Philippine Institute for Development Studies
Working Paper 81-13

SURVEY AND REVIEW OF FORECASTING MODELS IN INTERNAL GOVERNMENT REVENUES

Rosario G. Manasan

March 1981
Philippine Institute for Deve[illegible]ment Studies

SUMMARY DESCRIPTION OF PROJECT:

The need for good tax revenue forecasts cannot be overemphasized. It is a major input in budget planning and programming as well as a necessary guidepost against which to assess the tax collection effort of the government.

The purpose of this paper is to review and evaluate the existing works on tax revenue forecasting in the Philippines. This endeavor represents the initial step in an attempt to develop a tax revenue forecasting model for the Bureau of Internal Revenue. The review is intended to provide insights to the particular problems involved in the formulation and estimation of a revenue forecasting model and thus, set the stage for the development of such a model.

Earlier studies made on revenue forecasting are as follows: (1) the Kintanar-Mijares work; (2) the Jurado-Encarnacion government sector sub-model; (3) the Diokno public sector model; and (4) the various Bureau of Internal Revenue models.

The Kintanar-Mijares work suggests a procedure for forecasting corporate and individual income tax at a fairly high level of disaggregation. However, it has limiting assumptions arising from its use of a particular year sample data. Its forecasts for the other kind of taxes are aggregative in nature and are based on a simple time-trend.

Likewise, the Jurado and Encarnacion is rather aggregative using only six types of taxes. Nevertheless, this study was among the first to relate different tax groupings to different explanatory variables.

The Diokno study considered three types of taxes and used one variable, GNP, to explain the variations in said taxes. In this sense, the work can be said to be very limited for BIR purposes.

Two of the BIR's forecasting approaches, the compound growth rate technique and the time-trend analysis, assume that tax collections are influenced only by time. Furthermore, the compound growth rate technique utilized only bench-mark figures of the data base. The Gompertz curve time-trend analysis, on the other hand, has the tendency to give conservative forecasts in the long-run. The tax elasticity approach – the third used by the BIR-has the disadvantage of relating tax receipts to only one explanatory variable. Its advantage over the time-trend analysis, however, lies in the fact that the explanatory variable most appropriate to the type of tax is used. In this last approach, four types of taxes are considered.

The above models leave much to be desired in terms of the level of disaggregation and the use of explanatory variables. For future modelling works, there appear five basic research directions, viz., the spatial or regional; categorical or into particular taxes; technical or methodological; behavioral, and the macroeconomic view, which will all be explained in the conclusion.

TECHNICAL REPORT: (see attached copies)

PROBLEMS ENCOUNTERED AND RECOMMENDATIONS

As typical in many situations, there is lack of documentation on forecasting done by the BIR prior to recent years. With the turnover of people, it becomes difficult to know, much more to validate, the forecasting methods employed in the past. This suggests the full documentation of models that may be developed in the future - to include the data base. This should eliminate re-gathering of data already established and make the model implementable by Bureau personnel.

Submitted by:

MA. ROSARIO GREGORIO-MANASAN
Principal Investigator

____

Date

T A B L E O F C O N T E N T S

SURVEY AND REVIEW OF FORECASTING MODELS ON INTERNAL GOVERNMENT REVENUES

---

1. INTRODUCTION

A lot of well-intentioned and feasible government projects have been stalled even in preliminary processes, not entirely because of the failure of the government to meet the cost of these programs per se, but because of its inability to forecast accurately the revenues that it can generate. Revenue forecasting then becomes a key factor in economic, social and political planning. For the Bureau of Internal Revenue, it is a tool to assess its tax collection efforts.

The purpose of this paper is to review and evaluate the existing works on tax revenue forecasting in the Philippines. This endeavor represents the initial step in an attempt to develop a tax revenue forecasting model for the Bureau of Internal Revenue. The review is intended to provide insights to the particular problems involved in the formulation and estimation of a revenue forecasting model and thus, set the stage for the development of such a model.

In sections 4, 5, 6 and 7, we review and synthesize the works that have been conducted in this area, namely: (1) the various Bureau of Internal Revenue (BIR) methods and models for forecasting; (2) the Mintanar-Mijares work; (3) the Jurado-Encarnacion government sector sub-model; and (4) the Diokno public sector model. Each section will include a summary of the functional specifications together with the respective definition of the variables used in each model. The different functions specified by each of models reviewed were re-estimated using data for the period 1961-1979. It should be noted that all studied (except those of the BIR's) were originally estimated using the data of the sixties. It is to be expected that parameter values derived from these exercises will not hold for the more recent years which are of greater interest to the Bureau, in so far as revenue forecasting is concerned. "Historical" simulation over the new estimation period as was conducted to help assess the forecating capability of the various model reviewed. The root mean square per cent error, RMSE%, were in evaluating the said models. These statistic are discussed more fully in Section 2.

Section 3 provides a brief description of the data used in the analysis.

2. METHODOLOGY

"In the case of the single equation regression model, there exists a set of statistic test (R^2, F-test, t-test, etc.) that be used to judge the sign in a statistical sense of the model and its individual estimated coefficients... The model's evaluation must also depend on the purpose for which the model was built.

not R

A model designed to test a specific hypothesis or to measure some elasticity should have high t statistics[1]. A model designed for forecasting purposes should have as small a standard error of forecast as possible while a model designed to test a specific hypothesis or to measure some elasticity should have high t-statistics. Thus, the R^2 and the RMSE% were used to assess the different tax revenue models reviewed. A greater weight is given to RMSE% since the present study is more concerned with the forecasting abilities of the said models.

The coefficient of determination, R^2, defined as the ratio of the regression sum of squares $[RSS = (\hat{Y}_i - \bar{Y})^2]$ to the total sum squares $[TSS = (Y_i - Y)^2]$,

1 Pindych and Rubinfeld, p. 315

measures the proportion of the variation in the dependent variable which is "explained" by the regression equation or by the variations in the independent variables. The coefficinet of determination ranges from zero to one. A computed R^2 close to unity is indicative of good ~~to~~ fit. This implies that the variations in the endogenous variable can be largely explained by the variations in its determinants.

The RMSE% evaluates the "fit" of the individual variables in a simulation context. First, a historical simulation is performed using the model. Then, the resulting or simulated figures of the endogenous variables are examined on how closely each tracks its corresponding historical data series. Finally, the RMSE% is computed. It is defined as follows:

$$\text{RMSE\%} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\frac{Y_t^s - Y_t^a}{Y_t^a}\right)^2}$$

, where Y_t^s is the simulated value while Y_t^a is the actual value of the dependent

mean

Thus, the RMSE% measures the deviation of the simulated variables from its actual path, in percentage terms. As a rule of thumb, an RMSE% value equivalent to 10 per cent or less ~~or less than 10%~~ is acceptable, 2/ ~~i.e., an equation with such an RMSE% value, for forecasting.~~ [1]

2/ The NBER set 10% RMSE% as the standard for evaluating the forecasting capability of an equation.

"A single-equation regressin model can have significant t-statistics and a high R^2 and still forecast very badly period after period. This might result from a structural change (in the economy) occuring during the forecasts period and not explained by the model. Good forecasts, on the other hand, may come from regression models which have relatively low R^2s and one or more insignificant regression coefficients. This may happen because there is very little variation in the dependent variable, so that although it is not explained well by the model, it is easy to forecast."[3]

3. DATA

The data for the tax bases were obtained from the National Income Accounts Statistics of NEDA and annual money wage rate figures attainable ~~for~~ from the Central Bank Statistical Bulletin. The tax collections data came from the BIR Statistical Division. Tax data were available in fical year ~~sries~~ series from 1961-1974 and in calendar year series from 1974-1979. Conversion of tax data from fiscal Year (FY) to Calendar Year (CY) follows the scheme given below:

Given:	FY 2 = July (Year 1) - June (Year 2)
Subtract:	July-December (Year 1)
To get:	January-June (Year 2)
Add:	July-December (Year 2)
=	CY 2 = Jan. (Year 2) - Dec. (Year 2)

3) Pindyck & Rubinfeld p. 161

To Illustrate: Income Tax Collection of

Given:	FY 1964 (July 1963-June 1964)	= ₱419.7 M
Subtract:	July-December 1963	= 145.24 M
To get:	January-June 1964	₱274.46 M
Add:	July-December 1964	₱173.7 M
=	CY 1964 (Jan.-Dec. 1964)	₱448.16 M

[1]The National Bureau of Economic Research set the 10% RMSE as the standard for evaluating the forecasting capability of an equation.

[2]Pindych & Rubinfeld, p. 161.

4. THE VARIOUS BIR TECHNIQUES MODELS

Over the years, the BIR has used various techniques to forecast tax collections namely: the compound growth rate technique, time-trend analysis and the tax elasticity approach. The first two of these approaches, though longer employed. Thus, simulations were not derived for these methods, but only for the elasticity approach.

4.1 The Compound Growth Rate Technique

The initial effort to predict tax collections was rather simplistic and crude. The compound growth rate of actual collections in the preceding ten years were computed fro income tax, business tax, specific tax and other taxes separately. The growth rates thus obtained were multiplied with the present years tax collections to yield the forecasts for the next year. This technique was used for the 1977-1978 projections.

4.2 Time-Trend Analysis

In order to forecast tax revenue collections in 1979, time-trend analysis was employed. The revenue collections for the period 1970-1979 was plotted against time to get a rough idea of the general trend. Based on this, the rate of increase was found to be non-linear. Tax collection was increasing at a decreasing rate and a Gompertz curve was fitted to the data. Income tax, business tax and other taxes were separately estimated with this curve. However, forecasts for the specific tax were based on a simple linear regression using consumption of commodities as explanatory variable.

This method shares the same basic drawback of the compound growth rate technique, i.e. it considers time as the sole factor explaining tax receipts. Its advantage over the latter lies in its ability to consider turning points in the pattern of growth.

4.3. Tax Elasticity Approach

Forecasts of tax collection for 1980 onwards were based on the tax elasticity approach. The predicted increment (in absolute terms) in tax revenues in any given year is the product of tax collections in the previous year, the elasticity of the tax with respect to its base and the projected growth rate of the base. The different tax categories were related with different variables reflecting the appropriate tax base. Thus, the individual income tax was related with personal income; the corporate income tax was related to corporate income; the specific tax was related to value of manufacturing domestic product; license, business and other taxes treated as one, were related to industrial and services domestic product. The functional form used was that of the power curve, i.e., $Y = bX^m$, where m is the slope, b is the y-intercept, when the function is estimated by the double-log transformation. The parameter m is then interpreted as the elasticity of the tax with respect to the explanatory variable considered.

The re-estimated equations using the BIR's tax elasticity approach are as follows:

$$\text{Individual Income Tax} = 0.000056\ (\text{Personal Income})^{1.482} \qquad (1)$$

$r^2 = 0.938$ RMSE % = 13.7%

$$\text{Corporate Income Tax} = 0.3336\ (\text{Corporate Income})^{1.0174} \qquad (2)$$

$r^2 = 0.960$ RMSE % = 18.0%

$$\text{Specific Tax} = 0.08179\ (\text{Manufacturing Domestic Product})^{1.16} \qquad (3)$$

$r^2 = 0.9696$ RMSE % = 17.1%

$$\text{License, Business and other Taxes} = 0.004989\ (\text{Industrial Service Domestic Product})^{1.133} \qquad (4)$$

$r^2 = 0.992$ RMSE % = 7.7%

Table 1 presents the actual and simulated values of the various tax categories for the period 1961 - 1979 while Figures 1, 2, 3 and 4 provide a pictorial view of the same.

TABLE 1

Actual and Simulated Values of Various Tax Categories
Using BIR's Elasticity Approach, 1961-1979

YEAR	INDIVIDUAL INCOME TAX		CORPORATE TAX		SPECIFIC TAX		LICENSE, BUSINESS & OTHER TAXES	
	Actual	Simulated	Actual	Simulated	Actual	Simulated	Actual	Simulated
1961	79.1	70.5	174.4	171.5	259.8	252.4	134.9	142.8
1962	74.9	82.2	211.2	260.2	287.0	282.9	169.8	160.6
1963	94.0	100.8	259.4	307.8	329.3	329.0	190.2	185.2
1964	126.6	113.4	271.6	325.0	371.3	341.9	219.9	208.0
1965	140.5	131.4	298.6	274.7	378.3	362.1	226.2	232.4
1966	124.5	147.1	288.4	374.3	438.1	397.3	266.6	257.6
1967	216.6	174.8	395.3	416.8	481.2	431.0	287.5	275.9
1968	175.1	200.1	527.9	540.2	544.5	470.0	332.2	298.4
1969	252.9	235.2	602.6	509.8	553.7	511.3	360.0	328.6
1970	286.7	291.1	763.9	725.5	579.7	665.15	373.6	411.5
1971	371.9	379.2	956.7	606.9	645.9	790.7	422.2	456.9
1972	508.3	452.0	867.1	648.6	663.4	924.6	464.2	516.3
1973	520.6	621.6	1857.7	1560.8	828.3	1217.3	599.4	657.8
1974	788.9	1007.1	2391.3	1905.0	1623.3	1681.1	940.4	950.1
1975	1119.7	1212.0	1954.1	2030.2	1935.1	1944.9	1033.7	1126.9
1976	1483.0	1511.1	2222.3	2479.7	2515.1	2212.3	1315.9	1375.5
1977	2473.3	1928.5	2048.6	2212.7	3030.1	2663.7	1597.7	1658.9
1978	2548.9	2397.1	2641.6	3008.3	3614.2	3065.8	2041.2	1969.4
1979	3185.0	3151.7	2872.3	3563.6	4072.9	3795.9	3068.3	2537.8

FIGURE 1

INDIVIDUAL INCOME TAX
(ACTUAL & SIMULATED VALUES)

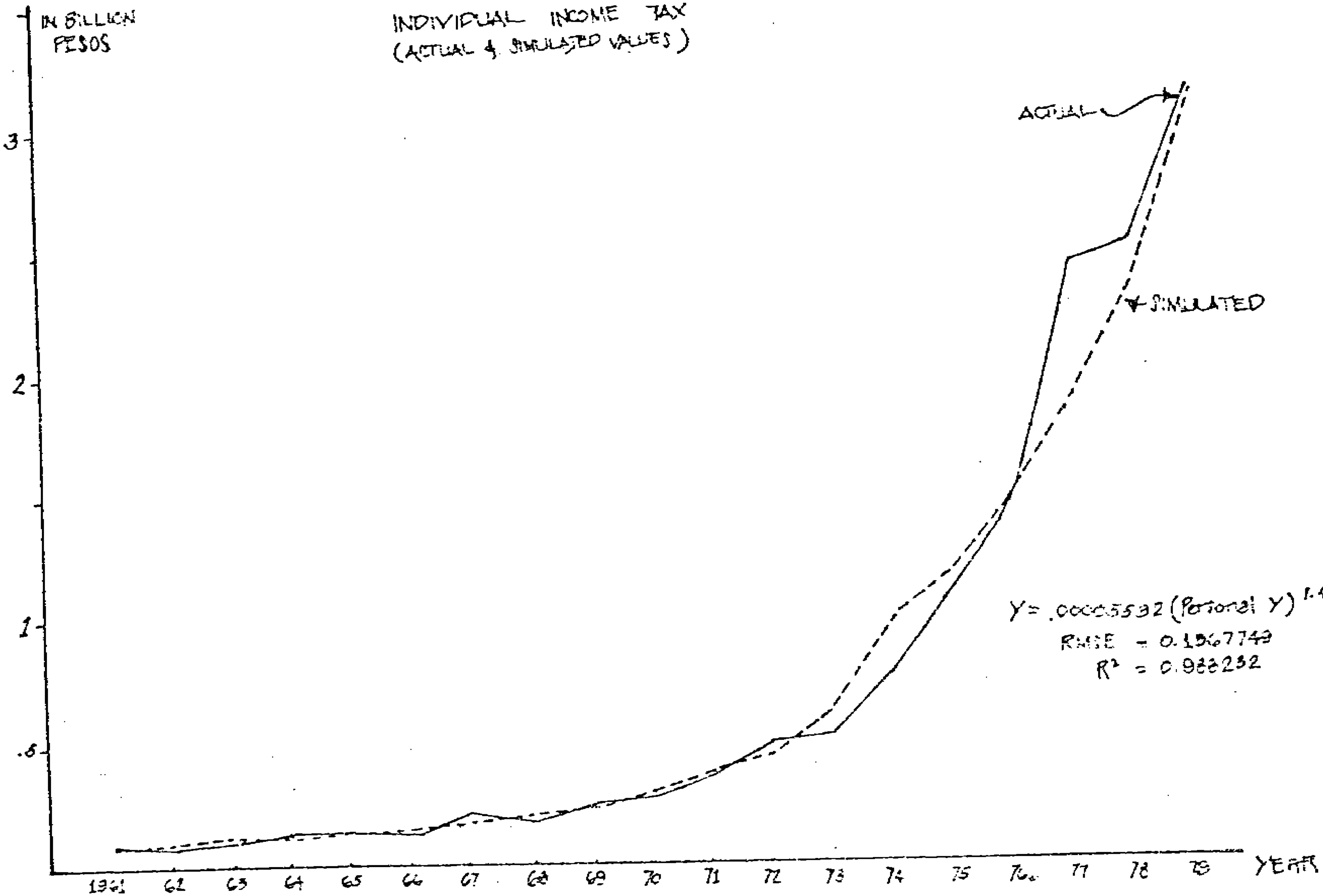


FIGURE 2

CORPORATE INCOME TAX
(ACTUAL & SIMULATED VALUES)

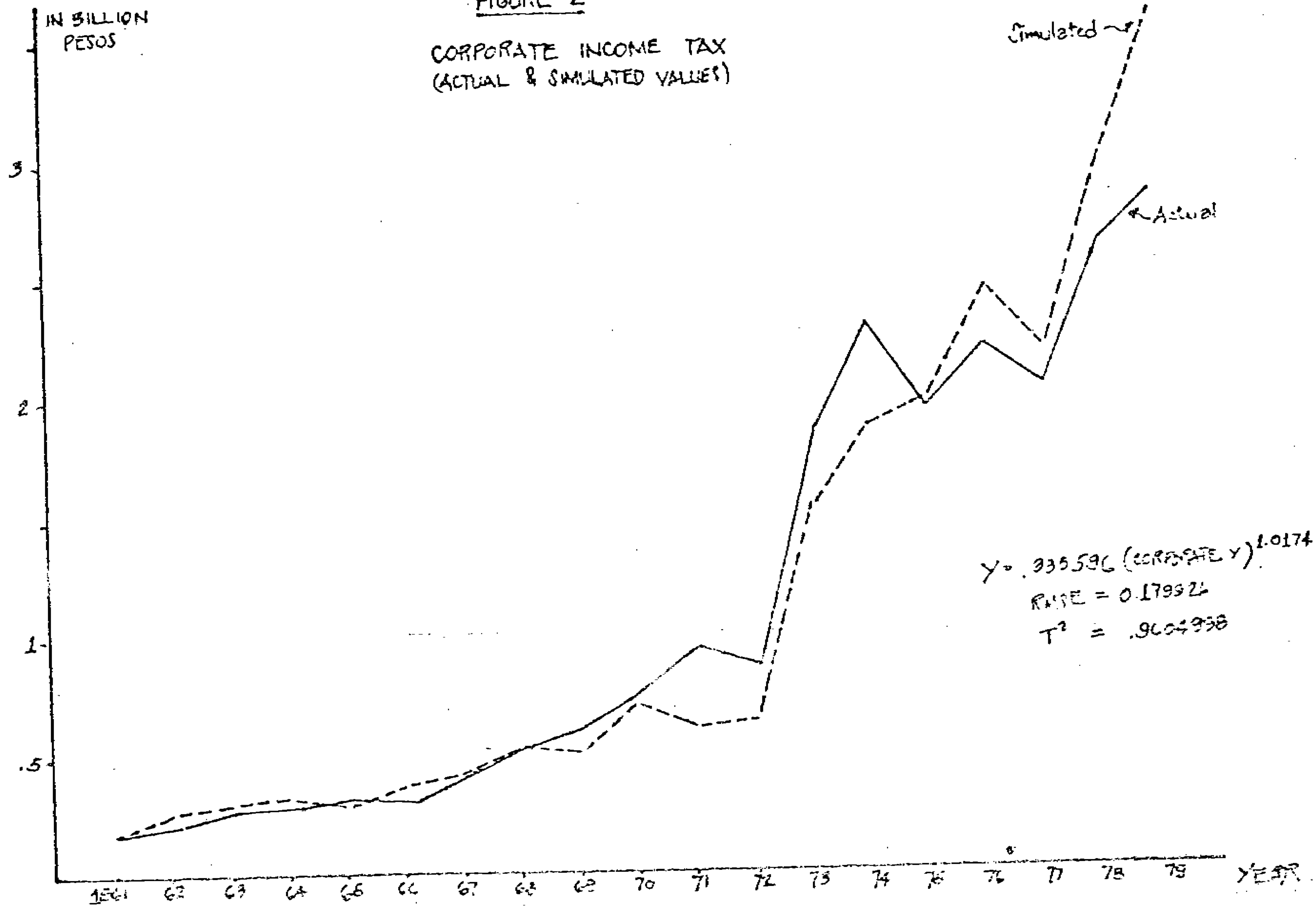


FIGURE 3

SPECIFIC TAX
(Actual & Simulated Values)

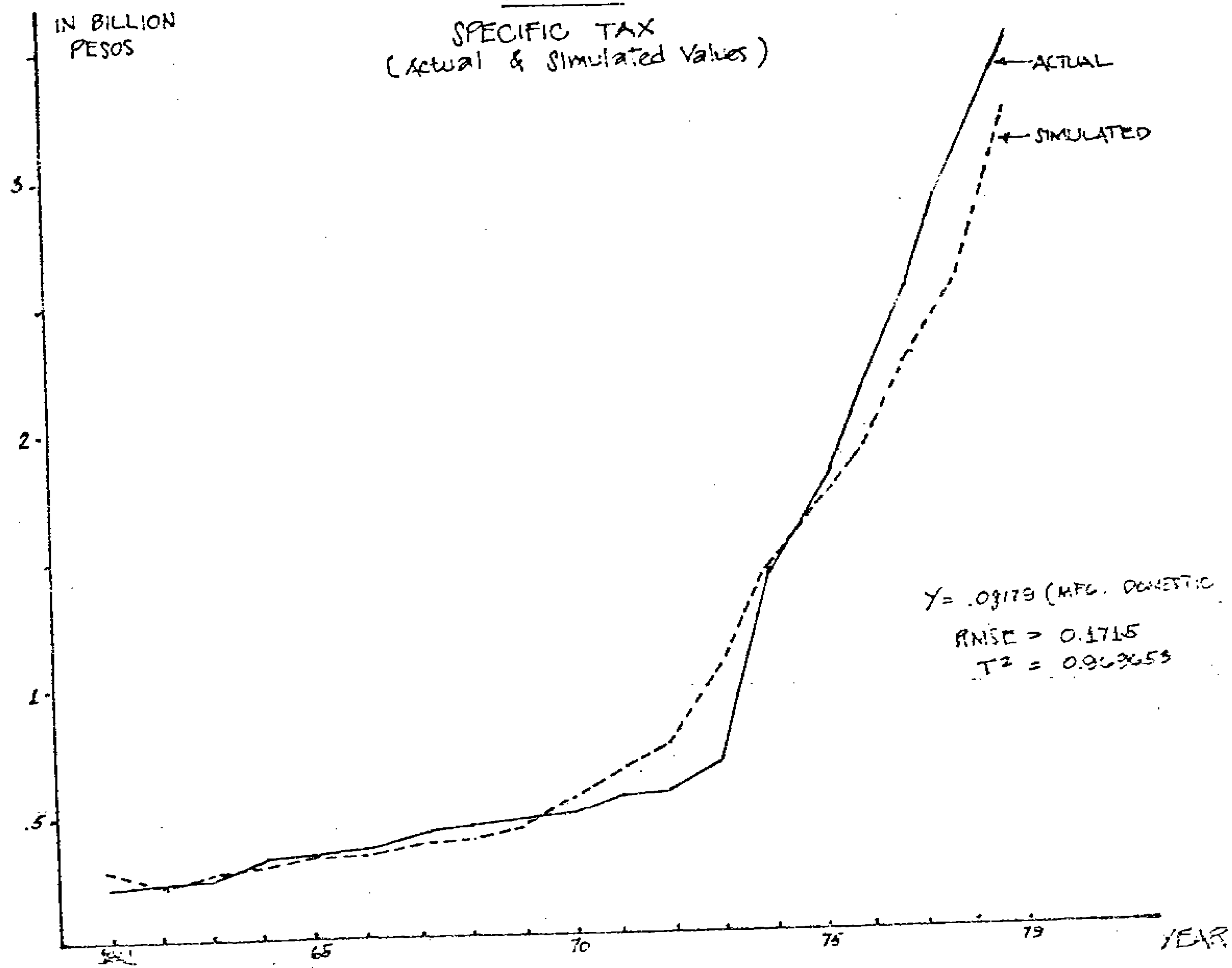
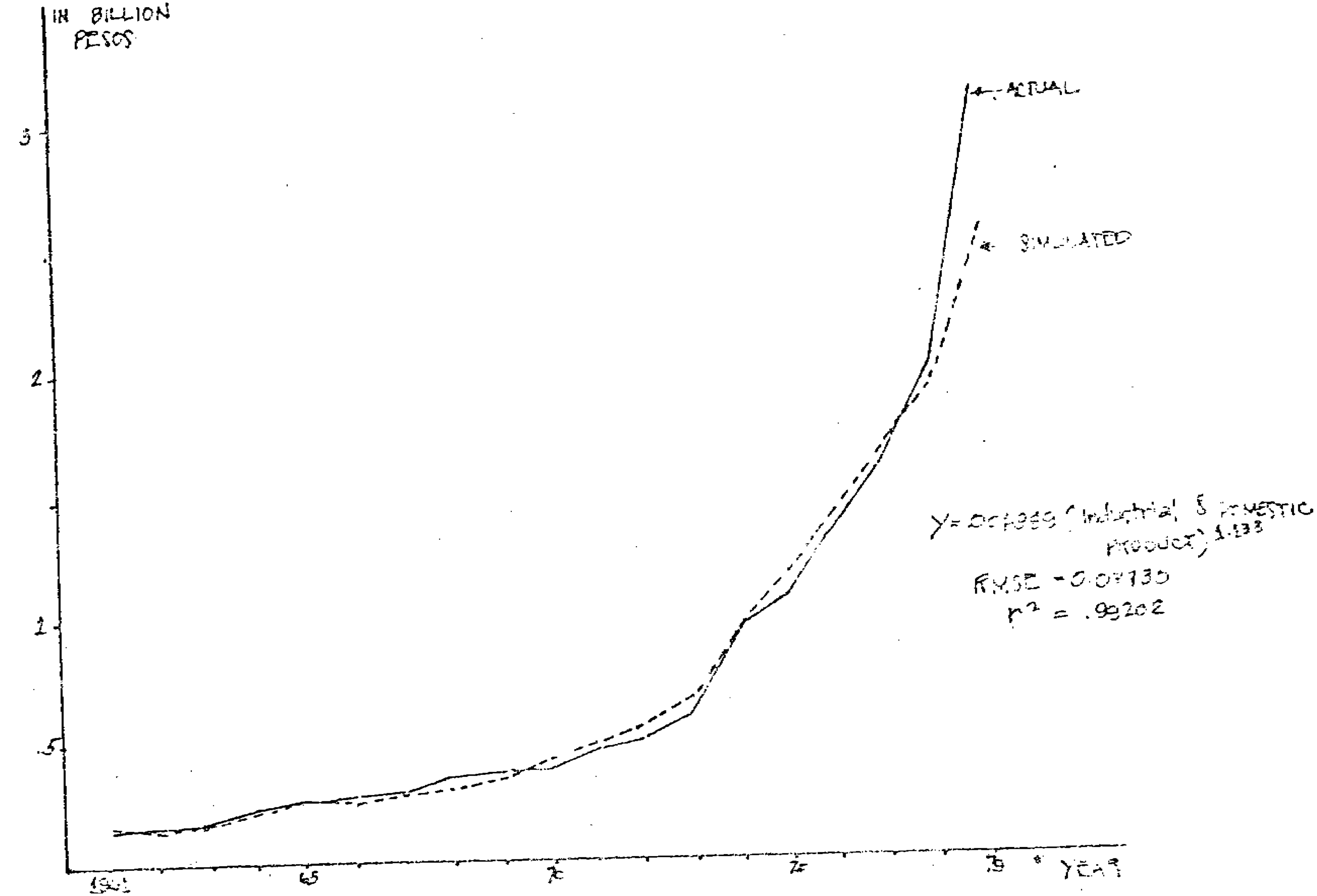


FIGURE 4

LICENSE & BUSINESS TAX
(ACTUAL & SIMULATED VALUES)



Of the four equations considered, only the simulations for license, business and other taxes resulted in a favorable root mean square percentage error of less than 10%.

5. KINTANAR-MIJARES TAX FORECASTING METHOD

Kintanar and Mijares (1965) suggested different frameworks for forecasting revenues from different tax categories. Based on these, they obtained revenue predictions for 1965-1970 using data from the earlier years.

In projecting revenue from the corporate income tax, the authors divided the work into three steps: (1) estimation of the total number of returns; (2) distributing the figure obtained in (1) to the various income taxe brackets; and (3) estimation of the mean net taxable income for the different tax brackets. The total number of returns was projected using a simple time-trend. This number was then apportioned to the different income brackets by assuming that percentage distribution follows the pattern exhibited by a subset of the total number of returns in 1962. The mean net taxable income for the various tax brackets was computed from the same subset mentioned above to come up with the middle estimate. Low and high estimates were obtained by utilizing the fiducial limits of the 95 per cent confidence interval of the mean estimate. These low, middle and high estimates of the mean taxable income of the different tax brackets were multiplied with the appropriate tax rates to arrive at estimates of the average tax assessment for each taxable bracket. The product of the tax assessment valuation and the projected number of returns in each income bracket yields the revenue forecasts for the corporate income tax by bracket.

The procedure followed in projecting the revenue from the individual income tax is similar to that of the corporate income tax. First, the total number of individual returns was projected by fitting a time trend to 1959-1962 data. Second, the percentage distribution of 1960 was applied to these projected values to yield the number of returns in each tax bracket. Third, estimates of the mean taxable income in each tax bracket was obtained from a sample of the 1960 data. Fourth, the appropriate tax rates were applied to the estimates obtained in the previous step to come up with the mean tax assessment for each bracket. Again, low and high valuation of the tax assessment were computed using the 95 per cent fiducial limits of the confidence interval for the mean tax assessment. Finally, the mean tax assessment for each income tax bracket was multiplied with the corresponding number of returns to estimate the total tax assessment. This five-step procedure is done for returns from both married and single individuals; then the total tax assessment for single individuals was added to that for married individuals to arrive at the aggregate estimate for these two groups for each income bracket.

Since the computation for the individual income tax and corporate income tax simulations over the estimation period 1961-1979 would require tedious effort, we simply estimated the RMSE % for the ex ante forecasts made by Kintanar and Mijares for 1963-1965. These estimates are presented in Table 2.

TABLE 2

Medium Projections from Kintanar

(PM)

	1963	1964	RMSE %
Projected Individual Tax Assessment	111.5	124.9	
Actual Individual Tax Assessment	94.0	126.6	13.2 %
	1964	1965	
Projected Corporation Tax Assessment	299.02	322.17	
Actual Corporation Tax Assessment	271.6	298.6	9.1 %

In forecasting the various components of license, business and other taxes, the only explanatory variable considered was time. Tax collection on sale of non-essential or luxury goods as described under Section 184 of the National Internal Revenue Code (NIRC) denoted by T_L, compensating tax collections under Section 204 of the NIRC denoted by T_c and specific tax collected on sale of cigars and cigarettes denoted by T_s, were directly regressed on time. However, in forecasting the tax receipts from sales of semi-luxury and non-luxury items, their tax bases denoted by B_s and B_n respectively, were regressed on time and the projected tax bases were then multiplied by the relevant tax rates.

To arrive at B_s and B_n values, the ratios of each to the actual total sales tax were taken:

(a) $\frac{\text{Tax on Semi-Luxury}}{\text{Total Sales Tax}}$; $\frac{\text{Tax on Non-Luxury}}{\text{Total Sales Tax}}$

then multiplied by the manufacturing component of Gross Domestic Product (GDP)

(b) $\frac{\text{Tax on Semi-Luxury}}{\text{Total Sales Tax}} \times \text{GDP} = B_s$;

$\frac{\text{Tax on Non-Luxury}}{\text{Total Sales Tax}} \times \text{GDP} = B_n$

(c) $B_s \times .35$ (average of 30% and 40%, the tax rates prevalent during the estimating years) $= S_L$

$B_n \times .07 = N_L$

where S_L represents tax receipts from semi-luxury items and N_L represents tax receipts from non-luxury items.

B_s was multiplied by 35% (average of 30% and 40%, the tax rates prevalent during the estimating years); and B_n was multiplied by 7%

In re-estimating the specification of Kintanar and Mijares using more recent data, the data on sales taxes, compensating tax and specific tax on cigars and cigarettes from 1961 to 1976 were obtained from the Statistical Division of the BIR. Data for years later than 1976 were not available because the amendments made in the NIRC resulted in unspecified and questionable tax reports which have yet to be reconciled.

The resulting equations are as follows:

$$T_L = 9.37655 e^{-.2164084t} \qquad (5)$$

$$r^2 = .406 \qquad \text{RMSE \%} = 103.58$$

$$B_s = -48.2251 e^{159.7906t} \qquad (6)$$

$$r^2 = .2067 \qquad \text{RMSE \%} = 174.4\ \%$$

$$B_n = 5475.94 - 1347.94t + 204.58t^2 \qquad (7)$$

$$r^2 = .977 \qquad \text{RMSE \%} = 23.36$$

$$T_c = .14976 + .0185711t \qquad (8)$$

$$r^2 = .082 \qquad \text{RMSE \%} = 89.28$$

$$T_s = 84.728 + 21.684t \qquad (9)$$

$$r^2 = .689 \qquad \text{RMSE \%} = 20.09$$

Note the very high RMSE % for equations (5), (6) and (8) indicating the inadequacy of these specifications in forecasting.

Table 3 presents the actual and simulated values of the non-income tax categories considered by Kintanar and Mijares while a graphical picture is given in Figures 5, 6,

Actual and Simulated Values of Various Tax Categories
Based on the Specifications of Kintanar & Mijares 1961-1976

YEAR	S_L=Tax on Semi-Luxury Items		N_L=Tax on Non-Luxury Items		B_S=Tax Base of Semi-Luxury Tax Collections		B_n=Tax Base of Non-Luxury Tax Collections	
	Actual	Simulated	Actual	Simulated	Actual	Simulated	Actual	Simulated
1961	1.79	-16.8788	53.02	383.3158	75.33	-48.23	3352.59	5475.94
1962	1.29	5.7975	67.82	303.2806	68.56	16.56	3611.76	4332.58
1963	(.73)	21.8867	80.5	251.8865	(40.68)	62.53	4476.79	3598.37
1964	2.63	34.3664	89.11	229.1331	132.76	98.19	4501.68	3273.3
1965	1.37	44.5630	92.23	235.0215	72.69	127.32	4895.66	3357.4
1966	.83	53.1841	113.46	269.5504	39.09	151.96	5341.43	3850.72
1967	.96	60.6521	116.13	332.7205	48.63	173.29	5857.71	4753.1
1968	[illegible]	67.2393	139.44	424.5311	47.73	192.11	6512.95	6064.7
1969	8.16	73.1318	141.37	544.9836	383.10	208.95	6635.72	7765.48
1970	1.53	78.4622	182.88	694.0766	77.55	224.18	9261.89	9915.38
1971	1.0	83.3284	206.09	871.8108	54.80	238.08	11266.30	12454.4
1972	1.76	87.8050	213.44	1078.1855	105.77	250.87	12774.83	15402.6
1973	2.41	91.9496	257.03	1313.2021	161.21	262.71	17183.55	18760.0
1974	10.53	95.8081	418.24	1576.8592	600.44	273.74	23813.16	22526.5
1975	11.04	99.4175	388.59	1869.1575	787.81	284.05	27707.66	26702.23
1976	1.66	102.8081	472.62	2190.097	113.91	293.74	32346.48	31287.1

TABLE [illegible]

Actual and Simulated Values of Various Tax Categories
Based on the Specifications of Kintanar & Mijares 1961-1976

YEAR	T_L = Tax on Luxury Items (184)		TC = Compensating Tax		TS = Cigars and Cigarettes	
	Actual	Simulated	Actual	Simulated	Actual	Simulated
1961	2.25	9.3765	.14	.16833	143.82	106.412
1962	6.18	8.4149	.18	.17761	146.80	117.254
1963	4.30	7.5519	.15	.18691	159.74	128.096
1964	4.52	6.7774	.26	.19619	145.08	138.938
1965	3.53	6.0824	.39	.20548	115.76	149.78
1966	6.06	5.4886	.06	.21476	136.59	160.622
1967	4.93	4.8988	.38	.22405	147.69	171.464
1968	3.46	4.3964	.26	.23333	158.89	182.306
1969	6.53	3.9455	.09	.24262	167.98	193.148
1970	4.63	3.5409	.30	.25191	167.28	203.99
1971	1.76	3.1777	.14	.26119	190.11	214.832
1972	8.48	2.8518	.14	.27046	212.26	225.674
1973	5.55	2.5594	.23	.27976	231.98	236.516
1974	3.42	2.2969	.16	.02891	285.61	247.358
1975	.71	2.0613	.62	.29833	325.00	258.200
1976	1.23	1.8499	.37	.30762	466.66	269.042

FIGURE 5

TAX ON LUXURY ITEMS (SECTION 184)
(Actual & SIMULATED VALUES)

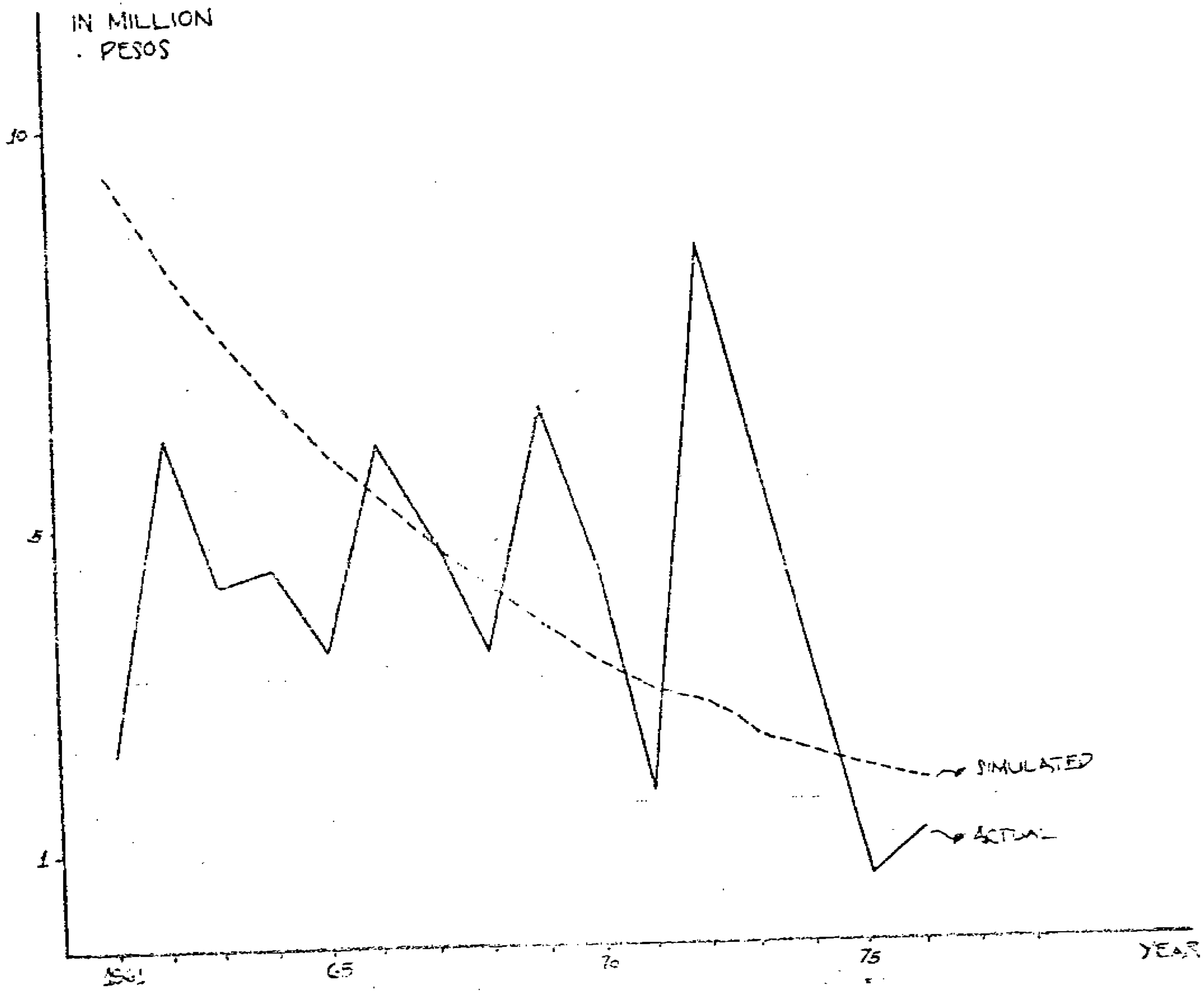


FIGURE 6

TAX ON SEMI-LUXURY ITEMS (SL)
(.35 X Bn)
(ACTUAL & SIMULATED VALUES)

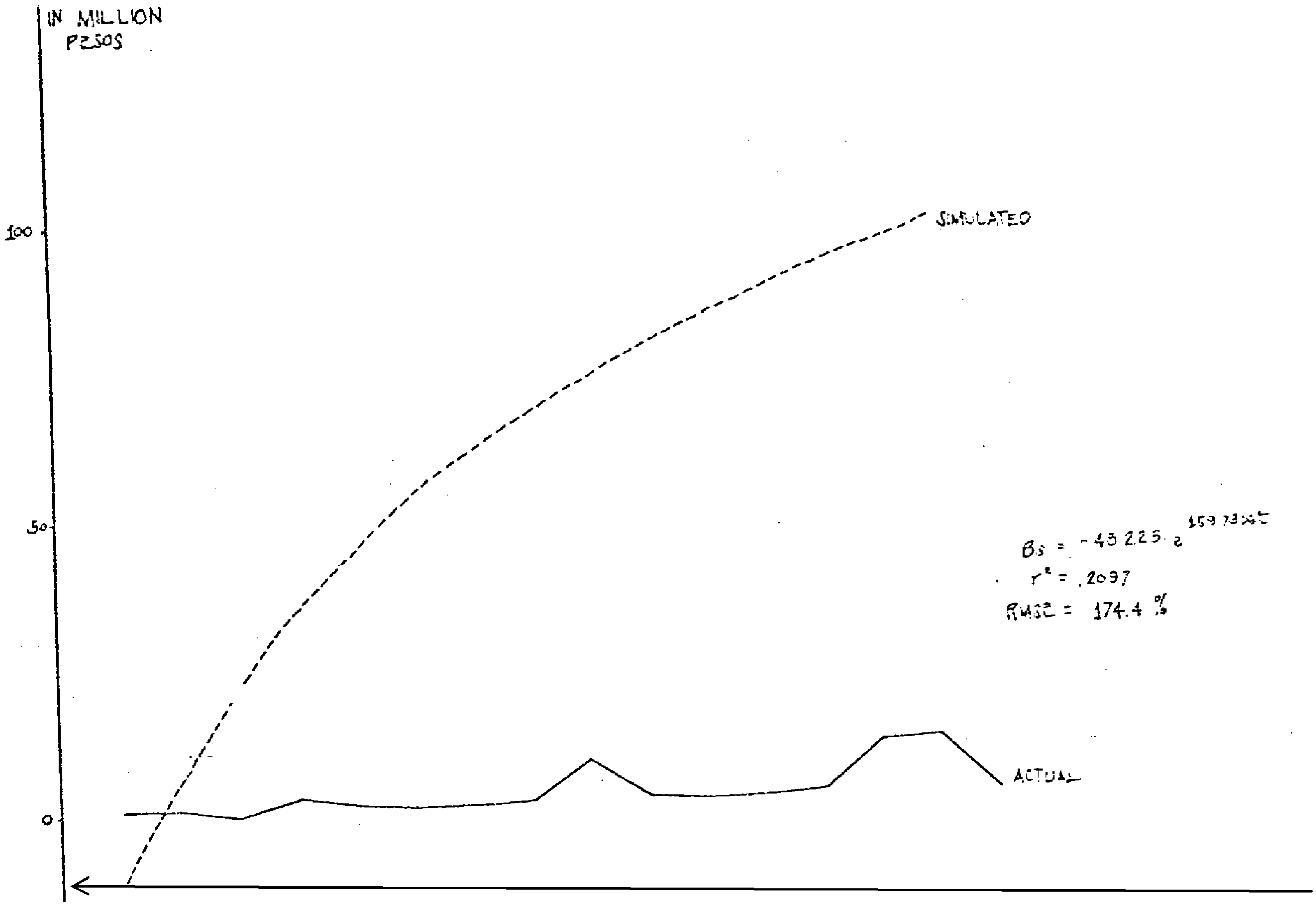


FIGURE 7

TAX ON NON-LUXURY ITEMS (NL)

(.07 X Bn)

(ACTUAL & SIMULATED VALUES)

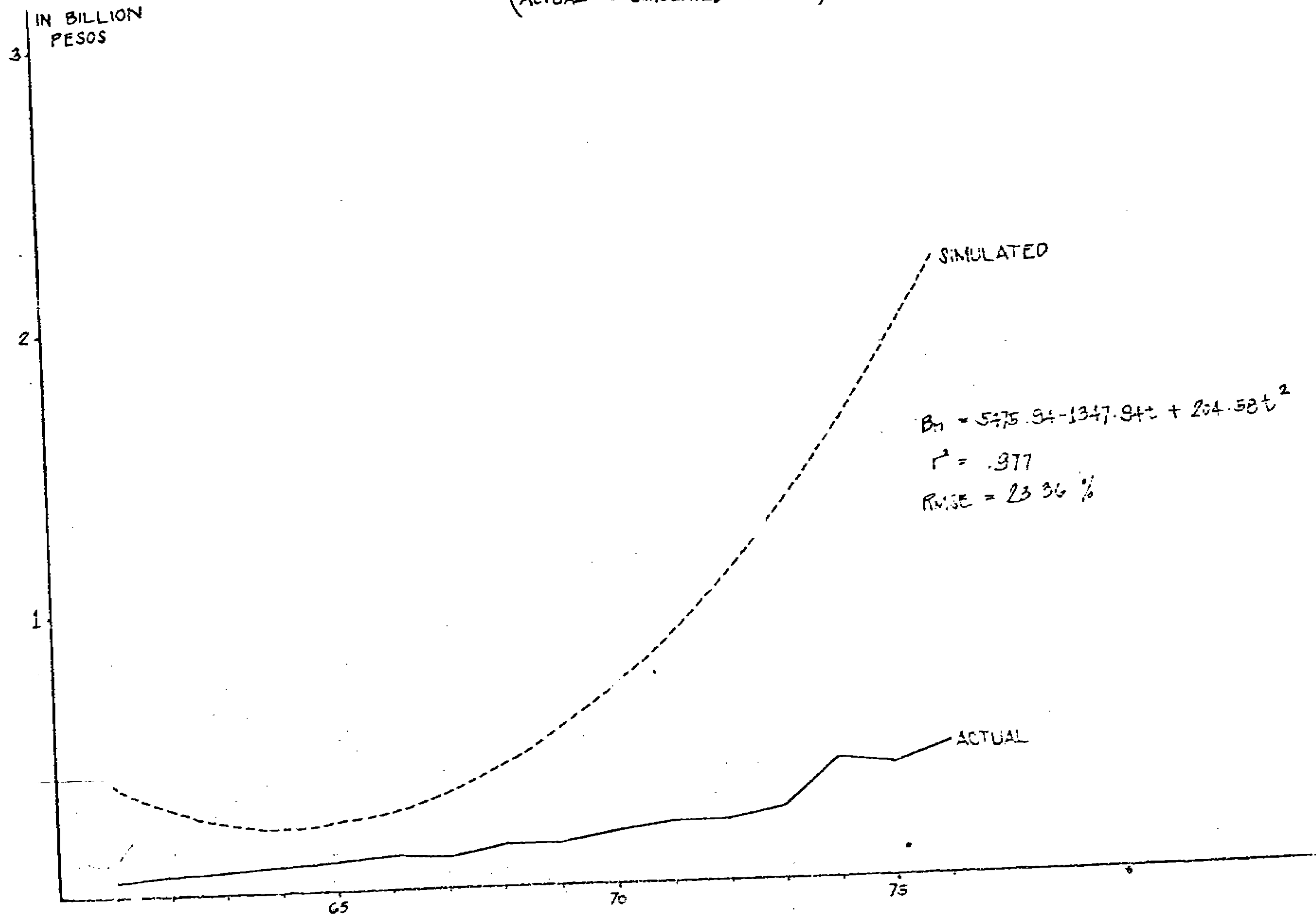


FIGURE 8

COMPENSATING TAX
(Actual & Simulated Values)

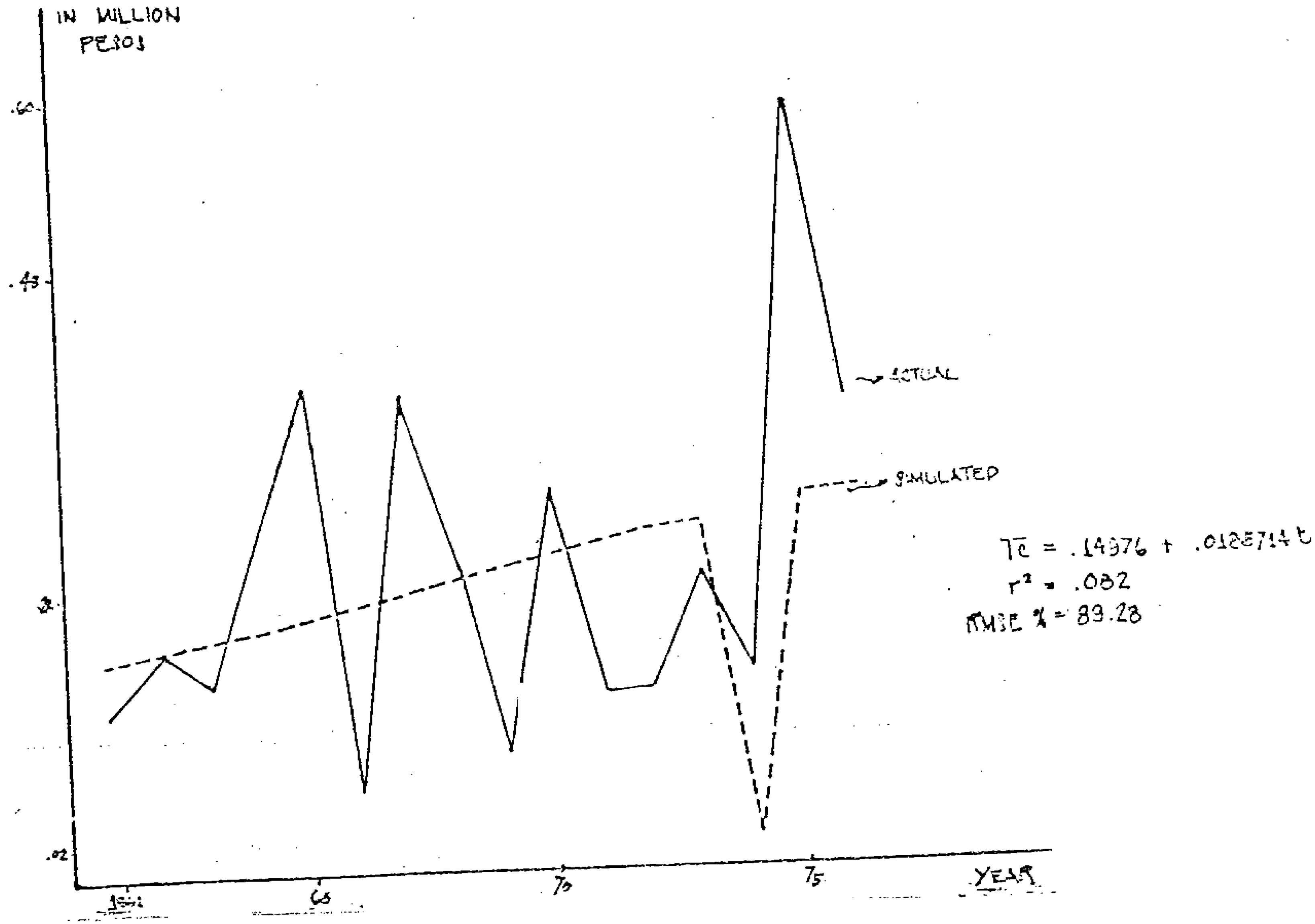
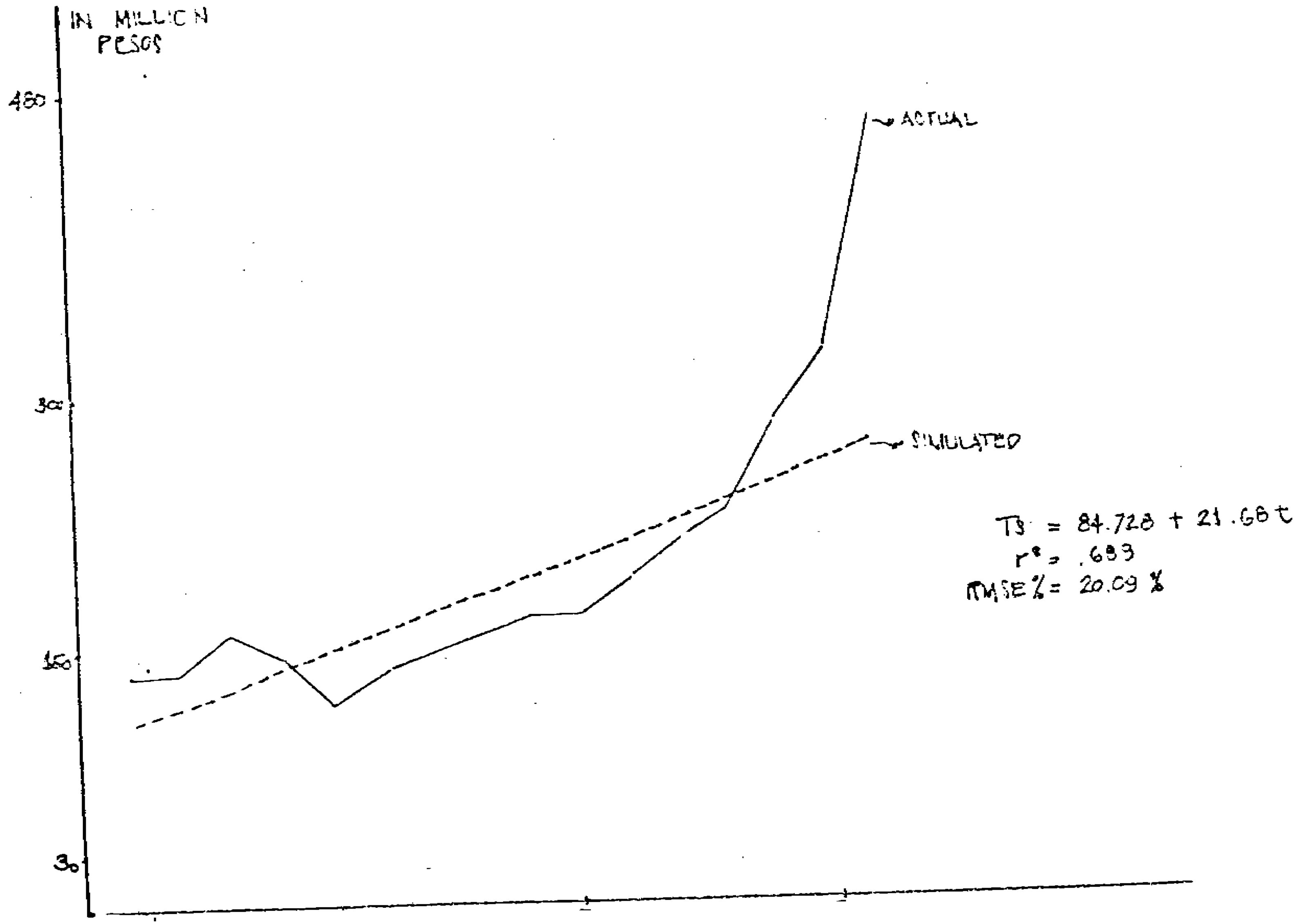


FIGURE 9

SPECIFIC TAX ON CIGARS & CIGARETTES
(Actual & Simulated Values)



7, 8 and 9

6. JURADO-ENCARNACION GOVERNMENT MODEL

Jurado and Encarnacion (1972) constructed a fiscal sector model employing ordinary least squares on data from 1955 to 1969. This model consists of eleven structural equations (ten of which are related to taxes) and two identities (one of which is tax related). The authors broke down total tax revenues into six tax categories. Revenue from corporate income tax was related with gross national product and the wage rate, as a proxy for corporate income, and a dummy variable to reflect tax rate changes. On the other hand, individual income tax receipts were regressed on personal income lagged one year. Indirect business tax collections (primarily from the sales or the percentage tax) were made a function of GNP. Tariff duties were related to the total value of imported goods and services and a variable that measures imports from the United States. Excise taxes, fees and penalties, charges on forest products, the franchise tax, wharfage and other fees were all classified under "other indirect taxes" and were explained by gross national product sold domestically and exports. Total taxes collected by the national government were related with the aggregate of all taxes received by the government as a whole. Finally, all other taxes of the National Government was obtained as a residual. The rest of the equations were devoted to explaining government expenditures and its components.

The model may be written as follows:[2]

Tbd = f(Y, W, Utr) (10)

Tbi = f (Y) (11)

Tp = f (Y_p-1) (12)

Toi = f (Y-X), x) (13)

where,

Tbd - direct income tax receipts from business enterprises which otherwise is known as corporate tax;

Tpd - direct income tax receipts from persons

Tbi - indirect tax receipts from business enterprises - sales Tax

Toi - receipts from other indirect taxes excise Tax; fees & penalties, DST, Forest Prod, franchise tax, wharfage & other fees

Y - Gross National Product

Yp - compensation of employees plus entrepreneurial and property income of persons

X - exports of goods and services

W - annual money wage rate, computed as equal to the daily wage rate of unskilled industrial workers in Manila multiplied by 250; in pesos

Utr - dummy variable for a change in tax rates;
= 1 for years beginning 1968, = 0 for years below 1968.

The model above was reestimated using data from 1961 to 1979. The resulting equations are as follows:

Tbd = 1137.10 + .716 + .618W + .800 Utr (10a)

R^2 =

[2]The import tax equation was not included in this report because so far, we are only interested in checking the models using BIR's internal data. The other identitites were not also included in the reestimation.

Note that in (10a) the regression coefficient for wages is positive as opposed to the negative coefficient that Jurado and Encarnacion previously obtained. To resolve this error, we tried regressing (Y-W) as a proxy for corporate income, against Tbd with the following output:

$$Tbd = 18.39 + .013\ (Y\text{-}W) + 415.69\ Utr$$

$$r^2 = .921 \qquad RMSE\ \% = 20.86 \qquad (10b)$$

$$Tpd = (367.91) + .025\ Y_{p-1}$$

$$r^2 = .967 \qquad RMSE\ \% = 73.38 \qquad (11a)$$

$$Tbi = (33.63) + .005\ Y$$

$$r^2 = .887 \qquad RMSE\ \% = 20.93 \qquad (12a)$$

$$Tbi = (473.12) + .0399\ (Y\text{-}X) + .028X$$

$$R^2 = .986 \qquad RMSE\ \% = 30.01 \qquad (13a)$$

It can be observed that, not a single one of the four equations resulted in a computed root mean square percentage error lower than 10%. Although the equation for other indirect taxes (13a) yields a high r^2, the turning point error must have been large so as to produce, at the same time, a high root mean percentage error.

TABLE 4

Actual and Simulated VAlues of Various Tax Categories
in Jurado and Encarnacion Model, 1961-1979

YEAR	direct taxes Tbd-from Business		direct taxes Tpd-from persons		indirect taxes Tbi-from Business		other indirect taxes Toi	
	Actual	Simulated	Actual	Simulated	Actual	Simulated	Actual	Simulated
1961	174.4	192.4119	79.1	82.00	56.46	41.0637	362.32	111.4993
1962	211.2	215.4614	74.9	55.81	75.29	50.8793	411.74	175.8013
1963	259.4	249.3363	94.0	22.73	84.09	64.5917	473.28	277.5810
1964	271.6	269.2849	126.6	30.22	96.26	72.4826	530.54	338.8348
1965	298.6	293.0850	140.5	61.42	97.13	82.4083	547.99	411.7238
1966	288.4	321.3834	124.5	109.26	120.35	94.1306	632.76	498.0428
1967	395.3	358.0433	216.6	149.78	122.02	108.9649	697.05	615.0813
1968	527.9	809.7383	175.1	207.57	143.92	124.1361	488.69	739.3149
1969	602.6	849.1649	252.9	265.14	156.06	140.1241	861.82	869.4383
1970	763.9	931.2244	286.7	335.41	189.04	173.5561	1000.46	1096.5143
1971	956.7	1028.1420	371.9	441.27	208.85	212.5146	1127.95	1395.8010
1972	867.1	1101.0082	508.3	598.79	223.68	241.9295	1228.32	1624.9543
1973	1857.7	1302.9277	520.6	712.23	264.99	241.7819	2370.41	2194.9644
1974	2391.3	1657.2966	788.9	971.68	432.19	462.3896	2885.11	3250.1155
1975	1854.1	1834.6670	1119.7	1499.66	400.34	533.4429	2287.76	3833.1834
1976	2222.3	2055.3386	1483.0	1730.01	475.51	621.1515	4345.19	4514.8479
1977	2048.6	2334.3776	2473.3	2107.82	639.94	732.0366	5679.86	5334.2782
1978	2641.6	2639.2008	2548.9	2525.36	1144.67	852.6440	6005.76	6277.2100
1979	2872.3	3163.0132	3185.0	2985.25	1467.13	1060	8351.67	7829.4916

FIGURE 10
DIRECT INCOME TAX RECEIPT FROM
BUSINESS ENTERPRISES

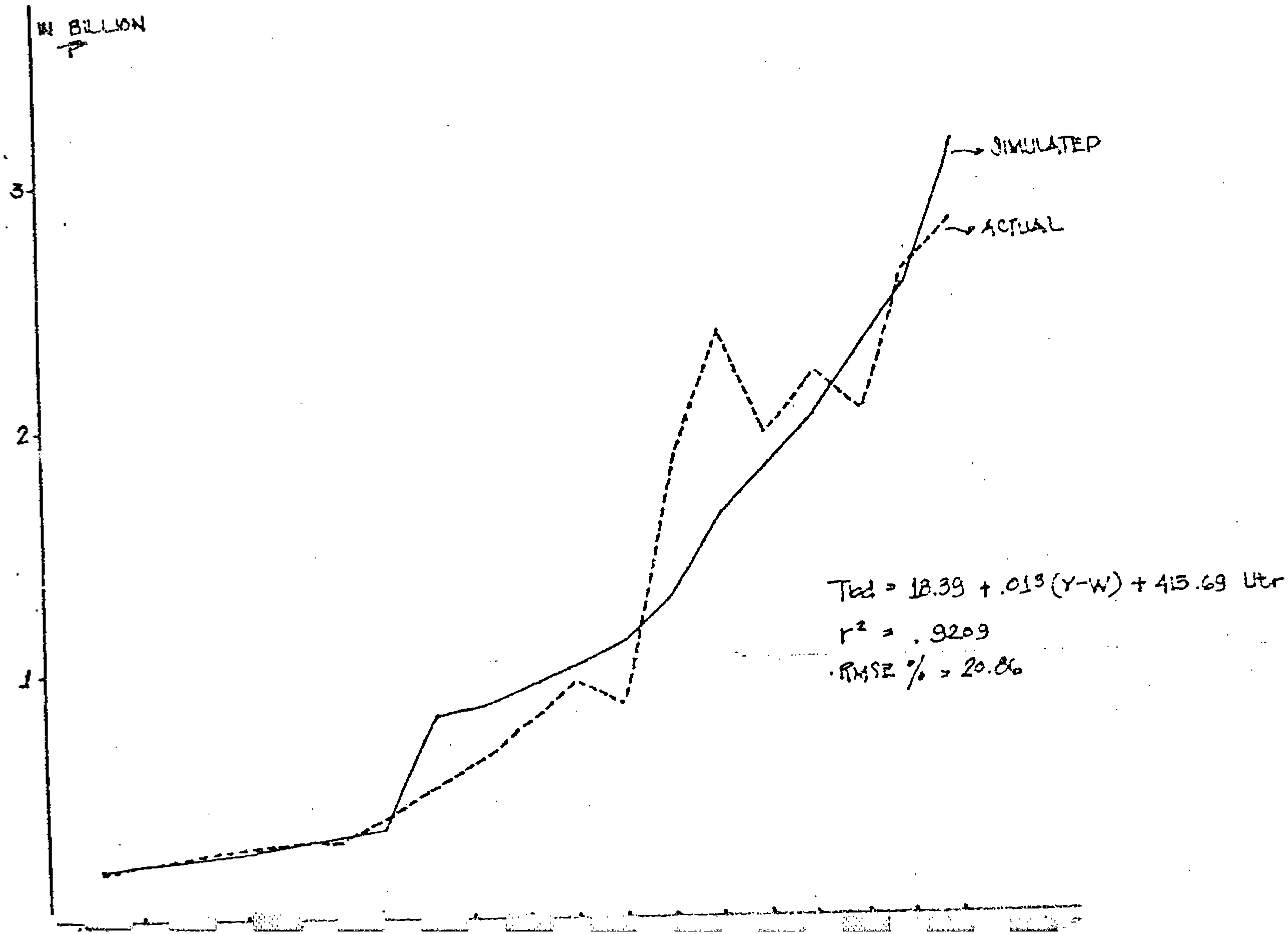


FIGURE 11

INDIVIDUAL INCOME TAX

(ACTUAL & SIMULATED VALUES)

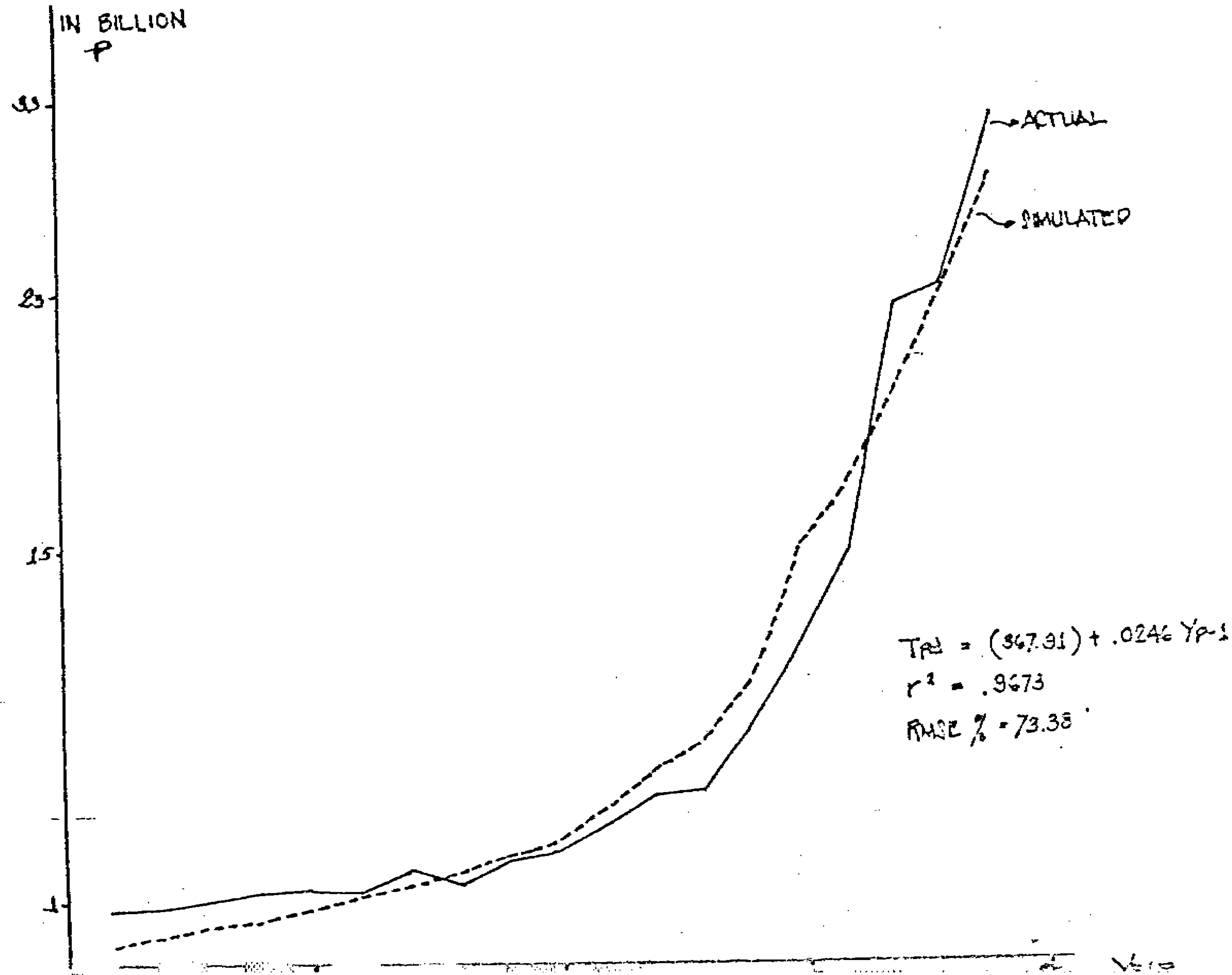


FIGURE 12

INDIRECT TAX RECEIPTS FROM BUSINESS (SALES)
(ACTUAL & SIMULATED VALUES)

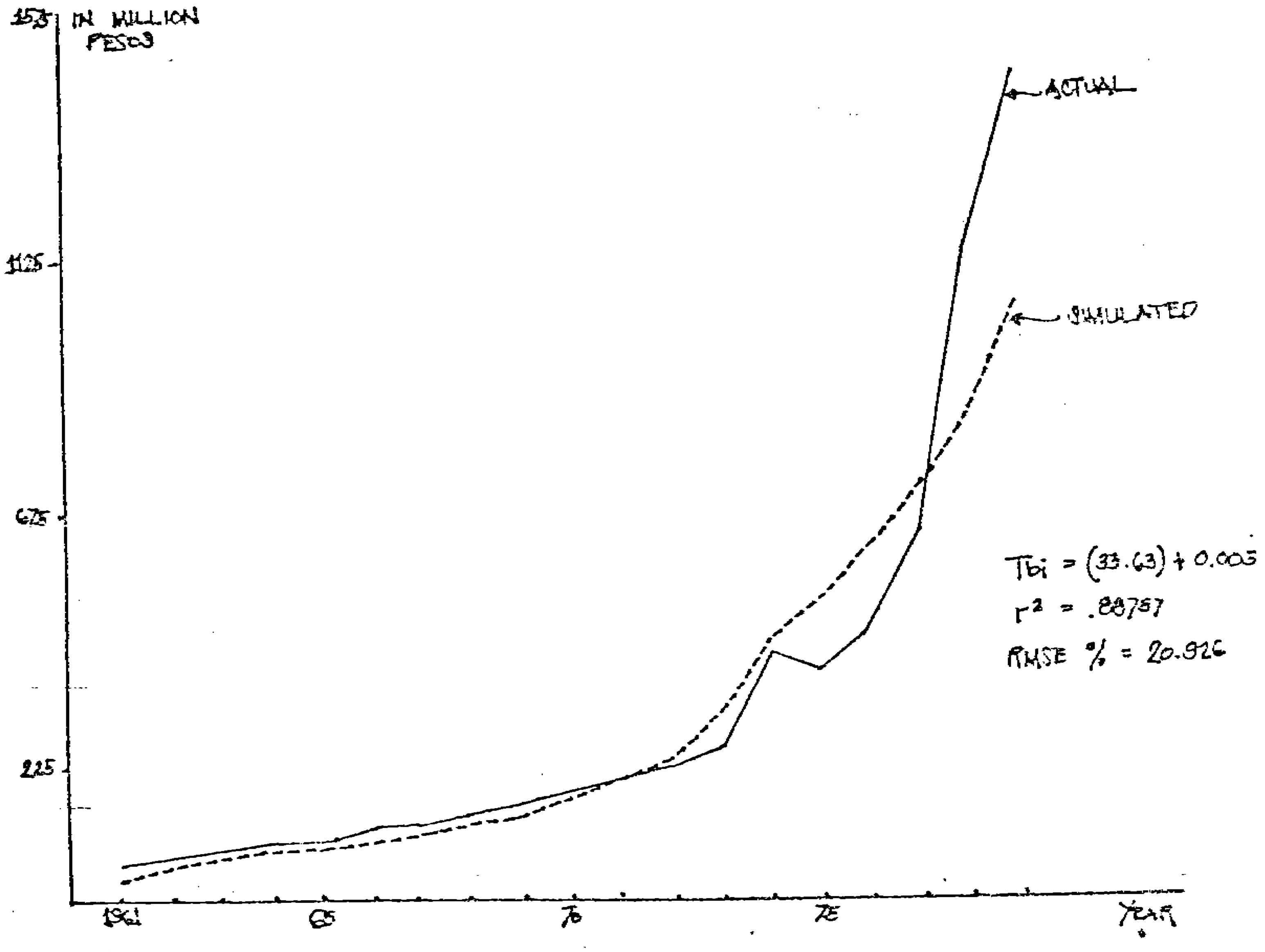
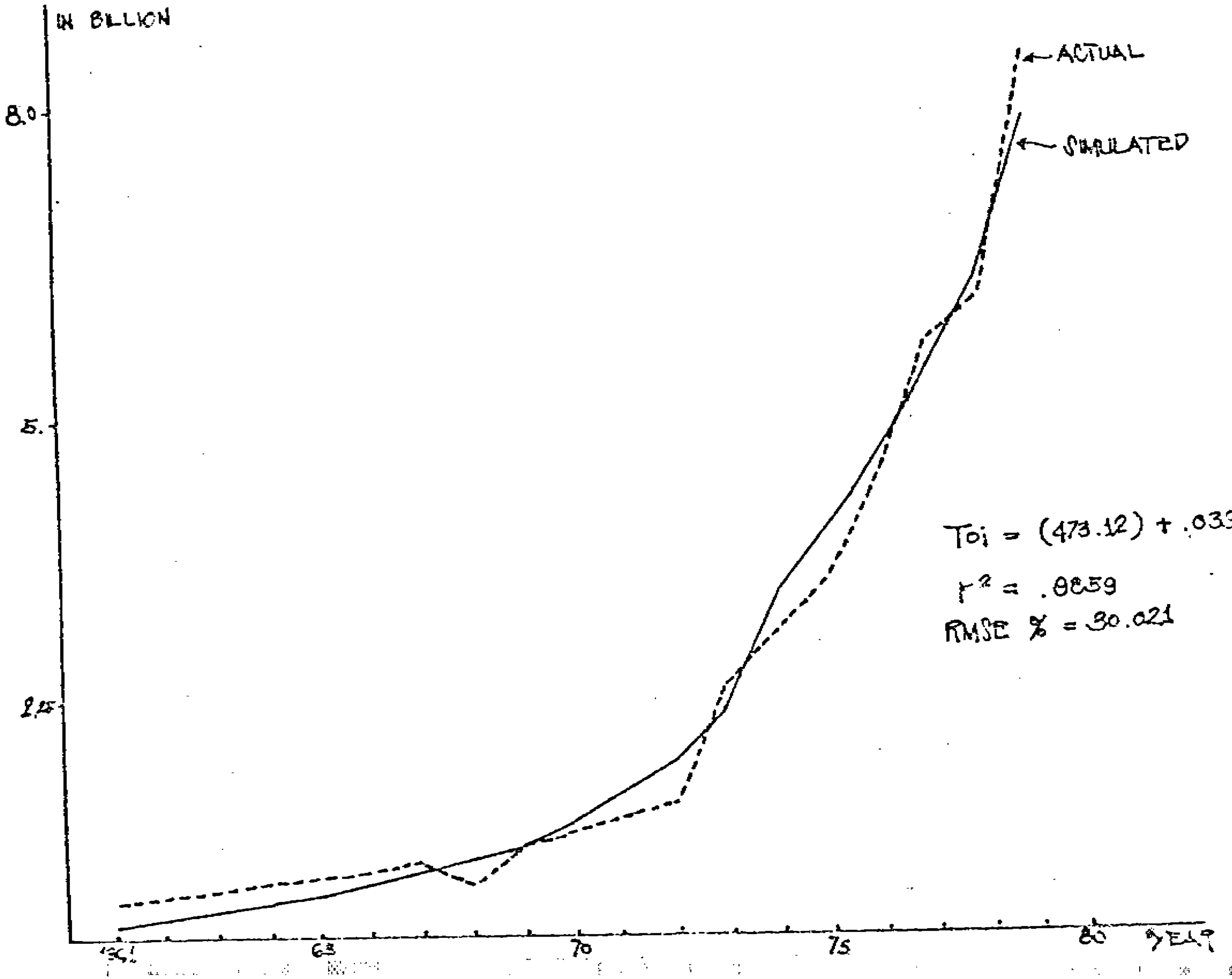


FIGURE 13

OTHER INDIRECT TAXES
(ACTUAL & SIMULATED VALUES)



7. THE DIOKNO PUBLIC SECTOR MODEL

Diokno (1972) constructed a fiscal policy model aimed at evaluating the adequacy of the Philippine tax system in meeting the public revenue needs of the Philippine economy. The model consists of six behavioral equations and six definitions. Four variables were considered as exogenous and values for the said variables were obtained from the National Economic Council's target in the FOUR-YEAR DEVELOPMENT PLAN. These variables are GNP, government investment expenditures, export tax revenue and the money wage rate. There are twelve endogenous variables; six of which are related to government revenue and the rest are related with government expenditures. The other two endogenous variables are government savings and government surplus.

Total government tax revenue was defined as the sum of direct taxes, internal indirect taxes and indirect taxes on the foreign trade sector. Each of these three tax categories was related to GNP. Similarly, non-tax government revenue was then defined as the sum of total tax collections and non-tax receipts of the government.

For our purposes, we made use of only two government tax revenues as endogenous variables:

(1) direct which include income taxes residence tax, transfer tax and real property tax, specific taxes on domestic products, license and business tax, DST, Franchise tax, charges on forest products.

(2) indirect taxes which include sales tax, fines and penlties, mining tax. These two tax classifications were related to GNP.

A simple model, i. e., $Y = b + mX_2$ for each tax classification was specified, where b is the y-intercept and m is the slope. The reestimated equations are as follows:

(1) Direct Taxes $= -219.0031 + 0.024$ (GNP)
$r^2 = .994$ RMSE% = 11.91 (14)

(2) Indirect taxes $= -390.25 + 0.0473$ (GNP)
$r^2 = .986$ RMSE% - 19.78 (15)

Note that RMSE% from (14) and (15) are relatively lower than those obtained from the specification of Kintanar and Mijares as well as those of Jurado and Encarnacion.

TABLE 5

Actual and Simulated Values of
Direct and Indirect Taxes, 1961-1979

YEAR	DIRECT TAXES Actual	DIRECT TAXES Simulated	INDIRECT TAXES Actual	INDIRECT TAXES Simulated
1961	289.7	242.35	421.8	228.7
1962	340.5	299.2	494.6	305.1
1963	415.2	383.3	568.6	418.0
1964	463.7	431.7	640.0	482.9
1965	517.8	492.5	656.6	564.5
1966	506.1	564.4	767.5	661.0
1967	638.1	655.4	835.2	783.1
1968	794.9	748.4	909.9	907.9
1969	885.0	846.4	1017.7	1039.4
1970	1082.3	1051.5	1189.5	1314.6
1971	1366.6	1290.3	1335.9	1635.0
1972	1409.2	1470.7	1452.0	1877.0
1973	1436.3	1960.3	2635.4	2534.0
1974	3216.0	2822.4	3320.3	3690.9
1975	3100.4	3258.1	3788.3	4275.5
1976	3735.8	3795.9	5021.6	4997.1
1977	4569.2	4476.0	6319.8	5909.8
1978	5237.4	5215.3	7150.5	6901.8
1979	6153.0	6489.0	9819.8	8610.8

FIGURE 14

DIRECT TAXES
(ACTUAL & SIMULATED VALUES)

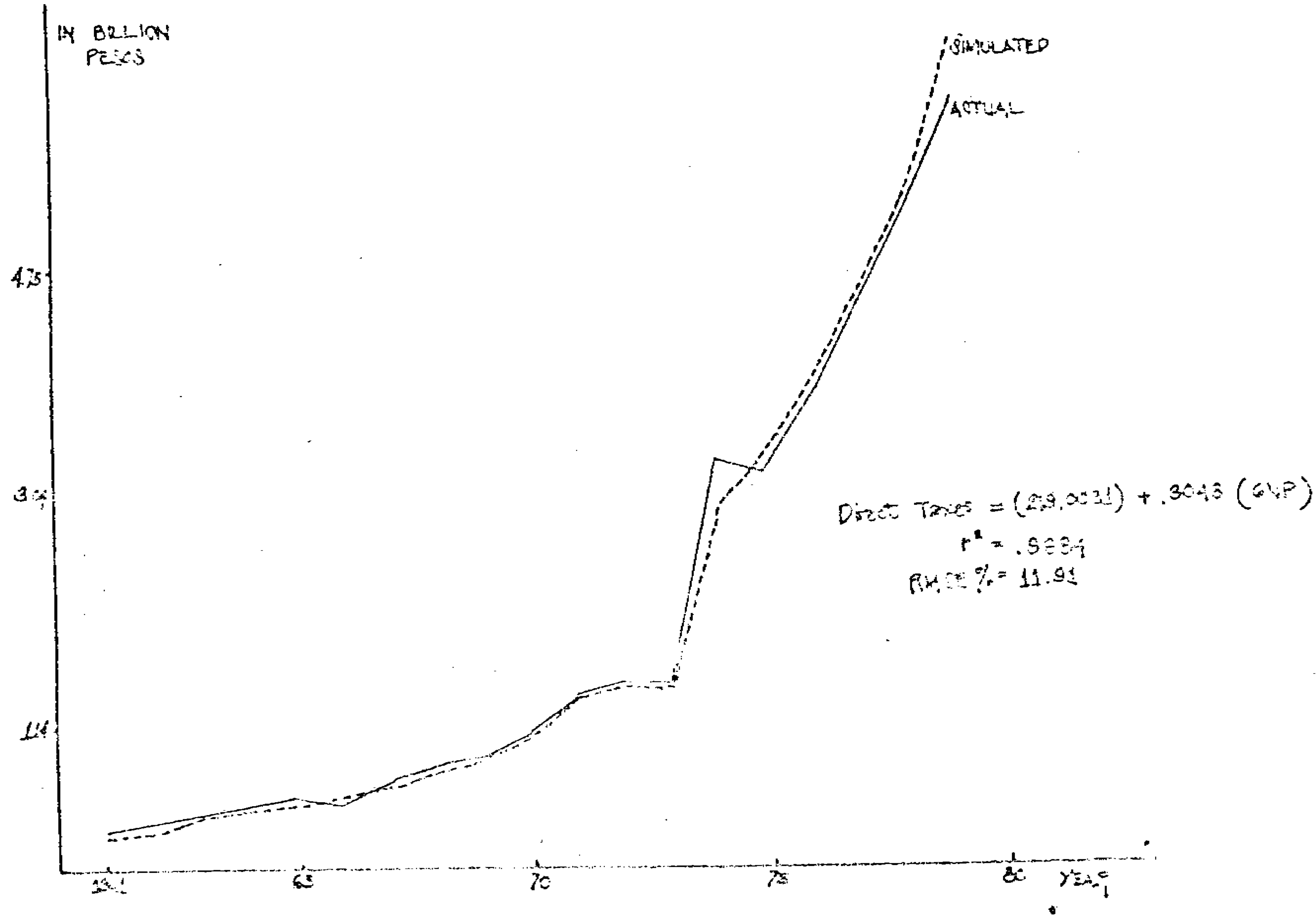
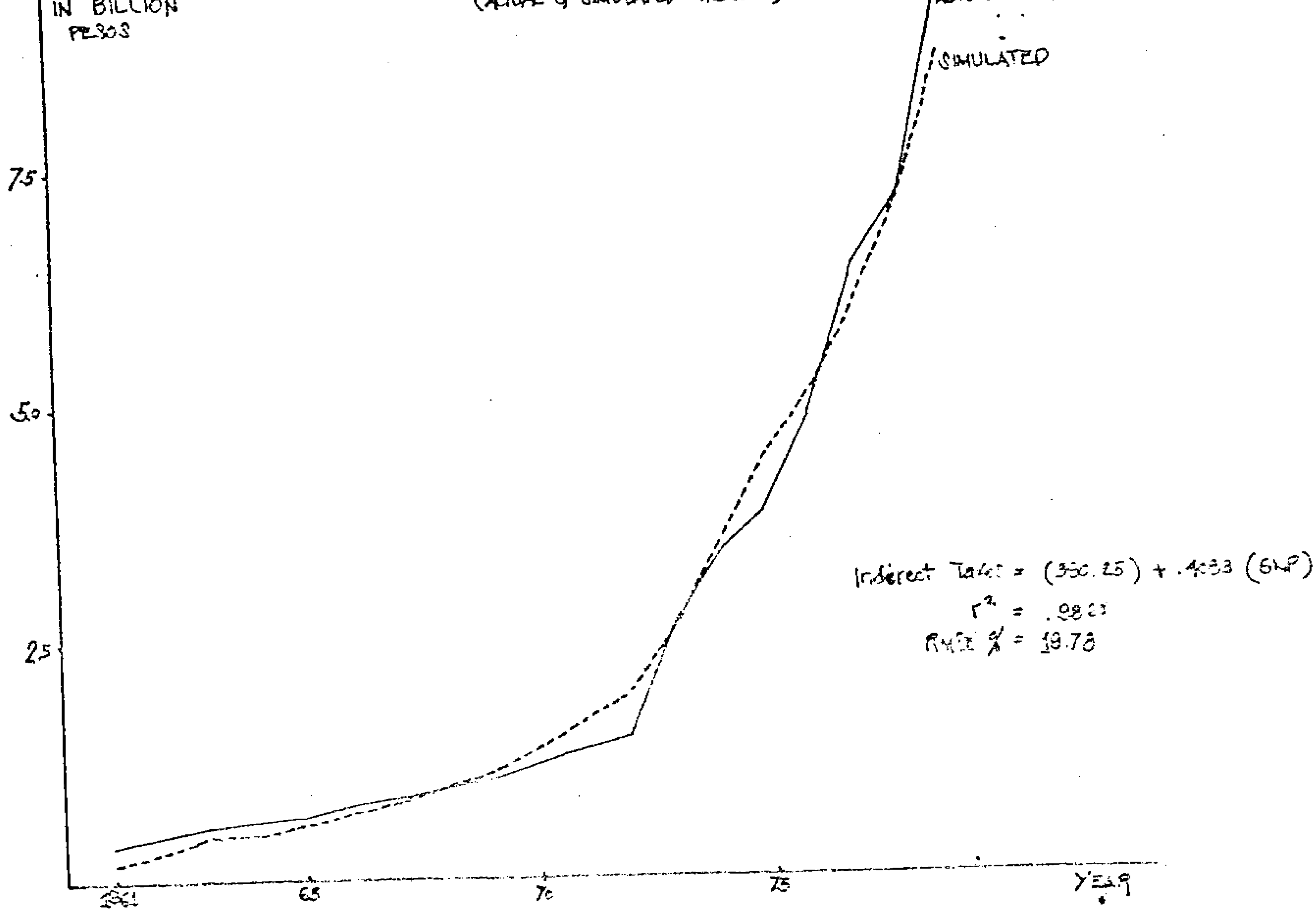


FIGURE 15

INDIRECT TAXES

(ACTUAL & SIMULATED VALUES)



8. SUMMARY AND CONCLUSION

We have reestimated and tested the models suggested by the different works undertaken in tax forecasting using:

(a) 1961-1979 as the estimation period;

(b) r^2 and RMSE % statistics to test the model's forecasting capability;

(c) BIR Statistical Division's data for the tax variables; and

(d) the National Income Accounts revised series for the indicators of the tax bases.

The results of this exercise are summarized in the following tables:

TABLE 6 SUMMARY OF TEST RESULTS

Table 6.1 R^2 and RMSE % for BIR Tax Elasticity Model

	Individual Income Tax	Corporate Tax	Specific Tax	License and Business Tax	Average
R^2	.988	.960	.970	.992	.978
RMSE %	13.68	17.99	17.15	7.73	14.14

Table 6.2 R^2 and RMSE % for Kintanar and Mijares Model

	Individual[3/] Income Tax	Corporate[3/] Tax	Tax on Luxury Items	Base of Semi-Luxury Items	Base of Non-Luxury Items	Compensating Tax	Cigars & Cigarettes	Average
R^2	-	-	.406	.209	.977	.082	.689	.473
RMSE %	13.2	9.1	103.58	174.4	23.36	89.28	20.1	61.86

/ no reestimation was done but used only the <u>ex ante</u> projections as presented in the paper

Table 6.3 R^2 and RMSE % for Jurado-Encarnacion Model

	Individual Income Tax (Tpd)	Corporate Tax (Tbd)	Specific Taxes (Toi)	License & Business Taxes (Tbi)	Average
R^2	.921	.967	.887	.986	.940
RMSE %	20.86	73.38	20.9	30.02	36.29

Table 6.4 R^2 and RMSE % for Diokno's Model

	Indirect Taxes	Direct Taxes	Average
R^2	.986	.988	.987
RMSE %	19.78	11.91	15.84

Based on these, we make the following observations:

1) Using tax bases as explanatory variable for the respective taxes is not a guarantee that the resulting forecasts will be highly accurate.

   Although Kintanar and Mijares' work virtually used time to explain the fluctuations in taxes, the RMSE% averaged at 39% which is only a little higher than the average RMSE% of Jurado and Encarnacion's model (36.3%), using indicators of the tax bases as the determinants. Of course, Kintanar and Mijares did a lot of tedious disaggregation in projecting taxable income by brackets for the individual and corporate categories which probably offset the large mean squared errors derived from the purely simple trend regressions done with the other specific taxes.

2) The more aggregate the endogenous variable, the better the simulations. In other words, as we try to explain particular taxes in detail, the more difficulties we meet.

   The low average RMSE% (15.85) in Diokno's work reflects this advangtage of aggregation more than anything else. On the other hand, the high RMSE% prevalent in Kintanar and Mijares' equations may be explained partly by the more volatile movements inherent in particular taxes, which are more difficult to capture.

3) The elasticities approach provided the lowest average RMSE% error (14.14); indicating that there is a practical and theoretical foundation in assuming that taxes move in a power function fashion against time and the tax bases, rather than in a simple linear one.

4) Lastly, the general underestimation of almost all historical simulations is to be expected because of the inability of simulations to capture other explanatory variables which may have caused an upward shift in the values of the endogenous variables through time.

The future direction of the Bureau's forecasting efforts should focus on the following facets:

1) Spatial or regional forecasting - An earlier exercise using linear programming, tried to derive the regional implication of the tax forecast and determine tax goals for the same. This was later substituted by a heuristic approach that was based on the regional gross domestic product. Even this approach has its problems knowing that the BIR has seventeen revenue regions as against the Philippines' administrative delineation which has only twelve regions.

2) Disaggregation forecasting - It was found out that the more detailed the analysis of taxes, the more

difficult the required forecasting techniques become. This very difficulty represents a challenge to econometric forecasting as a whole.

3) Refinements in estimation methodology - The methods used thus far, were generally limited because of the lack of cyclical analysis and the exclusion of lag variables. Any forecasting exercise is a search for the particular methodology most suited to give accurate projections.

4) Refinements in specification - Other variables aside from the tax bases, should be considered in terms of their ability to explain the movements of particular taxes. Since a tax revenue represents an income for the government, it is not totally unwarranted to consider costs of tax collection as one of the explanatory variables. In other words, a function specified as: Tax revenue = f (tax base, cost of tax administration), where both determinants are expected to relate positively to tax revenue, is worth looking into. Other dummy variables representing tax amnesty periods, new PD's should also be introduced.

5) Widening of scope - To the extent that tax receipts are determined by tax bases, it is imperative that we derive good forecast for the tax bases first before proceeding with the tax forecast proper.

Essentially, this would require an econometric effort in a macro-scale i.e., calling for a wider and more comprehensive view of the economy.

Of course, the common thread woven into these factors is the need to establish a definite and stable data base. This calls for a close working effort with the Data Processing Center and the Statistical Division of the Bureau.

It is in line with the preceding analyses that we approach the second ph   of this project. So far, we found out that the present BIR's elasticity approach produced the most acceptable historical simulations among all the other existing tax forecasting models and methods. This implies two things:

(a) While we are still in the process of developing a more suitable forecasting model with the minimum standard error of forecasts, the Bureau may use the elasticity approach for its immediate need to forecast tax revenues; and

(b) The RMSE % computed value for the BIR's elasticity approach should now serve as the benchmark in evaluating the structures to be formulated and estimated in the future.

We will also consider other means of evaluating the forecasting ability of the models particularly for those specifications whose forecasting errors do not deviate far from each other.

B I B L I O G R A P H Y

Bahl, Roy. "A Representative Tax System Approach to Measuring Tax Effort in Developing Countries", IMF Staff Papers, 1972

Evans, Michael. Macroeconomic Activity. University of Pennsylvania, 1967

Diokno, Benjamin E. The Philippine Tax System and the Public Revenue Needs of the Philippine Economy, FY 1972-1975 Quezon City, 1972

Jurado, Gonzalo, and Jose Encarnacion, Jr. Government Submodel of the Philippine Economy, 1955-1969. Philippine Economic Journal, Vol. XI, No. 2, 2nd semester, 1972.

Kintanar, Agustin Jr., and Mijares, Tito. Estimation and Forecasting of Tax Revenue in the Philippines, Q.C. IEDR, Discussion Paper p. 2-65.

Klein, Laurence. Econometrics. Maruzen Asian Edition 1953.

Pindyck R. and Rubinfeld, D. Econometric Models and Economic Forecasts, 1976

Sinay, Cream Gilda C. Buoyancy and Elasticity of the Philippine Tax System, CY 1961-1972. Q.C. 1974

Theil. Econometric Forecasts and Policy, Amsterdam, 1961

BIR Forecasting Models.