

## THE PUERTO RICO REVENUE PROJECTION MODEL: AN ECONOMETRIC APPROACH

John R. Newton and Victor Sepulveda\*

As is the case with most other state level governments, Puerto Rico's public revenue projections have been generated by traditional estimation methods with varying degrees of reliability. With the increasing demands of credit markets for believable estimates, especially during periods of unstable fiscal climates, many jurisdictions have seen the need for a more reliable and consistent means of projecting revenues and preparing budgets. In response to this need, the Puerto Rico Treasury has developed a new set of econometric equations to predict aggregate revenues by sources. These equations form part of a larger model of the Puerto Rican economy. This paper presents the proposed forms of the tax equations, the statistical results of estimating them, and projections made using them.

### The Scope of the Model

Statistical projection techniques apply best to aggregate variables with identifiable patterns or relationships to other factors. In the present case, experience has shown that statistical projections are reliable for neither most small revenue categories nor for erratic sources, but only for larger or more stable sources. For this reason alone, statistical projections are unlikely to completely supplant traditional, non-statistical estimation methods.

The most successful statistical projections have been those for total excise tax collections and income tax collections. Though some subcategories of excise taxes also yield usable projections, breaking most of these categories down further leaves a series of smaller, more erratic revenue sources. Consequently, the scope of statistical projections has been limited to the two basic revenue sources mentioned and certain other small but stable categories.

The revenue projection model discussed below has yielded more accurate projections than traditional methods for periods of one to five years.

### Data Sources

Data for the dependent tax variables of this model, including the revenue sources mentioned above, are from the Puerto Rico Treasury. Most of them require some adjustment due to changes in the tax laws redefining either the tax base or the tax rates in particular categories.

Data for the independent variables was obtained from three sources. Historical data on economic aggregates, such as gross income and employment, are from the Puerto Rico Planning Board and the Puerto Rico Department of Labor. Projections of these variables are from the Wharton/IAU econometric model for Puerto Rico. Data on the U.S. economy are from the Wharton (WEFA) data banks, available to the Treasury through a contract with Interamerican University. For most variables, annual observations are available from 1955 through 1983.

\*This paper was prepared while the authors were with the Puerto Rico Department of the Treasury as Assistant Secretary for Economic Affairs and Consultant in Econometrics, respectively. Newton is now Senior Economist with SRI International-Washington. Sepulveda died in November, 1984, at the age of 34, a victim of hemophilia.

### Structure of Equations and Estimations

We have used two different types of econometric estimations, depending on the revenue category involved. The bulk of the projections are based on standard linear regression, using combinations of Ordinary Least Squares (OLS), Two Stage Least Squares (2SLS), and the Limited Information Maximum Likelihood Method (LIML), as described below. Then, for smaller but temporally stable items, with less readily available explanatory variables, we have used a simple stepwise autoregressive times series projection. All estimations were made using the SAS (Statistical Analysis System) software package.

The simplest of our regression equations are those for income tax collections. We assume a nondifferentiated system in which total income tax collections (TIC) for a sector (individual, corporate or offshore) are related to the base, total income (Y), for that sector and to legal changes represented by dummy variables (D) in the following manner:

$$(1) \text{ TIC}_i = a + bY_i + cD_i + e_i$$

where  $e_i$  is the statistical disturbance term and the subscripts indicate the year of the observations.

Expressed in this manner, the estimate does not take account of differing tax rates by income bracket. The coefficient,  $b$ , becomes the average marginal tax rate for all income categories. Initial experiments with elaborations indicate that the predictive power of the equations would not be greatly expanded by including differentiation by brackets. On the other hand, such a specification would have other analytical advantages, as mentioned in a later section.

The income tax equations were estimated using OLS. In keeping with our initial purpose of producing a usable vehicle for revenue projection, we allowed no significant elaboration of the simple model. Significant autocorrelation occasionally appeared, however, and was dealt with by the SAS procedure AUTOREG, which estimates a multiple order autoregressive process and accordingly corrects the regression equation. Forecasting is generally improved by this process, since elements of the time series itself enter as explanatory variables for the first few forecasts.

The estimation of excise tax collections is slightly more challenging, since a simultaneous system is involved. In general we assume that consumption, price, cost, and taxation are related in the following manner:

$$(2) \text{ TC}_i = a_1 + b_1 YD_i + c_1 P_i + e_{1i}$$

$$(3) P_i = a_2 + b_2 \text{TC}_i + c_2 \text{AE}_i + d_2 C + F_2 D + e_{2i}$$

$$(4) \text{ TEC}_i = \text{TC}_i \times \text{AE}_i$$

where TC represents total consumption of the product involved, YD is disposable income, P is the average selling price of the product, C is an index of the relative cost of production of the product, D is a dummy variable for tax law or policy changes, AE is the average excise per unit, and TEC is the total excise collection. These three equations are equivalent to an equilibrium Supply-Demand system, with quantity supplied and demanded equal to TC, and the supply equation expressed as a price equation. The advantage of the above formulation is that it emphasizes the variables of concern for revenue projection and policy making.

In an interdependent system of this sort, with dependent variables from one equation appearing as regressors in another equation, standard OLS estimates can be inconsistent and biased. Techniques developed to resolve these problems include 2SLS, 3SLS, and LIML. All these methods are available automatically with SAS, allowing us to use and compare them all initially. Since the differences among the methods were insignificant, we have generally used 2SLS.

Even using the adjustment methods mentioned above, however, the coefficients and their  $t$ -values for a system of equations may be biased and inconsistent, due to multicollinearity or specification error. In such cases, estimating the reduced form of the system will provide better projections. For the system above, the reduced form equation is:

$$(5) \text{ TEC}_i = a_3 + b_3 \text{AE}_i + c_3 \text{AE}_i(YD_i) + d_3 \text{AE}_i(C_i) + f_3 \text{AE}_i(D) + g_3 \text{AE}_i^2 + e_{3i}$$

with all constant terms collected and combined in order to simplify exposition. This equation can be used for projection when econometric problems appear in the estimation of equations (2) through (4).

A much simpler method of projections is utilized for several smaller, but stable, categories of revenue.

In these cases, data for explanatory variables is usually lacking, and a full modeling effort would hardly be worthwhile. Fortunately a reliable and cheap (in terms of computer time) method exists for combining time trend projections and time series projections of the autoregressive sort. This method consists of first estimating a time trend and then estimating autoregressive parameters for observations, considered as deviations from that trend. No particular statistical efficiency claims are made for the method; it simply works and costs very little. Though we have not included the many small equations estimated by this method here, their projections are presented below.

All three estimation methods described here produced usable projections, as shown in the next section. The obvious emphasis is on usability and simplicity, not on statistical sophistication. For that reason, many adjustments have not been performed on the equations. From our perspective, elaborate specifications would be useless if forecasting were not improved. Also, as we describe below, many elements of traditional forecasting will never be formally included in the model, so an overly elaborate model would be wasted.

### Statistical Results and Projections

Table 1 presents the estimates of the equations whose specifications were described above. Definitions for all variables included in the model appear in Table 2. After discussing these estimates briefly, we will turn to the projections obtained.

Applying the specifications described in Equation (1) to domestic individual, domestic corporate, and offshore income tax collections yields the results shown in Equations (a) through (c) of Table 1. As indicated by the t-statistics, both income and the dummy variables for legal changes exert a statistically significant effect on total tax collections from each source. Notice that in the case of corporations and offshore collections, gross product is used as a proxy for income, since data for the relevant sector incomes were not available.

TABLE 1

#### Definitions of Variables for Revenue Projection Model

---

AEA	: Average excise tax on all goods, in percent.
AEB	: Average excise tax on beer, per unit.
AEC	: Average excise tax on cigarettes, per unit.
AED	: Average excise tax on distilled spirits, per gallon.
AEDI	: Average excise tax on imported distilled spirits, per gallon.
AEDL	: Average excise tax on local distilled spirits, per gallon.
Ca	: Absolute cost index for machinery and equipment as proxy for industry costs.
CR	: Relative cost index for machinery and equipment as proxy for industry costs.
DEA	: Dummy variable for change in excise tax law or policy.
DEB	: Dummy variable for change in beer excise law or policy.
DEC	: Dummy variable for change in cigarette excise law or policy.
DED	: Dummy variable for change in distilled spirits excise law or policy.
DIC	: Dummy variable for change in corporate income tax law or policy.
DII	: Dummy variable for change in individual tax law or policy.
DIO	: Dummy variable for change in offshore income tax law or policy.
PA	: Consumer price index for beer.
PB	: Relative price index for beer.
PC	: Relative price index for cigarettes.
PD	: Relative price index for rum, as proxy for price of distilled spirits.
TCA	: Personal consumption expenditures, as proxy for total consumption of products subject to excise tax, in dollars.
TCB	: Total consumption of beer, in gallons.
TCC	: Total consumption of cigarettes, in millions.

---

TABLE 1 (continued)

Definitions of Variables for  
Revenue Projection Model

TCDI	: Total consumption of imported distilled spirits, in gallons.
TCDL	: Total consumption of local distilled spirits, in gallons.
TECA	: Total excise tax collections, all sectors.
TECB	: Total excise tax collections for beer.
TECC	: Total excise tax collections for cigarettes.
TECD	: Total excise tax collections for distilled spirits.
TIC	: Total income tax collections, sector unspecified.
TICC	: Total income tax collections for corporations.
TICI	: Total income tax collections for individuals.
TICO	: Total income tax collections for offshore residents.
YDM	: Disposable income, in current dollars.
YDR	: Disposable income, inflation adjusted.
YGM	: Gross product, in current dollars.
YIM	: Personal income, in current dollars.

TABLE 2

Selected Estimated Equations  
From the Revenue Model

(a) TICC	= -15903 (5.92)	+ 2.346 YGM	+ 32792 DIC (8.25)						
(b) TICI	= - 45944 (109.0)	+ .5969 YIM	+ 15450 DII (4.70)						$R^2 = .998; d = 1.4$
(c) TICO	= -1613 (1.80)	+ .3086 YGM	+ 18513 DIO (6.20)						$R^2 = .964; d = 1.58$
(d) TCA	= -6971 (2.0)	+ 79.9 PA (1.8)	+ .091 YDM (10.5)						$R^2 = .999$
(e) PA	= 37.93 (2.7)	+ .386 TCA (2.2)	- 3.19 AEA (1.3)	+ 72.24 CA (9.6)	+ 6.52 DEA (3.0)				$R^2 = .998$
(f) TECA	= -106984 (2.0)	+ 28849 AEA (1.62)	+ .1062 + AEA(YDM) (22.7)	+ 2108 AEA(CA) (1.05)	- 1952 AEA(DEA) (3.39)	- 2380 (1.43)			$R^2 = .999; d = 1.57$
(g) TCB	= 35369 (3.32)	- .4587 PB (3.02)	+ .0125 YDR (26.9)						$R^2 = .967$
(h) PB	= 87255 (9.03)	- .074 TCB (.587)	+ 2929 AEB (.595)	- .16598 CR (1.76)	- .1263 DEB (.646)				$R^2 = .132$
(i) TECB	= 10047 (1.46)	- 32276 AEB (1.97)	+ 12.157 AEB(YDR) (6.46)	+ 10326 AEB(CR) (1.21)	- 554 AEB(DEB) (-.43)	+ 14783 AEB <sup>2</sup> (2.83)			$R^2 = .991; d = 1.86$
(j) TCC	= 2026 (9.1)	- .017 PC (5.2)	+ 3153 YDR (11.4)						$R^2 = .864$
(k) PC	= 74626 (4.0)	- 6.8 TCC (1.3)	+ 601 AEC (7.5)	- 4136 cr (.28)	+ 2580 DEC (1.3)				$R^2 = .822$
(l) TECC	= 5684 (.36)	+ 495 AEC (.48)	+ .4079 AEC(YDR) (3.17)	+ 43.9 AEC(CR) (.12)	+ 26.98 AEC(DEC) (.59)	- 11.02 AEC <sup>2</sup> (1.7)			$R^2 = .985; d = 1.71$
(m) TCDI	= 1947 (4.88)	- 315.9 PD (5.87)	+ .5073 YDR (20.4)						$R^2 = .951$
(n) TCDL	= -2678 (1.26)	+ 689 PD (2.41)	+ .6725 YDR (5.09)						$R^2 = .522$
(o) PD	= 10.01 (5.34)	+ .7826 TCDI (3.32)	+ .14 T CDL (.998)	+ .005 AEDI (.134)	+ .131 AEDL (1.02)	- 3.60 CR (2.97)	- .139 DED (1.1)		$R^2 = .754$
(p) TECD	= 45983 (3.6)	- 2733 AED (1.3)	+ 2.067 AED(YDR) (5.6)	- 2757 AED(CR) (1.89)	+ 38.54 (.20)	- 361 AED <sup>2</sup> (1.37)			$R^2 = .976; d = 2.21$

Note: Equations are presented in original form, without correction for autocorrelation. All statistics were brought within the acceptable range after correction.

For the excise tax estimates shown in Equations (d) through (p) of Table 1, the results were similar. Income variables normally proved significant, as did price and policy variables. Price is negative in the demand equations. Total consumption generally did not tend to exert much influence on price, indicating either a flat supply function or some mis-specification of the price equation. Our feeling is that the latter explanation is more likely, since the "cost" variable utilized in the price equations is a distant proxy and is highly colinear with the consumption variables. For this reason the coefficients of the price equations are suspect, and the reduced form equations are consequently selected projections. The reduced form estimates are shown in equations (f), (i), (1), and (p) of Table 1.

For most of the equations, potential predictive power as measured by  $R^2$  is relatively high. The real test of these estimates, however, is whether they predict future collections well. Table 3 compares the model's projections for fiscal year 1983 with actual collections. As indicated by the "difference" column, the results are favorable. The highest percentage error was 2.1 percent for total excise collections. Most projections were within one percent, and two projections were precise to within less than 0.1 percent of total collections.

TABLE 3

Econometric Projections Compared  
with Actual Collections  
for Selected Revenue Items  
Fiscal Year 1983  
(In Thousands)

	<u>Projected</u>	<u>Collected</u>	<u>Difference</u>	<u>Percent</u>
Income Taxes:				
Corporate	\$352,521	\$354,330	\$ 1,809	0.5 %
Individual	744,903	752,415	7,512	1.0
Offshore	56,729	56,729	0	0
Excise Taxes:				
Total Excises	\$512,081	\$522,733	\$10,692	2.1 %
Beer	88,066	87,206	- 860	0.1 %
Cigarettes	88,811	88,950	139	0
Distilled Spirits	62,111	61,016	- 1,095	0.02

## Use of the Results

The most important use to which this model has been put is the construction of the full General Fund revenue projection for the Commonwealth Budget. Table 4 presents the full projection for fiscal years 1984 and 1985. This estimate supplies the budget constraint for all unrestricted Central Government expenditures as well as the greater part of the expenditures of the municipalities. In addition, most Special Fund expenditures are directly proportional to certain items of the General Fund, so that in effect this single-projection determines most public spending on the Island.

**TABLE 4**  
General Fund Revenue Projections  
Fiscal Years 1984 and 1985  
(In Millions)

Items	:	1984	:	1985
Total		<u>\$2,510.5</u>		<u>\$2,737.7</u>
		<u>2,029.5</u>		<u>2,224.7</u>
Tax Revenues		<u>1,927.3</u>		<u>2,122.5</u>
Property Taxes		112.0		100.0
Income Taxes, Total		<u>1,228.3</u>		<u>1,357.1</u>
Individuals		721.3		739.9
Corporations		388.0		443.2
Partnerships		7.0		3.0
Withheld from non-Residents		32.0		55.0
"Tollgate Tax"		80.0		116.0
Inheritance and Gift Taxes		8.0		7.2
Excise Taxes, Total		<u>542.0</u>		<u>615.3</u>
Alcoholic Beverages, Total		<u>167.0</u>		<u>180.9</u>
Distilled spirits		69.0		78.0
Beer		91.0		93.6
Others		7.0		9.3
Other Taxable Goods, Total		<u>375.0</u>		<u>434.4</u>
Cigarettes		108.0		108.0
Petroleum Products		5.0		5.0
Motor Vehicle and Accessories		72.0		108.7
Tires and Tubes		4.0		6.1
Electrical and Gases Appliances		33.0		39.0
Horse Races		14.0		14.6
Insurance Premiums		14.0		14.0
General Excise Taxes		98.0		112.0
Petroleum Imports		—		—
Sugar		7.0		7.0
Others		20.0		20.0
Licenses		<u>73.0</u>		<u>42.9</u>
Motor Vehicles		25.0		30.9
Alcoholic Beverages		2.0		2.0
Other		10.0		10.0
Non-Tax Revenues		<u>102.2</u>		<u>102.2</u>
Lottery Proceeds		46.0		46.0
Miscellaneous, Total		<u>56.2</u>		<u>56.2</u>



TABLE 4 (continued)

General Fund Revenue Projections  
Fiscal Years 1984 and 1985

(In Millions)

Items	1984	1985
Fines and Penalties	17.0	17.0
Registration and Certification Fees	28.2	28.2
Others	11.0	11.0
From non-Commonwealth Sources	<u>481.0</u>	<u>513.0</u>
Customs Duties	62.0	65.0
U.S. Excises on Offshore Shipments	419.0	448.0

Most of the items listed in Table 4 will be familiar by now, as will the quantities involved. A simple calculation reveals that the econometric method is involved in some 90 percent of the total projected amount of the General Fund. Even so, it should be emphasized that more traditional methods are still necessary, both as a check and modification of the econometric approach and as the principal method for certain recalcitrant revenue items.

The output of statistical models is often adjusted to account for events or information which cannot be adequately included in the statistical equations. The differences between econometric and final revenue projection shown in Table 5 are attributable to such factors. One-time-only events or political considerations—such as a partial tax amnesty approved during 1983—can hardly be predicted econometrically. Thus traditional input is used to adjust econometric output. On the other hand, the statistical models determine the trends and interrelationships of all revenue items, to the exclusion of the traditional methods. We feel that this combined approach takes advantage of the strengths of both projection methods.

TABLE 5

Comparison of Econometric Projections with Final Projections  
for Selected Revenue Items Year 1985

(In Millions)

Items	Econometric Projection 1984	Final Projection 1985
Total	-	<u>\$2,737.7</u>
From Commonwealth Sources		<u>2,224.7</u>
Tax Revenues		<u>2,122.5</u>
Property Taxes	120.3	100.0
Income Taxes, Total		<u>1,357.1</u>
Individuals	817.0	739.9
Corporations	435.4	443.2
Partnerships	6.5	3.0
Withheld from non-Residents	63.3	55.0
"Tollgate Tax"	119.6	116.0
Inheritance and Gift Taxes	16.6	7.2
Excise Taxes, Total		<u>615.3</u>
Alcoholic Beverages, Total		<u>180.9</u>
Distilled spirits	67.2	78.0

TABLE 5 (continued)

Comparison of Econometric Projections with Final Projections  
for Selected Revenue Items Year 1985

(In Millions)

Items	Econometric Projection 1984	Final Projection 1985
Beer	94.3	93.6
Others	9.3	9.3
Other Taxable Goods, Total		<u>434.4</u>
Cigarettes	103.1	108.0
Petroleum Products	2.1	5.0
Motor Vehicle and Accessories	111.7	108.7
Tires and Tubes	4.0	6.1
Electrical and Gases Appliances	39.0	39.0
Horse Races	15.6	14.6
Insurance Premiums	12.0	14.0
General Excise Taxes	135.0	112.0
Petroleum Imports	—	—
Sugar	7.5	7.0
Others	13.8	20.0
Licenses		<u>42.9</u>
Motor Vehicles	27.4	30.9
Alcoholic Beverages	2.1	2.0
Other	5.7	10.0
Non-Tax Revenues		<u>102.2</u>
Lottery Proceeds	61.3	46.0
Miscellaneous, Total		<u>56.2</u>
Fines and Penalties	14.7	17.0
Registration and Certification Fees	24.9	28.2
Others	9.2	11.0
From non-Commonwealth Sources		<u>513.0</u>
Customs Duties	70.2	65.0
U.S. Excises on Offshore Shipments	275.0	448.0

Prospectively, we would like to use the revenue projection models for tax impact analysis and general policy simulation. Obviously the structure of taxation is not yet closely enough specified in our equations for such analysis. In order to simulate the effects of tax rate changes, for instance, both rates and bases would have to be included in the equations. Simulation should be the next step in the development of the revenue model.

### Conclusion

We have constructed a simple but reliable revenue projection model, using varied standard econometric and statistical techniques. Projections based on the estimated equations compare well with actual collections and with traditional projections for periods of one to five years. With certain elaborations of the structure of the equations, the model should also be useful for tax impact analysis and policy simulation.