção III.2 vimos que no caso em que os fatores de produção são capital e trabalho e a oferta de trabalho é suficientemente inelástica, o melhor para os trabalhadores seria que toda a arrecadação fosse sobre a folha de salários. Entretanto, se mais um fator de produção com oferta inelástica é considerado (subseção III.3) esse resultado pode ser totalmente invertido. Se pensarmos a função de produção da economia contendo vários tipos de mão-de-obra com qualificações diferenciadas, onde talento empresarial pode ser considerado um tipo de qualificação, e se estivermos interessados primordialmente com os trabalhadores menos qualificados, substituir os impostos desses trabalhadores por impostos sobre os demais tipos de mão-de-obra, ou sobre o valor adicionado, poderia se mostrar

uma medida eficaz. Isto poderia se dar reduzindo a tributação dos trabalhadores com menores salários ou, alternativamente, das firmas que os empregam. Como os trabalhadores pouco qualificados estão super-representados nas pequenas empresas, medidas como o SIMPLES poderiam atuar nesse sentido.

Referências bibliográficas

Almeida, S. C. F. *As contribuições sociais de empregados e trabalhadores: repercussões sobre o mercado de trabalho e grau de evasão*. IPEA, Documento de Política nº 8, fevereiro, 1992.

Amadeo, E. J. O real peso dos encargos sobre salários. *Folha de São Paulo*, 04 de janeiro, 1994a.

_____. O consenso sobre encargos trabalhistas. *Folha de São Paulo*, 06 de junho, 1994b.

Belluzzo, W. J. *Valorização de bens públicos: o método de avaliação contingente*. Dissertação de Mestrado, São Paulo: Departamento de Economia da USP, 1995.

Fernandes, R. Contratos informais de trabalho: uma análise a partir da experiência brasileira. *Anais do XXIV Encontro Nacional de Economia - ANPEC*, Águas de Lindóia, SP, dezembro de 1996.

Hart, R. A. *The economics of non-wage labour costs*. London: George Allen and Unwin, 1984.

Killingsworth, M. R. & Heckman, J. J. Female labor supply: a survey. *In: Ashenfelter, D. & Layard, R (eds.), Handbook of labor economics*. North-Holland, 1986.

Krugman, P. *Pop internationalism*. MIT Press, 1996.

Pastore, J. *Flexibilização dos mercados de trabalho e contratação coletiva*. São Paulo: Ed. Ltr, 1994.

Pencavel, J. Labor supply of men: a survey. *In: Ashenfelter, D. & Layard, R. (eds.), Handbook of labor economics*. North-Holland, 1986.

Pessoa, S. A. *O Impacto da redução do custo Brasil sobre a defasagem cambial*. IPE-USP, Programa de Seminários Acadêmicos nº 08/97, 1997.

Santos, A. L. Encargos sociais e custo do trabalho no Brasil. *In: Oliveira, C. E. B. & Mattoso, J. E. L. (orgs.), Crise e trabalho no Brasil: modernidade ou volta ao passado?* São Paulo: Ed. Página Aberta Ltda, 1996.

An econometric analysis of the effects of macroeconomic variables on interest rates in Brazil

Fernando Blumenschein[§]

ABSTRACT

This paper examines the importance of liquidity, credit, and Fisher effects on the interest rate in Brazil during the period 1975-1991, for five different segments of the financial market. The positive relationship between money and the interest rate prevailed in most of our results. This finding may represent a departure from both monetarism and rational expectation approaches. Price had a significant effect in the way money affected interest rate, especially under high rates of inflation. Some of our results confirmed the existence of segmentation among different segments of the financial market. Mainly the overnight market reacted differently. The implications of these results for the monetary transmission mechanism, and for the conduct of monetary policy in Brazil during the period 1975-1991, may have been substantial.

Key words: interest rate, money supply, credit, Fisher effect, Brazil.

RESUMO

Este artigo examina a importância do efeito liquidez, do efeito crédito e do efeito Fisher sobre a taxa de juro no Brasil durante o período 1975-1991 para cinco segmentos do mercado financeiro. Uma relação positiva entre a oferta de moeda e a taxa de juro prevaleceu na maioria das análises efetuadas. Este resultado pode representar uma modificação das previsões esperadas pelas correntes monetaristas e das expectativas racionais. O nível de preços teve um efeito significativo na maneira como a oferta monetária afetou a taxa de juro, especialmente quando prevaleceu altas taxas inflacionárias. Alguns resultados confirmaram a existência de segmentação entre diferentes segmentos do mercado financeiro. Principalmente o mercado overnight reagiu de forma diferenciada. As implicações destes resultados para o mecanismo de transmissão monetária e para a condução da política monetária no Brasil durante o período 1975-1991 podem ter sido substanciais.

Palavras-chave: taxa de juro, oferta de moeda, crédito, efeito Fisher, Brasil.

§ Professor of economics at the Faculdade de Economia e Administração (FEA) of the University of São Paulo (USP), Brazil.

I wish to thank Barclay Jones, Tom Davis, and Erik Thorbecke from Cornell University and two anonymous referees for comments and suggestions in an early version. The views expressed in this paper are those of the author.

Recebido em janeiro de 1998. Aceito em julho de 1998.

1 Introduction

The price of money is perhaps the most important of all prices. Nevertheless, a historical view of the theories of the interest rate shows how controversial and continual the debate on interest rate determination has been.¹ Although the complexity of the theme is indisputable, some justifications have been given for the existence of that controversy. Most of the controversy arises from the lack of explicit recognition of three important factors that underlie both the empirical and the theoretical studies of the interest rate.

The first factor is methodological. It is related to the type of approach used in the analysis of the interest rate. The choices between partial equilibrium versus general equilibrium, or static versus dynamic model, involve some judgments that must be made explicitly. Each framework precludes some insights due to its limitations. The second factor is related to time. Innocuous dispute may stem from the conflict between variables that explain short-run phenomena and those that explain long-run relations. The third, and less explored factor, relates to the type of economic system being investigated. Some countries present distinct features on their economic system that usually is not taken into consideration. This is the case of many developing countries in which the characteristics of their financial markets are, in many ways, distinct from those of the more developed nations. These characteristic features may bear considerable influence on the behavior of the interest rate.

Investigation about the factors influencing the rate of interest in economies of the developing countries has been the subject of numerous studies.² These studies have shared a common concern with respect to the high market real rates of interest that prevailed in those countries during most of the 1980s. On the other hand, other studies have found that the government real rates of interest have been negative for most developing countries.³ This asymmetric behavior between government rates and market rates seems to have been characteristic of many developing countries.

-
- 1 For a historical and analytical view of theories and models of the interest rate determination, see Lutz (1968) as well as Conard (1959).
 - 2 Sundararaja and Balino (1991) reviewed the studies on the interest rate for seven developing countries in Asia and South America. For a study of the interest rate in Mexico, see Khor and Rojas-Suarez (1991).
 - 3 See Guidotti and Kumar (1991), for an analysis of the interest rate in fifteen developing countries. In their text, government rates include the deposit rate, the discount rate, and the interest rate on government securities.

For the case of Brazil, a study that implemented correlations and cointegration analyses has shown that the link between interest rates, in different segments of the market, has been very weak, both in the short run and in the long run.⁴ For this reason, a thorough investigation of factors influencing the interest rates should consider each segment of the financial market separately. This fact contrasts with studies of interest rate determination in the developed countries, where a perfectly integrated financial market allows the impact of a given variable to have a roughly uniform effect in all segments. Moreover, the Brazilian financial market has some features which distinguishes it from the majority of the developing countries. On the one hand, it has some characteristic features, which are encountered in repressed financial markets. These characteristic features include curb market, capital controls, and currency substitution, among others. On the other hand, Brazil has a very developed banking system that manages the great majority of the financial intermediation, a sophisticated primary and secondary markets for government securities, and a fast growing stock market have been in place for more than two decades.⁵

Because of the degree of sophistication attained by its financial market, it is difficult to include Brazil in the category of financially repressed economies of the McKinnon-Shaw tradition. The study of the interest rate in Brazil has to be carried out within the context of the more traditional theoretical frameworks.⁶ This traditional framework includes three theoretical approaches. First, the Keynesian/Monetarist approach of liquidity preference (Liquidity Effect). This approach allows us to measure the effect of changes in the money supply and income on the level of the interest rates. Second, the theory of loanable funds which gives emphasis to the supply and demand for credit as the determinants of the interest rate (Credit Effect). Third, the Fisher approach, which emphasizes the influence of changes in the price level on the interest rate (Fisher Effect).⁷

The main objective of this paper will be to evaluate how the three macroeconomic approaches we mentioned above relate to the behavior of the interest rate in Brazil during the period 1975-1991, for five different segments of the financial market. The paper intends to investigate the issue through the implementation of econometric analyses.

4 See Blumenschein (1995), for an analysis of the interest rate in Brazil.

5 A detailed description of the sophistication of the Brazilian financial market was given in Lees, Boots and Cysne (1991).

6 In our text, the term "traditional" excludes, for example, the theories of term structure of interest rate, which do not apply to Brazil, and excludes real business cycle models.

7 See Friedman and Schwartz (1982) for a description of the theoretical background of both the liquidity effect and the credit effect. For the Fisher effect, see Fisher (1930).

Despite the fact that the effects of some other specific variables may help to explain short-run movements of the interest rate, a broader understanding of the behavior of the interest rate can be achieved only by considering the effect of those macroeconomic variables over the long run. The motivation to study the interest rate in Brazil derives from the unique features of its financial market that we have described, and from the lack of previous studies to investigate the issue.⁸

In the Brazilian case, the study has to take into consideration the various segments of the formal financial market, for reasons we outlined above. As a consequence of the asymmetric behavior of the interest rate among the various segments that comprise the formal financial market, our basic hypothesis is that the macroeconomic variables may influence different segments in different ways. Under these circumstances, the use of interest rate as an instrument of monetary policy in Brazil may be questioned. The paper will be structured as follows. The data we used and the features of the five different segments of the financial market will be briefly explained in section 2. Section 3 will outline the econometric techniques we applied, and the results we found. Some complementary discussion and interpretation of the results will be presented in section 4. The paper closes with concluding remarks in section 5.

2 Data and characteristics of the markets

The time series for both government and market interest rates have been constructed and made public only recently. This paper has taken advantage of these recent data, which allowed us to cover the period from 1975 to 1991.⁹ We will investigate five segments of the financial market. These segments are the primary and the secondary market for government bonds, and three segments for short-term borrowing and lending in commercial banks.

For the primary market for government bonds we used the index of monthly quotations for national treasury bonds (TB), published by the Ministry of the Economy (see Apecão (1991)). This index is constructed by taking into consideration the yield of government bonds in the primary market. Since there are several different types of bonds and the

⁸ For one of the few studies about the interest rate in Brazil, see Garcia (1991).

⁹ The period 1975-1991 has been chosen because of the availability of the data at the time that this paper had been written.

denomination of these bonds have changed over time, the use of the official index mentioned above seemed more appropriate. More importantly, this index is used as the basis for indexation of other segments of the financial market. This fact makes this index a representation not only of the interest rate paid in the primary market for government bonds, but also in some segments of the financial market that are subject to monetary correction (i. e. agricultural credit, household savings, and mortgages, among others).

For the secondary market, we used the interest rate paid on the overnight market (NG). This market is a repurchase contract market, which pays overnight interest rates. The holders of government securities use this market to obtain capital by using these securities as collateral.

The time series for market interest rates were the borrowing and the lending rates that prevailed in the banking system. The borrowing rate was the one the banking system paid on short-term certified deposits (CDB). Two lending rates were used to measure the cost of money for the private sector in the short term. One was the short-term rate for working capital (CG) and, the other, the rate used to discount private notes (DD), issued in short-term transactions. It is important to add that the interest rates in these segments are the result of market forces, since the levels of interest rate paid and charged are not subjected to governmental regulation. Moreover, these interest rates are good proxies for the curb market. The curb market was not included in our analysis because data for this market are not available in Brazil.

With exception of the quotations for national treasury bonds (TB), all the time series were obtained from the publication *Análise Financeira* (1991). They are monthly rates charged by and paid to the banking system, taken as an average for the whole country.

The maturity of the contracts for the market interest rates varied from one to six months. The average maturity of government securities varied during the period, but remained essentially small and never stayed greater than one year. All the interest rates we described above were nominal. We calculated the real interest rate by correcting the nominal interest rate for the official rate of inflation, for tax withholdings and, when pertinent, for other financial fees.¹⁰

10 For the average maturity of the primary market and of the overnight market during the eighties, see Jorgensen (1990). For the maturities of contracts and calculations of the market interest rate, see *Análise Financeira* (1991).

3 Econometric results

We will proceed in this section by investigating first the behavior of the nominal interest rates. This part will consist of regression analysis to investigate the effects of liquidity, credit, and inflation upon the rates of interest within the context in which the money supply is independent of the interest rate. Since most of the monetary policy stances in Brazil have not been constrained by concerns to promote stable interest rates, money-growth is likely to be uncorrelated with the error term.

Nevertheless, in the case of Brazil, the endogeneity of the money supply is related to the financing of the public deficit. Since part of the public deficit has been financed through monetary expansion, the money supply becomes an endogenous variable of the whole economy, through its direct link to the way the deficit is financed.¹¹ This form of endogeneity is an important point in the way the econometric analysis should be conducted. For this reason, we will inquire into the factors that may explain the behavior of the real interest rate in a scenario of endogenous money supply. In this inquiry, we will use the vector autoregression (VAR) technique, in which linear dynamic models allow us to analyze the issue with all the variables being endogenously determined. We expect that this combination of econometric approaches will permit us to check the robustness of our findings.

3A The nominal interest rate

3A1 The liquidity effect

The most traditional empirical approach to study the effects of monetary changes on the rate of interest was first proposed by Cagan and Gandolfi (1969). This approach consists of the OLS estimation of the following equation:

$$\Delta i r_t = \alpha + \sum_{i=0}^k \beta_i \Delta M_{t-i} + \varepsilon_t, \quad (1)$$

where ir is the level of interest rate, M is a measure of money-growth, ε_t is the iid $(0, \sigma^2)$ random error, and Δ is the first difference operator.

11 Dornbusch, Sturzenegger and Wolf (1991) discussed the endogeneity of the money supply in the context of money financing.

Cagan and Gandolfi (1969) found that the cumulative coefficients in the regression above were negative for the short term. They asserted that this finding was the evidence of the liquidity effect. After the publication of the Cagan and Gandolfi's results, several studies followed and the controversy of the existence and nature of the liquidity effect became central in monetary economics.

The most important source of controversy lies in the fact that different methodologies and different econometric techniques have been employed to appraise the issue. Because of their complexity and scope, we will not comment on them. For the purpose of this paper, it is important to clarify the reasons why we selected the classical OLS Cagan-Gandolfi approach in our analysis. The motives for this choice were threefold. First, the issue of the interest rate in Brazil should be investigated by the more traditional methodology, which can serve as a benchmark for future studies. Second, data limitations for inflationary expectations for Brazil preclude the implementation of other methodologies. Third, given the conditions under which the monetary policy is conducted in Brazil, it may not justify the use of other econometric techniques.¹²

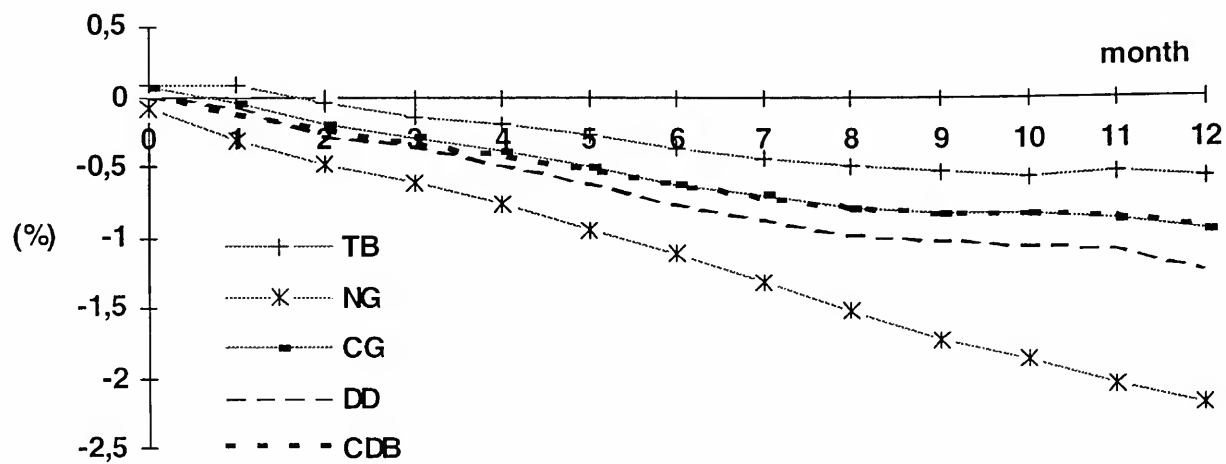
Figure 1 depicts the results for the twelve unconstrained cumulative lag effect¹³ of monthly changes of monetary growth of M1, on changes of monthly interest rate, in five different segments of the financial market. The coefficients are expressed in terms of monthly percentage changes in the interest rate per one percentage point change in the monthly rate of money-growth. The analysis of Figure 1 shows very clearly that the presence of the liquidity effect, due to changes of growth rate of M1, was very strong for all segments of the financial market. It is important to add that the great majority of the coefficients was significant at the 5% level.¹⁴ Nevertheless, the classical Cagan-Gandolfi curve response was not found for the five segments of the financial market. The upward sloped tail-end was absent in our results.

12 Besides the Cagan and Gandolfi approach, two other approaches have been used. They are the IS-LM-Aggregate Supply models and the Efficient Markets-Rational Expectation theory. See Mehra (1985), and Thornton (1988), for discussions of those approaches. For examples of different econometric techniques, see Stokes and Neuburger (1979), for a Box-Jenkins ARIMA analysis, Cochrane (1989), which used band-pass filters, and Mehra (1985), where instrumental variable for money-growth was employed.

13 The number of lags in the regression varied in different studies from twelve to thirty six. Although not reported here, the results for the twenty four lags were very similar to those of Figures 1-3, and most of the coefficients after the thirteenth lag were not significant. Also, the use of annualized data instead of monthly figures, did not change our results in any substantial way. The results for DW test and for multicollinearity can be found in Blumenschein (1994), Appendix F.

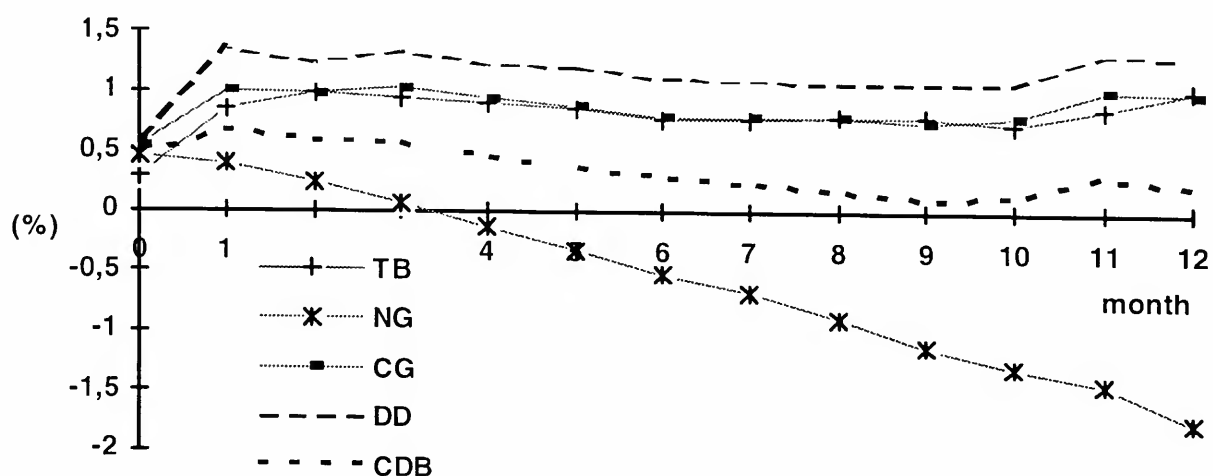
14 For the sake of brevity we will not report here the statistical significance of each coefficient.

Figure 1
Unconstrained Cumulative Lag Effects of Growth of M1
on Change of The Interest Rate. 1975-1991



Because it was never clear that the monetary authorities in Brazil were targeting a specific monetary aggregate, we examined for alternative measures of money in Equation (1). For this purpose, we used changes in the growth rate of M2.¹⁵ Figure 2 presents the results for changes in the growth rate of M2. Approximately half of the coefficients were statistically significant at the 5% level. With the exception of the primary market for government bonds (TB), and in some degree for the overnight market, all the segments did not show the presence of the liquidity effect. On the contrary, the increase on the rate of growth of M2 was translated into positive cumulative lag coefficients for the whole period. The pattern of response for the overnight market clearly reflects a departure from the other segments.

Figure 2
Unconstrained Cumulative Lag Effects of Growth of M2 on
Change of The Interest Rate. 1975-1991



¹⁵ In our analysis, the stock of public debt is part of M3 but not of M2.

The differences in pattern of responses for the monetary aggregates M1 and M2 prompted us to test which specification best represents the cumulative lag effects of money on interest rate. With that objective, we implemented the J test developed by Davidson and MacKinnon (1981).¹⁶ The results are presented in Table 1. These figures indicate that, for the period 1975-1985, the lag effects of money-growth on interest rates were best specified with the monetary aggregate M1, for three segments of the financial market (primary market, secondary market, and working capital). The results for the segments DD and CDB were inconclusive. For the period 1986-1991, the specification with M2 was favored for all three market interest rates (working capital, short-run credit, and certified deposits), but was inconclusive for the other segments (primary and secondary markets).

To examine for changes in the way the money-growth affected the interest rate, we conducted the F test for the constancy of the coefficients for successive break points. The results are presented in Table 2. These numbers corroborate that, after 1986, there was a statistically significant change in the relationship between money-growth of M2 and the interest rate in all segments of the financial market.¹⁷ On the other hand, the results for M1 did not confirm any change in the pattern of the coefficients, except for the overnight market. Although not reported here, F tests for M1 in alternative break points before or after 1986 were not statistically significant.

Table 1
J Test for Specifications of Lag Effects of Monetary Variables

INT. RATE	1975-1985		1986-1991	
	M1/M2	M2/M1	M1/M2	M2/M1
TB	0.42	6.71**	5.81**	3.06**
NG	-0.11	6.71**	5.85**	2.88**
CG	1.06	6.90**	6.09**	1.68
DD	3.06**	4.08**	7.25**	2.16*
CDB	2.64**	5.78**	7.55**	1.88

Note: * means significance at the 5% level. ** significance at the 1% level. M1/M2 represents the truth of M1 specification against the falsity of the M2 specification.

16 Consider two alternative hypotheses: $H_0: v_i = f_i(X_i, \beta) + \varepsilon_i$, and $H_1: v_i = g_i(Z_i, \gamma) + \varepsilon_i$, where X_i and Z_i are the lagged values of M1 and M2, respectively, β and γ are the vector of parameters and ε is assumed to be NID(0, σ^2). They assumed that H_0 and H_1 are not nested on each other. To test the truth of H_0 against the falsity of H_1 , the following regression was estimated: $v = (1 - \alpha)f_i(X_i, \beta) + \alpha\hat{g}_i + \varepsilon_i$, where $\hat{g}_i = g_i(Z_i, \hat{\gamma})$, and $\hat{\gamma}$ is the ML estimate of γ . If H_1 is true, then the true value of α is zero. The reverse holds for H_1 against H_0 . The results in Table 1 are the t test for α .

17 Before 1986, the F tests for M2 were not statistically significant.

Table 2
F Tests for the Constancy of Coefficients for Two Measures of Money:
Break-Point at 1986.01

INT RATE	M1		M2	
	F test	Prob.(%)	F test	Prob.(%)
TB	1.59	8.60	2.29	0.70
NG	1.91	2.93	2.87	0.07
CG	0.46	94.97	1.94	2.60
DD	0.82	63.70	3.48	0.00
CDB	0.41	96.90	1.93	2.60

In summary, the results for the F test and for the J test indicated that the specification with M2 more accurately represented the relationship between interest rate and money, especially during the second half of the 1980s. To more closely investigate the nature of the changes that occurred after 1986, we constructed the cumulative lag effects of M2 for two sub-periods, 1975-1985 and 1986-1991. The results are shown in Figures 3 and 4.

Figure 3
Unconstrained Cumulative Lag Effects of Growth of M2 on
Change of the Interest Rate. 1975-1985

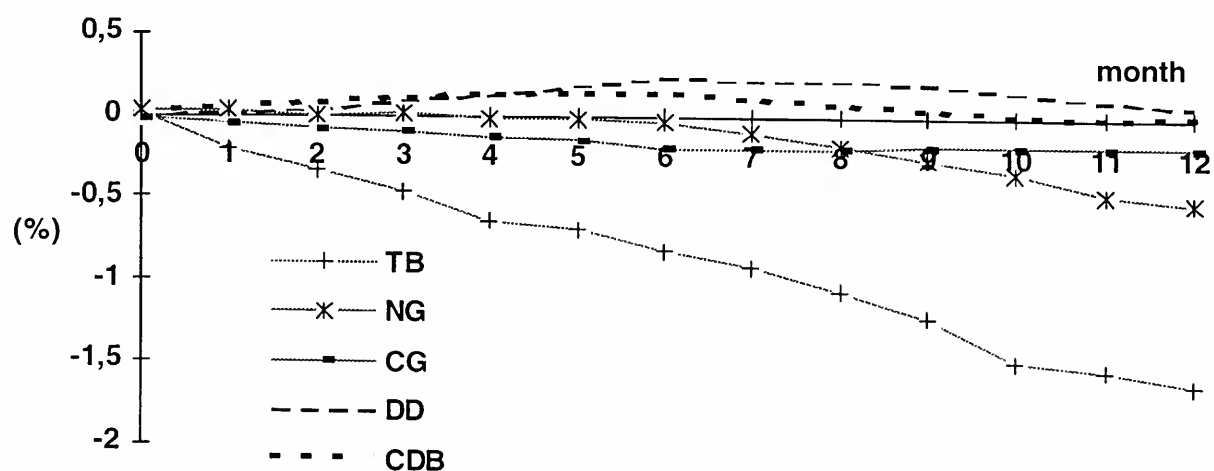
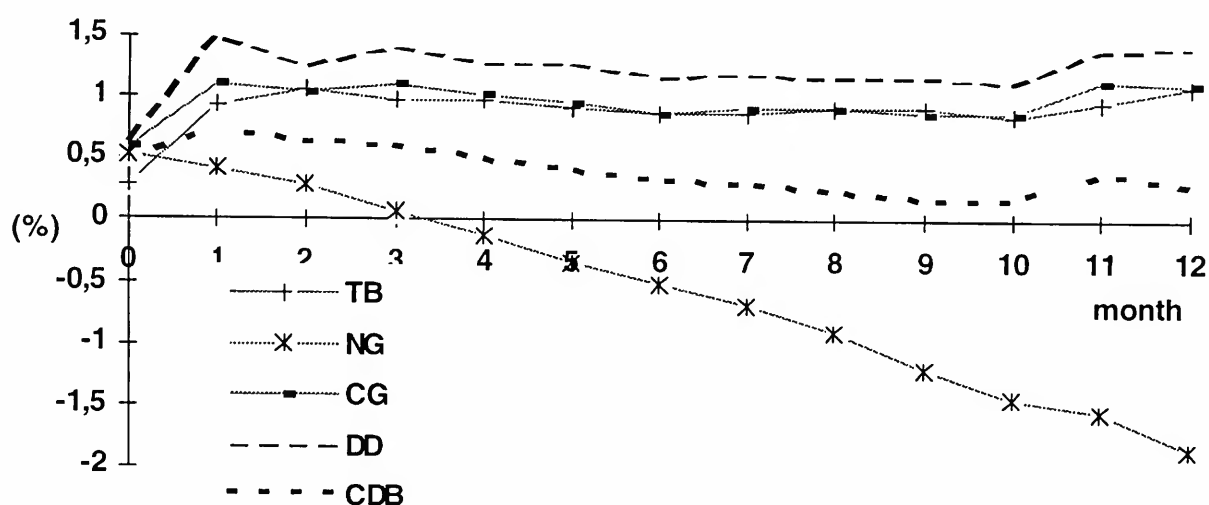


Figure 4
Unconstrained Cumulative Lag Effects of Growth of M2 on
Change of the Interest Rate. 1986-1991



Before 1986, there was a liquidity effect for three segments of the financial market (primary market, overnight market, and working capital), and a positive, but not pronounced, cumulative lag pattern for the two other segments (short-run credit and certified deposits). The liquidity effect was particularly strong for the primary market. Since the quotations in this market were used to calculate the monetary correction, some liquidity effect may have spilled over to segments of the financial market subject to indexation. The extent of this effect would be determined, among other things, by the relative importance of these segments.

The analysis of these figures very clearly shows that there was a remarkable weakening of the liquidity effect after 1986. For the second half of the decade of the 1980s, and with exception of the overnight market, there was a pronounced positive cumulative lag pattern between M2 growth and nominal interest rate. This assertion is corroborated by the statistical significance of the coefficients for the two periods. For the period 1975-1985, almost all coefficients were not statistically significant at the 5% level, whereas for the period 1986-1991 nearly half of the coefficients were statistically significant. One possible reason for the presence of a direct positive relation between money growth and interest rate after 1986, may be related to the presence of the Fisher effect. The inflation rate, measured by the I2 (FGV), increased from an average of 6% per month during the period 1975-1985, to 19% for the period 1986-1991. Moreover, the volatility of the inflation rate, measured by the standard deviation, increased from 3.3 to 16.2 between those periods. These figures may indicate that the formation of higher expectation of the inflation rate was translated into higher interest rates even in the short run. It must be pointed out that the positive responses of the interest rates in the short term were an extreme version of the Fisher effect, in which

the expectation of higher inflation rates preceded money-growth itself.¹⁸ It would be better named “anti-liquidity effect.”

Besides the Fisher effect, it is also important to assert that, during the 1980s, substantial changes occurred in the demand function for money itself, and in the arguments of the demand function. Financial market innovations and money substitutes had an important quantifiable effect on the demand for money and consequently on the relationship between inflation and the interest rate.¹⁹

The most important example of financial innovations was the payment of interest on money. Combined with high rates of inflation²⁰ that prevailed in the second half of the decade of the 1980s, this innovation dampened the desires to hold money. The collapse of the demand for money and the increase in the velocity of money sped up the process of formation of inflationary expectations, with a direct impact on interest rates. Money followed rather than led, and this behavior has a close association with the fact that the money supply was endogenously determined, and lacked the feature of anticipation. The Cagan-Gandolfi curve response of interest rate may hold for Brazil, but in an upside down form as revealed in Figure 4.

3A2 The credit effect

According to the loanable funds theory, variations on the total amount of credit allocated to the economy are the primary source of variations on the interest rate. A higher rate of monetary growth, through an increase of the supply of loanable funds by the banking system, would produce a single drop in the rate of interest.

Although, on the theoretical level, there is little controversy about the existence of the credit effect, some debate exists regarding the empirical difficulty in breaking down liquidity and credit effects. Hence, empirical analyses of the credit effect are scanty. We will present the results for the two empirical approaches encountered in the literature. These are Cagan’s approach and an alternative method proposed by William Poole.²¹

18 See Hanson (1980), for a confirmation that, in Brazil, changes in price can predict changes in money but the reverse is not true.

19 See Calomiris and Domowitz (1989), for a study of the money demand in Brazil, that takes into account financial innovations and money substitutes.

20 It is important to say that this combination is a direct cause of the financing of the public deficit under poor governmental credibility.

21 See Cagan (1972) and Poole (1974). The rationale behind the specification of equation 2 and 3 without lags was explained in details by Cagan (1972).

The analysis of the credit effect carried out by Cagan consists of investigating the following equation:

$$\Delta i = \alpha + \theta(\Delta gC) + \vartheta[\Delta(gC - gM)] \quad (2)$$

In Equation (2), i is the interest rate, gC is the monthly growth rate of credit of the banking system, Δ is the growth rate of money stock, gM is a constant term, and α is the first difference operator.²² The coefficient θ measures the effect attributed to credit expansion alone, or the loanable funds effect, while ϑ measures the effect of all other sources of monetary expansion. According to Cagan (1972), if the two coefficients have negative signs, the credit and the liquidity effects have the same outcome on the interest rate, and no separate role can be attributed to credit alone. For credit to have a specific role, $(\theta - \vartheta)$ must be statistically significant and negative. To test this hypothesis, we estimated Equation (3):

$$\Delta i = \alpha + (\theta - \vartheta)(\Delta gC) + \vartheta(\Delta gM). \quad (3)$$

The estimates of the coefficients of Equations (2) and (3), for five segments of the financial market, and for three periods, are presented in Table 3. The numbers in this table visibly reveal that the estimates varied considerably when we compare the numbers of the period 1975-1985 with those of the period 1986-1991. During the period 1975-1985, the credit effect had a significant influence only on one segment of the financial market, namely, working capital (CG). On the other hand, although small in values, the coefficient for non-credit sources $(\theta - \vartheta)$ for the short-term credit (DD) was negative, whereas for the overnight market (NG) it was positive. Moreover, both were statistically significant. This piece of evidence shows that the non-credit sources had an impact on these segments of the financial market.

An overall analysis of these results reveals that statistical significance is lacking for the majority of the coefficients. This fact indicates the neutrality, *strictus sensus*,²³ of the credit and non-credit sources of monetary expansions in affecting the interest rate for most segments of the financial market.

22 For the credit variable, we used the growth rate of total amount of credit to the private sector, and for the money variable, the growth rate of M2. All variables, including the interest rate, were expressed in percentage points. The credit variable excludes securities outside the government agencies. According to Cagan, this specification allows for more accurate estimates.

23 Neutrality, in this context, means the absence of either negative or positive effects.

Table 3
The Credit and Non-Credit Sources of Monetary Growth and
Their Relationship to the Interest Rates

INT.RATE	Cte. (α)	Credit (θ)	Non-Credit		R ²	DW	F
			(ϑ)	($\theta - \vartheta$)			
1975-1991							
TB	0.09 (0.2)	0.29 (4.6)**	0.01 (0.2)	0.28 (5.3)**	0.13	2.39	15.85**
NG	0.02 (0.1)	0.57 (13.9)**	0.42 (11.3)**	0.14 (4.0)**	0.50	2.05	93.91**
CG	0.07 (0.2)	0.61 (10.6)**	0.19 (3.8)**	0.43 (8.9)**	0.38	2.40	61.94**
DD	0.05 (0.1)	0.74 (9.1)**	0.08 (1.2)	0.65 (9.7)**	0.35	2.81	56.51**
CDB	0.05 (0.2)	0.51 (11.9)**	0.21 (5.7)**	0.30 (8.5)**	0.42	2.27	73.55**
1975-1985							
TB	0.08 (0.3)	-0.08 (-1.1)	0.02 (0.3)	-0.11 (-2.2)*	0.01	2.91	2.16
NG	0.07 (1.0)	0.05 (2.1)*	0.02 (1.0)	0.03 (1.9)*	0.02	2.32	2.62
CG	0.05 (1.4)	-0.02 (-1.9)*	-0.01 (-1.5)	-0.008 (-1.1)	0.01	1.06	1.84
DD	0.09 (1.9)	-0.02 (-1.5)	0.002 (0.2)	-0.024 (-2.4)*	0.03	1.92	2.92*
CDB	0.06 (2.3)*	0.01 (1.5)	0.007 (1.1)	0.005 (1.0)	0.003	1.49	1.20
1986-1991							
TB	0.13 (0.1)	0.40 (4.0)**	-0.03 (-0.3)	0.44 (4.9)**	0.25	2.28	13.0**
NG	0.06 (0.1)	0.69 (9.1)**	0.53 (7.8)**	0.16 (2.4)*	0.60	1.99	46.5**
CG	0.18 (0.2)	0.79 (8.0)**	0.19 (2.3)*	0.59 (6.8)**	0.49	2.55	35.6**
DD	0.04 (0.04)	0.96 (7.0)**	0.03 (0.28)	0.92 (7.7)**	0.48	3.02	34.5**
CDB	0.10 (0.2)	0.63 (8.8)**	0.23 (3.6)**	0.41 (6.4)**	0.52	2.43	40.3**

Note: * significance at the 5% level. ** significance at the 1% level.

The results for the period 1986-1991 present a different pattern from those for the period 1975-1985. The credit expansion had a strong, statistically significant, and positive effect on the interest rate for all segments. The coefficients for non-credit sources of monetary expansion, although smaller than those of credit, had a positive and statistically significant impact on the interest rate in the majority of the segments of the financial market. We can interpret these results as a valuable signal that credit has had a much stronger effect on rising the interest rates than the other sources of monetary expansion. Moreover, the size of these coefficients reveals also that the effect can be quite substantial and uniform across all segments.

According to Poole (1974), in the short term, the test of the credit effect would appear as a relationship between the rate of change of the interest rate and the rate of change of the money stock.²⁴ In Poole's point of view, the use of credit variables in the regression analysis

²⁴ Poole's argument took the line that "if money demand differs from money supply primarily because of changes in the later, then the credit effect should appear as a relationship between the level of interest rate and the rate of change of money." Friedman and Schwartz (1982) undertook their empirical analysis of the credit effect on these lines.

as proposed by Cagan was not fully satisfactory. Following Poole's method, we found that the correlation between the level of interest rate and the growth rate of money strongly confirmed the results of Table 3. Indeed, for growth of M1, the averaged correlation for the five segments increased from 0.20 during the period 1975-1985, to 0.60 for the period 1986-1991. For the monetary aggregate M2, the numbers were 0.70 and 0.86 for the two periods respectively.²⁵ These numbers confirm the nonexistence of the credit effect in both periods, and that the positive relationship between credit and interest rate was substantially higher after 1986. Nevertheless, they do not confirm the neutrality of credit, *strictus sensus*, before 1986.

Some rationale for the results of Table 3, and of the correlations suggested by Poole may be given by the factors we explored in analyzing the liquidity effect. Indeed, if prices preceded money-growth, a once-and-for-all increase of the interest rate in the short run would mask any credit effect. Moreover, an increase in the issuing of government securities that was required to finance the growing public deficit constituted a continuous drain on the availability of credit to the private sector. The fast growing repurchase contract in the secondary market may also have helped to offset the increase in any private holdings of new credit, by disposing of these assets.

3A3 The Fisher effect

The evidence for the nonexistence of either the liquidity effect or the credit effect prompted us to focus attention on the analysis of the Fisher effect. For economies with high inflation rates, Friedman (1968) suggested that the monetary effects on the interest rate would be very weak or even absent, due to the speed at which the economic agents would adjust to expectations of higher rates of inflation that follow an increase in the money supply.²⁶ For the specific case of the Brazilian experience, recent work by Garcia (1991) has found that, after 1975, the Fisher equation is a good approximation of the behavior of CDB (indexed and non-indexed) interest rate in Brazil. Although this evidence was important, it was partial owing to the fact that it focused only on one segment of the financial market.

25 The correlation differentials between segments were very small.

26 For studies in developing countries supporting the Friedman view, see Blejer (1978).

As a complementary analysis, we will investigate the hypothesis that the inflation rate influences the way money growth affects interest rates. To test this hypothesis, Equation (1) has to incorporate a variable for inflation. This will be achieved by implementing a weighted distributed lag. To this end, we will make the b coefficient of Equation (1) a function of the rate of inflation, i.e., $\beta_t = f(I)$. The function f will assume a first degree polynomial given by Equation (4).²⁷

$$\beta_t = \rho_t + \varphi_t I_{t-j} \quad (4)$$

In Equation (4), the coefficient ρ_t corresponds to the fixed effect of money on the interest rate, and $\varphi_t I_{t-j}$ corresponds to the variable effect, which depends on the inflation rate I .

A twelve-weighted lag was estimated for five segments of the financial market for the two periods, 1975-1985 and 1986-1991. For the sake of brevity, in Table 4 we present only the coefficients of the variable weights. It is important to point out that the cumulative lag for the weighted version showed a pattern very similar to those of Figures 3 and 4.

For the period 1975-1985, and with exception of the lending rates (CG and DD), the majority of the coefficients was negative. On the other hand, for the period 1986-1991, almost all the coefficients were positive. This pattern suggests that increasing prices attenuated the liquidity effect, and that this attenuation was more pronounced under higher rates of inflation.²⁸

The tests for the joint significance of the coefficients in Table 4, reported in Table 5, indicate that, before 1985, the hypothesis that monetary policy was affected by inflation was confirmed only in the segments working capital (CG) and certified deposits (CDB).²⁹ After 1985, only the overnight market (NG) did not confirm this hypothesis. As it will be discussed in the next section, this non-uniform behavior should be taken into consideration in analyzing the effects of monetary expansion.

27 This is a variation of the Almon lag scheme. In his scheme, ρ , and the polynomial, can take any degree. For examples of empirical works that used this methodology, see Tinsley (1967), Tanner (1979), and Melvin (1983). This methodology is distinct to the one used by Garcia (1991).

28 One of the difficulties in interpreting the results of Table 4 was that only the coefficients of the first two lags were significant at the 5% level. Nonetheless, the analysis of the coefficients of the first two lags confirms all the results of the paragraph above.

29 It was specified on page 16 and footnote 25 the rationale for this hypothesis.

Table 4
Variable Effect Captured by the Coefficients j in Weighted Distributed Lag

	TB	NG	CG	DD	CDB
lag	1975-1985				
0	0.013(0.56)	0.008(0.88)	-0.010(-2.24)	-0.0096(-1.67)	0.003(0.97)
1	-0.026(-1.02)	0.009(0.85)	-0.008(-1.71)	-0.007(-1.13)	0.006(1.60)
2	0.026(0.99)	-0.016(-1.50)	-0.002(-0.49)	-0.001(-0.17)	0.004(1.02)
3	0.011(0.40)	0.0015(0.13)	0.006(1.04)	0.007(1.02)	0.0005(0.14)
4	0.017(0.57)	0.011(0.90)	0.006(1.01)	0.008(1.08)	0.006(1.46)
5	0.023(0.80)	0.006(0.50)	0.002(0.35)	0.014(1.97)	0.011(2.67)
6	0.037(1.25)	0.003(0.26)	-0.019(-0.33)	-0.001(-0.16)	0.009(2.38)
7	0.004(0.012)	0.0005(0.04)	-0.009(-1.58)	-0.004(-0.64)	0.005(1.15)
8	-0.002(-0.007)	-0.006(-0.54)	-0.001(-0.31)	0.004(0.58)	-0.007(-1.71)
9	0.031(1.04)	-0.017(-1.50)	0.016(2.75)	0.006(0.90)	-0.004(-1.02)
10	-0.129(-4.38)	0.021(1.76)	0.013(2.29)	-0.010(-1.45)	0.002(0.44)
11	0.075(2.52)	0.005(0.45)	0.005(0.85)	-0.013(-1.82)	0.005(1.11)
12	-0.010(-0.32)	-0.019(-1.50)	-0.002(-0.33)	0.006(0.78)	0.009(1.90)
	1986-1991				
0	-0.008(-2.39)	0.012(3.47)	0.011(2.77)	0.010(2.04)	0.012(3.47)
1	0.009(2.02)	0.017(4.10)	0.026(5.11)	0.030(5.70)	0.017(4.10)
2	0.013(3.16)	0.004(0.92)	0.005(1.04)	0.003(0.57)	0.004(0.92)
3	-0.003(-0.78)	0.000(0.00)	0.00(0.00)	-0.0015(-0.26)	0.00(0.01)
4	-0.0002(-0.05)	0.002(0.62)	0.002(0.36)	0.001(0.21)	0.003(0.62)
5	0.004(0.82)	0.004(1.06)	0.008(1.41)	0.003(0.59)	0.004(1.06)
6	0.006(1.28)	0.002(0.59)	0.003(0.58)	0.002(0.36)	0.003(0.59)
7	0.004(0.91)	0.002(0.44)	0.0001(0.02)	0.002(0.29)	0.002(0.44)
8	0.002(0.43)	0.003(0.64)	0.002(0.36)	0.003(0.49)	0.003(0.64)
9	0.004(0.81)	0.002(0.38)	0.003(0.46)	0.002(0.31)	0.002(0.38)
10	0.005(1.23)	0.004(0.93)	0.008(1.55)	0.006(1.09)	0.004(0.93)
11	0.006(1.35)	0.003(0.62)	0.004(0.73)	0.004(0.62)	0.002(0.62)
12	0.01(3.22)	0.002(0.07)	0.00(0.02)	0.001(0.38)	0.0002(0.07)

Table 5
F Tests for the Joint Significance of the Coefficients

INT. RATE	1975-1985		1986-1991	
	F test	Prob. (%)	F test	Prob. (%)
TB	1.00	44.7	3.61	0.1
NG	0.90	53.5	1.55	16.8
CG	2.07	3.5	3.94	0.0
DD	1.21	29.0	6.04	0.0
CDB	1.94	4.9	2.28	2.9

In summary, two important qualifications should be emphasized regarding the power of the interaction between money and price in affecting the interest rate. First, the degree of the effect seems to be positively related to the level of inflation that prevailed in Brazil during two different periods. By international standards, although Brazil experienced high rates of inflation before 1985, monetary policy still had some bearing on moderating upward movements in the interest rate during this period, at least for some segments of the financial market. After 1986, with higher inflation rates prevailing, this moderating effect vanished. Second, during the second half of the 1980s, the overnight market seemed more immune to the effect of inflationary expectations than other segments. The increasing importance of this market in allocating savings to the public sector, at the end of the decade of the 1980s, may have reduced the cost of debt financing.³⁰

3B The real interest rate: vector autoregression analysis

The real rates of interest, in all segments of the financial market, were very high and volatile, especially after 1985. The investigation of factors that may have affected the real portion of the nominal interest rates seems mandatory to understand the behavior of the interest rate in Brazil. The analysis of the real rates of interest will be conducted by using the Vector Autoregression (VAR) technique.

The seminal paper by Sims (1980) introduced the methodology of Vector Autoregression (VAR) in the study of macroeconomic problems. This sub-section will use the VAR approach to analyze the behavior of the real interest rates. We will maintain the same theoretical

³⁰ For the importance of the overnight market, see Jorgensen (1991). The idea that the government could ameliorate the burden of debt financing by issuing nominal debt was rejected by Garcia (1991).

apparatus we employed in the study of the nominal rates. Our main purpose in this subsection is to investigate how the macroeconomic variables (liquidity, income, and price level) influenced the real component of the rate of interest in different segments of the financial market.³¹

In the vector autoregression approach all the variables are assumed to be endogenous. For each interest rate, consider the vector $Y_t = (i2_t, m2r_t, y_t, ir_t)$ of four variables that comprise the system being investigated. In this vector, $i2$ is the rate of inflation, $m2r$ is the real level of the money supply, y is the real level of income, and ir is the real rate of interest.³² With the ordering given by the vector above, we estimated the following system of equations which defines the unrestricted VAR approach.

$$Y_t = \sum_{k=1}^4 B_t(k) Y_{t-k} + u_t \quad u_t \sim N(0, \Sigma) \quad (5)$$

The linear dynamic system of equations above states that the model is joint modeled, and that the present value of a variable is determined by four lags of all the variables.³³ A variable for trend was not included because the variables did not present any trend pattern. Since the coefficient matrix $B_t(k)$ is difficult to interpret, the analysis of the VAR was conducted by looking at the variance decomposition of the system and at the impulse response functions.

The decomposition of variance quantifies the relative importance of the variables in explaining the forecast error variance. The results for decomposition of variance were presented in Table 6. The following observations are worth noting. First, among the macroeconomic variables, the inflation rate was the most important factor in explaining monthly forecast variance in the real interest rate in all segments.³⁴ Second, the importance

31 Simulations with the inclusion of a variable for credit resulted in lack of explanatory power for the decomposition of the variance, and for that reason credit was not included in the system.

32 In the vector $Y_t = (i2_t, m2r_t, y_t, ir_t)$ $i2_t$ = rate of inflation measured by I2 (FGV), $m2r_t$ = real level of M2, y_t = index of industrial electrical consumption (IIEC), and ir_t = real rate of interest. The index of industrial production (IIP) in Brazil is poorly surveyed. Also, unlike IIEC, IIP does not capture the informal economy, which is substantial in Brazil. Nonetheless, either choice gave very similar results.

33 Results for a 4 lags and an 8 lags VAR were very analogous. The orthogonalization of the error covariance matrix was made by the Cholesky method. The results we found were very robust for various different ordering of the variables. For some authors, the fact that the results are insensitive to the ordering of the variables does not justify causal interpretations among these variables. See Cooley and LeRoy (1985). See also Sachs (1982) and Sims (1982), for further discussion of this point.

34 With exception of the overnight market for the period 1986-1991, when real M2 was the main factor.

of inflation as an explanatory factor was considerably smaller after 1985. Third, the importance of changes in real M2, as a variable that explains decomposition of forecast error variance, increased after 1985 for all segments. Fourth, the overnight market shows a pattern in the decomposition of the variance that was substantially different from the other segments.

Table 6
Variance Decomposition of the Real Interest Rates

PERIOD	1975-1985					1986-1991				
	S.E.	I2	M2R	Y	I.R.	S.E.	I2	M2R	Y	I.R.
1	2.72	26.24	5.10	0.01	68.65	4.25	7.93	2.00	2.07	87.97
2	3.61	21.97	4.77	0.01	73.25	5.19	5.97	1.41	1.52	91.09
3	3.67	22.18	4.63	0.04	73.14	5.46	5.79	10.06	1.64	82.49
4	3.74	23.90	4.65	0.90	70.55	5.61	6.61	10.20	1.99	81.19
5	3.79	23.34	4.55	2.96	69.16	5.69	7.42	11.26	1.99	79.31
6	3.99	24.55	4.45	4.40	66.59	5.79	7.38	11.59	2.08	78.93
1	1.36	71.20	0.64	0.30	27.85	5.34	4.11	26.72	0.04	69.11
2	1.45	73.81	0.71	0.62	24.84	7.33	2.40	38.95	4.49	54.14
3	1.52	72.43	0.82	0.58	26.17	8.42	2.82	33.73	14.66	48.77
4	1.56	71.77	2.48	0.63	25.11	8.75	2.82	33.27	14.44	49.45
5	1.60	69.39	4.60	0.62	25.37	8.85	2.77	33.85	15.04	48.32
6	1.61	68.20	4.55	0.77	26.47	8.88	2.76	33.62	15.59	48.01
1	0.25	0.93	0.21	4.46	94.39	3.32	12.62	0.73	4.06	82.56
2	0.32	21.84	0.17	4.57	73.41	3.56	16.13	3.75	3.58	76.54
3	0.33	22.29	0.16	4.20	73.35	4.25	12.85	9.96	3.58	73.59
4	0.34	22.31	0.17	4.25	73.27	4.34	12.38	9.73	3.65	74.22
5	0.34	22.35	0.32	4.41	72.91	4.37	12.41	9.75	3.77	74.06
6	0.34	22.70	0.36	4.41	72.54	4.48	13.52	10.52	4.29	71.64
1	0.67	0.63	0.00	0.40	98.97	4.48	1.75	1.67	3.36	93.20
2	0.75	8.86	1.67	0.91	88.54	5.23	4.85	1.37	2.73	91.03
3	0.76	10.45	1.71	1.59	86.23	5.48	9.23	2.40	4.23	84.12
4	0.77	11.14	2.27	1.57	85.00	5.48	10.82	2.64	4.93	81.59
5	0.78	11.72	2.21	2.89	83.17	5.70	10.97	3.37	5.66	79.97
6	0.79	12.07	2.55	2.90	82.46	5.95	12.19	3.12	5.69	78.98
1	0.36	3.89	0.01	1.67	94.42	2.74	12.87	0.72	3.57	82.82
2	0.47	30.19	0.01	1.14	68.65	2.99	11.50	4.69	3.03	80.77
3	0.49	34.45	1.16	1.43	62.96	3.48	8.82	6.35	7.06	77.74
4	0.49	35.62	1.26	1.40	61.72	3.61	8.58	5.92	6.85	78.64
5	0.50	35.15	1.26	1.91	61.67	3.66	8.52	5.75	7.04	78.67
6	0.50	35.16	1.38	1.91	61.54	3.71	8.93	5.91	6.85	78.28

Note: S.E.=standard error, I2=inflation, M2R=money supply, Y=income, and IR=interest rate.

The impulse response function traces the response (current and future) of a variable of separated one standard deviation shock of each variable that comprises the system. Figure 5 depicts the impulse response functions for six months ahead, of separated shocks on inflation, growth of real money stock of M2, and real income. The response patterns following a shock were as follows. First, the responses for the first month after the shock had been implemented were as expected. A shock in prices and income increased the real interest rates while a shock in real M2 decreased the real interest rates. Second, although the volatility of the response was common to all five segments, the responses after the first month varied considerably in intensity and direction depending on the segment. Third, by looking at the y-axis, it is clear that the intensity of the response was substantially higher after 1985. Some interpretations of these results and of the results for the nominal interest rate will be given in the next section.

4 Interpretations and implications

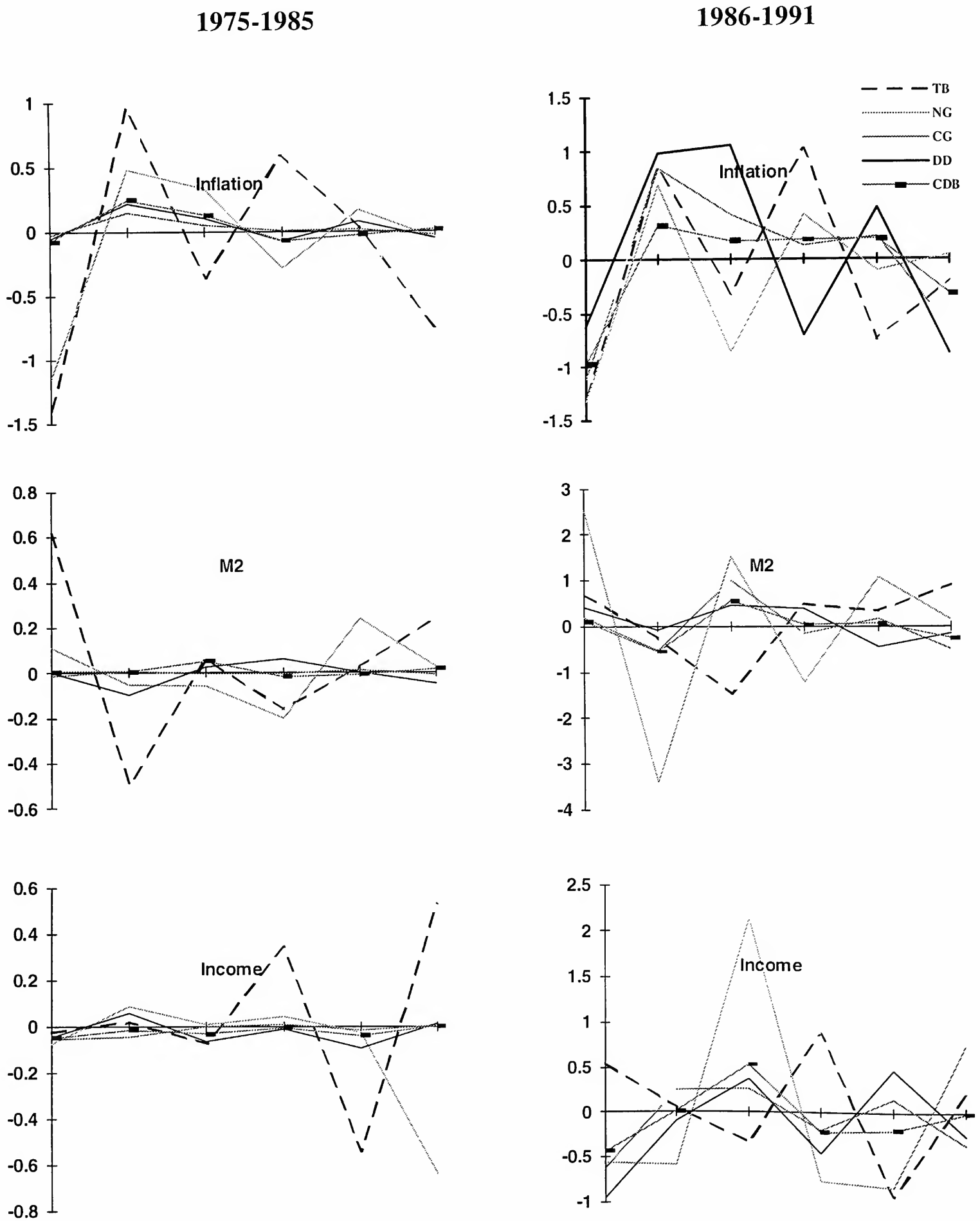
This section will elaborate on three different issues. Firstly, there is an attempt to compare the results we obtained in Section 3 with the prediction of the monetarists and views of the rational expectation school. Secondly, the section pursues the interpretation of those results in the light of two 'ad hoc' factors. It brings the role played by risk and segmentation in explaining some of the main findings. These two issues are presented in sub-section 4A. Thirdly, we elaborate on the implications of these results for the efficacy of the monetary policy in general, and for the functioning of the monetary mechanism in particular. This is briefly stated in sub-section 4B.

4A Interpretations

The understanding of the relationship between the interest rate and macroeconomic variables in the long run, in an environment of a closed economy, was the main concern of the preceding sections. Although the money supply implicitly incorporates external capital movements, the effective roles that these movements and the exchange rate policy may have played on the determination of the interest rate, have not been addressed in this paper. They were excluded not because they were considered secondary, but for methodological reasons.³⁵

35 A positive net capital inflow would exert a downward pressure on the interest rates by increasing the supply of credit. The exchange rate policy would determine the return of foreign exchange denominated assets, and would serve as a benchmark for the interest rate in other segments of the financial market.

Figure 5
Impulse Response Functions of the Real Interest Rates to Shock
on Inflation, Real M2, and Real Income



The understanding of the timing pattern of money and the interest rate in an environment of changing inflationary regimes was considered one of the central themes of our analysis. Moreover, we emphasized that the interpretations of our findings should take into consideration some features of the financial market in Brazil.

A positive relationship between money supply and the interest rate prevailed in most of the segments of the financial market. It is worth noting that this upside down Cagan-Gandolfi curve without a tail end cannot be fully explained by the monetarist view. If there is a positive relationship between money growth and the interest rate, the notion that monetary policy is non-neutral in the short run should not be accepted. This non-neutrality comes from the fact that higher market interest rates following monetary expansion, either from credit and non-credit sources, cannot have a positive effect on income.

On the other hand, the rational expectation amendment that only unexpected money-growth is non-neutral has its pitfalls in interpreting our results. First, the concept of anticipation of money supply is difficult to hold in the Brazilian financial system. If the money supply is endogenous, as we have described, it may not have any characteristic of anticipation, in the way anticipation is defined in a setting with an independent central bank. For this reason, the characterization of monetary policy as anticipated or not, in the context of the Brazilian economy, may be inappropriate. Second, it is also difficult to accept that the economic agents, including the banking system, have the sophistication (technical and informational) to monitor the behavior of the money supply. Third, and more importantly, the positive pattern curve response does not allow room for non-neutrality even for the unexpected money growth.

Some understanding of the interest rate puzzle was achieved by our partial results regarding the effects of inflation on the relationship between money and the rate of interest. Although monetarism has emphasized that the liquidity effect would be weakened, or even irrelevant, in an economy with high rates of inflation, the positive pattern we found is an overshooting of this view. Prices were moving faster than money.³⁶ Under these conditions, the economic agents will consider inflation rather than money growth as the relevant information. Besides, in economies of the developing world, information regarding prices is more credible and easier to obtain and to interpret. The banking system will not raise interest rates if they expect forthcoming monetary tightness. Instead, they will look at the expected inflation. This pattern would be more prevalent if the money supply adjusted passively, like in Ultra-Keynesian models. In Brazil, the passivity of the money supply has been recognized in several studies.³⁷

36 For an empirical analysis of this point see Rossi (1991).

37 See Tobin (1970), for an elaboration on Ultra-Keynesian models. For the studies about the passivity of the money supply see Footnote 10.

The effect of prices on the relationship between money and the interest rate became more acute after 1986, when the inflation rates were higher and more volatile. Hence, the money supply was even more ineffective in influencing the interest rate, thereby causing the real rates to become remarkably higher. One channel in which increasing prices may undermine liquidity stances is through the risk associated with macroeconomic conditions.³⁸ Under the scenario of high and accelerating prices, expectation of higher prices would increase not only the nominal rate, but the real rate as well. An explanation for higher real rates could be found in the concept of risk premium, which is closely associated with the studies of the term-structure of interest rates. High and accelerating inflation would increase the overall risk of the economy, forcing the financial system to incorporate it in the form of higher risk premium and, consequently, higher real rates.³⁹ This component of risk would be incorporated mostly because of higher prices, but with a loose connection to the tightness of the money supply. Nevertheless, the results for the primary and secondary markets did not give total support for this argument. After 1986, on average, the real rates of interest were lower and the correlation between inflation and the real rates smaller. The market interest rates were much higher after 1986, but only the segment of short-run credit showed a considerable increase in the correlation between inflation and the real interest rate. The incorporation of a risk premium associated with inflation into the real rates varied considerably among different segments. In general, the market lending rates were more prone to this process.⁴⁰

The existence of weak linkages between government interest rates and market rates in the Brazilian financial market was confirmed in many of our results. One important result was the asymmetric response of the secondary market for government securities or the overnight repurchase market. We found substantial liquidity effect in the overnight market which somehow resulted from the fact that this market was shielded from the Fisher effect. The reasons for this unique behavior seem complex. Some light was shed by our analysis of the real rates of interest where we showed that, after 1986, a large part of the variance

38 Irving Fisher explored the effect of risk on the interest rate in what he called "Third Approximation" of the interest rate theory. He also recognized the difficulty in measuring economic risks.

39 Borrower's creditworthiness collapses owing to deteriorated macroeconomic conditions, which will set a wedge between cost of funds for the banks and the collateral. See Gertler (1991), which cited Irving Fisher on this line of thought.

40 For the period of moderate rates of inflation (1975-1985), the contemporary correlations between real rates of interest and inflation were the following: TB = -0.176, NG = -0.299, CG = 0.299, DD = 0.254, and CDB = -0.145. For the period of high and volatile inflation rates (1986-1991), the correlations switched to the following values: TB = -0.498, NG = -0.259, CG = 0.105, DD = 0.440, and CDB = -0.119. Some scholars argue that we should not accept timing evidence as an empirical proof of propositions about causation. See Tobin (1970), and Friedman (1970). Moreover, any empirical test linking risk premium with higher real rates has to come up with a comprehensive measure of risk associated with various macroeconomic conditions. Here we have just considered inflation rates.

decomposition was explained by the real growth of the money supply. With a fast growing money supply, and higher inflation rates, lending to the overnight market would produce the best combination for reduced risk with higher liquidity, which can be translated into lower real rates.⁴¹ A sharper impulse response function of the overnight market, for changes in real money, also confirmed the findings for the variance decomposition. The belief that the targeting of the real interest rate in the overnight market would constitute an adequate instrument of monetary policy may be misleading under these circumstances.⁴²

4B Implications

The set of results we have encountered may have some bearing on the monetary transmission mechanism. By linking the monetary and the real side of the economy, the interest rate is considered a key variable in the monetary transmission mechanism. The effectiveness and nature of the transmission mechanism should be taken into account for the optimal conduct of monetary policy. The implications of our findings for the monetary transmission mechanism may be sizable. First, the overnight repurchase market represents a close substitute for money. As Tobin and Brainard (1963) argued, “*the possibility of substituting intermediary liabilities for currencies offers a partial escape from monetary restriction. But so long as the intermediary liability is an imperfect substitute for currency, the escape is only partial.*” In the case of Brazil, the overnight market is almost a perfect substitute. The presence of this quasi-perfect money substitute dampens the liquidity effect considerably. Second, the process of investment in the private sector is distorted by the existence of segmentation. Portfolio decisions will reflect concerns only with the short term and capital formation is discouraged. Third, the adjustment of interest rate structure among all segments, which constitutes the core of the monetary transmission mechanism, will also affect wealth allocations between real and financial assets. As we mentioned before, the shrinking pool of loans to the economy due to disintermediation for the overnight market may allocate resources to real assets sub-optimally. Moreover, the variability of results for impulse response functions for different segments may suggest that the adjustment of the interest rate structure is slow and perhaps imperfect, especially if the economy is experiencing high inflation rates.

41 For a discussion of the argument that lower rates in those markets were due to a premium for higher liquidity, see Guidotti and Kumar (1991).

42 See Clifton (1990), for a macro model of interest targeting for Brazil. His model concluded that although the targeting of the real interest rate in the overnight market would allow the economic agents to calculate the conditional expectation of the next period price level, it would not permit the monetary authority to exercise control over the price level over the long run. Our results indicated that this short-run effectiveness has not influenced the real interest rates in other segments of the financial market.

One important point not yet investigated in the Brazilian economy relates to the role that the overnight repurchase market for government securities has played in the liquidity of the economy in the short term, and in absorbing private savings, and consequently reducing credit to the private sector in the long term. It seems that the total amount transacted in this market has been growing exponentially and may have constrained lending in other segments of the market.⁴³ The reduced quantity of credit available in the banking system, due to the increased role played by the repurchase overnight market, may have raised the market interest rates relative to the overnight rate. The gap would prevail due to the process of segmentation, which we analyzed in previous studies.⁴⁴

Our approach did not investigate the demand side of the economy and its possible impact on the real rates. Basic tastes and technology may have changed considerably during the 1980s. This would exert an upward pressure on the demand side, which may have had considerable effect on the interest rates beyond those determined by money and prices alone. Our own evidence has pointed in the direction that change in real income was a trivial proportion of the real interest rate variance decomposition in all segments of the financial market. Nevertheless, these findings do not discharge the importance of changes in preferences and production conditions.

5 Concluding remarks

The high levels of market interest rates that prevailed in Brazil during the 1980s had a positive relationship with lagged values of monetary expansion. This was documented by an inverted Cagan-Gandolfi curve. This occurrence represents a departure from the rational expectation approach and an overshooting of the monetarist view. We found that this behavior was more prevalent under higher inflationary regimes, which is an indication of the presence of the Fisher effect. It is important to assert that although the directions of the responses were uniform among the segments of the private financial market, they varied quantitatively.

43 See Jorgensen (1990), for some figures about the size of this market.

44 See Blumenschein (1995).

On the other hand, the responses of the overnight repurchase contract market have confirmed the presence of the liquidity effect. Moreover, inflation had little or no influence on the timing pattern of money and the interest rate in this segment. These results indicate that the linkages between the overnight market and other segments are far from perfect. The market rates of interest can react differently and consequently behave differently at least in the short term.

It is important to notice that, besides the existence, as well as the strength, of each type of effect we investigated, the speed of the response may determine the neutrality of a monetary shock. In this regard, some considerations have to be given to factors, or scenarios, that affect the pace of the response. Some theories give emphasis to the formation of rational expectation as a factor accelerating the response. Nonetheless, little emphasis has been given to high rates of inflation, or to the structure of the financial market, as important causes of overreaction. These factors may contribute substantially to the neutrality of money, especially for economies of the developing countries.

One implication of our investigation is that the government lost one important policy tool. Monetary policy stances, either by monitoring monetary aggregates or by targeting the interest rate in the overnight market, have low effectiveness in affecting market interest rates. Hence, stabilization programs cannot count on monetary policy, at least in the short term.

Another important implication is related to the monetary transmission mechanism and, consequently, capital formation and economic growth. Imperfect and slow adjustment in the structure of the interest rate coupled with flight from productive real assets may have had substantial effect in diminishing investment in the long term.

References

Análise Financeira. *A taxa de juros no Brasil*. Edição Especial. São Paulo: Nova Análise Editora Ltda, 1991.

Apecão. *A economia brasileira e suas perspectivas*. Rio de Janeiro: Edições Apec No. XXX, 1991.

Blejer, Mario. Money and the nominal interest rate in an inflationary economy. *Journal of Political Economy* 86, p. 529-34, June 1978.

Blumenschein, Fernando. *Essays on macroeconomic policy and the institutional framework of the financial market in Brazil*. Ph.D. Dissertation, Cornell University, 1994.

_____. A integração do mercado financeiro brasileiro durante o período 1978/90. *Pesquisa e Planejamento Econômico* 25, p. 449-78, dezembro 1995.

Cagan, Phillip. *The channels of monetary effects on interest rate*. New York: Columbia University Press, 1972.

Cagan, Phillip and Gandolfi, Arthur. The lag in monetary policy as implied by the time pattern of monetary effects on interest rate. *American Economic Review, Paper and Proceedings* 59, p. 277-84, May 1969).

Clifton, Eric V. Real interest rate targeting: an example from Brazil. *International Monetary Fund, Working Paper No. 111*, 1990.

Cochrane, John H. The return of the liquidity effect: a study of the short-run relation between money growth and interest rates. *Journal of Business and Economic Statistics* 7, p. 75-83, January 1989.

Colomiris, Charles W. and Domowitz, Ian. Asset substitution, money demand, and the inflation process. *Journal of Money, Credit, and Banking* 21, p. 79-89, February 1989.

Conard, Joseph W. *An introduction to the theory of interest rate*. Los Angeles: University of California Press, 1959.

Cooley, Thomas F and LeRoy, Stephen F. Atheoretical macroeconomics. A critique. *Journal of Monetary Economics* 16, p. 283-308, November 1985.

Davidson, Russell and MacKinnon, James G. Several tests for model specification in the presence of alternative hypothesis. *Econometrica* 49, p. 781-93, May 1981.

Dornbusch, Rudiger, Sturzenegger, Federico and Wolf, Holger. Extreme inflation: dynamics and stabilization. *Brookings Papers on Economic Activity*, p. 2-84, 1991

Fisher, Irving. *The theory of interest rate*. New York: The Macmillan Company, 1930.

- Friedman, Milton. The role of monetary policy. *The American Economic Review* 58, p. 1-17, March 1968.
- _____ Comments on Friedman. *Quarterly Journal of Economics* 84, p. 318-27. May 1970.
- Friedman, Milton and Schwartz, Anna J. *Monetary trends in the United States and the United Kingdom. Their relations to income, prices, and interest rates, 1867-1975*. Chicago and London: The Chicago University Press, 1982.
- Garcia, Marcio G. P. *The formation of inflation expectation in Brazil*. Ph.D. Dissertation. Stanford University, 1991.
- Gertler, Mark. Financial structure and aggregate economic activity: an overview. *Journal of Money, Credit, and Banking*, 20, p. 559-88, August 1988, Part 2.
- Guidotti, Pablo E. and Kumar, Manmohan S. Domestic public debt of externally indebted countries. International Monetary Fund, *Occasional Papers* N° 80, June 1991.
- Hanson, James A. The short-run relation between growth and inflation in Latin America: a quasi-rational or consistent expectation approach. *American Economic Review* 70, p. 972-89, December 1980.
- Jorgensen, Erika. Domestic public debt in Brazil: a descriptive analysis with policy implications. The World Bank, May 1990, *mimeo*.
- Khor, Hoe E. and Rojas-Soarez, Liliana. Interest rate in Mexico. The role of exchange rate expectations and international creditworthiness. International Monetary Fund, *Staff Papers* 38, p. 850-71, December 1991
- Lees, Francis A., Botts, James M. and Cysne, Rubens P. *Banking and financial deepening in Brazil*. New York: St. Martin's Press, 1990.
- Lutz, Friedrich A. *The theory of interest rate*. Chicago: Aldine Publishing Company, 1968.
- Mehra, Yash. Inflationary expectations, money growth, and the vanishing liquidity effect of money on interest: a further investigation. Federal Reserve Bank of Richmond, *Economic Review* 71, p. 23-35, March/April 1985.
- Melvin, Michael. The vanishing liquidity effect of money on interest: analysis and implications for policy. *Economic Inquiry* 21, p. 188-202, April 1983.

- Poole, William. Book Review. *Journal of Political Economy* 80, p. 665-68, May/June 1974.
- Rossi, J. W. Oferta monetária, nível de atividade econômica e inflação. *Revista Brasileira de Economia* 45, p. 91-103, janeiro/março 1991.
- Sachs, Jeffrey D. Comments. *Brookings Papers on Economic Activity*, p. 157-62, 1982.
- Sims, Christopher A. Macroeconomics and reality. *Econometrica* 48, p. 1-48, January 1980.
- _____ Policy analysis with econometric models. *Brookings Papers on Economic Activity*, p. 107-67, 1982.
- Stokes, Houston H. and Neuburger, Hugh. The effect of monetary changes on interest rates: a Box-Jenkins approach. *Review of Economic and Statistics* 61, p. 534-48, November 1979.
- Sundararajan, Vasudevan, and Balino, Tomas J. T. *Banking crises: cases and issues*. Washington D.C.: International Monetary Fund, 1991.
- Tanner, J. Ernest. Are the lags in the effect of monetary policy variable? *Journal of Monetary Economics* 5, p. 105-21, January 1979.
- Thornton, Daniel L. The effect of monetary policy on short-term interest rate. Federal Reserve Bank of St. Louis, *Review* 70, p. 53-72, May/June 1988.
- Tinsley, P. A. An application of variable weight distributed lags. *Journal of the American Statistical Association* 62, p. 1277-89, December 1967
- Tobin, James. Money and income: post hoc ergo propter hoc? *Quarterly Journal of Economics* 84, p. 310-17, May 1970.
- Tobin, James and Brainard, Willian. Financial intermediaries and the effectiveness of monetary controls. *American Economic Review, Papers and Proceedings* 53, p. 383-400, 1963.

A crise asiática e o papel do FMI

Ruben Dario Almonacid[§]

“...o montante das correções cambiais que se verificaram (nos países asiáticos) excedem qualquer estimativa razoável do que seria necessário para corrigir uma supervalorização inicial do Baht etc. Nesse ponto os mercados reagiram exageradamente...”

Stanley Fischer, Vice Diretor Gerente do FMI, em
O Estado de São Paulo 25/01/98

“O FMI não quer ser o banqueiro de último recurso, a quem todos pedem ajuda se há um problema.”

M. Camdessus, Diretor do FMI - Em *El Pais* 02/
12/97

1 Introdução

O México e os Tigres Asiáticos são apenas alguns exemplos recentes de como o mundo acadêmico, os organismos internacionais, a comunidade financeira internacional e a imprensa podem mudar abruptamente de opinião, transformando os que se consideravam modelos de política econômica, a serem seguidos pelo mundo em desenvolvimento, em culpados das maiores ineficiências e merecedores do castigo que receberam: Instabilidade e Crises Econômicas.

Duas interpretações possíveis para esta mudança de opinião são: ou os países eram de fato modelos e em algum momento mudaram a política deixando de sê-lo, ou eles nunca o foram, havendo apenas um erro de diagnóstico e o reconhecimento do erro. Neste caso,

[§] PhD em Economia pela Universidade de Chicago. Livre Docente e ex-professor da FEA-USP.

porém, como explicar que todos errem ao mesmo tempo e na mesma direção e que estes erros sejam tão grandes e se repitam com tanta frequência? E se fosse assim, de onde surge a “luz” para que o reconhecimento do erro aconteça de uma vez?

Uma interpretação alternativa seria a de que o mundo acadêmico e os organismos internacionais não têm um diagnóstico correto de qual é a fonte dos problemas. Isto explicaria por que as soluções recomendadas têm sido contraditórias e mutantes no tempo. O FMI, o responsável mor pela estabilidade dos mercados mundiais, tem mudado diversas vezes seu ponto de vista e suas recomendações de política econômica. A sua recente defesa de maior liberdade para os movimentos de capitais parece demonstrar que o Fundo não percebe que a origem de todos os problemas de instabilidade está precisamente na grande mobilidade de capitais, sem um órgão que a regule e controle, e que esteja disposto a funcionar como fornecedor de liquidez de última instância quando os movimentos atingem dimensões de corrida.

Durante o século passado e começo deste até a grande depressão, particularmente, inúmeras casas bancárias nos EUA e no resto do mundo fecharam suas portas, vítimas do pânico e corridas financeiras. Muitas delas eram instituições sólidas que não mereceram a sorte que lhes coube. Elas sofreram as conseqüências da existência de um marco institucional inadequado e do predomínio de idéias econômicas equivocadas.

Geralmente essas corridas financeiras evoluíam para crises econômicas e recessões e no processo levavam para a falência uma infinidade de empresas do setor produtivo, sendo que, também neste caso, muitas delas eram sólidas, eficientes e produtivas. O desemprego, a fome, a miséria e as doenças eram apenas alguns efeitos colaterais do comportamento de pânico que se difundia nessas circunstâncias (qualquer semelhança com a atual situação da Indonésia ou da Rússia não é mera coincidência).

Com a evolução da Ciência Econômica descobriu-se que estas catástrofes não eram males necessários e, portanto, que poderiam ser evitadas. Os Bancos Centrais passaram a regular a liquidez e servir de emprestadores de última instância (*lenders of last resort*), dispostos a ajudar as instituições financeiras que sofriam ataques especulativos e, dessa forma, minimizar todas as conseqüências nocivas do ciclo econômico. Décadas ininterruptas de prosperidade como as vividas pelos EUA pareciam o prêmio merecido pela melhor compreensão dos fenômenos econômicos.

No entanto, aquilo que se tornou obsoleto no mundo desenvolvido continua assolando ainda hoje os países emergentes (e até o Japão). Com a globalização, as crises são cada vez mais rápidas e ocorrem em maior escala e frequência. As mais recentes, do México,

dos Tigres Asiáticos, da Rússia e as outras várias a caminho, são apenas alguns exemplos de um fenômeno que já vivemos no passado.

Parece totalmente absurdo e paradoxal que o maior credor do mundo e que tem também o maior superávit em transações correntes há bastante tempo, o Japão, esteja tendo uma corrida na sua moeda, tenha sido obrigado a desvalorizá-la em mais de 50% e tente - em vão - defendê-la todos os dias dos ataques especulativos. Na nossa interpretação, o porquê desse fenômeno e sua solução parecem claros.

O fenômeno econômico do mundo globalizado de hoje é idêntico ao que levava às corridas bancárias e às crises do mundo desenvolvido até a década de 40. A sua solução, que se faz cada dia mais urgente, já existe. Faltam, talvez, apenas o reconhecimento da semelhança do problema e a vontade política para implementá-la.

Este trabalho apresenta duas proposições principais. A primeira pretende demonstrar que é da natureza das economias de mercado exagerar no valor que atribuem a seus ativos. Quando o exagero é generalizado e para cima, surgem as bolhas especulativas; quando para baixo, advêm as corridas e as crises, que, por analogia, poderíamos chamar de “bolas murchas”. Atualmente estamos vivendo ambas as situações simultaneamente em diferentes partes do mundo. A segunda analisa este modelo, mais especificamente quando aplicado ao mercado de câmbio, e tenta mostrar como a tendência ao exagero das economias de mercado afeta o balanço de pagamentos dos países e contribui para gerar a situação de instabilidade cambial pela qual passam hoje diversos países.

2 A tendência das economias de mercado para os exageros¹

Numa economia de mercado, o preço de todo tipo de ativo é definido como o valor presente dos fluxos futuros de retorno que o ativo promete pagar. Existem dois tipos de fluxos de retorno: (1) um fluxo de rendas - pecuniárias, como aluguéis, dividendos, juros etc. ou não-pecuniárias, como os serviços que um carro ou uma casa oferecem a quem se utiliza do bem, e (2) um fluxo de ganhos de capital, que resulta da mudança do preço do ativo ao longo do tempo. Todo ativo (que tem um preço maior do que zero) necessariamente vai oferecer algumas dessas formas de fluxo de retorno. Assim, por exemplo:

¹ As principais conclusões desta seção estão formalizadas no item A.1 do Apêndice Matemático.

- a) Um imóvel que rende R\$ 1.000,00 de aluguel por mês, e que se espera que seu preço permaneça constante ao longo do tempo, terá um preço igual ao aluguel mensal dividido pela taxa mensal de juro real. Se esta for de 1% ao mês, o imóvel valerá R\$ 100.000,00.
- b) Uma ação que paga dividendos e tem seu preço esperado, valorizando-se ao longo do tempo, terá um preço igual à soma do valor presente dos dividendos e do preço esperado para o futuro.
- c) Um bônus de cupom zero (que não paga nada durante sua existência, mas promete pagar um valor fixo no resgate) terá como preço de mercado o valor presente desse valor fixo, sempre descontado à taxa de juros de mercado.

Mas, vamos supor que um imóvel que rende um aluguel de R\$ 1.000,00 encontre-se num bairro que esteja se valorizando e o mercado espera que seu preço cresça à taxa ρ por mês. Esse imóvel, além do aluguel, estará dando um ganho de capital e seu preço de mercado será maior que os R\$ 100.000,00 do exemplo "a". Se $\rho = 0,5\%$ ao mês, o preço do imóvel estaria subindo aproximadamente R\$ 500,00 por mês e, conseqüentemente, seu preço de mercado incorporaria essa valorização esperada e hoje poderia valer R\$ 150.000,00, por exemplo.²

Analogamente, um ativo cujo preço vai se deteriorando ao longo do tempo, como um automóvel ou um computador, que perdem valor pelo desgaste ou pela obsolescência produzida pelas novas tecnologias, ou, no caso do imóvel, um localizado em um bairro decadente (como alguns de centro de diversas cidades), terá seu preço de mercado caindo e dando um ganho de capital negativo, uma perda, portanto, que também deverá ser incorporada ao preço. Nesse caso, o preço deverá ser menor que o valor presente do fluxo esperado de retornos. Assim, o imóvel que rendia R\$ 1.000,00 por mês poderia valer R\$ 70.000,00 ao invés dos R\$ 100.000,00 que valia quando não havia nem ganhos nem perdas de capital.

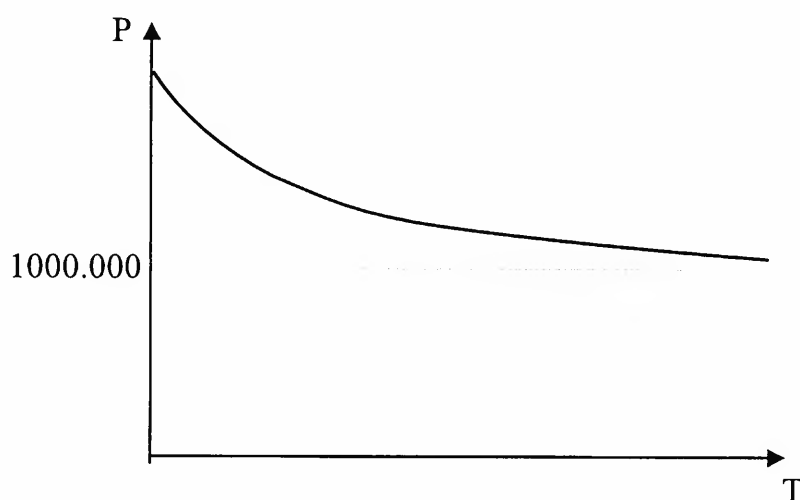
Continuando com o exemplo do imóvel, por ser um ativo com um preço de mercado mais simples de ser determinado, vamos mostrar como se calcula esse preço quando existem ganhos de capital e qual a relação desse preço com o que prevalecia quando não havia estes ganhos.

2 Este exemplo supõe que a valorização do ativo não se reflita na sua renda, como é usual no caso do mercado acionário. Se o aumento de preço afeta a renda, como freqüentemente acontece com os imóveis, o aumento de renda pode ser permanente ou temporário. Desvios de preços ocorrerão tanto quando a renda não muda como quando essas mudanças são temporárias.

Simplificadamente, o preço do imóvel é igual ao valor presente dos aluguéis a serem recebidos, por exemplo, nos próximos 2 anos, mais o valor presente do preço esperado para daqui os mesmos 2 anos. Se o imóvel se valoriza 0,5% ao mês, o preço esperado para daqui a 24 meses será 12,72% maior que o preço atual. Este preço, porém, não poderá ser hoje de R\$ 100.000,00, senão o ativo estaria rendendo mais que os outros (R\$ 1.000,00 de aluguel + R\$ 500,00 de valorização); todo mundo iria querer comprá-lo, elevando seu preço.

O preço do imóvel, neste caso, teria que ser R\$ 188.716,78, de forma que os R\$ 1.000,00 de aluguel mais os 0,5% de valorização dêem 1% de retorno, como todos os outros ativos. Realizando este mesmo experimento para todos os possíveis prazos, de 1 mês a 1.000 meses, obtemos uma curva com a relação entre o preço do imóvel e o prazo para quando é esperado o preço futuro. Esta curva é mostrada na Figura 1.

Figura 1



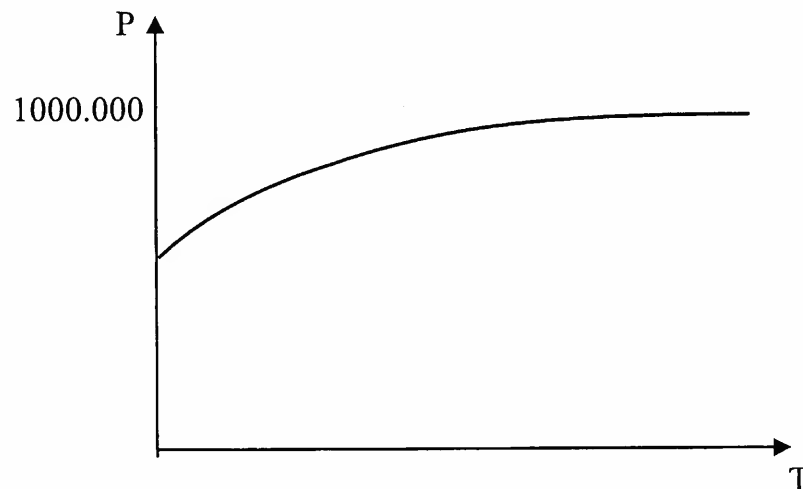
O preço varia entre R\$ 200.000,00 para o primeiro mês e R\$ 100.669,95 para o milésimo mês. Na verdade, à medida que o prazo aumenta, o preço se aproxima de R\$ 100.000,00 continuamente:

$$\lim_{t \rightarrow \infty} P(\rho = 0.5\% \text{ a .m.}) = \text{R\$ } 100.000,00,$$

nossos velhos R\$ 100.000,00 de quando não havia ganhos de capital!

A mesma conta para um imóvel que se desvaloriza 1% ao mês é apresentada na Figura 2.

Figura 2



Neste caso, o preço varia entre R\$ 50.505,00 para o primeiro mês e R\$ 99.995,46 para o milésimo mês, novamente, e coincidentemente, o preço tende para o nosso velho conhecido: os R\$ 100.000,00;

$$\lim_{t \rightarrow \infty} P (\rho = -1\% \text{ a .m.}) = \text{R\$ } 100.000,00$$

e esta relação vale para qualquer ρ entre r (a taxa de juro de mercado) e $-\infty$. Como o valor presente dos fluxos esperados de retornos deve ser positivo, o ρ não pode ser maior que r , porque aí o preço do ativo ficaria indeterminado.

Vamos supor, como antes, que os agentes econômicos trabalhem com expectativas para 2 anos na frente ($T = 24$), que o preço inicial seja R\$ 100.000 e que $\rho = 0.5\%$. Como o preço não irá saltar imediatamente para os R\$ 188.716,78, haveria um período em que o ativo estaria rendendo mais do que 1%. Será a fase de expansão. Mas, tão pronto este preço seja atingido, talvez antes, talvez depois, o ativo deixará de oferecer ganhos de capital e, neste caso, qualquer preço acima de R\$ 100.000 implicará que ele estará rendendo menos do que 1% e iniciará sua descida (a fase de contração). Portanto, com os supostos aqui definidos, a expansão já leva consigo as sementes da próxima contração. A descida poderia ser tal que, quando se atingisse o preço de R\$ 100.000, as perdas de capital já estivessem próximas de zero e portanto o equilíbrio fosse restabelecido. Isto daria uma aterrissagem suave (*soft landing*). Mas é possível também que quando o preço

chegasse aos R\$ 100.000 ainda existissem perdas de capital significativas e portanto que o preço do ativo continuasse caindo (haveria exagero para baixo - *undershooting*) e a aproximação final ao equilíbrio seria oscilatória.³

O preço de mercado de um ativo que paga uma renda mensal de R\$ 1.000,00 em perpetuidade, com juro de 1% ao mês, poderá variar entre, aproximadamente, R\$ 50.000,00 e R\$ 200.000,00, dependendo do horizonte com o qual trabalham os agentes econômicos e de seu humor! Se eles estão pessimistas, valerá menos; se otimistas, valerá mais, ainda que o preço de equilíbrio de longo prazo seja sempre R\$ 100.000,00. Mais ainda, se além de deixar variar apenas ρ , como no exemplo, fossem permitidas variações na taxa de juros de mercado e no fluxo esperado de retornos (que também dependem das expectativas dos agentes econômicos), as oscilações de preço que se conseguem poderiam ser ainda maiores.⁴

Este tipo de comportamento é particularmente relevante para dois mercados macroeconômicos importantes: A Bolsa e o Mercado Cambial. O Mercado Imobiliário está se tornando crescentemente significativo, porque o efeito riqueza que ele gera tem implicações importantes para as decisões de consumo e investimento, bem como para o equilíbrio externo. Os movimentos de capitais, que podem variar mais rapidamente e em valores maiores que os fluxos de bens e serviços, jogam um papel desestabilizador nesse processo. É o que analisa a próxima seção.

3 Os movimentos de capitais e os processos especulativos⁵

Nos últimos anos temos observado que valorizações exageradas de três tipos de ativos imóveis, ações e câmbio - foram responsáveis por surtos de euforia ou pânico e movimentos de capitais em volumes elevados que levaram a bolhas e corridas especulativas. A origem dessa relativa instabilidade dos mercados estaria na dinâmica do setor externo. A explicação se baseia no seguinte raciocínio:

3 Num mundo de expectativas racionais, os agentes econômicos trabalhariam com horizontes infinitos e não poderia haver bolhas especulativas, nem "bolas murchas"

4 Aumentos no fluxo esperado aumentam o preço de equilíbrio; aumentos no juro, o reduzem.

5 A formalização matemática desta seção é desenvolvida no item A.2 do Apêndice Matemático.

O Balanço de Pagamentos (BP) é definido como:

$$BP = STC + SCK = \Delta R, \quad (1)$$

onde STC é o Saldo em Transações Correntes e abrange todas as transações com bens e serviços de um país com o resto do mundo. SCK é o Saldo da Conta de Capitais e inclui todas as suas transações financeiras com o exterior. O BP também é igual à variação das reservas internacionais (ΔR).

Quando o BP está em equilíbrio, STC é igual a $-SCK$ e as reservas internacionais estão constantes.

O STC é uma função da taxa real de câmbio (π), da renda do país (y) e da do resto do mundo (y_w).

$$STC = f(\pi, y, y_w) \text{ com } f_1 > 0, f_2 < 0 \text{ e } f_3 > 0. \quad (2)$$

Aumentos da taxa real de câmbio e da renda do resto do mundo melhoram o STC, aumentos de y o pioram.

O SCK é uma função do diferencial entre o juro na moeda local (i) e o juro de aplicações na moeda estrangeira ($\theta + \hat{\pi}^e$):

$$SCK = g(i - \theta - \hat{\pi}^e), \text{ com } g' > 0, \quad (3)$$

onde θ é o juro válido no resto do mundo e $\hat{\pi}^e$ a taxa esperada de desvalorização da moeda local. Um aumento do juro local tende a gerar uma entrada de capitais; um aumento da desvalorização esperada, uma saída de capitais.

Postulamos também que a taxa de câmbio tende a variar além das expectativas, desvalorizando-se quando o BP é deficitário ($\Delta R < 0$) e valorizando-se quando o BP é superavitário ($\Delta R > 0$):

$$\hat{\pi} = h(BP) + \hat{\pi}^e = h(STC + SCK) + \hat{\pi}^e = h(\Delta R) + \hat{\pi}^e \text{ com } h' < 0. \quad (4)$$

Finalmente, postulamos que a taxa esperada de desvalorização é uma função positiva da taxa observada de desvalorização, por exemplo, nas expectativas adaptativas:

$$\frac{d\hat{\pi}^e}{dt} = \ell(\hat{\pi} - \hat{\pi}^e), \text{ com } \ell' > 0. \quad (5)$$

Como no caso da demanda de moeda numa economia fechada, a estabilidade deste processo dinâmico depende de que $f_1 > \ell'g$.⁶ O impacto de uma desvalorização da moeda no STC tem que ser maior que seu impacto sobre a taxa esperada de desvalorização da moeda multiplicado pelo impacto desta sobre o SCK.

Um aumento do juro interno gera um influxo de capitais [eq. (3)]. Este, por sua vez, tende a revalorizar o câmbio [eq. (4)], o que afetará as expectativas quanto à taxa de câmbio [eq. (5)], levando à sua valorização ($\pi \downarrow$). Isto aumentará novamente o juro recebido pelos estrangeiros e induzirá novos influxos de capitais. Este círculo vicioso é agravado porque a entrada de capitais gera uma expansão monetária e, se o Banco Central quiser esterilizá-la, irá colocar no mercado dívida interna, pressionando os juros internos para cima, fortalecendo ainda mais o influxo de capitais. Esta é a situação que o Banco Central do Brasil viveu diversas vezes no passado, sendo a última vez quando da introdução do Real.

Apresentamos, a seguir, quatro exemplos recentes de como essa dinâmica perversa tem contribuído para desestabilizar as economias de alguns países emergentes e do Japão.

1. (STC < 0, SCK > 0, |STC| < SCK $\Rightarrow \pi \downarrow$) Países como o Brasil e a Argentina, que sofriam processos hiperinflacionários e introduziram planos de estabilização, tiveram inicialmente que manter juros elevados para quebrar as expectativas inflacionárias. Estes juros reais elevados induziram influxos maciços de capitais, o que levou a uma apreciação do câmbio, reforçando ainda mais os juros reais positivos para os

6 Cagan, Philips. *The Dynamics of Hyperinflations*.

O modelo do mercado monetário em uma economia fechada seria: (1) $m^d = f(y, i)$ com $f_1 > 0$ e $f_2 < 0$, onde m^d é a demanda real de moeda, y a renda e i o juro nominal, definido como $i = \theta + \hat{p}^e$ onde θ é o juro real e \hat{p}^e a taxa esperada de inflação; (2) $\hat{p} = h(m - m^d) + \hat{p}^e$, $h' > 0$, a inflação é uma função do excesso de oferta de moeda e da taxa esperada de inflação; e (3) $\frac{d\hat{p}^e}{dt} = \ell(\hat{p} - \hat{p}^e)$, $\ell' > 0$, a inflação esperada é uma função positiva da taxa observada de inflação.

Neste caso, para que o processo dinâmico seja estável, a elasticidade preço da demanda de moeda vezes a elasticidade das expectativas inflacionárias (quanto aumenta a taxa esperada de inflação quando a inflação atual aumenta 1%) deve ser menor do que 1 em valor absoluto. Também neste caso o processo dinâmico impõe que haja exageros (*overshooting*), pois a inflação acumulada deve exceder a expansão monetária acumulada, isto é, para uma maior inflação, o estoque real de moeda deve cair.

investidores externos. Como resultado deste processo, o câmbio atrasou e a balança comercial, que era altamente positiva, e o STC se tornaram crescentemente negativos, agravando a situação da dívida externa e a necessidade de manter juros reais elevados. Este processo aparentemente se encaminha para um desenlace traumático.

2. ($STC > 0$, $SCK < 0$, $STC > |SCK| \Rightarrow \pi \downarrow$). O Japão no começo desta década, apresentava elevados superávits em transações correntes e saídas de capitais, mas estas, em ritmo menor que o saldo em transações correntes. Como consequência, a moeda se apreciava, o que reforçava a remuneração esperada já elevada em iens e reduzia o incentivo à saída de capitais e reforçava o desequilíbrio entre STC e SCK. A moeda chegou a valer 85¥/US\$.
3. ($STC < 0$ e $SCK < |STC| \Rightarrow \pi \uparrow$). SCK inicialmente positivo e, depois, negativo. O Tigres Asiáticos de 97-98 e a Rússia de 98. Com o STC negativo e SCK positivo, mas não suficientemente grande para financiá-lo, existe uma tendência para se desvalorizar a moeda. Isto reduz a remuneração esperada para as aplicações na moeda local e incentiva a saída de capitais. A desvalorização que acaba acontecendo para tentar melhorar o STC confirma as expectativas de desvalorização e favorece ainda mais a saída de capitais. O processo de reequilíbrio é um longo e sofrido caminho de recessão até que o STC atinja valores compatíveis com o SCK. Normalmente, só quando o STC se torna positivo, e deixa de ser necessário desvalorizar a moeda, é que o SCK deixa de pressionar o câmbio. A esta altura, o câmbio já terá ido muito além do que seria necessário para restabelecer o equilíbrio externo.
4. ($STC > 0$ e $STC < |STK| \Rightarrow \pi \uparrow$) Japão 98. O maior credor do mundo enfrenta uma situação recessiva e dificuldades no setor financeiro. Abaixam-se os juros para induzir aumentos da demanda agregada, mas isto torna a remuneração em iens muito inferior à remuneração em outras moedas. Há um aumento da saída de capitais acima do necessário para equilibrar o BP e isto põe pressão no câmbio. A desvalorização do câmbio confirma as expectativas de que manter recursos em iens é um mau negócio. Há novamente maior saída de capitais e maior desvalorização da moeda. A moeda perde o valor e chega a ¥146/US\$ e ainda existem fortes pressões para a saída de capitais. Os juros já caíram tudo o que poderiam (atualmente eles estão em 0,5% ao ano) e a saída de capitais supera os US\$ 150 bilhões. A economia se desorganiza, o consumo, o investimento, a renda, o nível de emprego caem, gerando recessão. As recomendações dos especialistas vão desde expandir a oferta de moeda (Krugman) até aumentar despesas e cortar impostos ou aumentar a inflação (FMI). Todas seriam, no nosso entender, medidas contrárias às necessárias para restabelecer o equilíbrio entre

SCK e STC, que seria diminuir a saída de capitais e a necessidade de desvalorizar a moeda (aumentando os juros, se necessário).⁷

Fica claro com esses exemplos que para os fins da dinâmica deste processo o que importa é a relação entre o STC e o SCK e não o valor absoluto de um ou de outro. Portanto, o Japão, que é altamente superavitário, e a Rússia, que é altamente deficitária, no STC, podem ser ambos vítimas da mesma corrida especulativa.

As alternativas tecnicamente corretas para resolver este problema parecem ser duas: ou a) se cria uma moeda única, e com isso desaparece a possibilidade de especular contra a moeda de um país - a política monetária pode se concentrar, mais uma vez, em evitar ou, pelo menos, minimizar exageros nos mercados de outros ativos (ações, imóveis e bens e serviços), como numa economia fechada; ou b) se cria um órgão supranacional, como o FMI, que tenha como objetivo evitar ou pelo menos minimizar as corridas contra a moeda dos países, da mesma forma que os Bancos Centrais nacionais minimizam as corridas contra o seu sistema bancário.

Para que esta função seja corretamente executada, isto é, para que o BC (e por analogia o FMI) tenha sucesso em conter uma corrida especulativa, é necessário que: (1) ele atue com rapidez, para evitar que a especulação ganhe forças; (2) ele aja com a intensidade necessária para convencer os especuladores que a intenção é mesmo cortar o processo; (3) ele tenha um sistema de liberação de recursos automático (sem condicionalidade) e (4) a corrida que se quer evitar tenha seus fundamentos corretos. Não se evita a corrida de um banco falido nem se defende uma moeda que está desajustada (primeiro se corrige o desajuste e depois se defende). O FMI normalmente não atende a nenhuma destas exigências. Tem agido com morosidade, tem prometido liberar recursos em volumes insuficientes, tem imposto condições e, quando estas não são satisfeitas, segura os recursos e tem defendido países que nem sempre estavam com suas condições macroeconômicas em ordem. Com esta política, em muitos casos, o FMI tem exacerbado os processos especulativos. Não é de se estranhar, então, que sua atuação como organismo responsável por evitar corridas especulativas tenha, até agora, sido um fracasso.⁸

7 Qualquer expansão monetária adicional, redução de impostos ou aumento de gastos apenas dariam munição adicional para novas saídas de capitais. O governo acaba de anunciar uma nova redução da taxa de juros, agora para 0.25% ao ano, e o ien fica novamente atacado.

8 No caso da crise asiática, por exemplo, primeiro deveria ter sido dada uma desvalorização de 10 a 15% que era a magnitude aproximada do desajuste, e depois defendido a moeda com todos os instrumentos à disposição. Isto teria resolvido o problema e cortado a especulação. Por outro lado, os mais de 80% de desvalorização da rúpia, ou do rublo, não serão suficientes para resolver o problema, além de todos os custos sociais, em termos de renda, emprego, destruição de investimentos, saques, situações de pânico, fome e doenças que irão acontecer por conta da desorganização institucional que se segue a toda corrida generalizada, como as que estão sofrendo Indonésia e Rússia, entre outros. É claro também que nem a economia mais forte do mundo suportaria uma desvalorização de 80% de sua moeda sem sofrer todas as conseqüências que estas duas economias estão sofrendo.

Mas, enquanto a solução definitiva para as corridas especulativas não chega, os países têm que usar todos os instrumentos de política econômica disponíveis para evitar desequilíbrios no BP (STC muito diferente de SCK), que geram pressão sobre o câmbio e fluxos de capitais elevados, especialmente se eles se acumulam sob a forma de dívida externa elevada, porque a dívida torna o país dependente das expectativas e da vontade dos estrangeiros (ela retira a liberdade de usar a política econômica para atingir objetivos de interesse do país).

4 Conclusões

Este trabalho parte do pressuposto de que o preço dos ativos em uma economia de mercado é obtido calculando-se o valor presente do fluxo esperado de retornos e dos ganhos (positivos ou negativos) de capital. O preço, assim obtido, excede (para ganhos positivos) e é menor (para ganhos negativos) que o preço que definiria o equilíbrio de longo prazo, isto é, o que exclui os ganhos de capital. Simulações realizadas para valores plausíveis da taxa de juros e de ganhos de capital mostram que as discrepâncias entre esses dois preços podem ser grandes, implicando que os exageros factíveis também poderão ser grandes. Há, portanto, espaço considerável para bolhas especulativas (quando o preço de mercado excede o valor de equilíbrio) e para “bolas murchas” (quando o preço fica abaixo do valor de equilíbrio). Nesta interpretação, a volatilidade e as grandes flutuações de preços observadas, particularmente em tempos recentes, nos mercados acionário, imobiliário e de câmbio, seriam reflexos normais do comportamento dinâmico nos mercados de ativos, isto é, elas já pressupõem que as condições de estabilidade estão sendo satisfeitas.

Quando esta idéia é aplicada ao mercado de câmbio, obtêm-se resultados para o STC, o SCK e o BP compatíveis com os que se observam na crise atual. Todas as corridas especulativas, tanto as que produzem bolhas especulativas como as que levam às “bolas murchas” são, em última instância, altamente destrutivas de riqueza e de bem-estar e deveriam ser evitadas.

À luz dessas conclusões, existem duas formas alternativas de amenizar a situação de turbulência observada: (1) criando um sistema monetário com uma moeda única e utilizando o instrumental de política monetária para evitar exageros no mercado de outros ativos e (2) criando um órgão supranacional, talvez o FMI, mas com quantidade suficiente de ativos disponíveis e vontade de intervir nos mercados para minimizar flutuações desequilibrantes.

Enquanto estas soluções, que exigem uma grande movimentação política e mudanças no marco institucional, não acontecem, o governo dos países ricos deveriam unir-se e combater ativamente os movimentos especulativos. Deixar, como até agora, que a situação se agrave, pode implicar que o problema se generalize, afetando uma grande quantidade de países, transformando uma crise especulativa em uma grande recessão e até gerar respingos também para eles.

Apêndice matemático

Este apêndice matemático mostra formalmente algumas das proposições analisadas no texto:

A.1. O modelo de determinação do preço de um ativo:

A fórmula geral para a determinação do preço de um ativo qualquer é:

$$P_0 = \int_0^T D(t) e^{-\int_0^t r(s) ds} dt + P_T e^{-\int_0^T r(t) dt} \quad (1)$$

onde P_0 , o preço definido para o momento 0, tem dois componentes: o 1º representa o valor presente (VP) do fluxo de retornos, $D(t)$, descontado à taxa de juros de mercado, $r(t)$, ambos esperados para prevalecer entre o momento 0 e um momento futuro T qualquer. O segundo termo é o VP do preço do ativo esperado para esse momento T

Esse segundo termo, por sua vez, pode ser expresso como:

$$P_T = P_0 e^{\int_0^T \rho(t) dt} \quad (2)$$

o preço atual corrigido por uma taxa de ganhos de capital $\rho(t)$ esperada para prevalecer entre os momentos 0 e T

Se o fluxo de retornos, $D(t)$, a taxa de juros, $r(t)$, e a taxa de ganhos de capital, $\rho(t)$, são constantes, e utilizando (2), a expressão (1) simplifica para:

$$P_0 = D_0 \int_0^T e^{-r_0 t} dt + P_0 e^{(\rho_0 - r_0)T} \quad (3)$$

Resolvendo o integral e agrupando os valores em P_0 , temos:

$$P_0 [1 - e^{(\rho - r_0)T}] = \frac{D_0}{r_0} (1 - e^{-r_0 T}) \quad (4)$$

$$P_0 = A \cdot B^{-1} \quad (4')$$

onde $A = \frac{D_0}{r_0} (1 - e^{-r_0 T})$ é o VP do fluxo esperado de retornos e $B = [1 - e^{(\rho - r)T}]$ é a diferença entre a VP do Preço do Ativo no presente e no momento futuro T

Tomando diferentes valores para T e usando $r_0 = 1\%$, calculamos a Tabela 1, que mostra os valores de P_0 apresentados nas Figuras 1 e 2 do texto.

Tabela A.1
Cálculo do Preço de um Ativo que Inclui Ganhos de Capital

T	A	$B (\rho = 0,5\%)$	$P_0 (\rho = 0,5\%)$	$B (\rho = -1\%)$	$P_0 (\rho = -1\%)$
1	1000	,0050	200.000	,0198	50.505
10	9516	,0488	195.118	,1813	52.488
24	21.340	,1131	188.683	,3812	55.981
50	39.346,93	,2212	177.880	,6321	62.248
100	63.212,05	,3935	160.640	,8647	73.103
150	77.686,98	,5276	147.246	,9502	81.759
200	86.466,47	,6321	136.792	,9817	88.078
1000	99.995,46	,9933	100.669,95	1,0000	99.995,46

Observe-se que para que o objeto em análise seja um ativo seu preço deve ser positivo ($P_0 > 0$). Isto implica que:

dado $\rho \begin{matrix} \geq \\ < \end{matrix} r$, $B \begin{matrix} \leq \\ > \end{matrix} 0$, portanto, A também deverá ser $\begin{matrix} \leq \\ > \end{matrix} 0$ e vice-versa.

Vejam agora alguns casos particulares:

1. De (4), segue-se que se $\rho_0 = 0$ (i.e., se $P_T = P_0$, ou seja, não se esperam ganhos de

capital), $P_0 = \frac{D_0}{r_0} = P_0^E$ para todo T , que dá o que chamamos de preço de equilíbrio.

2. De (4), segue-se também que:

$\lim_{T \rightarrow \infty} P_0 = \frac{D_0}{r_0} = P_0^E$ para todo ρ_0 , i.e., as séries de preços que incluem uma taxa de

ganho de capital ρ_0 tendem para o mesmo preço de equilíbrio P_0^E qualquer que seja o valor de ρ .

3. Para que $A < 0$, é necessário que $D < 0$ por algum período entre 0 e T . Sendo $A < 0$, B deverá sê-lo também, o que implica que ρ deverá ser maior que r . Este caso, que corresponde ao de um imóvel em construção, deve necessariamente ter um ganho de capital superior à taxa de juros. Reciprocamente, um ativo que tem uma valorização superior à taxa de juros deve apresentar um fluxo de retorno negativo (fluxo de pagamentos) para que o preço seja positivo e finito.

4. No caso de um bônus de cupom zero, $A = 0$, portanto de (1), segue-se que:

$P_t = P_T e^{-r(T-t)}$, onde P_T é o valor fixo de resgate e P_t o preço do bônus em qualquer instante entre 0 e T

5. Se o ativo em questão é uma ação, $D(t)$ será o fluxo de dividendos pagos. Se o dividendo é uma fração dos lucros por ação, $[E(t)]$, e se estes, espera-se, crescem à taxa θ , teremos $E(t) = E_0 e^{\theta t}$ e $D(t) = \alpha E_0 e^{\theta t}$. Introduzindo estes valores em (4), teremos:

$$P_0 (1 - e^{(\rho - r)T}) = \frac{\alpha E_0}{r - \theta} (1 - e^{-rT})$$

e tomando limites, temos:

$\lim_{T \rightarrow \infty} P_0 = \frac{\alpha E_0}{r - \theta}$ que seria o preço de equilíbrio de uma ação, que distribui uma fração α dos lucros e tem estes crescendo à taxa θ .

6. Finalmente, utilizando a expressão (4') mais uma vez, temos:

$$\frac{1}{P_0} \cdot \frac{\partial P_0}{\partial \rho} = \frac{\partial A}{\partial \rho} \frac{1}{A} - \frac{\partial B}{\partial \rho} \frac{1}{B} > 0 \quad e$$

$$\frac{1}{P_0} \frac{\partial P_0}{\partial r} = \frac{\partial A}{\partial r} \frac{1}{A} - \frac{\partial B}{\partial r} \frac{1}{B} < 0,$$

isto é, P_0 é uma função crescente de ρ e decrescente de r .

A.2. A dinâmica do setor externo:

O modelo do balanço de pagamentos é:

$$BP = STC + SCK = \Delta R, \quad (1)$$

$$STC = f(\pi, y, y_w) \quad \text{com } f_1 > 0, \quad (2)$$

$$SCK = g(i - \theta - \hat{\pi}^e) \quad \text{com } g' > 0, \quad (3)$$

$$\hat{\pi} = h(BP) + \hat{\pi}^e \quad \text{com } h' < 0 \text{ e} \quad (4)$$

$$\frac{d\hat{\pi}^e}{dt} = \ell(\hat{\pi} - \hat{\pi}^e) \quad \text{com } \ell' > 0 \quad (5)$$

O problema consiste em encontrar as trajetórias dinâmicas para π e $\hat{\pi}^e$. Para este fim, linearizamos o sistema de equações diferenciais formado pelas equações (4) e (5):

$$\hat{\pi} = h' [f_1 (\pi - \pi_0) - g' (\hat{\pi}^e - \hat{\pi}_0^e)] + \hat{\pi}^e + \dots \quad e \quad (4')$$

$$\frac{d\hat{\pi}^e}{dt} = \ell' [(\hat{\pi} - \hat{\pi}_0) - (\hat{\pi}^e - \hat{\pi}_0^e)]. \quad (5')$$

Para calcular as raízes características deste sistema fazemos:

$$(\lambda - h' f_1) (\pi - \pi_0) + (h' g' - 1) (\hat{\pi}^e - \hat{\pi}_0^e) = 0 \quad e \quad (4'')$$

$$-\ell' \lambda (\pi - \pi_0) + (\lambda + \ell') (\hat{\pi}^e - \hat{\pi}_0^e) = 0. \quad (5''')$$

As raízes características λ_1 e λ_2 são obtidas da solução do determinante:

$$\begin{vmatrix} \lambda - h' f_1 & h' g' - 1 \\ -\ell' \lambda & \lambda + \ell' \end{vmatrix} = 0, \text{ ou seja:}$$

$$(\lambda - h' f_1) (\lambda + \ell') + \ell' \lambda (h' g' - 1) = 0 \quad \text{o que dá:}$$

$$\lambda^2 - h' (f_1 - \ell' g') \lambda - \ell' h' f_1 = 0,$$

que é uma equação quadrática da forma $\lambda^2 + b.\lambda + q = 0$

E para que haja duas raízes diferentes, ambas com sua parte real negativa, o que garante estabilidade ao processo dinâmico, é suficiente que:

$$\frac{b}{2} = -\frac{h'}{2} (f_1 - \ell' g') > 0 \quad e \quad q = -\ell' h' f_1 > 0$$

Dadas as hipóteses do modelo: f_1, g' e $\ell' > 0$ e $h' < 0$, temos que $q > 0$ sempre, e $\frac{b}{2} > 0$ se $f_1 > \ell' g'$. Isto é: o impacto de uma desvalorização da moeda no STC tem que ser maior que

seu impacto sobre a taxa esperada de desvalorização multiplicado pelo impacto desta sobre o SCK. Se $\left(\frac{b}{2}\right)^2 > q > 0$, ambas as raízes serão reais (não haverá parte imaginária) e, conseqüentemente, não haverá *overshootings*. Se $\left(\frac{b}{2}\right)^2 < q$, as raízes características terão parte imaginária e a aproximação ao equilíbrio será oscilatória.

No caso do modelo do BP, não poderemos transformar a condição de estabilidade em elasticidades, como no caso do modelo monetário, porque tanto BP, STC e SCK como $\hat{\pi}^e$ podem ser $\begin{matrix} \geq \\ < \end{matrix} 0$.

A arte do direito e da economia: um artigo autobiográfico*

William M. Landes[§]

I Introdução

Em seu artigo “Como Trabalho” Paul Krugman mostra que o formalismo crescente da moderna economia leva, hoje, a maioria dos estudantes de pós-graduação em economia a adquirir as habilidades matemáticas necessárias antes de entrar na escola de pós-graduação.¹ Suspeito fortemente de que também aconteça o inverso: o estudante que necessita de uma formação técnica será desencorajado a escolher uma carreira em economia. Não foi sempre assim. Como Krugman, entrei para a economia com uma formação em artes liberais, adquirindo habilidades técnicas quando necessário, tanto durante como depois da pós-graduação. Minha jornada, entretanto, foi mais tortuosa e sem planejamento que a de Krugman. Terminei um professor de economia e direito como resultado de uma improvável sucessão de eventos.

Comecei com uma especialização em arte na *High School of Music & Art* em New York. Embora especialização em arte também fosse exigida para atender ao modelo padrão dos cursos acadêmicos, não era um programa acadêmico árduo e foi possível cursá-lo razoavelmente bem sem muito esforço. A ênfase era claramente sobre artes e muitos pós-graduandos continuaram cursando arte e música em faculdades especializadas na área de New York. Desisti, pois era apenas um estudante médio. Também tentei arquitetura no curso secundário. Mas, neste ponto, não fui melhor e decidi abandoná-la, em parte mais porque meus amigos mais próximos tinham muito mais talento que eu.²

* Este artigo está sendo publicado com a autorização do próprio autor e da revista *The American Economist*. Originalmente este texto, “The Art of Law and Economics: An Autobiographical Essay”, foi publicado em Michael Szemberg (ed.), *Passion and Craft, Economists at Work*. Ann Arbor: Michigan University Press, 1997.

§ Professor de Direito e Economia *Clifton R. Mussel* da University of Chicago Law School. Gostaria de agradecer a Elizabeth Landes, Martha Nussbaum e Richard Posner pelos muitos comentários úteis. Como o leitor observará, o termo “arte” no título relaciona-se com o assunto desse ensaio em vários sentidos.

Tradução do inglês: Sylvia M. S. Cristóvão dos Santos.

1 Paul Krugman, How I work, 37. *American Economist* 25, 1993.

2 Aquele amigo, Charles Gwathmey, prosseguiu para tornar-se um dos principais arquitetos nos Estados Unidos, hoje.

Quando ingressei no *Columbia College*, aos dezessete anos, não estava bem preparado para seu exigente programa acadêmico (que permanece amplamente intacto até hoje). Tinha boa habilidade em artes, mas hábitos subdesenvolvidos de estudo. Jogar tênis e tocar piano, freqüentar clubes de jazz e só vadiar por *Greenwich Village* com meus amigos do curso secundário era mais interessante para mim do que estudar civilização ocidental e humanidades. Mas, em um pormenor, Música e Arte me ensinou uma lição valiosa. Impressionou-me a importância de ser criativo em uma profissão. Tenho levado esta lição comigo por toda a minha carreira acadêmica. Esforço-me para ser criativo tanto em minha escolha de tópicos como em minha abordagem a eles. Raramente tenho descoberto um tópico pelo exame minucioso da literatura econômica ou pela exploração das notas de rodapé, esperando encontrar tópicos desarrumados para pôr em ordem. Muitas vezes tenho encontrado casualmente um bom tópico enquanto preparo minhas aulas, participo de seminários e *workshops*, assisto às aulas na escola de direito, converso com colegas ou apenas leio o jornal. O truque é reconhecer o que se encontra casualmente ou, como Robertson Davies escreve em sua última novela: “*ver o que está bem debaixo do seu próprio nariz; essa é a tarefa*”³

II Aprendizado precoce como um economista

Fiz meu primeiro curso de economia no penúltimo ano de faculdade. Dois fatos ainda continuam gravados em minha mente sobre aquele curso. Um, foi o pouco esforço feito para mostrar que microeconomia poderia esclarecer problemas do mundo real. Meus colegas de classe e eu concluímos o curso acreditando que as suposições de microeconomia eram tão irrealistas que economia não poderia ter qualquer postura quanto aos problemas do mundo real. O outro foi a condenação da publicidade pelo professor como um desperdício social monstruoso, uma visão partilhada pela maioria dos profissionais de economia naquela época. Por negligência, tornei-me um economista especializado em meu

3 Ver “*The cunning man*” at 142. O doutor que diz estas palavras adiciona, entretanto, que não é uma tarefa tão fácil porque a citação completa diz: aprender a ver o que está bem debaixo do seu próprio nariz: esta é a tarefa e uma tarefa difícil. Martha Nussbaum aponta que Robertson Davies não foi o primeiro a fazer esta consideração. Foi feita anteriormente pelos filósofos gregos. Por exemplo, em um artigo sobre Heráclito, David Wiggins escreve: “*Mas o poder de Heráclito - sua reivindicação para ser o pensador mais adulto de sua idade e um homem crescido entre crianças e adolescentes baseou-se precisamente na capacidade para especular; na teoria do pensamento, exatamente em física, não onde a especulação precisa de todas as observações úteis, ou onde necessita de mais teoria disponível para abocanhar, mas onde os fatos fossem tão grandes e familiares quanto o céu e tão óbvios que tomam o talento real para prestar atenção neles.*” Ver David Wiggins, “Heraclitus’ conceptions of flux, fire and material persistence” at p. 32 em *Language and logos: studies in ancient greek philosophy*, apresentado ao G.E.L. Owen, ed. M. Schofield e M. Nussbaum (Cambridge: Cambridge University Press, 1982).

último ano na *Columbia* e assisti aos cursos em finanças públicas, moeda e banco e um seminário para especialização em economia. Depois da graduação fui trabalhar em Wall Street, em uma firma de corretagem, produzindo gráficos coloridos (minha habilidade em arte ajudou-me), trilhando os movimentos em rendimentos das ações, capital líquido etc. de diferentes companhias, na esperança de que eu, ou um dos membros do departamento de pesquisa, pudesse detectar prováveis tendências nos preços dos títulos. Logo observei que estudar era mais divertido e polêmico do que trabalhar. Assim, depois de quatro meses em Wall Street, retornei à *Columbia*, em tempo parcial. Minha intenção era obter grau de mestre em economia e, em última análise, trabalhar para alguma agência do governo. Tornar-me um professor em economia ou mesmo obter o Ph.D. não estava na tela do meu radar.

Ao contrário de muitas escolas de pós-graduação seletivas, *Columbia* tinha uma considerável política de portas abertas, admitindo grande número de estudantes e deixando os princípios de sobrevivência darwinianos operar. Havia sempre uns poucos estudantes excepcionais na *Columbia* que conseguiam concluir seu doutorado em quatro ou cinco anos, mas a maioria não sobrevivia. Ou eles obtinham grau de mestre ou perdiam interesse depois de um ano ou dois e desistiam. (No outro extremo, *Columbia* era também a casa para um número de estudantes profissionais que estavam por lá há dez ou quinze anos, trabalhando em uma tese que eles provavelmente nunca concluiriam). Depois de meu primeiro ano na pós-graduação, durante o qual continuei trabalhando tempo parcial em Wall Street, observei que tinha talento para economia e solicitei minha admissão para o programa de doutorado. O chefe do departamento examinou minhas notas e declarou que “meu prognóstico era bom” e, assim, tornei-me um estudante de doutorado em tempo integral.

Sucesso no curso de pós-graduação requer inteligência, esforço sistemático e trabalho duro. Sucesso excepcional na *Columbia* requeria um pouco de sorte também. Sorte para ser arrancado da massa de estudantes por um grande economista e colocado sob sua asa. Tive sorte.

Na primavera de meu segundo ano na pós-graduação assisti, como ouvinte, ao curso de Gary Becker sobre capital humano, que cobria seu manuscrito sobre o assunto, até então não publicado (naquela época). Desde que Becker tinha estado de licença no *National Bureau of Economic Research*, durante meu primeiro ano, eu não tinha assistido ao seu curso “teoria de preço” ou, o que era agora mais comumente chamado, microeconomia. Em aula, Becker me chamava regularmente (algumas vezes, pensava, “cruelmente” porque eu era somente um ouvinte) e se referia a mim como “uma pessoa excessivamente zelosa”. Se

eu não sugerisse a resposta primeiro, Becker arrancava-a de mim. Tendo sido membro de uma faculdade de direito por vinte anos, ainda fico impressionado com a diferença entre o estilo de ensinar de Becker (incomum mesmo no departamento de economia) e aquele do típico professor da escola de direito. Como nas de aulas de direito, Becker pedia a cooperação de estudantes que não se apresentassem voluntariamente e trabalhava com ele por uns poucos minutos até que (confiantemente) ele sugerisse a resposta correta. Em contraste, o professor da escola de direito, praticando o método socrático, chama diferentes estudantes em uma sucessão rápida, jogando um contra o outro. (“Senhor Y, você concorda com a resposta do Senhor X?”). Para impelir o estudante a pensar mais claramente, o professor, muitas vezes, diversificará o hipotético até o raciocínio atrás da primeira resposta desmoronar. Um prêmio é dado para a agilidade verbal e para o raciocínio rápido. Uma série de questões pode ficar sem uma resposta definitiva e o professor se dirige para o próximo caso. De fato, o método socrático incute no estudante a idéia de que não há nenhuma resposta correta ou fácil em direito. Contudo, o fato de que o método sobrevive (mas, de uma forma um pouco mais tolerante hoje) é um tributo ao seu valor em treinar estudantes para se tornarem advogados práticos.

Assistir, como ouvinte, ao curso de Becker em capital humano marcou o início de meu aprendizado como um autêntico economista. Na verdade, já havia estado na pós-graduação por um ano. Contudo, pela primeira vez, comecei a avaliar que economia não era só um conjunto de instrumentos formais, mas uma maneira de pensar a respeito de problemas interessantes do mundo real. Comecei a entender as vantagens de simplificação e suposições irrealistas descritivas, e como uma pessoa criativa poderia desenvolver um modelo simples para esclarecer um problema do mundo real. Estes modelos forneciam uma proposta para pensar sistematicamente sobre política pública e direito. Além de dizer que a política X era boa ou satisfatória, poder-se-ia usar princípios econômicos para explicar detalhadamente as conseqüências daquela política.

Durante meu terceiro ano na pós-graduação completei meu curso de teoria de preço de Becker e fui aprovado em meus exames orais de qualificação. Por três ou quatro meses antes dos exames orais fiz parte de um pequeno grupo de estudantes (nós nos chamávamos “Becker Bombers”) que se encontrava regularmente para rever questões dos exames anteriores de Becker e problemas do livro de texto brochura de Milton Friedman.⁴

4 O livro de texto, *“Price theory: a provisional textbook”* foi baseado no curso de pós-graduação de Friedman, e um número de problemas naquele livro foi sugerido por Aaron Director, um professor de Economia na University of Chicago Law School.

Trabalhar este material do começo ao fim deixou claras, para mim, as diferenças entre saber economia e pensar como um economista. O primeiro vem do domínio da linguagem e dos princípios formais da economia que são encontrados em livros de textos de pós-graduação e artigos em periódicos profissionais. O último, da aplicação desses instrumentos com graus variados de sofisticação para resolver problemas. O problema particular deve ser um problema econômico convencional (por exemplo, estabelecer um preço máximo menor para materiais de construção usados do que o preço de uma construção nova), ou um problema não ordinariamente visto como um problema econômico (por exemplo, fazer leis protegendo a privacidade leva a comportamento mais impróprio). Qualquer problema envolvendo objetivos e escolhas alternativas, reprimidas por recursos e oportunidades disponíveis limitadas, é jogo limpo para a economia. O problema não necessita envolver mercados explícitos ou preços observáveis para que se possa derivar preços imaginários que funcionem como preços de mercado. Frequentemente, simples conceitos econômicos aplicados de maneira criativa produzem discernimentos sutis. Tudo isso deve parecer lugar comum hoje, mas há trinta anos não era. Como um tributo aos esforços pioneiros de Gary Becker, que aceitamos como verdadeiros, o domínio da economia não está confinado aos mercados explícitos, mas é uma “maneira de olhar a vida.”⁵

Meu lance seguinte de boa sorte foi decidir rapidamente sobre um tópico de dissertação. Becker propôs que eu estudasse se a condição das leis de justo emprego melhoraram a situação econômica dos não-brancos nos Estados Unidos.⁶ Animadamente, concordei porque queria trabalhar com Becker e porque o tópico era intrinsecamente interessante. Para que eu comesse, Becker deu-me uma cópia de um artigo não publicado de George Stigler e Claire Friedland que usava análise de regressão para estimar os efeitos da vantagem da regulação do estado.⁷ Àquele tempo, era altamente novo para um economista empregar análise de regressão múltipla para estimar empiricamente os efeitos reais de uma lei ou regulação. O artigo de Stigler e Friedland foi um dos primeiros. Antes de incumbir-me da análise empírica, comecei com a intenção de desenvolver um modelo para explicar

5 Ver Becker's Nobel Lecture intitulado "*The economic way of looking at life*."

6 Em 1963 (quando comecei minha dissertação) treze Estados tinham aprovado a legislação do justo emprego. A legislação federal dos direitos civis não foi decretada até 1964.

7 George Stigler e Claire Friedland, What can regulators regulate? The case of electricity, 5 *J. Law & Econ.* 1, 1962.

os prováveis efeitos das leis do emprego justo. Aqui, adicionei sanções contra firmas que discriminavam os não-brancos para a teoria de discriminação de Becker. Assumi que um empregador, violando a lei do emprego justo, enfrentava uma probabilidade, ao invés de uma certeza, de ser capturado, e uma sanção, se capturado. Quanto maior aquela probabilidade e quanto maior aquela sanção maior o custo de discriminar e maior a probabilidade de o empregador aumentar sua demanda por trabalhadores não-brancos em relação aos brancos. Assim, eu tinha uma tese que se prestava não somente para a modelagem criativa (usando o modelo de utilidade esperada para analisar execução de uma lei), mas também era capaz de responder empiricamente uma importante questão de política pública.

Desenvolver um modelo era uma tarefa fácil comparada à realização da análise empírica. Adquirir habilidades empíricas requer muito do “aprender fazendo”. A escola de pós-graduação não me havia preparado para os muitos meses que eu deveria gastar acumulando dados meticulosamente, estado por estado, dos volumes do censo, calcular médias do estado por classes de salários, anos de escolaridade e outras variáveis, adaptar as distribuições de Pareto para explicar intervalos fechados do censo, colecionar dados das comissões estaduais do emprego justo sobre o número de acusações, gastos com a execução judicial, sanções, e assim por diante. Eu era meu próprio assistente de pesquisa e realizei muitos desses cálculos em uma calculadora mecânica que, freqüentemente, falhava. Felizmente, hoje os computadores tornam possível evitar este tipo de trabalho tedioso, embora não tenha a impressão de que isso tenha aumentado a freqüência de dissertações empíricas em economia.

III Dando a partida em direito e economia

Em seu sentido mais amplo, “direito e economia” se aplica a uma grande parte do campo da organização industrial. Ambos cobrem, entre outras coisas, o estudo da regulação legal de mercados, incluindo análise econômica das práticas comerciais descritas em casos de antitrustes. Esses casos fornecem uma rica fonte de material sobre estas práticas como comércio, distribuição exclusiva, restrições verticais e trocas de informações entre competidores. Ambos os campos incluem também pesquisa sobre as conseqüências teóricas e empíricas de diferentes tipos de regulações e leis governamentais. Assim, um recente artigo do *Journal of Law & Economics* (um periódico em organização industrial e direito e economia) inclui artigos de economistas sobre os efeitos anticompetitivos das cláusulas de nação mais favorecida, os efeitos sobre preços dos títulos de *recalls* reguladoras de remédios, a performance da indústria de aviação sob desregulação e o impacto da

legislação do dissídio coletivo sobre as disputas trabalhistas no setor público. Para meus propósitos, entretanto, quero definir direito e economia mais limitadamente. Quero limitá-lo ao que é chamado o “novo” direito e economia, um campo que começou essencialmente com o artigo de Ronald Coase sobre custo social há trinta anos⁸ e onde a maioria do trabalho foi realizada em escolas de direito e não em departamentos de economia.

O “novo” direito e economia aplica os instrumentos da economia para o próprio sistema legal. Usa a economia para explicar e esclarecer doutrinas legais em todos os campos do direito, incluindo os campos de delitos do direito comum, contratos e propriedade, propriedade intelectual, direito corporativo, direito de falência, direito criminal e os próprios processos legais (por exemplo, os efeitos do honorário da substituição de estatutos, regras de descoberta e precedente legal sobre litígio). O “novo” direito e economia não está limitado a áreas de direito que consideram somente mercados explícitos. É uma teoria tanto das próprias regras legais como de suas conseqüências para conduta. O primeiro é o mais polêmico dos dois. Considera regras e doutrinas legais como “dados” para testar a hipótese de que o direito é melhor explicado como esforços pelos juízes, muitas vezes implicitamente, para decidir casos como se eles estivessem tentando promover a eficiência econômica.

Iniciei casualmente no “novo” direito e economia e não por meio de qualquer plano concebido profundamente para trabalhar nesta área.⁹ Rapidamente, depois de terminar minha tese sobre os efeitos das leis do justo emprego (o “velho” direito e economia), deparei com um artigo de jornal sobre acordo em casos criminais. O artigo mostrava que somente uma pequena fração (provavelmente menos de cinco por cento) dos acusados criminais realmente iam a julgamento. Os demais confessavam-se culpados, muitas vezes para reduzir substancialmente as acusações. Investigando um pouco mais, aprendi, similarmente, que somente uma pequena fração de casos civis terminavam em julgamento. A maioria era julgada improcedente antes do julgamento. Essas decisões não somente pareciam um tema natural para análise econômica (um exemplo de “ver o que está bem debaixo do seu próprio nariz”), mas, do ponto de vista econômico, ninguém os tinha examinado previamente - talvez porque economistas acreditassem que o mais provável era que as pessoas se comportassem mais emocionalmente do que racionalmente em uma situação de litígio.

8 R. H. Coase, “The problem of social cost,” 3. *J. Law & Econ.* 1, 1960.

9 Mais adiante eu uso a frase “direito e economia” também para denotar o “novo” direito e economia.

Como primeiro estudante de Becker, não tive dificuldade em assumir que os litigantes comportavam-se racionalmente em situações de não-mercado. O teste básico era saber se o comportamento racional era uma suposição útil e não se ele era descritivamente realista. Raciocinei (usando o teorema de Coase) que o promotor público e o acusado deveriam chegar a um acordo sobre uma sentença se ambos pudessem estar em melhor situação comparado ao risco de um resultado incerto do julgamento. Similarmente, os litigantes deveriam preferir compor um processo civil improcedente desde que se pudesse chegar a um acordo que deixasse ambos em melhores circunstâncias do que seus resultados esperados do julgamento. Assumindo que os julgamentos implicavam maiores custos que acordos ou doações, mostrei que se os litigantes chegassem a um acordo sobre as probabilidades de ganhar e perder o julgamento eles sempre conciliariam (a menos que tivessem fortes preferências para o risco) porque, em um acordo, a utilidade de cada litigante seria maior que sua utilidade esperada de um julgamento. Além disso, seria mais provável ocorrer julgamentos quando os litigantes estivessem mutuamente otimistas (i.e., cada litigante acreditava que tinha maior probabilidade de vencer um julgamento do que seu oponente), quando riscos fossem preferidos, quando o custo dos julgamentos fosse menor em relação ao custo de se chegar a um acordo e quando os interesses em litígio fossem maiores (porque aumentaria a diferença nos resultados esperados para as partes mutuamente otimistas). Meu artigo também tem implicações para execução judicial pelo fato de ter mostrado que criminosos, como uma classe, poderiam estar em pior situação por meio dos acordos, e que até mesmo qualquer infrator particular esteve em melhor situação evitando um julgamento (um dilema do verdadeiro prisioneiro) porque doações liberavam recursos que capacitavam o promotor público a perseguir mais criminosos.

Apresentei uma versão preliminar deste artigo¹⁰ em 1967 para o *workshop* na *University of Chicago*. Naquela época, eu era professor assistente no departamento de economia em Chicago. Minha conferência não foi recebida com muito entusiasmo. Mais tarde, um de meus colegas sênior no departamento chamou-me de lado para um conselho amigável. Ele disse que eu estava cometendo um equívoco em minha carreira dedicando-me a pesquisa sobre problemas próprios dos tribunais, somente de interesse marginal para outros

10 O artigo foi inicialmente intitulado "Rationing the services of courts." Uma versão substancialmente revisada, que continha uma análise empírica da frequência tanto dos casos criminais quanto civis testada em diferentes jurisdições nos Estados Unidos, foi eventualmente publicado no *Journal of Law & Economics* em 1971. Este artigo e os artigos de Richard Posner, Jack Gould e Steven Shavell têm estimulado uma volumosa literatura em direito e economia sobre a resolução de disputas legais. Em um artigo-resenha um tanto datado, Robert D. Cooter & Daniel L. Rubinfeld, *Economic analysis of legal disputes and their resolution*, 27. *J. Econ. Lit.* 1967, 1989 discutem mais de 100 desses artigos. Suspeito que o número de artigos tenha pelo menos dobrado desde o ano do artigo-resenha.

economistas. Sucesso profissional, enfatizou ele, requer trabalho sobre problemas de interesse mais recente para outros economistas. Perguntei-lhe como se sabia o que era de “interesse mais recente” Ele respondeu que se poderia avaliar o interesse observando em que problemas outros economistas estavam trabalhando atualmente e tentar dar um passo adiante. Decidi perguntar a Gary Becker o que ele pensava (embora suspeitasse que já soubesse o que ele deveria dizer). Becker tinha acabado de concluir seu artigo sobre economia do crime, e um dos estudantes de Becker, Isaac Ehrlich, estava completando uma tese na *Columbia* sobre os efeitos repressivos das taxas de condenação e sanções sobre crimes. Becker discordou do meu colega de Chicago. Seu conselho foi simples: execução judicial e litígio são temas sociais importantes e interessantes que podem ser esclarecidos pelos economistas; não se aflija muito se seu trabalho faz parte da novidade recente em economia; e um bom trabalho finalmente será reconhecido. Felizmente, escutei Gary Becker.

Em 1968, mudei de Chicago para New York City, aceitando uma bolsa de estudos no *National Bureau of Economic Research*, e um ano mais tarde associei-me ao *staff* de pesquisa do NBER. Naquela época, os membros do *staff* do NBER residiam quase todos em New York City. Embora também tivesse compromissos acadêmicos no departamento de economia da *Columbia* e, posteriormente, no *Graduate Center* da *City University*, minha vida intelectual se centrava no NBER. O NBER me proporcionava liberdade para escolher projetos interessantes e para evitar as perturbações associadas à desordem dos estudantes da *Columbia* durante este período. O *Bureau* tinha uma atitude profissional e nenhum contra-senso com respeito à pesquisa - projetos eram empreendidos com a expectativa de que eles deveriam ser concluídos, prazos finais rígidos eram impostos, relatórios de desenvolvimento eram requeridos e diretores de pesquisa tinham um forte interesse no trabalho sob sua direção. Mas havia também a troca de idéias e a informalidade de uma universidade que eu nutria. O *Bureau* era um lugar ideal para conduzir pesquisa empírica séria.

Quando me associei, o NBER era melhor conhecido por sua pesquisa empírica em temas econômicos tradicionais, tais como ciclos de negócios e contabilidade da renda nacional, mas estava começando a estender-se para outras áreas da economia. Becker e Jacob Mincer conduziram projetos sobre economia da educação e capital humano, e Victor Fuchs dirigiu um programa em economia da saúde. O *Bureau* estabeleceu formalmente um programa em direito e economia em 1971, que foi subsidiado pela *National Science Foundation*. O Programa incluía Becker, eu, Isaac Ehrlich e Richard Posner (então professor na *University of Chicago Law School* e agora juiz chefe da *U.S. Court of Appeals of Seventh Circuit*). A inclusão de Posner cobriu um defeito crítico no programa. Para

aplicar economia em outras áreas de direito além de crime e os tribunais, nós necessitávamos de alguém perito em direito. Posner parecia ideal. Ele tinha um forte interesse em economia, já havia publicado vários artigos em antitruste amplamente respeitados e estava começando a aplicar economia em delitos e administração judicial.

Deve ser mencionado que as primeiras aplicações de economia para direito no NBER (pré-Posner) e em qualquer outro lugar não requeriam quase nenhum conhecimento de direito. Isto era verdade para o artigo de Becker sobre crime, para os estudos pioneiros de Ehrlich sobre repressão e execução judicial e para o meu próprio trabalho sobre os tribunais, acordos e o sistema de fiança. Que isso deveria ser assim não é surpresa. Nós éramos economistas, aplicando os instrumentos teóricos e empíricos da economia para o estudo sistemático da execução judicial.¹¹ Na verdade, tínhamos que desenvolver um entendimento básico sobre os termos legais relevantes e instituições sob investigação, mas isso requer muito menos conhecimento de direito do que se tornar familiar e à vontade com regras e doutrinas legais para analisá-las de um ponto de vista econômico.

IV Um economista em uma escola de direito

Embora o *Bureau* estabelecesse um ambiente de pesquisa soberbo, não poderia competir com o estímulo intelectual da *University of Chicago*. Chicago era a casa dos economistas que eu mais admirava - Becker, Coase, Friedman e Stigler. Pelo menos me oferecia a oportunidade de trabalhar mais intimamente com Posner. Assim, em 1973, aceitei ansiosamente um cargo na *University of Chicago Law School*. A *Law School* tinha como longa tradição manter um economista em sua faculdade, começando com Henry Simons, Aaron Director e Ronald Coase. Quando cheguei, Coase ainda era um membro ativo da faculdade, mas lecionava somente um curso ocasional. Ainda, meu cargo era qualquer coisa incomum. Eu estava sinceramente interessado em explicar regras e doutrinas legais de uma perspectiva econômica. Coase não. Ele acreditava que conhecimento de direito e de instituições legais era valioso porque ajudava a entender como os mercados explícitos trabalhavam verdadeiramente. Mas ele tinha pouco interesse em mostrar, por exemplo, que as várias doutrinas legais regulando compromissos financeiros para acidentes ou custos contratuais tinha uma lógica econômica implícita. É uma das ironias do direito e economia que a pessoa, cujo trabalho pioneiro (citado pelo comitê do prêmio Nobel) forneceu a base para essa área, estivesse pouco entusiasmado com seu desenvolvimento. Coase acreditava

11 Coase era uma exceção. Ele assitiu alguns cursos de direito comercial, e sua teoria do custo social discute um bom número de importantes casos de infração na Inglaterra.

que muito de direito e economia estava fora do domínio da economia e que, de qualquer maneira, advogados, em vez de economistas, estavam melhor preparados para a iniciativa. A maioria dos professores de direito ia mais além. Eles acreditavam que advogados também falhavam, explicando direito de uma perspectiva econômica.

Nas poucas escolas de direito com um economista em sua faculdade em 1970 (em contraste com um professor de direito que tinha a oportunidade de ter pós-graduação em economia), o economista era contratado para ensinar teoria do preço, co-ensinar com um professor de direito um curso sobre regulação de comércio, tais como antitruste, e servir como um meio para os poucos professores de direito que achavam que economia poderia ter alguma coisa a contribuir com sua área particular de direito. O economista não se imiscuía com direito, nem era esperado que ele assim procedesse. E mesmo quando ele se mantinha no campo da economia os resultados poderiam ser alterados. Tem-se somente que relembrar a velha história do curso sobre antitruste em Chicago nos anos 50, co-ministrado pelos Professores Edward Levy, mais tarde Secretário de Justiça dos Estados Unidos, e Aaron Director. Durante as quatro primeiras aulas de cada semana Levy deveria revisar cuidadosamente os casos e trabalhar para que as argumentações econômicas dos juízes fizessem sentido. No quinto dia, Director deveria explicar por que tudo que foi visto nas últimas quatro aulas estava errado.

A. Interesses de pesquisa

Há vinte anos, havia duas opções abertas para um economista que queria contribuir para o “novo” direito e economia. Ele poderia colaborar com um professor de direito interessado em economia ou aprofundar-se em direito e, dados tempo e esforço suficientes, tornar-se suficientemente à vontade com os assuntos legais para trabalhar por si mesmo. (Hoje há um terceiro caminho. Estudando lei e literatura econômica pode-se ser capaz de encontrar problemas promissores, mas, muitas vezes, técnicos para trabalhar.) Escolhi fazer ambos. Colaborei com Posner e aprofundei-me no estudo do direito. Não que eu pretendesse ser um advogado, mas queria saber o bastante sobre diferentes áreas do direito para ver onde a economia poderia ser mais útil. Diferente da maioria de outros economistas, eu realmente apreciava estudar processos de advocacia. Entretanto, lia-os com os olhos de um economista. Procurava e, muitas vezes, encontrava uma lógica econômica implícita no resultado de um processo. E, se não atingisse completamente a prerrogativa legal ou interpretasse mal o que o juiz disse, resultados nada incomuns, sempre tinha Posner ou um de meus outros colegas na escola de direito para me corrigir.

Meu primeiro artigo com Posner começou com um comentário teórico sobre o artigo de Becker e Stigler sobre execução judicial privada. Mostramos que execução judicial privada poderia levar para a superexecução em relação à (ótima) execução judicial pública porque a multa mais alta levaria os executores a progredir ao invés de reduzir suas atividades de execução judicial.¹² Mas o artigo transformou-se rapidamente em um projeto mais ambicioso. Testamos as predições da análise contra observações do mundo real. Explicamos por que há uma maior segurança na execução judicial privada em contrato, delitos e outras áreas do “direito privado” comparado ao direito criminal; por que vítimas ao invés de outros têm os direitos (prerrogativas) exclusivos para processar e corrigir violações; por que o orçamento das agências de execução judicial pública tende a ser menor em relação ao que gastam executores privados, maximizadores de lucros; e por que executores judiciais públicos invalidam leis particulares, recusando-se a instaurar processos enquanto os executores judiciais privados não. Aplicamos também o modelo para extorsão e suborno como formas de execução judicial e as normas legais que regulam recompensa por propriedade perdida ou roubada - também um método de compensar executores judiciais privados.

Em um aspecto importante o artigo sobre execução judicial representou um nítido desvio do meu trabalho anterior. Este trabalho aplicava sistematicamente a economia para um grande número de normas legais e mostrava como essas normas promoviam eficiência econômica. Naturalmente, isso foi devido a Posner, porque me faltava o conhecimento necessário de direito. Mas estava determinado a corrigir essa deficiência assistindo, como ouvinte, aos cursos de direito - particularmente, os cursos básicos do primeiro ano, tais como procedimento civil, contratos e delitos - e ministrando aulas e seminários de direito juntamente com professores de direito.

Pelos vinte anos seguintes, Posner e eu fizemos, em co-autoria, mais de 25 artigos e um livro sobre direito do delito. Nosso trabalho era verdadeiramente um esforço conjunto e continua até hoje. Tenho tido maior responsabilidade pelo modelo econômico e Posner pelo direito, mas cada um de nós tem contribuído substancialmente tanto para economia como para direito. Realmente houve ganhos substanciais de troca porque cada um trouxe diferentes habilidades para o empreendimento e o produto final excedeu em muito a soma das partes individuais. Cada um de nós contribuiu para aumentar o produto marginal do outro. Examinando os artigos, seria ilusório dizer “Posner fez isto” ou “eu fiz aquilo”

12 Execução ótima de uma lei (de acordo com o artigo anterior de Becker) envolvia uma pequena probabilidade de apreensão e condenação e uma alta multa que produzia o mesmo nível de repressão que o de uma alta probabilidade e baixa multa, mas a um custo menor.

porque as idéias, escolha de tópicos, enfoques para eles e execução eram sempre esforços conjuntos. Os tópicos sobre os quais colaboramos cobriam uma grande esfera de assuntos legais, incluindo antecedentes legais, a resolução de disputas legais, leis regulando auxílio, tais como salvamento em direito marítimo, antitruste, delitos, o papel de um judiciário independente, segredos de comércio, marcas registradas e direitos autorais. Nosso trabalho mais conhecido, o livro “*The Economic Structure of Tort Law*”, mostrou como um modelo econômico relativamente simples de maximização de riqueza poderia explicar e organizar o que, à primeira vista, parecia um arranjo incompreensível de regras e doutrinas não relacionadas, determinando responsabilidade de delito. Cobrimos todas as áreas de delitos, desde problemas simples como a escolha entre negligência e regras rígidas de responsabilidade para acidentes ordinários, até questões mais complicadas envolvendo proteções para responsabilidade, causação de delitos conjuntos (dois ou mais lesados), danos catastróficos (muitas vítimas) e delitos intencionais.

Embora também tenha trabalhado em vários projetos meus, incluindo artigos sobre litígio e lei de direitos autorais, continuo a fazer trabalhos em colaboração com Posner e, mais recentemente, com Larry Lessig, um professor de direito recentemente nomeado em Chicago.¹³ Surpreende-me que a colaboração entre advogados e economistas não seja mais comum porque os ganhos de intercâmbio parecem muito substanciais. Além de Posner e eu, a única outra colaboração de longo prazo envolve Charles Goetz, um economista, e Robert Scott, um professor de direito e agora Reitor da *University of Virginia Law School*.¹⁴ Por outro lado, um fenômeno crescente nas escolas de direito é o advogado que também tem Ph.D. em economia. A maioria deles está nas recentes escolas de pós-graduação de direito. Seus trabalhos são uma forma de colaboração entre um advogado e um economista, mas envolve uma pessoa.

B. A mudança de papel do economista na escola de direito

Com o passar dos anos, fiquei muito mais à vontade com direito e muito mais integrado à cultura da escola de direito. Isto é também verdade para outros economistas que ocupam

13 Lessig e eu estamos completando um projeto em grande escala, estimando empiricamente a influência e reputação dos juízes do tribunal federal, enumerando citações para suas opiniões. Considerando citações como um “*output*,” estávamos seguindo a literatura do capital humano, e estimamos equações de citações sobre experiência e uma variedade de outras variáveis. Não somente classificamos juízes, mas examinamos fatores que devem explicar diferenças de influência entre eles (por exemplo, raça, sexo, qualidade da escola de direito, performance da escola de direito, experiência anterior etc.).

14 Como uma medida grosseira dos benefícios da colaboração, Landes e Goetz foram responsáveis por mais de 45% das intimações em jornais de direito para os artigos e livros de economistas de quinze escolas de direito de topo. (Ver Landes e Posner, *The influence of economics on law: a quantitative study*, 36. *J. Law & Econ.* 384, 1993).

cargos em tempo integral nessas escolas. Gastamos muito mais tempo com nossos colegas nas escola de direito do que com economistas do departamento de economia ou de administração. Proximidade é uma das razões, mas há outras influências mais fundamentais operando.

Uma delas é que os departamentos de economia têm-se tornado menos interessados em áreas aplicadas, tais como o “novo” direito e economia. Economia tornou-se mais formal e teórica. A pesquisa está cada vez mais interessada em demonstrar habilidades técnicas e resolver problemas técnicos do que em analisar problemas sociais. Conseqüentemente, o economista da escola de direito sente-se menos à vontade intelectualmente com o outro lado do campo. Felizmente, isto é menos verdade em Chicago. Eu continuo acompanhando economia e *workshop* em escolas de administração com Becker, Sam Peltzman, Sherwin Rosen e outros, mas Chicago é incomum.

Outra é que economistas nas escolas de direito têm mais em comum com professores de direito hoje do que há vinte anos porque economia transformou o conhecimento legal em delitos, contratos, seguros, antitruste, corporações, lei ambiental, propriedade intelectual e outras áreas de comércio relacionadas. Há um grande número de professores de direito que se consideram membros do movimento de direito e economia.¹⁵ Outra indicação da crescente importância da economia em escolas de direito é a nomeação de economistas (mas, nenhum outro não-advogado) para cargos em tempo integral em toda especialização. Vinte anos atrás o economista na escola de direito era uma figura periférica. Hoje ele ocupa uma posição central.

Um fator relacionado é a crescente importância da economia no ensino do direito. Escolas de direito são escolas profissionais que consideram como sua primeira missão formar futuros profissionais. Para que a economia adquira importância mais que marginal, deve demonstrar sua relevância na formação de futuros advogados no exercício da profissão. Isso tem que ser feito mediante contribuições significativas para a prática do direito. A economia tem alterado o antitruste; desempenha um importante papel em seguros, pensão, ambiente, competição injusta e litígio de discriminação; e é importante em avaliação e cálculos de indenização em, potencialmente, todo processo comercial em grande escala. Estudantes de direito estão prontos para reconhecer o valor da economia na

15 Uma boa medida disso é que os advogados compreendem cerca de 50% de aproximadamente 1 200 membros da recentemente formada American Law & Economics Association.

prática do direito. Saber economia lhes dá uma vantagem sobre seus concorrentes. Como resultado, os cursos de direito e economia estão cada vez mais populares nas escolas de direito quando ministrados em conjunto por advogados e economistas sobre uma variedade de assuntos. Além disso, não é raro hoje um economista ministrar sozinho um curso de direito, o que não acontecia há trinta anos. Considere minhas próprias responsabilidades de professor: embora participe de *workshop* em direito e economia, ensino direitos autorais, marcas registradas e competição injusta e (meu favorito) direito da arte. Estes não são cursos de direito e economia, mas cursos regulares da escola de direito. Na verdade, adiciono uma dose pesada de economia, não só porque sou um economista, mas também porque os casos discutem explicitamente e reconhecem a importância dos fatores econômicos e porque o uso da economia (pelos advogados) em assuntos de direito privado tem-se tornado lugar comum. De fato, tenho-me tornado tão integrado ao mundo do direito acadêmico que sou agora um professor de direito e economia, não só um professor de economia (meu título original na escola de direito).

C. O futuro

Impressiona-me os comentários que me fazem, em várias ocasiões, jovens estudantes que estão se iniciando hoje em direito e economia. O ponto essencial desses comentários é “quando você começou havia muitas áreas de direito abertas para economia, mas você e os outros esgotaram todos os problemas interessantes e agora nada restou.” Há, naturalmente, um elemento de verdade nisto, mas é muito exagerado. Antes, um economista, assistindo a um curso em delitos ou contrato na escola de direito, era como uma criança em uma loja de doces - havia um tópico interessante a ser descoberto em quase toda aula. De fato, a dificuldade não era encontrar tópicos, mas decidir sobre qual deles trabalhar. Meu livro de delitos com Posner é um bom exemplo. Enquanto assistia ao seu curso de delito, avancei gradualmente nas notas econômicas sobre casos de doutrinas discutidos em aulas e em arquivos. Então refinei e expandi esse material em conexão com um curso que ministrei em direito e economia. Essas notas tornaram-se o ponto de partida para nosso livro sobre delitos. Mas, hoje, a análise econômica sobre campos do direito comum como delitos e contratos tem sido tão visada que seria um engano para um jovem estudante concentrar-se neles. O mesmo provavelmente é verdade para modelos sobre litígio, embora aqui eu esteja menos seguro porque aplicações recentes da teoria dos jogos para litígio têm produzido algum conhecimento novo interessante.

O que sobrou? Estudantes de direito e economia só recentemente aplicaram os instrumentos da teoria dos jogos para entender como doutrinas legais devem dominar

comportamento estratégico e informação assimétrica.¹⁶ Esta é uma área promissora para trabalho futuro. Voltando aos campos particulares do direito, observa-se que direito constitucional tem sido apenas tocado pela análise econômica. E direito da família, direito criminal (tão distinto de estudos empíricos de repressão), procedimento legal e propriedade intelectual têm sido relativamente negligenciados comparativamente a delitos, contratos e direito coletivo. Esses campos também permanecem promissores para trabalho futuro. Mas o lado mais negligenciado do direito e economia é o empírico. Em muitas áreas do direito e economia há uma escassez de estudos empíricos que certamente são dignos de ser trabalhados. Recentemente examinei todos os artigos publicados no *Journal of Legal Studies* (o principal periódico do “novo” direito e economia) durante os últimos cinco anos e descobri que somente cerca de 20% tinham algum conteúdo empírico, ao contrário do *Journal of Political Economy*, onde mais de 60% dos artigos publicados continham análise empírica substancial.¹⁷ Esta diferença não pode ser levada em conta somente por diferenças na disponibilidade de dados. Há conjuntos substanciais de dados sobre o número e a disposição de casos criminais e civis tanto em julgamento como em níveis de apelação, laudos em casos civis, sentenças em casos criminais, rendimento de advogados, taxas de acidente, e assim por diante. Além disso, bancos de dados legais computadorizados tornam possível, a custos relativamente pequenos, extrair quantidades significantes de informação de casos para desenvolver conjuntos de dados relevantes para o problema em questão.

Finalmente, há diferentes enfoques para pesquisa. Pode-se trabalhar produtiva e criativamente tanto intensiva quanto extensivamente. O primeiro enfoque é ilustrado pelo trabalho de Coase sobre problemas tais como custo marginal, a organização de firmas, custo social e monopólios de bens duráveis. Antes de Coase, economistas trabalharam nesses problemas por muitos anos. Contudo, Coase foi capaz de dizer alguma coisa nova e

16 Para um excelente ponto de partida nesta direção, ver Douglas Baird, Robert Gertner e Randal Picker, *Game theory and the law*. Harvard University Press, 1994.

17 A razão de existir relativamente poucos artigos empíricos em direito e economia é uma questão interessante em si mesmo. Recentemente, fiz menção a este tema em uma apresentação sobre direito e economia no encontro anual da American Economics Association em 1994. Adiantei várias explicações, incluindo o fato de que o sucesso inicial do direito e economia em escolas de direito não vinha de estudos empíricos, mas da luz que a economia lança sobre doutrinas legais; que a cultura da escola de direito valoriza a perspicácia verbal e habilidades analíticas, mas não a análise empírica cuidadosa; que direito e economia tem estado centrado nas escolas de direito ao invés de em departamentos de economia ou administração; e que os professores de direito, os maiores colaboradores para direito e economia, são selecionados por habilidades verbais e não quantitativas. Igualmente embaraçoso é por que economistas em faculdades de direito também tendem a evitar análise empírica, mas isto novamente está relacionado tanto à estrutura de recompensa nas escolas de direito e ao tipo de economistas que têm sido atraídos para direito e economia.

recente sobre esses problemas e, essencialmente, mudar a maneira de pensar dos economistas. Becker, por outro lado, trabalhou principalmente extensivamente, mostrando a relevância da economia para uma ampla gama de questões sociais consideradas geralmente fora do alcance da economia. Essas questões incluem casamento, divórcio, educação infantil, educação, altruísmo, crime, vício e formação preferencial. Como Becker e Coase mostraram, pode-se ganhar Prêmios Nobel com qualquer tipo de abordagem. O fato de existir agora um corpo substancial de literatura em direito e economia torna simultaneamente mais difícil trabalhar extensivamente, mas mais fácil trabalhar de forma intensiva.

D. Consultoria ou direito e economia em ação

A descrição de minha carreira em Direito e Economia estaria seriamente incompleta se não considerasse consultoria ou o que chamo “direito e economia em ação”. Em 1977, Posner, Andrew Rosenfield, então um estudante do terceiro ano na faculdade de direito, e eu fundamos a firma Lexecon Inc.. Economia estava, então, começando a entender litígios, antitrustes e procedimentos regulatórios. Nós estávamos confiantes que isso iria desempenhar um enorme papel no futuro. Ao mesmo tempo, firmas de direito e seus clientes expressavam constantemente insatisfação com a qualidade dos serviços de consultoria econômica que estavam recebendo. A principal reclamação era que, ao final das contas, eles não estavam obtendo um bom retorno pelo seu dinheiro. Análise econômica e os estudos empíricos eram caros e, raramente, proporcionavam muita ajuda. Mas parte do problema estava com os advogados, que tinham um entendimento tão ínfimo de economia que não sabiam como utilizá-la efetivamente.

A idéia por detrás da Lexecon era muito simples. Existia um nicho de mercado para uma firma que fornecesse serviços de consultoria econômica de alta qualidade que seriam relevantes e úteis em assuntos litigiosos e regulatórios. Nós reunimos qualidades únicas nessa empreitada. Posner era um advogado que conhecia economia, eu era um economista que sabia como explicar economia para advogados, e Rosenfield, que tinha treinamento de pós-graduação em economia para seguir em frente com sua graduação em direito e que, diferentemente de nós dois, tinha disponibilidade de tempo para se dedicar integralmente a Lexecon. Juntos pudemos selecionar que tipo de estudo econômico deveria ser feito, dirigi-lo e supervisioná-lo eficientemente e, quando necessário, convocar outros economistas acadêmicos que tivessem experiência e conhecimento especializado nas áreas pleiteadas.

Lexecon também desempenhou um papel educacional. Muitos advogados, excepcionalmente talentosos e experientes, ficavam à deriva quando se tratava de economia

e estatística. Mas eles eram aprendizes rápidos e ávidos. Nós explicávamos economia básica (e mesmo econometria) e mostrávamos a eles como poderiam usar economia para estruturar e reforçar seus argumentos legais. Com esta gama de serviços, éramos capazes de convencer as firmas de advocacia a nos entregar o lado econômico de vários grandes casos. Tínhamos um outro atrativo de vendas. Não nos constituíamos uma ameaça às firmas de advocacia. Não estávamos disputando seus clientes porque não praticávamos advocacia. Na verdade, a Lexecon tornou-se um instrumento competitivo nas mãos das firmas de advocacia porque lhes permitia oferecer aos seus clientes um produto superior.

Como se costuma dizer, o resto é história. A Lexecon tornou-se imensamente bem-sucedida e gerou várias imitações. É uma fonte de grande satisfação pessoal para mim ter ajudado a criar e a desenvolver a Lexecon. Atualmente, a Lexecon tem cerca de cento e vinte e cinco empregados em tempo integral em Chicago¹⁸ (embora Posner a tenha deixado em 1981, quando se tornou juiz e eu reduzi significativamente o meu papel nos últimos anos), incluindo uma ampla equipe de economistas com Ph.D. extraordinariamente capazes e parceria com um grande número de economistas acadêmicos de ponta, incluindo ganhadores de prêmios Nobel.¹⁹

Consultoria econômica tornou-se uma opção cada vez mais atrativa para muitos dos mais brilhantes Ph.Ds. em economia. Ela oferece uma perspectiva de remuneração maior do que na academia (mas não a perspectiva de uma efetivação formal) e uma amplitude de problemas do mundo real para trabalhar porque o papel da evidência econômica, anteriormente limitada aos casos antitrustes e cálculo de prejuízos em casos de injúrias pessoais, tem-se expandido para abranger potencialmente todas as espécies de litígios comerciais em grande escala. Economistas são rotineiramente empregados em áreas tais como seguros e lei coletiva, lei de pensão, regulamento ambiental e de segurança e litígio de discriminação. De fato, seria imperícia não usar um economista nessas áreas.

Há naturalmente uma importante diferença entre acadêmicos e consultores. Um acadêmico organiza sua própria agenda. Tem o luxo de escolher qualquer problema que atrai sua imaginação e o ritmo para executá-lo. Não é assim em consultoria. Aqui, o problema é colocado diante de você, que enfrenta a urgência do tempo, a tensão do litígio,

18 Devo adicionar que os escritórios da Lexecon foram projetados por Charles Gwathmey (ver nota 2).

19 Rosenfield é agora Presidente da Lexecon e um professor sênior na University Chicago Law School, onde ensina antitruste.

longas horas, viagens longe de casa. Entretanto, milhões de dólares podem estar em jogo e seu papel deve ser crucial. Não é surpresa que seja envolvido pelo calor da disputa e sinta aquela emoção da satisfação de ter feito um trabalho de primeira linha. Recompensas vêm mais vagarosamente, se é que vêm, para o economista acadêmico.

Há um equívoco comum sobre litígio entre economistas acadêmicos que têm pouca ou nenhuma experiência de consultoria. Eles assumem que a pressão da questão judicial compele um economista que presta declarações como uma testemunha pericial a relatar de maneira injusta e imparcial sua análise, a apresentar somente resultados favoráveis e a manipular os dados a fim de propor as respostas que o cliente deseja. A falha neste argumento é que ele ignora como funciona a questão judicial. O perito conta com ambos os fatos e sua análise muda para o lado do litigante oposto antes que qualquer depoimento seja dado no tribunal. A oposição, armada com seu próprio economista, checará os cálculos do perito oposto, reestimarão suas equações, analisará a sensibilidade das estimativas para especificações alternativas, examinará como os resultados se alteram quando outras variáveis são adicionadas, e assim por diante. A combinação de altos riscos e os trabalhos do sistema adversário indicam que há uma probabilidade muito alta de que quaisquer enganos, se intencionais ou inadvertidos, serão desmascarados. O mesmo se mantém para apresentações econômicas perante agências reguladoras, tais como a *Antitrust Division* do *Department of Justice* ou do *Federal Trade Commission*. Eles têm sua própria equipe de economistas profissionais para analisar o trabalho do perito. Compare isto com o trabalho acadêmico. Um periódico com arbitragem perceberá frequentemente enganos. Mas é muito mais fácil escapar impunemente com análises empíricas descuidadas e intencionalmente enganosas em estudos acadêmicos do que em questões judiciais porque é raro que outros economistas se darão ao trabalho de checar o trabalho anterior.

V Comentários finais

Descrevendo sua evolução como um economista, Ronald Coase escreveu: “*Eu vim a observar para onde estava indo somente depois que cheguei. A emergência de minhas idéias em cada estágio não foi parte de algum grande esquema.*”²⁰ Esta frase capta também minha jornada. Eu não tinha nenhum caminho particular de carreira em mente quando comecei a pós-graduação. Escolhi economia em lugar de outra carreira

20 R. H. Coase, *My Evolution as an Economist* (1994), versão não publicada de uma conferência em “Lives of the Laureates” series, proferida na Trinity University, San Antonio, Texas em Abril 12, 1994.

principalmente porque me ocupei ao máximo com aulas de economia quando era um aluno de graduação. Comecei em direito e economia acidentalmente porque encontrei por acaso um artigo de jornal sobre acordo. Realmente, eu pretendia aplicar economia em importantes questões sociais, mas direito era só uma das muitas possibilidades. Trabalhei amplamente tópicos em direito que, relembrando, evidenciam uma aproximação comum, mas não um esquema total, para transformar-se em conhecimento legal. Nunca pensei que fizesse parte de um movimento, mas agora é lugar comum ouvir sobre como o “movimento de direito e economia” tem transformado o conhecimento legal e o ensino.

Fui também extraordinariamente feliz por ter trabalhado com Becker e Posner. Becker abriu meus olhos quando estudante para o poder da economia no esclarecimento de questões sociais, e desde então tem sido uma fonte de inspiração. Posner é, provavelmente, o acadêmico em leis mais influente e certamente o mais produtivo neste século. Se ele tivesse escolhido outra carreira, ficaria difícil imaginar se o campo direito e economia teria sido bem-sucedido.²¹ Também tive outro extraordinário lance de sorte. Há mais de vinte e cinco anos casei-me com uma economista que tem sido minha melhor crítica e a fonte de incontáveis idéias. Conheci Lisa quando era professor assistente, e ela uma estudante do primeiro ano de pós-graduação no departamento de economia em Chicago. Se as regras e políticas comuns que regulam assédio sexual nas universidades e qualquer coisa parecida tivessem em voga em 1968, eu nunca teria saído com uma estudante de pós-graduação.²² Embora políticas de assédio sexual imponham custos de transações e outros custos sobre mercados de encontros e casamentos, muitos acreditam que os benefícios (por exemplo, reduzindo custos de coerção pelos homens) excedam estes custos extras. Em meu caso, eu teria sido um grande perdedor. Mas o tema geral de assédio sexual é um grande tópico para futuros trabalhos em direito e economia.

21 Seria mais correto dizer “não tivesse ele escolhido economia como uma de suas carreiras.” Posner é também um juiz no tribunal federal de apelações (cujas opiniões são citadas com mais frequência que as de qualquer outro juiz) e um colaborador significativo em outros campos, tais como direito e literatura e jurisprudência.

22 Embora políticas de assédio sexual não proibam um professor de encontrar-se com uma estudante de pós-graduação no mesmo departamento sob quaisquer circunstâncias, não acredito que seja possível elaborar e implementar uma política que somente previna conduta prejudicial e claramente censurável (por exemplo, abusos de estudantes femininas por supervisores de dissertação masculinos), mas não condutas benéficas também. Por que as relações sociais têm resultados incertos, um membro da faculdade não efetivo pode deparar com custos substanciais relacionados ao seu emprego e penalidades (algumas vezes merecidamente, mas outras vezes, não) se a relação se torna amarga. Sabendo disso, uma sábia escolha é não iniciar a relação.

Dedicação à pesquisa

Juan Hersztajn Moldau[§]

Minha formação

Minha dedicação à economia e ao desenvolvimento de uma carreira acadêmica não foi uma decisão resultante de uma longa predisposição nessa direção. Quando me inscrevi no concurso vestibular não sabia exatamente no que estava me metendo. Naquela época, poucas pessoas na minha idade tinham uma noção exata das funções do economista. As carreiras de maior prestígio eram medicina e engenharia. Economia era uma área nova, bem diferente daquilo que a maioria das pessoas costumava fazer. A Faculdade de Economia da USP, na época, era talvez uma das melhores do Brasil em métodos quantitativos. O professor catedrático era o Luiz de Freitas Bueno, e fui aluno do professor Ikeda, que foi escolhido como professor homenageado. Era um curso muito bom em estatística e econometria. Não havia macroeconomia, como a gente conhece hoje. O que existia era um curso de moeda e bancos. Uma outra área forte era a de história econômica, liderada pela professora Alice Canabrava, que formou uma escola significativa, deixando discípulos até os dias de hoje. Em microeconomia tínhamos o professor Carvalho, que possuía uma didática muito boa.

O IPE veio, de certa forma, satisfazer as expectativas que eu havia formado com relação ao curso de graduação. Na pós-graduação, eu senti um interesse muito grande por parte de meus colegas, diferentemente do que ocorria na graduação. Havia professores estrangeiros que trabalhavam em conjunto com os professores brasileiros na organização do curso. Dentre eles, posso citar o professor Sahota, em econometria; e o professor Maneschi, em finanças públicas. Dos professores brasileiros, devo mencionar o Affonso Celso Pastore, em econometria, e a Beth Mindlin, que introduziu o Ackley. Com isso tomamos um primeiro contato com o que era realmente a macroeconomia. Os cursos de Teoria de Preços eram dados pelo Carvalho. Se não me engano, o Werner Baer também chegou a dar alguns cursos.

[§] Professor Titular do Departamento de Economia da FEA-USP.

O programa era bastante pesado, principalmente quando comparado com o curso de graduação. Eu diria que essa dicotomia que existe ainda hoje entre a graduação e a pós-graduação pode ser aplicada também àquela época. Quero me referir especificamente ao grande salto de exigências e nível de abordagem na passagem da graduação para a pós-graduação.

A minha principal motivação em fazer o curso do IPE foi a possibilidade de concorrer a uma bolsa de estudos nos Estados Unidos. Naquela época, o único curso que concedia o mestrado era a FGV, no Rio de Janeiro. O meu objetivo era fazer o mestrado no exterior. Eu ainda não tinha interesse de seguir carreira acadêmica. O meu objetivo era trabalhar na iniciativa privada. Eu não me sentia motivado em atuar como professor, não me sentia ainda suficientemente confiante para assumir tal posição. O fato de na época não ter ainda me decidido a seguir uma carreira acadêmica acabou representando um custo muito alto para mim. Em primeiro lugar, por eu não ter-me candidatado a uma posição na própria faculdade e, em segundo lugar, também pelo fato de não ter-me empenhado em me candidatar a uma vaga em uma universidade de maior prestígio do que a universidade de Vanderbilt, a qual mantinha, na ocasião, uma relação estreita com o IPE.

Minha idéia era ficar fora no máximo 18 meses e voltar para atuar na iniciativa privada. Eu não tinha interesse em entrar para o doutorado. No entanto, após 12 meses, eu já havia completado boa parte dos requisitos necessários para fazer o PhD. Foi então que resolvi seguir o programa de doutoramento.

Influências intelectuais

Desde a época da graduação, sempre tive um grande interesse em teoria econômica, principalmente em microeconomia. O professor Georgescu-Roegen foi a pessoa que mais me influenciou, tanto como professor como, e principalmente, como autor. Ele realmente teve um impacto desproporcional dentro do departamento de economia de Vanderbilt. Georgescu questiona o *mainstream* porque ele sempre teve uma preocupação muito grande em analisar os fundamentos. Eu acho que um dos trabalhos mais interessantes do Georgescu é justamente aquele em que ele faz uma crítica à função de produção neoclássica. Ele não é propriamente agressivo em relação à Teoria Neoclássica. Ele não se insurge contra as regras e os métodos da Teoria Neoclássica. O que ele faz é discutir os fundamentos que são utilizados, talvez de forma inadequada, pelo *mainstream*. Ele questiona seriamente o tratamento do tempo na função de produção. Ele procura fazer uma distinção clara entre os insumos que são transformados no processo de produção e os agentes do processo de produção, que são o trabalho e o capital. Ou seja, capital e trabalho não são transformados

no processo, eles não são incorporados ao produto final. Eles são agentes do processo atuando sobre os insumos que são transformados em produto final. Portanto, a forma de você tratar analiticamente a participação daqueles elementos deve ser totalmente diferente do tratamento analítico que você deve dar à participação dos fatores de produção que são transformados no processo e, para isso, o tempo deve ser considerado de forma explícita. Eu acho que esse foi um trabalho fundamental do Georgescu e que não recebeu o devido reconhecimento. Ele apresentou um resumo desse trabalho, em 1969, numa reunião da *American Economic Association*, sendo a versão completa publicada em 1971 como capítulo de um livro importante – *The Entropy Law and Economic Process*.

No início da década de 70 tínhamos alguns outros professores importantes em Vanderbilt. Dentre eles posso citar o Fred Westfield e o Nichols, que é uma pessoa muito conhecida. No entanto, o Georgescu era uma pessoa que se impunha em relação aos demais pelo seu prestígio acadêmico e por sua personalidade. Minha visão da teoria econômica foi muito influenciada por ele. Meu principal programa de pesquisa, até os dias de hoje, tem sido fortemente influenciado por ele.

Foi a partir da leitura de dois de seus artigos que tomei conhecimento de uma crítica que ele fazia à Teoria Convencional da Utilidade, no sentido de que a existência de uma função utilidade constituía-se numa restrição muito forte, em alguns casos uma restrição muito pouco razoável, como fundamento da Teoria da Escolha. Nesses dois artigos¹ ele argumenta que a existência de uma função utilidade teria que ser assumida implicando considerar as necessidades dos agentes econômicos redutíveis a um denominador comum. Em realidade, a noção de utilidade acabaria por encobrir as motivações primárias de comportamento dos agentes econômicos. Isto representou, para mim, uma visão totalmente diferente a respeito dos fundamentos da Teoria da Escolha. O Georgescu defendia também a idéia de que Menger não deveria ser considerado um precursor da Teoria de Utilidade, pois sua teoria era radicalmente diferente das teorias de Jevons e Walras. Menger, segundo Georgescu, pensava em uma Teoria da Escolha não em termos da existência de uma coisa que poderia ser chamada de utilidade, mas sim considerando necessidades irreduzíveis às quais o agente estaria sujeito. O artigo “*utility*” que é um *survey* muito grande, muito extenso do Georgescu sobre a Teoria da Utilidade começa desde os gregos e tenta encontrar alguns fundamentos da Teoria de Utilidade já a partir do Aristóteles. Ele formulou, como desafio, que a análise da escolha talvez devesse partir da possibilidade da existência de necessidades irreduzíveis e não de uma noção abstrata como a de utilidade.

¹ “Choice Expectation and Measurability” (*Quarterly Journal of Economics*, 1954) e “Utility” (*Encyclopedia of Social Sciences*, 1968).

Com vinte e poucos anos, eu achei que esse desafio do Georgescu poderia ser meu tema de tese. Mais tarde percebi que foi uma demonstração de imaturidade. Cometi dois erros que muitas vezes são cometidos por pessoas nessa fase da vida. O primeiro deles é querer fazer, na tese de doutoramento, o trabalho mais importante da vida. O segundo é querer abordar um tema acima das suas possibilidades técnicas. Na época, eu não estava em condições de abordar um tema como este. Mas, curiosamente, o Georgescu aceitou me orientar e jamais me desestimulou. No entanto, ele possuía uma personalidade um tanto difícil para acomodar a atividade de orientação. Creio que o Ibraim Eris foi o único a concluir uma tese com o Georgescu. Depois de algum tempo percebi que não seria uma deliberação sábia prosseguir na tentativa de desenvolver esse tema, naquele momento. Então tomei a decisão de voltar para o Brasil, mesmo não tendo ainda um outro tema de tese para desenvolver. Eu já havia passado no *qualifying* e só me restava a tese, o que não era pouco, como estava então percebendo.

Pesquisa aplicada

Ao voltar ao Brasil, engajei-me na FIPE e, algum tempo depois, na FEA-USP. Para isso, recebi todo o apoio do Affonso Celso Pastore, que era o principal executivo da FIPE. Na FEA-USP eu ajudei o Guilherme Dias no programa de mestrado com especialização em bancos de desenvolvimento. Fui obrigado a me aprofundar em análise custo-benefício, que eu nunca havia estudado. Assim, de forma quase fortuita, acabei por enveredar por uma linha de pesquisa que, para mim, foi importante por muitos anos e que era representada pela análise de projetos. Eu tive, na verdade, muita sorte em me envolver com essa área porque o início dos anos 70 foi talvez um dos períodos mais gloriosos do desenvolvimento da Teoria da Análise Custo-Benefício. Havia autores desenvolvendo trabalhos importantes nessa área, como o Harberger e outros membros da escola de Chicago, como o Sjaastad. Nesta época começaram a sair os primeiros manuais importantes sobre elaboração de projetos: o manual da OCDE, que incluía o volume de Little e Mirrless sobre análise social de projetos, e o manual da UNIDO. Alguns anos depois, em 1975, foi publicado o livro de Squire e Van Der Tak. Mesmo no Brasil havia pessoas que estudavam esses assuntos. Eu poderia destacar o Edmar Bacha, que produziu trabalhos importantes. Posso citar um artigo seu em co-autoria com o Lance Taylor, que saiu no *Quarterly Journal of Economics*, em 1971. Foi um artigo que teve grande impacto. O professor De Faro publicou um livro sobre engenharia econômica que usávamos no curso. Havia ainda o Cláudio Contador, que escreveu um importante livro de texto sobre a avaliação social de projetos.

Os grandes bancos de desenvolvimento, principalmente o BID e o Banco Mundial, passaram a formar grupos de estudo sobre metodologia de análise custo-benefício. Havia grupos envolvidos no estudo de metodologias genéricas e, principalmente, grupos de estudo preocupados em analisar a aplicação da metodologia de avaliação custo-benefício para o estudo de problemas específicos como, por exemplo, avaliação de projetos de estradas, projetos de saúde. O Cristóvão Buarque trabalhava com o BID e participava desses grupos de estudo. Mais tarde ele também escreveu um livro sobre o assunto. Realizavam-se simpósios internacionais e já começava o desenvolvimento de cursos de mestrado pleno com especialização em bancos de desenvolvimento, financiados pelo BID. Foi um período glorioso de desenvolvimento desta área e eu tive muita sorte de me envolver com essa linha de pesquisa, que me prendeu por muitos anos.

Eu não havia abandonado a idéia de obter o PhD nos Estados Unidos. O meu envolvimento com essa área de pesquisa acabou me estimulando a escolher um tema na área de análise custo-benefício. Eu estava muito insatisfeito com a dicotomia entre análise privada e social. Eu achava que essa dicotomia estava colocada de forma muito imprecisa na literatura econômica. Na tese, defendida em Vanderbilt em 1976, desenvolvi a idéia de que você não deveria se restringir a essa dicotomia entre análise privada e social. A definição da metodologia de cálculo dos preços sombra de bens e serviços deveria partir de uma definição muito clara do ponto de vista de análise. Em minha tese eu procurei mostrar que a metodologia para calcular os preços sombra de bens e serviços como capital, mão-de-obra, divisas e bens intermediários poderia nos levar a resultados diferentes se o projeto fosse avaliado, por exemplo, do ponto de vista de uma entidade nacional ou sob a ótica de uma entidade internacional. Ou seja, a existência de externalidades do país em que seria realizado o projeto com relação a outros países poderia implicar a necessidade de um ajuste no cálculo desses preços. Por exemplo, se um determinado projeto implantado no Brasil fosse avaliado pelo BNDES e pelo BID poderia haver resultados diferentes, pois, do ponto de vista do BID, você teria que incorporar as possíveis externalidades que o projeto poderia causar com relação a outros países, que não interessariam em uma análise feita do ponto de vista do país. Minha tese foi uma tentativa de especificação da metodologia de ajuste dos preços sombra dos principais elementos que entram na avaliação de um projeto do ponto de vista nacional e do ponto de vista internacional. Havendo uma discrepância no resultado da avaliação do ponto de vista nacional e internacional, a entidade internacional deveria, eventualmente, estabelecer um esquema de compensação pelas externalidades positivas ou negativas. Cheguei a propor, inclusive, mecanismos de compensação. As entidades internacionais poderiam adotar mecanismos de financiamento

mais favoráveis para os países que gerassem externalidades positivas. Poder-se-ia, inclusive, incorporar pesos redistributivos, pois era uma preocupação muito grande dos bancos internacionais privilegiar projetos implantados em países menos desenvolvidos. O Brasil, que era considerado um país relativamente mais desenvolvido do que a Bolívia, por exemplo, deveria receber financiamentos em condições menos favoráveis.

Naquela época ainda não havia a obsessão que há hoje com relação às publicações. Eu não tive nenhuma preocupação em tirar um artigo importante a partir dos capítulos centrais da tese. Isso nunca passou pela minha cabeça. A única coisa que fiz foi traduzir para o português um apêndice da tese e publicá-lo na *Revista Estudos Econômicos*, em 1976, sobre a definição desses esquemas compensatórios. Eu não tive qualquer preocupação em pegar o capítulo central e tentar publicá-lo em uma revista internacional. O meu envolvimento com a área de análise custo-benefício incluía ministrar cursos na pós-graduação, na graduação e cursos para executivos. Eu ia ao Rio de Janeiro dar cursos na FUNCEX e orientava teses em análise de custo-benefício. Eu me tornei coordenador do programa de mestrado com especialização em bancos de desenvolvimento. No âmbito do convênio da FIPE com o BID, escrevi um texto sobre aspectos específicos da análise custo-benefício que, posteriormente, foi publicado como livro.

Nessa ocasião empenhei-me, fundamentalmente, em pesquisas sobre temas relacionados com economia aplicada. Em 1985 publiquei um artigo na *Revista Brasileira de Economia* tratando do custo doméstico das exportações. Posteriormente, fiz uma pesquisa contratada pelo IPEA também sobre o custo das exportações brasileiras com dados mais atualizados. Essa pesquisa foi feita junto com o Eli Pelin, sendo que nós publicamos um artigo na PPE contendo seus principais resultados. Participei também de algumas pesquisas com o professor Campino na área de nutrição, sempre cuidando da parte de avaliação econômica desses programas. Depois eu participei de alguns projetos de pesquisa com o Guilherme Dias, José Juliano, Martone, Simão Silber e outros pesquisadores da FIPE. Em pesquisa aplicada trabalhei frequentemente em equipe. Já em pesquisa teórica meu trabalho foi, na maioria das vezes, individual.

Relação com a pesquisa

O ponto de partida é ter interesse por um algum assunto específico para o qual você imagina poder eventualmente trazer algum tipo de contribuição. A escolha do tema de tese é muito complicada. Eu tive muita dificuldade em escolher meu tópico de doutoramento e menos dificuldade em escolher o tema da tese de livre docência, que acabou sendo o que

eu havia escolhido inicialmente como tópico da tese de PhD. Quando você tem autonomia de decidir o assunto de suas pesquisas, você tem que procurar aquilo que o Paulo Yokota, que foi meu professor na faculdade, uma vez mencionou aos alunos: “*you need to find a good vein. One time you find a good vein, you have everything to go forward.*” É difícil achar um veio bom. Tem pessoas que nunca encontram.

Há um aforismo que pode ser eventualmente muito útil para se compreender a questão da escolha de um tema interessante de pesquisa, que, se não me engano, é devido ao Mário Henrique Simonsen. “*um resultado teórico ou é óbvio ou está errado.*” Ou seja, quando você encontra um resultado que agride o senso comum, você tem razões para desconfiar dele. Isso é uma coisa que deveria sempre chamar a nossa atenção. No meu caso, ao ler um artigo do Sonnenschein de 1971, achei absolutamente estranha a sua tentativa de mostrar a possibilidade da existência de funções de demanda sem fazer a hipótese de transitividade, bastando apenas fazer a hipótese de convexidade das preferências. Desde o início aprendemos que uma hipótese crucial para a definição de uma função de demanda é a hipótese de transitividade. Entretanto, o resultado do Sonnenschein passou a ser aceito por diversos autores. Em função disso, eu fiz uma conjectura de que talvez a hipótese de convexidade implicasse uma forma, pelo menos mais branda, de transitividade, para então tentar provar a existência de funções de demanda sem uma hipótese explícita de transitividade. Parece uma coisa simples, mas eu levei anos trabalhando em cima disso até que o trabalho estivesse em condições de ser submetido a uma revista internacional. Em 1996, ele acabou sendo publicado pela revista *Journal of Mathematical Economics*. Esse, portanto, é um exemplo de um tema de pesquisa que partiu da idéia de que você deve desconfiar de resultados que ferem o senso comum.

Teoria econômica

Minha tese de livre docência foi uma tentativa de aceitar o desafio do Georgescu, ou seja, de elaborar um modelo de escolha a partir do reconhecimento da possibilidade de existência de critérios irreduzíveis como motivadores do comportamento dos agentes. Isso significava apresentar um modelo sem a hipótese de continuidade das preferências que pode implicar a existência de uma função utilidade. A tese foi defendida em 1986 e agraciada com o prêmio Haralambos Simeonidis no mesmo ano. Posteriormente, comecei a elaborar um artigo em cima do capítulo central da tese. Minha intenção era submetê-lo a uma revista internacional. Essa tarefa acabou se revelando um verdadeiro projeto de pesquisa de muitos anos. Uma primeira versão do artigo eu apresentei, em 1987, na reunião da Econometric Society. Foi somente em 1991 que cheguei a uma versão que pudesse ser

publicada. Fiquei em dúvida sobre qual revista seria mais apropriada para publicação. Terminei por decidir-me pela *Journal of Economic Theory*, que é uma revista voltada para teoria pura. O artigo não foi rejeitado, sendo que com poucas alterações foi finalmente publicado em 1993.

Atualmente estou procurando aplicar essa linha de pesquisa para o caso da firma. A idéia é olhar a Teoria da Firma como aplicação da Teoria da Escolha. É uma tentativa de mostrar que, ao contrário do que acontece corriqueiramente, você deveria enfatizar também hipóteses comportamentais que estariam por detrás do comportamento dos tomadores de decisão. Estou elaborando um modelo que admite objetivos múltiplos, de tal forma que a maioria dos modelos existentes poderiam eventualmente ser tomados como casos particulares. Tento também explorar a noção de racionalidade limitada para mostrar que quando você admite critérios múltiplos você pode eventualmente encontrar definições mais precisas para conceitos utilizados, por exemplo, pelos *behavioristas*, que são muito criticados pela apresentação daqueles de forma pouco rigorosa.

Publicação

É preferível publicar no exterior. No começo, eu não via isso de forma tão clara. Para atrair atenção é preciso publicar em uma revista boa, de primeira linha. Principalmente em teoria, é importante publicar no exterior porque, no Brasil, o público é limitado. Além disso há pouca gente trabalhando em teoria no Brasil. A obsessão de publicar dois ou três artigos por ano prejudica, muitas vezes, a possibilidade de se estudar e publicar artigos menos superficiais. Para publicar um trabalho significativo em uma revista importante é necessário estar preparado para gastar muito tempo. O trabalho que estou desenvolvendo agora já leva alguns anos, sem nenhuma garantia de que venha a ser bem-sucedido. Acho que sou meio improdutivo. Evidentemente, há pessoas mais produtivas que publicam um artigo por ano em revistas importantes. Nos EUA, há pessoas com menos de 40 anos e que já publicaram mais de 40 artigos, mesmo admitindo que cada artigo não seja uma coisa totalmente diferente dos demais. Mesmo assim trata-se de casos excepcionais. De qualquer forma, existe no EUA uma pressão maior do que aqui. Devido à longa espera até que o artigo seja publicado, muitas vezes, quando o artigo sai, ele já está obsoleto, principalmente quando se trata de um assunto que é muito debatido na literatura. Quando é um assunto mais próximo do tradicional, onde não existe um dinamismo tão grande, o risco dele ficar obsoleto é bem menor. Há linhas de pesquisas que podem ser exploradas e que são coisas já mais antigas. Por exemplo, o tema da racionalidade limitada, que começou a ser tratado

nos anos 50 pelo Simon, está experimentando atualmente certo “*revival*”. Outro tema muito debatido, hoje, é o dos bens públicos locais, que foi um assunto que começou também na década de 50.

Linhas de pesquisa

Hoje em dia é difícil abrir uma revista de teoria importante que não tenha um artigo sobre Teoria dos Jogos. A Teoria dos Jogos permite entender questões de estratégia, quer dizer, quando há um número limitado de agentes que poderiam adotar um comportamento em que as decisões de um agente levam em conta as decisões dos outros, em que cada um é obrigado a formular expectativas e hipóteses a respeito do comportamento de outros agentes. Há muitas situações desse tipo. No caso de uma firma, por exemplo, quando se trata de determinar o preço interno, ou seja, o preço de transferência de um insumo de uma determinada divisão para outra, nos deparamos com um caso de comportamento estratégico entre os executivos das duas divisões. Existe a tentativa de tratar situações desse tipo por meio da Teoria dos Jogos, embora, segundo autores como o Radner, essas tentativas não tenham sido muito bem-sucedidas. Existem muitos problemas teóricos que ainda não foram resolvidos na Teoria dos Jogos. Frequentemente são feitas hipóteses para viabilizar certas demonstrações que determinados autores consideram muito fortes como, por exemplo, a hipótese de que cada jogador julga racional o comportamento do oponente. Isto pode explicar, em parte, por que a teoria dos jogos tem atraído o interesse de tantos autores.

Uma outra linha de pesquisa importante diz respeito à assimetria das informações. Há alguns trabalhos, por exemplo, que usam informações assimétricas para explicar o desemprego. Foram desenvolvidos também diversos modelos de agente principal para explicar, por exemplo, a relação entre gerentes e proprietários de firmas e entre acionistas e gerentes. Neste caso, trata-se de um caso típico de agente principal em que o acionista poderia induzir o gerente a atingir resultados que maximizem a sua função objetivo.

Finalmente, existem trabalhos que têm a sua origem em *papers* básicos publicados há mais de quarenta anos. Ainda hoje é possível encontrar muitos artigos tratando do teorema de Arrow, discutindo formas de resolver o teorema da impossibilidade. A própria Teoria dos Jogos fez com que ganhassem novamente proeminência coisas do século passado, como o modelo de Cournot. Há trabalhos sobre esse modelo que utilizam as novas técnicas de Teoria dos Jogos. Eu tenho uma certa preferência para tratar assuntos básicos. É difícil em teoria econômica você poder afirmar que determinado assunto esteja liquidado. Há muitos

assuntos que talvez pouca gente se interesse em estudar hoje porque se considera que estejam já resolvidos. Em certo sentido, seria uma coisa desestimulante imaginar-se que não haveria a possibilidade de progredir em direção à resolução de novas questões, a partir da solução definitiva de problemas já formulados. De certa forma, este tipo de evolução acontece. No entanto, creio que existem assuntos básicos que ainda não estão totalmente resolvidos. No caso dos modelos básicos de escolha e da firma, por exemplo, há algumas hipóteses que poderiam ser relaxadas ou que poderiam ser tratadas de outra forma.

ERRATA

Artigo: **Wage gender discrimination and segmentation in the Brazilian labor market**

Autora: **Ana Lúcia Kassouf**

Publicado em: *Economia Aplicada*, v. 2, n. 2, abril-junho, 1998

1) p. 246, 1ª linha, **onde se lê:**

“where x_s and z_s are exogenous variables and $E(u_s | x_0, x_1, x_2, z_0, z_1, z_2) = 0$ and $E(\eta_s | x_0, x_1, x_2, z_0, z_1, z_2) = 0$.”

leia-se:

“where x_s and z_s are exogenous variables and $E(u_s | x_0, x_1, x_2, z_0, z_1, z_2) = 0$ and $E(\eta_s | x_0, x_1, x_2, z_0, z_1, z_2) = 0$.”

2) p. 146, 12ª linha, **onde se lê:**

“Suppose that F_j ($j=0,1,2$) have cumulative distribution function”

leia-se:

“Suppose that η_j ($j=0,1,2$) have cumulative distribution function”

3) p. 246, 16ª linha, **onde se lê:**

“Then, it can be shown that

$$Prob(\varepsilon_s < z_s \gamma_s) = Prob(I = s) = \frac{\exp(z_s \gamma_s)}{\sum_j \exp(z_j \gamma_j)}$$

Thus, the distribution function of I_s is given by

$$F_s = Prob(\varepsilon_s < \varepsilon) = Prob\left[\left(\max_{j=0,1,2, j \neq s} I_j^* - \eta_s\right) < \varepsilon\right] = \frac{\exp(\varepsilon)}{\exp(\varepsilon) + \sum_{j=0,1,2, j \neq s} \exp(z_j \gamma_j)}$$

leia-se:

“Then, it can be shown that

$$\text{Prob}(\varepsilon_s < z_s \gamma_s) = \text{Prob}(I = s) = \frac{\exp(z_s \gamma_s)}{\sum_j \exp(z_j \gamma_j)}$$

Thus, the distribution function of ε_s is given by

$$F_s = \text{Prob}(\varepsilon_s < \varepsilon) = \text{Prob}\left[\left(\text{Max}_{j=0,1,2, j \neq s} I_j^* - \eta_s\right) < \varepsilon\right] = \frac{\exp(\varepsilon)}{\exp(\varepsilon) + \sum_{j=0,1,2, j \neq s} \exp(z_j \gamma_j)}$$

4) p. 247, 3ª e 4ª linhas, **onde se lê:**

“where the dependent variable w_s is observed if and only if the category s is being chosen, i.e., $\varepsilon_s < z_s \gamma_s$. Consider the following transformation to normality”

leia-se:

“where the dependent variable w_s is observed if and only if the category s is being chosen, i.e., $\varepsilon_s < z_s \gamma_s$. Consider the following transformation to normality:”

5) p. 247 5ª linha, **onde se lê:**

“where $\Phi(\cdot)$ is the distribution function of the standard normal.

leia-se:

“where $\Phi(\cdot)$ is the distribution function of the standard normal.”

6) p. 247, 7ª linha, **onde se lê:**

“The condition $\varepsilon_s < z_s \gamma_s \Leftrightarrow \varepsilon_s^* < J_s(z_s \gamma_s)$, and if u is normal distributed we have that,”

leia-se:

“The condition $\varepsilon_s < z_s \gamma_s \Leftrightarrow \varepsilon_s^* < J_s(z_s \gamma_s)$, and if u is normal distributed we have that,”

7) p. 247. 11ª linha, **onde se lê:**

“where ϕ is the density function of the standard normal, $\sigma_s^2 = \text{Var}(u_s)$, and ρ_s is the correlation coefficient between u_s and ε_s^* ”

leia-se:

“where ϕ is the density function of the standard normal, $\sigma_s^2 = \text{Var}(u_s)$, and ρ_s is the correlation coefficient between u_s and ε_s^* ”

8) p. 247 14ª linha, **onde se lê:**

“can be estimated by ordinary least squares after substituting the estimated values of γ_s from the multinomial logit model where I_s is regressed on z_s by maximum likelihood.”

leia-se:

“can be estimated by ordinary least squares after substituting the estimated values of γ_s from the multinomial logit model where I_s is regressed on z_s by maximum likelihood.”

9) p. 259, 16ª linha, **onde se lê:**

“experience, i.e., it is expected a positive coefficient for β_1 and negative for β_2 . Besides”

leia-se:

“experience, i.e., it is expected a positive coefficient for β_1 and negative for β_2 . Besides”

10) p. 264, 5ª linha, **onde se lê:**

$$= 0.101 - 0.00266 \text{ age} + 0.00266 \text{ educ} - 0.00171 \text{ educ}$$
$$= 0.101 - 0.00266 \text{ age} + 0.00095 \text{ educ}$$

leia-se:

$$= 0.101 - 0.00266 \text{ age} + 0.00266 \text{ educ} - 0.00171 \text{ educ}$$
$$= 0.101 - 0.00266 \text{ age} + 0.00095 \text{ educ}$$

SUMÁRIO

STEPHEN HABER	379	The Efficiency Consequences of Institutional Change: The Political Economy of Financial Market Regulation and Industrial Productivity Growth in Brazil, 1866-1934
JOSÉ ROBERTO VICENTE	421	Determinantes da Adoção de Tecnologia na Agricultura Paulista
EVELISE NUNES DO ESPÍRITO SANTO	453	Agricultura no Estado de Santa Catarina, Período 1920-1985
RAUL CRISTOVÃO DOS SANTOS	475	Smith's Glasgow Lectures: A Foundation or an Obstacle for the Wealth of Nations?
DANTE MENDES ALDRIGHI	491	A Segmentação no Mercado de Trabalho do Setor Industrial de São Paulo (1889-1920)
RENATO LEITE MARCONDES	533	Padrão-Ouro e Estabilidade
JOSUÉ PEREIRA DA SILVA	561	<i>Sindicalismo e Política. A trajetória da CUT, de Iram Jácome Rodrigues (Resenha)</i>

Revista de
**Economia
Política**

Revista de Economia Política, vol. 18, nº 3 (71), julho-setembro/1998

ARTIGOS

PADRÕES DE ESPECIALIZAÇÃO, HIATOS TECNOLÓGICOS E CRESCIMENTO COM RESTRIÇÕES DE DIVISAS

Otaviano Canuto 5

Patterns of specialization, technological hiatus and foreign-exchange constrained growth

Artigo que apresenta um modelo Norte-Sul de interação entre comércio exterior e crescimento com restrição de divisas.

O EMPRESÁRIO BRASILEIRO: UM ESTUDO COMPARATIVO

Sérgio de Oliveira Birchal 20

The Brazilian entrepreneur: a comparative study - Exame do debate sobre o empresário na historiografia econômica brasileira, comparando origens sociais e fontes de capital dos empresários mineiros, paulistas e cariocas no século 19.

OS NOVOS CLÁSSICOS E A TEORIA DOS CICLOS DE HAYEK

Jorge Eduardo de Castro Soromenho 32

The new classics and Hayek's cycle theory Procura demonstrar que a teoria novo-clássica e a de Hayek apresentam diferentes concepções sobre o modus operandi das economias capitalistas.

PRODUTO, CAPITAL E TAXA DE LUCRO: PAÍSES INDUSTRIALIZADOS E AMÉRICA LATINA

Samuel Kilsztajn 55

Product, capital and profit rate: industrialized countries and Latin America - O objetivo deste artigo é comparar a margem do lucro, a relação capital/produto e a taxa de lucro de países industrializados e da América Latina, discutindo paralelamente a relação entre taxa de lucro e nível de produtividade e salário das economias mencionadas.

OS FUNDAMENTOS MICROECONÔMICOS DOS INDICADORES DE DESENVOLVIMENTO SÓCIO-ECONÔMICOS

Juan Hersztajn Moldau 70

Microeconomic elements of the socio-economic development indices Propõe o uso de modelos teóricos não convencionais para fundamentar a construção e utilização de indicadores, destacando aspectos conceituais.

PERSPECTIVAS PARA A IMPLANTAÇÃO DE UMA POLÍTICA INDUSTRIAL NO BRASIL

Robson Ribeiro Gonçalves 84

Perspectives for the implementation of an industrial policy in Brazil Este artigo analisa alguns aspectos relacionados à implementação de uma política industrial no Brasil coerente com os contextos interno e externo da atualidade.

A AGENDA DE MÉDIO PRAZO NO BRASIL E O FUTURO DA PETROBRÁS	
<i>Fabio Giambiagi e Adriano Pires Rodrigues</i>	100
<i>The medium-term agenda in Brazil and teh future of Petrobrás</i> Ao discutir a possível privatização da Petrobrás, este paper procura contribuir para organizar a discussão e os argumentos em torno da proposta de privatização.	

DOCUMENTO

A INSERÇÃO EXTERNA E O DESENVOLVIMENTO	
<i>Gustavo H. B. Franco</i>	121
<i>External insertion and development</i> - Trabalho escrito em junho de 1996 pelo então Diretor de Assuntos Internacionais do Banco Central do Brasil, foi objeto de ampla discussão. Procurando prestar um serviço a nossos leitores, publicamos aqui a íntegra do artigo.	

RESENHAS DE LIVROS

André Lara Resende: <i>Bolhas e pêndulos</i> (<i>Paulo Roberto Arvate</i>)	148
Eli Diniz: <i>Crise, reforma do Estado e governabilidade</i> (<i>Maria Lucia Teixeira Werneck Vianna</i>)	149
John Kenneth Galbraith: <i>A sociedade justa</i> (<i>Luiz Antonio de Oliveira Lima</i>)	152
Rosa Fontes (ed.): <i>Estabilização e crescimento</i> (<i>Luis Eduardo Afonso</i>)	154

0

21

148

149

152

154