A Study on the Shadow Economy and the Tax-Gap: The case of CPMF in Brazil*

Claudio R. Lucinda† *FGV/EESP and EAESP*

Paulo R. Arvate‡ *FGV/EESP and EAESP*

Abstract

This paper aims to present the effects of the CPMF (Provisional Contribution on Financial Transactions), a Social Security Contribution created to finance the Brazilian health system. During the congressional debates surrounding its approval, the Brazilian conressmen have also used as a reason for its enaction its ability to fight tax evasion. In order to assess the real impact of this measure, a MIMIC (Multiple Causes-Multiple Indicators) model was estimated to determine the size of the Brazilian Shadow Economy. Our results indicate that, despite these points of view, the Brazilian Shadow Economy - whose definition includes tax evasion - has increased during the time CPMF was in place, and this Contribution has contributed for this increase. We also have measured the extent of the TAX-GAP (potential taxes and contributions revenue loss) resulting from the size of Brazilian Shadow Economy, both as total tax and contribution revenue and as a percentage of CPMF revenue.

JEL Codes: O17, D78, H2, H26

Category: International Public Choice/Political Economy or Public Goods and Public Finance

*The authors would like to thank the financing by GVPesquisa.

†Professor of Economics and Corresponding Author. Address: Av. 9 de Julho, 2029-São Paulo-SP-Brazil. Zip Code: 01313-902 Phone: +55-11-3281-7765 Fax: +55-11-3281-7813 E-mail: claudiolucinda@fgvsp.br.

‡Professor of Economics. Address: Av. 9 de Julho, 2029-São Paulo-SP-Brazil. Zip Code: 01313-902 Phone: +55-11-3281-7765 Fax: +55-11-3281-7813 E-mail: parvate@fgvsp.br.

1 Introduction

In order to generate funds for the Public Health system in Brazil, a Social Security Contribution was enacted during the first months of 1994. This contribution had a characteristic which set it apart from the former initiatives - its collection at every financial transaction. Due to this reason this contribution was called Provisional Contribution on Financial Transactions, or CPMF[1]. Since Brazil has a developed system of payments, this new contribution seemed destined to raise the desired funds. Furthermore, its proponents also announced its ability to act as a powerful instrument to counter tax evasion, because the information generated in collecting it at every financial transaction might generate an estimate of each taxpayer´s income independent from that presented for tax purposes. Thus, the CPMF not only would raise the desired funds, but also would help cut back on the size of the Shadow Economy[2].

On the likely effects of CPMF on the Shadow Economy, by contrast, the international literature has a very solid verdict: as with any new tax, the activities which comprise the Shadow Economy will comprise a greater part of the national income.

In order to ascertain the true effects of this measure, the present research was carried out. This task presented some difficulties, the most important of all being the measurement of the Shadow Economy. The first step on this course would be establishing a precise meaning for the concept, which was derived from the work of Schneider and Kinglmair(2004), quoted from Lippert and Walker (1997). This definition includes both legal and illegal activities, as shown in the following table:

Table 1: Shadow Economy - Activities

Type of Activity	**Monetary transactions**		**Non monetary transactions**	
Illegal Activities	Trade with stolen goods; drug dealing and manufacturing; prostitution; gambling; smuggling, fraud; etc.		Barter of drugs, stolen goods, smuggling, etc. Produce or growing drugs for own use. Theft for own use.	
	Tax Evasion	**Tax Avoidance**	**Tax Evasion**	**Tax Avoidance**
Legal Activities	Unreported income from self-employed worker; wages, salaries and assets from unreported work related to legal services and goods	Employee discounts, fringe benefits	Barter of legal services and goods	All do-it-yourself work and neighbor help

Source: Schneider e Kinglmair (2004).

The second step to be taken care of relates to the actual measurement of the Shadow Economy, measured as

[1]The acronym, which will be used troughout the paper, is derived from the Portuguese name of the measure, *Contribuição Provisória sobre Movimentações Financeiras*.

[2]See an excerpt from State Representative Ieda Salvatti in the Santa Catarina House in December 2000: "But, more important than benefiting 27 million Brazilian citizens, more important, even, than making the income tax issue more equitable in Brazil, this secrecy lifting measure by the CPMF, would amount to a final blow against money laundering, unreported earnings, and the drug trade." See an excerpt from Senator Roberto Requião who, despite voting against the CPMF when under appreciation at the Senate, made a point to indicate that, under different circumstances, he ".. would vote in favor of a tax similar to the CPMF, due to its comprehensiveness and its ability to prevent evasion, "

above. There is no data bank on it: hence, it must be estimated. And there are several estimation methods. Bajada and Schnieder (2003) have discussed the methods used to estimate it, dividing them into six categories:

1. **Qualitative estimat**e: qualitative research on the monetary cost of illegal activities, such as bribery, for example, generates information on the size of the Shadow Economy. This sort of study is very dependent on the quality of the primary data, and studies that generate such information are very important in the struggle against illegal activities.

2. **Discrepancy between income declared for tax purposes and that measured by selective checks:** The Shadow Economy is estimated by the discrepancy between the income declared for tax payment purposes and the income revealed by independent audit. The main criticism of this method is that it can generate a bias: the individuals selected for audit may not be representative of the population, maybe due to a cost-benefit analysis is involved in sample selection.

3. **Discrepancy between statistics of National Expenditure and Income:** In theory, on raising the national accounts, national expenditure and national income should be identical. Whenever they are not identical, the technical procedure to be adopted by the agency that computes National Accounts is to minimize this discrepancy by reviewing each component of expenditure and income. Where the discrepancy remains, errors and omissions are assumed to preserve the equality of expenditure and income. An attempt to estimate the Shadow Economy would not be rendered entirely impossible if there was an independent checking of expenditures. In the presence of such an estimate, comparing it with the official national income as calculated in National Accounts would certainly reveal the size of the Shadow Economy as equal to the discrepancy.

4. **Discrepancy between the official and the actual labor force:** the Shadow Economy may be estimated by calculations of informal - or unregistered - employment. Certain individuals are part of the labor force, but are not in formally registered positions. In Brazil, the institutions in charge of open unemployment statistics publish monthly this sort of data. According to Bajada and Schnieder (2003), this type of statistic as proxy for the Shadow Economy suffers from many drawbacks, because countless other factors can determine the existence of an informal labor market. In Europe, for example, unofficial employment might be the result of the decision to hold a second job, which certainly is not applicable to Brazil. The cost of officially employing a worker in Brazil is significantly higher. According to 2004 data published by IBGE[3] (Brazilian Statistics Buerau), half of the labor force is unregistered.

5. **Currency demand approach:** The Shadow Economy estimate can be constructed from the difference between the demand for currency, considering factors that might impact it, and the demand for currency without such factors. This approach assumes that the transactions composing the Shadow Economy are undertaken in the form of cash payments, leaving no observable traces for the authorities. With the same technique, there may be many differences in results among different studies concerning the same country. One of the papers most often quoted along these lines is Tanzi (1983).

6. **Latent Variable Modelling - Multiple Indicators Multiple Causes (MIMIC) Model**: An estimate of the size of the Shadow Economy is obtained from the interplay of several variables, some of them assumed to impact the evolution of the Shadow Economy, and others which are assumed to be influenced by it. These variables include some used in the previously discussed techniques: labor market, tax evasion, demand for currency, etc. Giles (1998) and others have produced a series of studies that estimate it using this methodology.

[3]Its portuguese name is: *Instituto Brasileiro de Geografia e Estatística*, the reason for the acronym.

This last technique, using the Latent Variable modelling, was used in this paper, following the recent work done by Ribeiro and Bugarin (2003) in estimating the size of the Shadow Economy. After obtaining this estimate, the next step was to ascertain the role the Tax Collection and, specifically, the CPMF collection had on the evolution of the size of the Shadow Economy. Another point analyzed in this research was to determine the loss of potential tax and social contribution revenue (called the TAX-GAP) due to these choices on Tax Revenue and CPMF.

This article is divided into six sections, including this introduction. The second section introduces the MIMIC model structure used in the works of Giles (1998), Bajada (1999, 2002) and Ribeiro and Bugarin (2003), as well as some recent results from estimates of the size of the Shadow Economy in Latin American countries including Brazil. The third section introduces our estimate of the Shadow Economy in Brazil. To obtain it, we need a point to calibrate the MIMIC model. Without the calibration, the MIMIC model gives us only an index, which will be given from an estimate using the Currency Demand approach. In the fourth section we discuss the relation between both Total Tax Revenue and CPMF with the size of the Shadow Economy. In the fifth section, we discuss the share of CPMF in the total TAX-GAP loss . The final section discusses the main results of the paper.

2 The Multiple Indicators Multiple Causes Model (MIMIC)

The structure of the Multiple Indicators Multiple Causes (MIMIC) model that was the basis for the estimates of Giles (1999), Barjada (1999 and 2002) and Klnglmair and Schnieder (2004) was as follows. Supposing the size of the Hidden Economy could determine the behavior of a matrix of $(T \times p)$ observed variables denoted $\mathbf{Y}$- which we can call the indicator variables - we can present the relationship between these indicator variables and the size of the Hidden Economy as follows:

$$\mathbf{Y} = \eta \times \mathbf{\Lambda} + \varepsilon \tag{1}$$

In which:

- $\mathbf{\Lambda}$ is a $(1 \times p)$ vector of coefficients.
- η is a $(T \times 1)$ vector representing the size of the Hidden Economy.
- ε is a $(p \times T)$ matrix of distrubances, normally distributed and uncorrelated with other variables.

The size of the Hidden Economy, in turn, which we denoted by η_t, could be influenced by a vector of $(T \times q)$ observed variables $\mathbf{X}$- which we can call the causal variables. We can present this relationship as follows:

$$\eta = \mathbf{X} \times \mathbf{\Gamma} + \xi \tag{2}$$

In which:

- $\mathbf{\Gamma}$ is a $(q \times 1)$ vector of coefficients.
- ξ is another vector $(T \times 1)$ of disturbances, also normally distributed and uncorrelated with other variables.

We also have to impose some conditions on the behavior of the error terms:

$$\begin{aligned} Var(\xi) &= \Psi \\ Cov(\varepsilon) &= \Theta \times \epsilon \end{aligned}$$

Substituting 2 in 1, we have a sort of multivariate regression model with a very complex error structure:

$$\begin{aligned} \mathbf{Y} &= \mathbf{X} \times \mathbf{\Pi} + z \\ \mathbf{\Pi} &= \mathbf{\Gamma} \times \mathbf{\Lambda} \\ z &= \mathbf{\Lambda} \times \xi + \varepsilon \\ Cov(z) &= \mathbf{\Lambda}^{'} \mathbf{\Lambda} \Psi + \Theta\epsilon \end{aligned} \tag{3}$$

There are many methods available for estimation of this sort of model, from the Maximum Likelihood presented in Jöreskog and Goldberger (1975), and the general approach presented by Watson and Engle (1983), which stresses the similarities of this sort of model with Kalman Filter Modelling. The most recent estimates of the Shadow Economy for Latin America were found in Kinglmair and Schnieder (2004) and are reproduced in Table 1:

Table 2: Size of the Shadow Economy in Latin American Countries

Country	**Shadow Economy in % of GNP 1999/2000**
Argentina	25.4
Bolivia	67.1
Brazil	**39.8**
Chile	19.8
Colombia	39.1
Costa Rica	26.2
Dominican Republic	32.1
Ecuador	34.4
Guatemala	51.5
Honduras	49.6
Jamaica	36.4
Mexico	30.1
Nicaragua	45.2
Panama	64.1
Peru	59.9
Uruguay	51.1
Venezuela	33.6
Average	**41.0**

Source: Kinglmair and Schneider (2004).

According to these estimates, Brazil is very close to the average. The country with the biggest Shadow Economy is Panama, while the smallest in found in Chile. Since these estimates are from a cross-section of data, they are not useful for the assessment of the likely effects of CPMF. The next sextion is the beggining of our attempt to do so.

3 Estimating the Shadow Economy in Brazil

The goal of this section is to estimate the size of the Shadow Economy in Brazil between 1995 and 2002. Our estimate will be based on the approach developed in the previous section, a MIMIC model. The results generated this model will not be easily interpreted because they are presented as an index. In order to convert the results to an index which represent this size of the Shadow Economy, we need to calibrate the series thus obtained with another estimate. In their estimate for Brazil, Ribeiro and Bugarin (2003) did not estimate the volume of currency

the Shadow Economy uses and, as a result, their results were presented as an index. The technique used in the determination of the calibration point is presented in the following section.

3.1 Shadow Economy: The Currency Demand Approach

As presented in the first section of the paper, Barjada (1999 and 2000) estimated the size of the Shadow Economy by using the difference between the amount of currency demanded in an economy and the amount that would be demanded if the factors responsible for the demand of currency related to the Shadow Economy were absent. In order to establish a point to calibrate the series obtained with the MIMIC approach, we started by carrying out a procedure similar to the one used by this author, estimating an Error Correction Model (ECM) with the following specification[4]:

$$\begin{aligned} \Delta Ln(PMPP_t) &= \delta_0 + \delta_1 \Delta Ln(PIB_t) + \delta_2 \Delta Ln(R_t) + \delta_3 \Delta Ln(T_t) + \delta_4 \Delta Ln(\pi_t) + \\ &+ \delta_5 D_t^2 + \delta_6 D_t^3 + \delta_7 D_t^4 + \delta_8 Ln(PMPP_{t-1}) + \psi_2 Ln(PIB_{t-1}) + \psi_3 Ln(R_{t-1}) + \\ &+ \psi_4 Ln(T_{t-1}) + + \psi_5 Ln(\pi_{t-1}) + \delta_9 TREND_t + \delta_{10} D_t^{REAL} + \varepsilon_t \end{aligned} \tag{4}$$

The definition of the variables is as follows:

- Currency ($PMPP_t$) - Denominated in millions of Reais.
- Gross Domestic Product at market prices (PIB_t) - Denominated in millions of reais. The GDP estimate is produced by the Brazilian Bureau of Economic Analysis - IPEA[5].
- Short-Term Interest Rate (R_t) - It is an interest rate negotiated between banks and the public, expressed as % per month.
- Revenue from the Social Contributions as percentage of GDP (T_t) - Sum of the revenue of six different social contributions, including CPMF[6]. Expressed in millions of Reais.
- Inflation Rate (π_t) - As measured by the change in the General Prices Index (IGP).
- D_t^2, D_t^3 and D_t^4- They are quarterly dummies relative to the second, third and fourth quarters of the year.
- $TREND_t$- Temporal trend;
- D_t^{REAL}- Dummy relative to Real Plan. Real Plan was the stabilization program that reduced inflation in Brazil. Up until the third quarter, 1994, when inflation was high, the variable will be assigned value zero. From that point onwards, with lower inflation rates, the variable will be regarded as equal to 1.

[4]After different tests, this was the best model estimated to represent the demand by money on the HE. Note that was determinant role played by the Social Contributions, rather than the Total Tax Revenue, in this determination. In fact, the growth of the Social Contributions as share of the Total Revenue is evident in the 1990s. Arvate and Silva (2004) showed that the cause of the growth of the Social Contributions in the 1990s is connected with the earmarking of federal government revenues pursuant to the 1988 Constitution. The New Constitution imposed sharing traditionally Federal taxes with State and Municipals governments, and to postpone the passage of expenditures to them with a complementary legislation (Laws with hierarchy fewer than the Constitution). Since this complementary legislation was late in coming, the Federal government responded by investing heavily in the collection of the Social Contributions (which are not shared with States and Municipalities). The growth of the Total Government Revenues in the 1990s is due to the appearance (in any cases) and growth of the Social Contributions.

[5]*Instituto de Pesquisa Econômica Aplicada.*

[6]Revenue from CIDE (Contribution on Fuels), FINSOCIAL/COFINS (Contributions to Social Security on the revenue of companies), CPMF, PIS/PASEP (Social Integration Program/Civil Servant's Wealth-Building Program, on the revenue of companies and on civil servants' pay) and CSLL (Contribution on Net Profits).

All of the data were taken from the IPEA web site[7] on a quarterly basis. The statistics describing the variables (average, standard deviation, minimum and maximum, etc.) are found on Appendix B. On specification 4, the subscripts denote the time periods, Δ denotes the first difference operator and Ln the Neperian Logarithm. Before we can carry out the actual estimation of specification 4, however, we must be sure this specification can be regarded as a proper approximation of the underlying data structure. Two points must be analyzed: the stationarity of the series used, and the existence of a long-run relationship between them. The stationarity of the series is considered by recourse to the ADF (Augmented Dickey-Fuller) test, whose results are presented below.

Table 3: ADF Test Results

Variables	**Test Statistic**	**Results (5% of Significancy)**	**Deterministics**
$Ln(PMPP_t)$	-9.882	I(0)	Constant and Trend
$Ln(PIB_t)$	-2.470	I(1)	Constant and Trend
$Ln(T_t)$	-3.193	I(1)	Constant and Trend
$Ln(R_t)$	-4.154	I(0)	Constant

The results in the preceding table show that two variables are stationary: $Ln(PMPP_t)$ and $Ln(R_t)$. However, and particularly in the case of variable $Ln(PMPP_t)$, the result as to whether or not a unit root exists is unclear because it changes with the deterministic terms included in the test specification. Therefore, we decided to consider both variables as I(1).

The next step is to carry out a cointegration analysis to determine whether a long-term relationship among variables do exist. The test results can be seen in the table that follows:

Table 4: Cointegration Test - Johansen

Data Trend:	None	None	Linear	Linear	Quadratic
Rank or Number of Coint. Eq.	No Intercept No Trend	With Intercept No Trend	With Intercept No Trend	With Intercept With Trend	With Intercept With Trend
Number of Cointegrating Relationships (5%) Selected by Model (columns)					
Trace	3	4	3	5	5
Max. Eigenvalue	2	2	2	3	3

Note that the cointegration test indicates the presence of more than one long-term relationship among variables. Since we are not interested in characterizing the long run relationship *per se* and the literature states that Ordinary Least Squares estimation is consistent according to these results, we could use the estimates obtained through equation 4 as appropriate for the calibration of the subsequent estimate. The results follows:

[7] `http://www.ipeadata.gov.br`

Table 5: Estimation Results - Error Correction Model

Variable	**Coefficient**	**t Statistic**	**P-Value**
Constant	-2.019	-1.882	0.075
$\Delta ln(PIB_t)$	0.806	5.059	0.000
$\Delta ln(R_t)$	-0.061	-1.096	0.286
$\Delta ln(T_t)$	0.094	0.914	0.371
$\Delta ln(p_t)$	0.019	1.545	0.138
D_t^2	0.009	0.212	0.834
D_t^3	-0.009	-0.167	0.869
D_t^4	0.263	4.340	0.000
$ln(PMPP_{t-1})$	-0.745	-4.230	0.000
$ln(PIB_{t-1})$	0.667	3.366	0.003
$ln(R_{t-1})$	-0.038	-0.880	0.389
$ln(T_{t-1})$	-0.076	-0.733	0.472
$ln(p_t)$	0.017	0.964	0.347
$TREND_t$	0.017	3.438	0.003
D_t^{REAL}	0.237	1.028	0.316
R^2	0.991		
Adjusted R^2	0.985		
Log likelihood	72.682		
Durbin-Watson stat	2.291		

We can use these estimated coefficients to compute the predicted value of the $PMPP_t$ variavle, which we will denote by $\overline{PMPP_t}$. The next step is to determine the amount of currency without the factors generating currency demand due to the Shadow Economy. This quantity, called $\overline{PMPP_t^e}$, was calculated as:

$$\begin{aligned}\Delta Ln(\overline{PMPP_t^e}) &= \widehat{\delta_0} + \widehat{\delta_1}\Delta Ln(PIB_t) + \widehat{\delta_2}\Delta Ln(R_t) + \widehat{\delta_4}\Delta Ln(\pi_t) + \\ &+\widehat{\delta_4}D_t^2 + \widehat{\delta_5}D_t^3 + \widehat{\delta_6}D_t^4 + \widehat{\delta_8}Ln(PMPP_{t-1}) + +\widehat{\psi_2}\Delta Ln(PIB_{t-1}) \\ &+\widehat{\psi_3}Ln(R_{t-1}) + \widehat{\psi_5}Ln(\pi_{t-1}) + \widehat{\delta_{10}}TREND_t + \widehat{\delta_{11}}D_t^{REAL}\end{aligned}$$

Circumflexes indicate coefficients estimated by OLS in equation 4. The next step is to define the money volume used in the Shadow Economy, called $PMPP_t^H$, calculated by $PMPP_t^H = \overline{PMPP_t} - \overline{PMPP_t^e}$. From the series thus computed, we obtained a share of the size of the Shadow Economy in monetary terms of about 24.89% of GDP from the whole period. It is important to note that the purpose of this series is just to provide a single point for calibration of the series obtained with the MIMIC model.

3.2 Shadow Economy: The MIMIC Approach

The final step on the construction of a series representing the evolution of the Shadow Economy is to apply the Latent Variable Modelling approach - MIMIC Model. Our estimate used the following data series, obtained monthly for the period between January 1995 and December 2002:

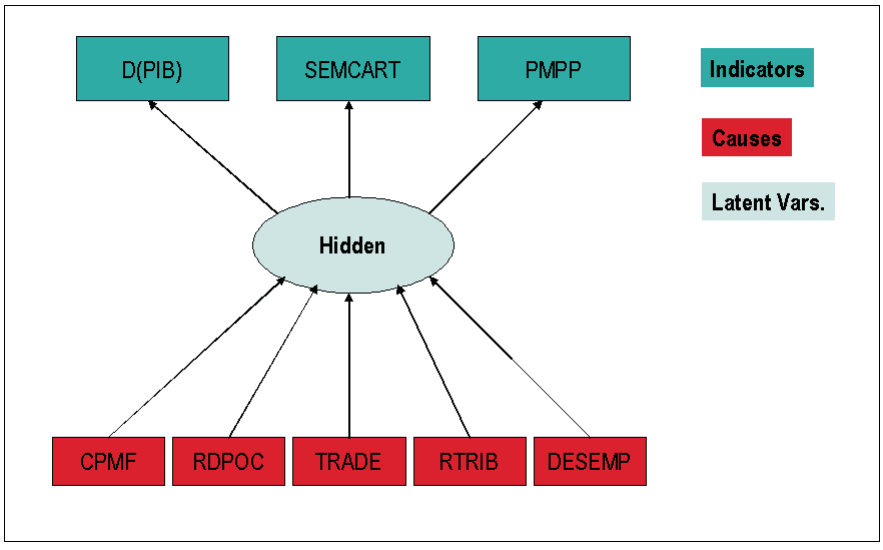
- PIB_t- Gross Domestic Product in current mullions of Reais. The data was deflated by IGP, and presented as a indez number, (base July 1994=100).
- $SEMCART_t$- Index number (base Jul/1994=100) representing the number of unregistered workers over 15 years old in the following metropolitan areas: Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo

and Porto Alegre. Source: Monthly IBGE Employment Survey[8] .

- $CPMF_t$- Net revenue from the Provisional Contribution on Financial Transactions - in current millions of Reais, deflated by General Prices Index, expressed as an index number (base July 1994=100).
- $RTRIB_t$- Gross Revenue Managed by the Receita Federal in current millions of Reais, deflated by the General Prices Index, expressed as an index number (base July 1994=100).
- $RDPOC_t$- Proxy for Available Income by Employed Population. Computed as the difference between Gross Domestic Product and the Gross Revenue Managed by the Receita Federal ($RTRIB_t$), divided by the number of employed individuals in the aforementioned metropolitan regions.
- $TRADE_t$- Represents the degree of openness of the economy. It is the sum of exports and imports of goods and services, as a share of GDP. These data were obtained from the IPEA database.
- $DESEMP_t$- Open Unemployment Rate in the Metropolitan Areas. Comprehends the following metropolitan regions: Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo and Porto Alegre. Source: Monthly IBGE Employment Survey.
- IGP_t - Inflation Rate as measured by the General Price Index[9]. Source: IPEA database.
- POC_t - Employed Population in Metropolitan Regions. Comprehends the following metropolitan regions: Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo and Porto Alegre. Source: Monthly IBGE Employment Survey.

The analyzed series were submitted to the Augmented Dickey-Fuller test for the detection on nonstationarity. Only the PIB_t series was nonstationary. Given this result, it was included in the first differences. The basic structure of the MIMIC model employed is portrayed in the upcoming figure.

Figure 1: MIMIC Model Structure



[8] *Pesquisa Mensal de Emprego do IBGE* - PME/IBGE.
[9] *Índice Geral de Preços - Versão Disponibilidade Interna.*

All of the models obtained from subsets of these variables were tested. In order to select specifications - from a total of over 120 models - the following criteria were used: Akaike Criterion (AIC), Bayesian Information Criterion (BIC), Browne-Cudeck Information Criterion (BCC), Least Discrepancy Function Value (C), Least Discrepancy Function Value minus Degrees of Freedom (C-df), Least Discrepancy Function Value divided my Degrees of Freedom (C/df). The best models according to each criterion are as follows.

Table 6: Selected Specification Results - MIMIC Model

Variables	**1**	**2**	**3**	**4**	**5**
INDICATORS					
$SEMCART_t$	0.198	0.196	0.191	0.187	0.187
	(0.027)	(0.027)	(0.028)	(0.026)	(0.027)
$PMPP_t$	1	1	1	1	1
ΔPIB_t					-0.005
					(0.018)
CAUSES					
$CPMF_t$	0.299	0.244	0.216	0.212	0.213
	(0.040)	(0.036)	(0.035)	(0.034)	(0.034)
$TRADE_t$	5.947	6.012	5.232	5.529	5.483
	(0.762)	(0.767)	(0.726)	(0.719)	(0.717)
$DESEMP_t$		6.046	5.792	5.47	5.474
		(1.612)	(1.639)	(1.633)	(1.633)
$RTRIB_t$			0.326	0.407	0.411
			(0.094)	(0.097)	(0.097)
$RDPOC_t$				0.337	0.322
				(0.207)	(0.207)
C	348.661	337.422	333.182	331.687	331.522*
C-df	323.661	313.422	310.183	309.687*	310.522
AIC	386.661	377.422	375.183*	375.687	377.522
BCC	390.637	381.608	379.578*	380.291	382.336
BIC	435.383	428.709*	429.034	432.102	436.502
C/df	13.946*	14.059	14.486	15.177	15.787

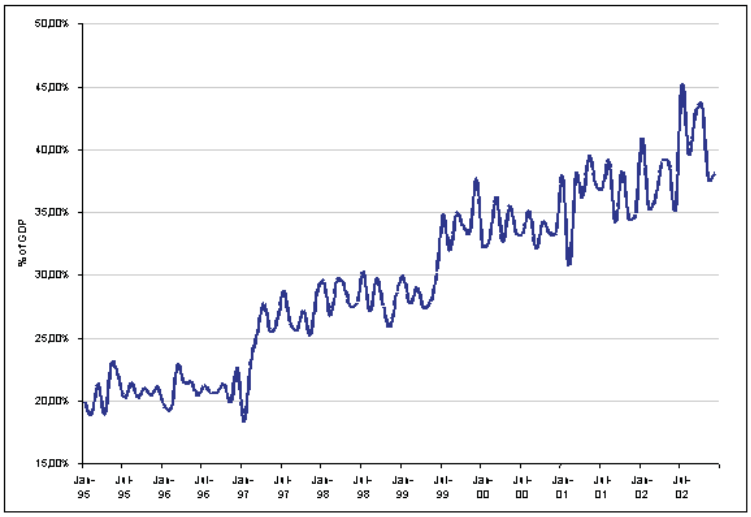
OBS: Asterisks denote the Most Adequate Model
OBS2: Standard Errors in Parentheses

As we can see, out of the five selected models, only Model 3 is the most appropriate under two of the model-selection criteria (AIC and BCC). In addition, it is rather close to the winners under both C-df and BIC, and has all variables relevant at 5%. Therefore, we considered Model 3 as the most appropriate to represent behavior of the Shadow Economy in Brazil.

From the results above, considering the effects of causal variables on the latent variable, we find that the CPMF had a positive effect, contrary to what its proponents said at the time of its inception. As regards the effects os other variables, we found the variable $TRADE_t$increases the size of the shadow economy. This is contrary to what we found in the previous literature on the subject for developed countries, since this literature points out that greater openness reduces economic regulation and the size of the Shadow Economy. However, this finding is consistent with those presented by Ribeiro e Bugarin (2003).We also find that the $DESEMP_t$ variable had a positive effect on the size of the Shadow Economy, which was consistent to the international literature on the subject. And finally, it was found that the $RTRIB_t$ variable had a positive effect on the size of the Shadow Economy.

With these values for the coefficients, with the 21.3% share of GDP obtained at the end of the first quarter of 1995, from the results of the previous section, we can build a series for the size and evolution of the Shadow Economy from this point on.The Figure below presents the results:

Figure 2: The Size of the Shadow Economy as a % of GDP



Based on the previous figure, the trajectory of evolution of the share of the Shadow Economy in Brazil goes upward. Except for both the years 1995 (with 20.71% of GDP on average) and 1996 (with 20.96% of GDP), the results results grew continuously. In 1997 there was a significant leap from the average 20% of GDP in previous years to 25.69% of GDP. This growth has continued. In 2002, the HE reached an average level of 39.40% of GDP.

From the results presented, we could conclude that the CPMF does, in fact, increase the size of the Shadow Economy. However, further analysis is required to cracaterize fully the relationship between the CPMF and the Shadow Economy as a percentage of GDP. This is the aim of the next section.

4 The relationship between CPMF and the Shadow Economy

Our goal in this section is to determine how much Total Tax Revenue (the $RTRIB_t$ variable used in the MIMIC estimate) and the CPMF, considered separately, in isolation does affect the size of Shadow Economy. To attain this goal, we will develop some models capable of explaining the relationship. Some initial points can be made about this relationship, in view of the international evidence on the topic:

- In general, the relationship between the size of Shadow Economy and these tax variables obeys a nonlinear relationship, that is, the marginal effect of a tax load increase on the size of the Shadow Economy should not be constant.
- The relationship between Shadow Economy size and these variables should contemplate a point above which the effect of a tax revenue increase on its size is null.

With these points in mind, several formal relationships were tested. In particular, the following distributions were tested, based on the seminal work of Giles and Caragata (1999) for the New Zealand economy:

Logistic Distribution:

$$\frac{SE_t}{PIB_t} = \frac{1}{1 + e^{\frac{\frac{RTRIB_t}{PIB_t} - \alpha}{\beta}}} + \gamma \left(\frac{\Delta PIB_t}{PIB_t} \right) + \varepsilon_t$$

Normal Cumulative Distribution[10]:

$$\frac{SE_t}{PIB_t} = \Psi \left(\frac{\frac{RTRIB_t}{PIB_t} - \alpha}{\beta} \right) + \gamma \left(\frac{\Delta PIB_t}{PIB_t} \right) + \varepsilon_t$$

Extreme Values Distribution[11]:

$$\left(\frac{SE_t}{PIB_t} \right) = 1 - EXP \left[-EXP \left(\frac{\frac{RTRIB_t}{PIB_t} - \alpha}{\beta} \right) \right] + \gamma \left(\frac{\Delta PIB_t}{PIB_t} \right) + \varepsilon_t$$

Gompertz Distribution:

$$\left(\frac{SE_t}{PIB_t} \right) = \alpha^{\beta^{\frac{RTRIB_t}{PIB_t}}} + \gamma \left(\frac{\Delta PIB_t}{PIB_t} \right) + \varepsilon_t$$

The definitions of the variables are the same as in the previous section, and the $\left(\frac{SE_t}{PIB_t} \right)$ the size of the Shadow Economy in terms of GDP. Let us now proceed to analyze these relationships. Initially, we analyze the presence of unit roots in each of the three series. To this end, we perform Augmented Dickey-Fuller tests with constant and trend, which indicated the stationarity of each one. The next table presents the final results after estimation of all nonlinear relationships between the Size of the Shadow Economy and Tax Revenue:

Table 7: Relationship Between $RTRIB_t/PIB_t$ and the Shadow Economy

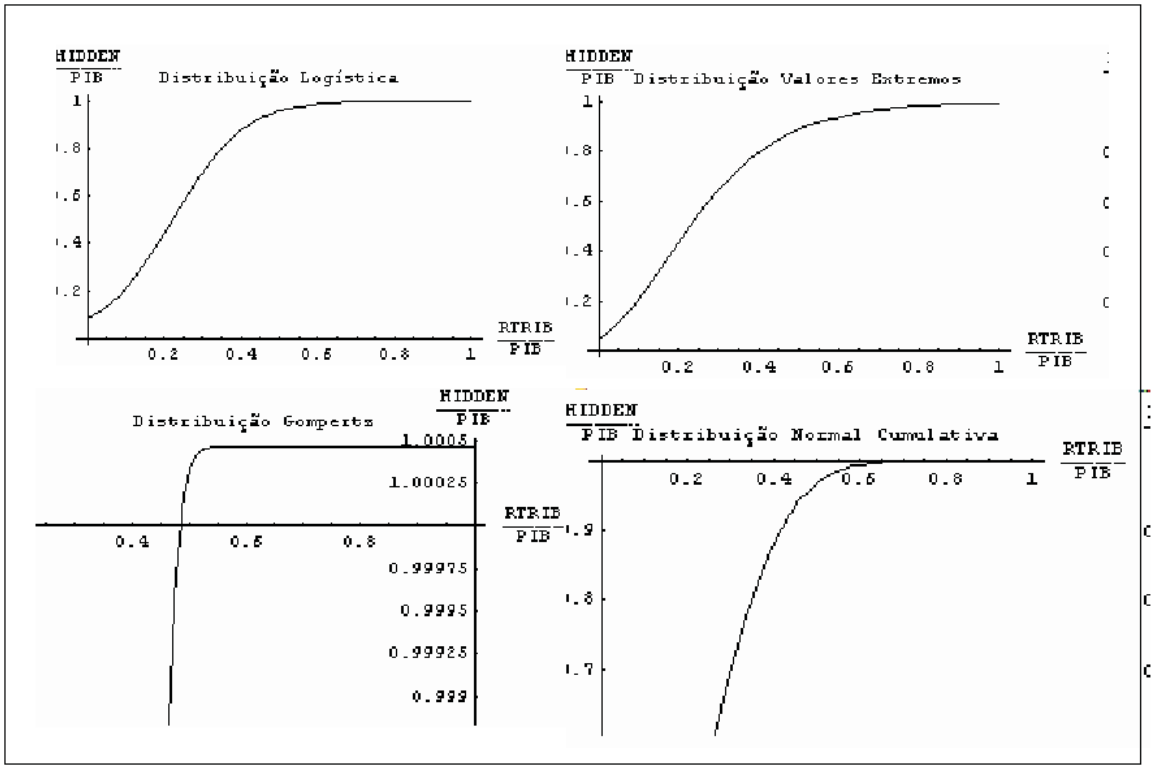
	Logistic	**Normal Cumulative**	**Extreme Value**	**Gompertz**
α	0.2210	0.2223	0.2587	0.0468
	(0.006)	(0.006)	(0.009)	(0.011)
β	0.0913	0.1502	0.1111	0.0015
	(0.007)	(0.012)	(0.009)	(0.001)
γ	0.2332	0.2351	0.2288	0.2393
	(0.094)	(0.094)	(0.095)	(0.093)
Adj. R^2	0.6349	0.6375	0.6285	0.6437
DW	0.7293	0.7368	0.7084	0.7569

Note: Standard Errors in Parentheses

All variables are significant at 5%. The next figure depicts the shape of rhe relationship assumed under the models above:

[10] In the specification below, Ψ denotes the normal cumulative distribution.
[11] EXP denotes the exponentiation operator. That is, $EXP(x) = e^x$.

Figure 3: Relationship Between $RTRIB_t/PIB_t$ and SE_t/PIB_t



We can see that, for all of the relationships analyzed, for an $RTRIB_t/PIB_t$ value over 60%, the share of Shadow Economy in GDP is close to 100%. For comparison purposes, the average level in the period for the $RTRIB_t/PIB_t$ variable is 15%, that is, the variable would have to quadruple for its influence to reach its upper limit - a size equivalent to that of the measured economy. In addition, we find that out of the four possible relationships stipulated between Total Tax Revenue and the Shadow Economy, two - Normal Cumulative and Gompertz - behave very differently from the other two - Logistic and Extreme Values. In the case of the latter two, the behavior of the relationship between Shadow Economy size and Total Tax Revenue volume is much smoother than in the case of the former two. Therefore, we could conclude that the Logistic and the Extreme Values Distributions better describe the relationship between these variables.

The same tests were performed using the revenue from the CPMF in lieu of Tax Revenue. The results are as follows:

Table 8: Relationship Between $CPMF_t$ and the Hidden Economy

	Logistic	**Normal Cumulative**	**Extreme Value**	**Gompertz**
α	0.0260	0.0265	0.0334	0.2718
	(0.001)	(0.001)	(0.001)	(0.008)
β	0.0198	0.0331	0.0233	0.0000
	(0.001)	(0.001)	(0.001)	(0.000)*
γ	0.0537	0.0544	0.0522	-0.0502
	(0.058)*	(0.059)*	(0.058)*	(0.100)*
Adjusted R^2	0.8556	0.8503	0.8554	0.5607
DW	0.5628	0.5501	0.5852	0.2337

Note: Standard Errors in Parentheses
Note: Starred Values Non-significant

We find that the effect of the CPMF Revenue on the size of the Shadow Economy is also positive, although general results are not as good as those obtained with Total Tax Revenue. This is due to the fact that some of the coefficients used in estimating distributions are not significant at 5%, as can be seen from the starred results in the previous table.

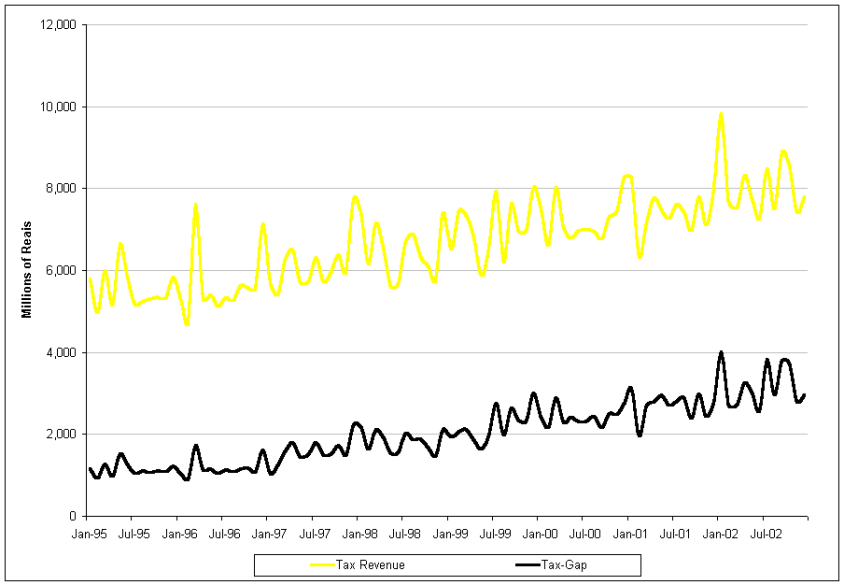
5 TAX-GAP and the CPMF's contribution to this process.

The loss of potential revenue, or TAX-GAP, is the share of potential revenue lost as a result of the presence of the Shadow Economy. The formula used to compute it, as found in Giles (1998a), is:

$$TAX - GAP_t = RTRIB_t \times \left(\frac{SE_t}{PIB_t} \right) \tag{5}$$

Two results will be presented. The first one computes the TAX-GAP based on Total Tax Revenue and the second one computes it based on CPMF. We chose this path in order to observe the CMPF's isolated effect in and determine its contribution to the process. Let us begin with the calculation using Total Tax Revenue.

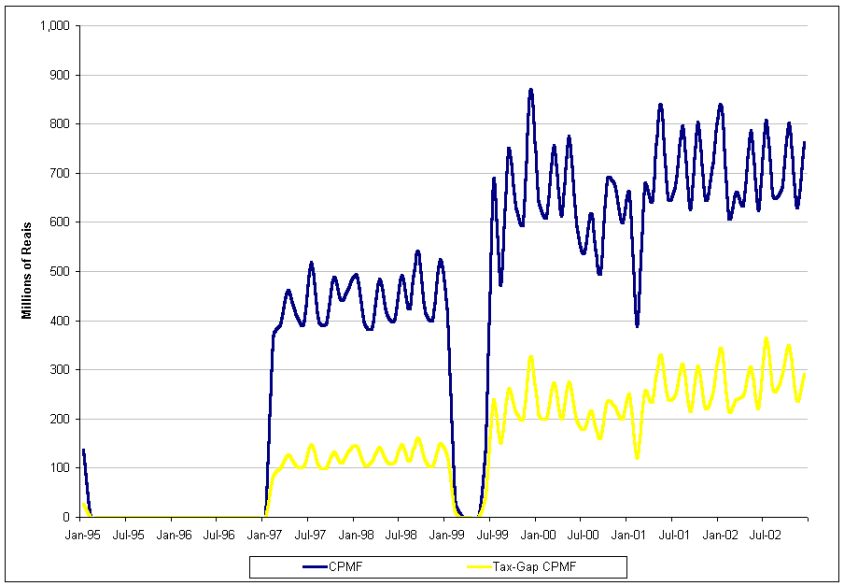
Figure 4: Tax Gap and Total Tax Revenue



From January 1995 to December 2002, out of an average total revenue of 6,704 million Reais, the presence of the Shadow Economy caused the loss of 2,046 million Reais. The trend is increasing.

The next graphic shows results for the CPMF alone:

Figure 5: CPMF and TAX-GAP of the Contribution



The CPMF revenue in the Figure above reflects the description in Appendix A. The IPMF (regarded here as CPMF revenue because of its similar characteristics), despite having stopped being collected in December 1994, still generated revenues in January 1995. This is why some collection is found and captured on the beginning of

the graphic. January 1997 marks the beginning of the collection of CPMF proper, which was interrupted for a few months in 1999 until Congress renewed its force (the CPMF is a temporary contribution that is periodically renewed at the Executive's request).

Considering the series from the first month of CPMF collection (Feb/97) until is near interruption in February 1999, the CPMF's TAX-GAP was responsible for an average 6.67% of the Total TAX-GAP. Soon after this interruption, from June 1999 until the end of 2002, this average rose to 8.74% of the total. Since both TAX-GAP series sport growth, this is yet another indication that the CPMF has been operating perversely.

6 Conclusions

In order to generate funds for the Public Health system in Brazil, a Social Security Contribution was enacted during the first months of 1994. This contribution had a characteristic which set it apart from the former initiatives - its collection at every financial transaction. Due to this reason this contribution was called Provisional Contribution on Financial Transactions, or CPMF. In order to ascertain the true effects of this measure, the present research was carried out. This task presented some difficulties, the most important of all being the measurement of the Shadow Economy. In order to establish a point to calibrate the series obtained with the MIMIC approach, we started by carrying out a procedure similar to the one used by this author, estimating an Error Correction Model (ECM).

The final step on the construction of a series representing the evolution of the Shadow Economy is to apply the Latent Variable Modelling approach - MIMIC Model. Out of all the possible model estimates through MIMIC, the one that was most appropriate under at least two criteria (AIC and BCC) and very close to the winners under both C-df and BIC, was the one that included the CPMF among causal variables. Mainly after 1997, when the CPMF entered into force (remember that there was IPMF revenue until January 1995; the IPMF was a tax that preceded the CPMF and had the same characteristics), the Shadow Economy grew yearly as a share of GDP (from an average 20% level in 1995 and 1996 to something close to 39% in 2002).

However, further analysis is required to cracaterize fully the relationship between the CPMF and the Shadow Economy as a percentage of GDP. This is the aim of the fourth section. To attain this goal, we will develop some nonlinear models capable of explaining the relationship. As regards the results, observing the Logistic, Normal, Extreme Values and Gompertz distributions, the maximum Shadow Economy value occurs when the total government revenue is at 60% of GDP.

The final step in the analysis was to ascertain the potential revenue loss due to the Shadow Economy. The results indicate the presence of the Shadow Economy caused an average loss in terms of revenue of around 2,046 Millions of Reais between January 1995 and December 2002. The share of the TAX-GAP computed with the CPMF in isolation was also found to increase as compared to the TAX-GAP computed with Total Tax Revenue. CPMF was responsible for an average 6.67% from entry into force until the brief interruption, in February 1999. Soon after this interruption, in June 1999, the average including data until December 2002 rose to 8.74% of the total. Concluding, the results of this paper provides evidence that the CPMF has had an influence on the growth of the Shadow Economy, and not the other way, as was assumed in arguments used by Legislative members at the time of implementation.

References

[1] Arvate,P.R. and da Silva, M.F.G. (2004). "Constitutional change and the federal government revenue administration in Brazil between 1982 and 2001". *Article submitted to the Constitutional Policy Economy Journal.*

[2] Bhattacharyya, D.K.(1990). "An econometric method of estimating the "hidden economy", United Kingdom (1960-1984): estimates and tests". *Economic Journal* 100.

[3] Cagan, P. (1958). "The demand for currency relative to the total money supply". *Journal of Political Economy*, vol.66.

[4] Barjada (1999) "Estimates of the Underground Economy in Australia" *The Economic Record*, vol 75, No 235, December.

[5] Barjada (2002) "How Reliable are Estimates of the Underground Economy?" *Economics Bulletin*, Vol. 3, No. 14 pp. 1-11.

[6] Barjada, C. and Schnieder, F. (2003). "The size and Development of shadow economies in the Asia-Pacific." *Paper presented at the Public Choice Meeting -Nashville/USA.*

[7] Bhattacharyya, D.K. (2004). "On the use of the hidden economy estimates". *Public Choice* 118: 169-181.

[8] Feige, E.L. (1979). "How big is the irregular economy?" *Challenge*, vol 22.

[9] Frey,B.S and Weck-Hannemann, H. (1984). "The hidden economy as an "unobserved" variable. *EER*, 26.

[10] Giles, D.E. (1997). "The hidden economy and tax-gap in New Zealand: a latent variable analysis". *Department of Economics, University of Victoria.*

[11] Giles, D.E. (1997a). "Causality between the measured and underground economies in New Zealand". *Applied Economics Letters*, vol.4.

[12] Giles, D.E. (1998). "Measuring the hidden economy: implications for econometric modeling". *Department of Economics, University of Victoria.*

[13] Giles, D.E. (1998a). "The underground economy: minimizing the size of government". *Department of Economics, University of Victoria.*

[14] Giles, D.E. (1999). "Modeling the hidden economy and the tax gap in New Zealand". *Department of Economics, University of Victoria.*

[15] Hamilton, J. D. (1994). *Times Series Analysis*. Princeton University Press.

[16] Jöreskog, K. G. and Goldberger, A. (1975) "Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable". *Journal of the American Statistical Association*, Vol. 70, nº351, pages 631-639.

[17] Lima, B.M.F. de (1985). "Criptoeconomia ou Hidden Economy". *IBRE/Editora da FGV*. Estudos Especiais No. 5.

[18] Loayza, N.A.(1997) "The economics of the informal sector: a simple model and some empirical evidence from Latin America". *Policy Research Working Paper*, 1727. World Bank.

[19] Lippert, O.and Walker, M. (eds) (1997). *The underground economy: Global evidences of its size and impact*, Vancouver, B.C. The Frazer Institute.

[20] Ribeiro, R.N. e Bugarin, M.N.S. (2003). "Fatores Determinantes e Evolução da Economia Submersa no Brasil". *Estudos Econômicos*, vol 33, No. 2.

[21] Schneider, F (2000) "Illegal Activities and the Generation of Value Added: Size, Causes and Measurement of Shadow Economies" *Bulletin on Narcotics*, Vol. LII, Nos. 1 e 2.

[22] Schneider, F and Klnglmair, R. (2004) " Shadow Economies around the world: What do we know? Paper presented at the Public Choice Meeting 2004 em Baltimore (USA).

[23] Soares (1994). Murilo Rodrigues da Cunha. Evasão fiscal: teoria do comportamento dos agentes públicos e privados. *Thesis presented at FGV/EAESP.*

[24] Tanzi,V. (1980). Underground Economy and tax evasion in the United States: estimates and implications. IMF.

[25] Tanzi, V. (1983) "The Cash Economy in the United States: Annual Estimates, 1930-1980" *International Monetary Fund Staff Papers* 30, 283-08.

[26] Watson, M. W. and Engle, R. F. (1983) "Alternative Algorithms for the Estimation of Dynamic Factor, MIMIC and Varying Coefficient Regression Models" *Journal of Econometrics* 23 385-400.

A Brief history of the origins, evolution and collection of the CPMF

The roots of the CPMF lie in the IPMF (Imposto Provisório sobre Movimentações Financeiras - Temporary Tax on Financial Transactions). The IPMF had the same format as the CPMF (levied on any financial transaction) and its initial rate was .25%. It remained in force between January and December 1994 and resulted in revenues, according to Receita Federal data, of 1.01% of GDP. The IPMF, according to governmental discourse at the time, was a necessary evil in view of the challenges of fiscal adjustment (the expression "necessary evil" is the government's express acknowledgement of the fact that the IPMF's design would greatly increase economic inefficiency)

In January 1997 the government starts collecting the CPMF, after a two-year interruption in IPMF collection. There was a good justification to resume collection: the federal government needed to cover expenditures in the Health area, whose budget was in turmoil after the 1988 Constitution (the New Constitution increased Social Security expenditures by raising the amount and number of benefits granted). Since Health and Social Security shared the same sources of funds prior to the constitutional reform, the promulgation of the New Constitution left fewer resources for the Health area. The solution was to create a new source of funds. Since the IPMF did not have a specific purpose, the tax was converted into a contribution. Its rate dropped from .25% to .20% (between Jun/17/99 and Jun/17/00 the rate rose again to .38%; between Jun/18/00 and Mar/17/01 it dropped back to .30%; and returned to .38% since Mar/18/01).

It is important to note, in view of the goal of this article, that the CPMF bill (passed in 1996 and with force in 1997) expressly prohibited the Receita Federal from using collection information generated by the CPMF to formalize tax credits against tax-evaders. The reason for this was an attempt to preserve taxpayers' confidentiality. With the surprising aggregated tax-evasion estimates prepared by the Receita Federal in late 2001, the government and Congress were convinced that the Receita Federal should use the data, imputing non-existent tax credits against tax-evaders. Likewise, even prior to the official use of this information by the Receita Federal, taxpayers feared that collection data might be put to this use, during both the IPMF and the CPMF phases. A summary of CPMF revenues as share of the GDP can be seen in the next table:

Table 9: Evolution of Receipts of CPMF

Years	**1997**	**1998**	**1999**	**2000**	**2001**
CPMF/PIB	0.79%	0.89%	0.83%	1.34%	1.45%

B Descriptive Statistics - ECM

	ΔPIB_t	$SEMCART_t$	$PMPP_t$	$RDPOC_t$	$CPMF_t$	$TRADE_t$	$RTRIB_t$	$DESEMP_t$
Mean	0.100	119.289	274.847	105.554	105.096	17.116	147.397	7.009
Median	0.373	116.682	282.268	103.912	121.469	15.699	149.118	6.999
Maximum	12.628	136.123	409.593	126.194	225.253	29.220	216.185	9.029
Minimum	-12.362	103.376	141.918	83.443	-1.036	9.877	104.418	4.250
Std. Dev.	5.429	9.896	59.000	9.557	75.922	4.260	22.944	1.315
Skewness	-0.158	0.178	-0.337	0.202	-0.334	0.690	0.223	-0.272
Kurtosis	2.441	1.685	2.842	2.295	1.653	2.666	2.491	1.897
Jarque-Bera	1.650	7.427	1.922	2.639	9.040	8.073	1.829	6.047
Probability	0.438	0.024	0.383	0.267	0.011	0.018	0.401	0.049
Sum	9.618	11451.760	26385.320	10133.180	10089.220	1643.142	14150.120	672.873
Sum Sq. Dev.	2800.181	9304.086	330698.500	8676.531	547592.500	1723.852	50011.890	164.215
Observations	96	96	96	96	96	96	96	96

C Descriptive Statistics - MIMIC Model

	CDB_t	PIB_t	$PMPP_t$	$RTRIB_t$	IGP_t
Mean	4.39	229734.00	17786.23	0.06	0.14
Median	1.64	233974.10	17177.92	0.05	0.03
Maximum	47.80	364306.20	42351.42	0.09	1.94
Minimum	1.17	18881.63	719.91	0.04	-0.01
Std. Dev.	10.10	75593.38	9394.03	0.01	0.44
Skewness	3.91	-0.58	0.42	0.46	3.86
Kurtosis	16.46	3.40	2.83	1.98	16.09
Jarque-Bera	373.49	2.34	1.13	2.91	355.80
Probability	0.00	0.31	0.57	0.23	0.00
Sum	162.37	8500160.00	658090.40	2.12	5.25
Sum Sq. Dev.	3669.80	2.06E+11	3.18E+09	0.01	6.92
Observations	37	37	37	37	37