THE STRENGTHS AND WEAKNESSES OF FACTOR ANALYSIS IN PREDICTING CUBAN GDP

Nicolás Sánchez and Miles Cahill[1]

This paper has a dual purpose. At the technical level, it updates the work of Gitanjali Joglekar and Andrew Zimbalist (1989) on the use of factor analysis to estimate per capita Gross Domestic Product (GDP) for Cuba. At the analytical level, it questions the validity of this approach when the underlying economic structure of factor analysis is based on the data of relatively unregulated economies such as those of other Latin American countries.

The main conclusion of the paper is that while factor analysis is indeed a powerful predictor of the economic performance of market economies, it tends to overestimate the level of economic activity of command economies, and it fails to take into account subjective (i.e., utility-related) values which are assumed to be an intrinsic part of any measure of economic well being. If this criticism stands, then it follows that other similar studies of Cuban GDP are failing to provide a good measure of economic performance and well being in this island nation.

This study is part of a more comprehensive attempt to compare individual Latin American economies, as of 1989-1990, with those of the various states of the United States for prior decades. This has dictated the type of data that has and has not been used: on the one hand, it has been necessary to find comparable statistics for the Latin American countries and the various states of the United States; on the other, it has been impossible to use easily accessible, standard data for international transactions and monetary variables. As a result, the data used in this study are not identical to those used by Joglekar and Zimbalist (1989).

Although the authors found some other per capita GDP estimates for Cuba, these presented at least one of the following problems: (a) the methodologies used in the calculations were not made explicit; (b) the numbers were inconsistent, over time, with well known trends; and (c) the validity of the figures were questioned by the scholars reporting them.

Central Intelligence Agency estimates (CIA, 1990-present) fall under (a) above. For both 1989 and 1990, per capita GNP for Cuba was given at $2,000, a figure that is about 10% below the figure for Argentina for 1989 ($2,217) and close to 20% below the figure for 1990 ($2,560). In contrast, per capita GNP for Haiti was $380 for 1989 and $440 for 1990. In those two years the Cuban figures fell between the two estimates for Chile: $1,970 for 1989 and $2,130 for 1990.[2] Years later, the CIA reported a 1996 per capita GDP figure for Cuba of $1,480; for

Argentina, $8,600; for Chile, $8,400; and for Haiti, $1,000.

These figures are not entirely comparable over time, since the earlier estimates referred to GNP and the later ones to GDP, and the earlier estimates used market prices while the later ones used purchasing power parity (PPP) prices for the foreign exchange rates. However, the CIA figures suggest that Cuba in 1990 had some catching up to do relative to Argentina, but was very close to Chile; then Cuba experienced a catastrophic drop in economic activity during the 1990s. The first two inferences require confirmation, while the third one is widely agreed to in the economics literature (Betancourt, 1996). Those who have carefully studied the history of the Castro regime have argued that the apparent health of the Cuban economy during the 1980s could be attributed to the massive Soviet subsidies during that period (Pérez, Jr., 1988); this contention will be addressed later in the paper.

United Nations estimates for Cuban GDP,[3] as reported in the various issues of *Human Development Report*, are inconsistent. According to this publication (UNDP, 1991-present), real per capita GDP for Cuba, using PPP$, was $2,500 in 1989, $2,200 in 1990, $2,000 in 1991, and then surprisingly $3,412 for 1992 and $3,000 for 1993. Caveats to the 1992 and 1993 figures state that these new estimates made use of the Penn World Tables, which began to be used widely in the literature in the 1990s.

The figures for 1992 and 1993 present some serious problems of interpretation. Since most scholars consider 1993 the trough of economic performance in Cuba (Mesa-Lago, 1996), would it be appropriate to assume that per capita GDP in 1989 was, say, 35%[4] above the 1993 figure? In other words, was Cuban per capita GDP in 1989 at the $4,050 level, close to that of Argentina ($4,310) for that year? The *Human Development Report* is not supportive of this assertion, since earlier issues noted a much wider disparity between the Cuban and Argentinean economic performances. The puzzle that the newer figures create makes it appropriate to try to find an independent, and hopefully transparent, estimate for Cuban GDP at the end of the 1980s.

A third set of figures for Cuban GDP comes from the *Banco Nacional de Cuba,* and this set has been widely reported by serious scholars interested in Cuban affairs. In addition to being reported in pesos, creating the need to argue about the appropriate exchange rate, their reliability have been seriously questioned. Pérez-López (1997) pointed out that the 34.8% contraction that the Cubans reported between 1989 and 1993 "still is not as large as the one which the Cuban media had reported earlier, one that was presumably also based on official statistics." Carranza Valdés (1993) estimated a cumulative decline of about 38% for the Global Social Product (which presumably would track GDP) for the shorter 1989-92 period. Ritter (1997) estimated a decline of "perhaps" 45% for per capita GDP between 1989 and 1993, while Pastor and Zimbalist (1998) estimated a cumulative decline of 35% to 45% for GDP between 1989 and 1993.

In light of these problems, then, the authors of this study searched for earlier estimates of Cuban economic performance in order to update them to the 1989-1990 period, which was the one of interest to them. This they found in the Joglekar and Zimbalist (1989) article, which had the added advantage of using the same methodology that was being used by the authors to compare the Latin American economies to those of the states within the United States. This methodology had also been used before to find per capita GDP estimates for the Eastern European command economies (Szilágyi, 1978).

THE STRENGTHS OF FACTOR ANALYSIS

Factor analysis is an extension of multiple regression analysis, including some modifications that make the approach less likely to be influenced by the personal biases of the researcher, and less dependent on the assumption of independence among variables. Techni-

---

3. Based on a wide variety of sources, including the World Bank.

4. This figure was chosen to match the estimates of the Cuban National Bank.

cal explanations of factor analysis at various levels of difficulty can be found in the books by Norušis (1994), Kline (1994), Yotopoulos and Nugent (1976), and Mulaik (1974).

Ideally, in multiple regression analysis, one would observe a dependent variable on the left hand side of an equation that is "explained" by a set of independent variables on the right hand side. The problem, however, is that the connecting link between dependent and independent variables needs a strong theoretical basis. Furthermore, many of the so called "independent" variables may not really be independent at all, but rather change or move together in response to some other unknown variables or "factors." Hence the researcher usually encounters criticism of both the proposed theoretical framework and the method of untangling the mutual dependence among the presumably independent variables.

In contrast, factor analysis is used to determine the underlying determinants of many variables, without the need to postulate causality. This is especially useful because economic and social indicators are closely intertwined, making it nearly impossible to find a set of economic and social variables that are not correlated in some way. This interdependence makes normal regression analysis problematic but, surprisingly, does not adversely affect factor analysis.

Factor analysis, then, is a formal mathematical procedure that estimates the unobserved independent factors (or components) that characterize the various so-called independent variables. Because they can be constructed to be independent, the factor estimates (called factor scores) can be used in regression analysis to find the correlation between the factors and the dependent variable. The dependent variable may or may not be incorporated directly into the procedure that estimates the independent factors. If one is looking to estimate a missing observation for a dependent variable, such as per capita GDP for a particular country (as is the goal of this study) then this variable is not made part of the procedure.

The mathematical procedure is less likely to be influenced by the personal biases of the researcher because he or she is supposed to include as many variables as possible, allowing it to dictate how the variables are grouped together into factors. It is also usually found that a much smaller number of factors than variables is able to explain most of the variance of any of the variables included in the procedure, or later used as a dependent variable.

Critics of factor analysis point out that in reducing the original variables into a smaller number of factors, some information is lost. Specifically, when the number of extracted factors is less than the original number of variables, the factors do not explain all the variance of the original variables. Further, it is suggested that the regression coefficients obtained by factor analysis are biased (Green, 1997).

However, when the variables are closely related, simple regressions are not possible. Multicollinearity (as this problem is called) results in high standard errors and low t-statistics despite high R^2 values. This makes it difficult to determine which of the explanatory variables do in fact significantly affect the dependent variable. Factor analysis, by extracting a few linearly independent factors, provides a method to allow the information in the correlated variables to be included in the regression (Green, 1997). Furthermore, this statistical method determines the importance of each of the variables in making up factors, rather than leaving it to the personal biases of the researcher. Thus, factor analysis is an acceptable method with which to address the multicollinearity problem.

Regrettably, the number of factor *sets* that may be constructed is infinite: this follows from the fact that the number of factors is usually less than the number of variables used to construct or discover the factors. Hence, it is customary to follow a small number of standard approaches, and then consider which set of factors provides the better economic interpretation of what has been found. Since this study was mainly interested in finding a set of factors that correlated well with GDP, it allowed the most common procedure, called principal components, to dictate the construction of the factors.

As an indication of the usefulness of factor analysis, this study will first apply it to 1990 social and eco-

nomic data of the Latin American economies.[5] The approach will extract underlying factors, and use these factors to estimate 1990 per capita GDP for each market economy (but not Cuba). Later the same approach will be used to estimate the per capita GDP for Cuba.

Per capita GDP is the most often used indicator of economic development and welfare for an economy. GDP measures the market value of all final goods and services produced within the borders of a nation in one year. It is a measure of national income because income is earned from the sale of goods produced. When divided by the population, GDP measures how much income each man, woman and child would receive if all income was distributed equally.

Since higher incomes lead to higher levels of consumption, per capita GDP is also used as an indicator of economic welfare. However, per capita GDP is by no means a flawless yardstick to compare the welfare of nations. One reason is that not all goods are sold on organized markets. Another is that exchange rates for the local currency must be chosen, and these may not be market rates. A third reason is that it does not take into account the actual distribution of income. Yet, despite these faults, it has become the standard with which to measure and compare the level of development and welfare.

A basic hypothesis of this study is that the level of per capita GDP is strongly correlated with, and thus may be predicted by, the combination of 37 social and economic indicators that relate to demographics, labor market, infrastructure, education, health, energy use, the agricultural sector, and the relative importance of other sectors in the economy.[6] The analysis employs the commonly used principal components method of factor analysis to extract seven (independent) factors from these observed indicators, and then uses a linear regression to determine the correlation between these extracted factors and per capita

GDP.[7] Since about 84% of the variance of the sample can be accounted for by the factors, it is appropriate to conclude that the regression equation is a good fit; this is confirmed by an F test, which is statistically significant at the 99.9% level. The formal procedure is detailed below. See now Figure 1.

The figure above compares actual and predicted levels of per capita GDP for 1990, where the countries are sorted by the actual level of per capita GDP for each country. As can be seen, factor analysis is indeed a good method of predicting the level of per capita GDP, despite the fact that it does not use production data.

Few students of the Latin American countries would dispute that *actual* per capita GDP is a good measure of economic development of these countries, yet the *predicted* per capita GDP values of factor analysis may offer an even more accurate ranking of development for the Latin American economies. The *predicted* values place the Southern-cone countries of Argentina, Uruguay and Chile (together with oil-rich Venezuela) ahead of Mexico and Brazil in terms of economic performance and possibly development. The *predicted* values for the poorest countries may also provide a more accurate ranking of economic performance than those based on *actual* GDP.

Given these results, then, it appears that factor analysis is a natural tool to estimate the per capita GDP of relatively unregulated market economies, and that it could conceivably be used to estimate the market value of per capita GDP in command economies, as others have done (e.g. Szilágyi, 1978). This conclusion will be seriously questioned in a subsequent section, after the study reports on the main findings for Cuba.
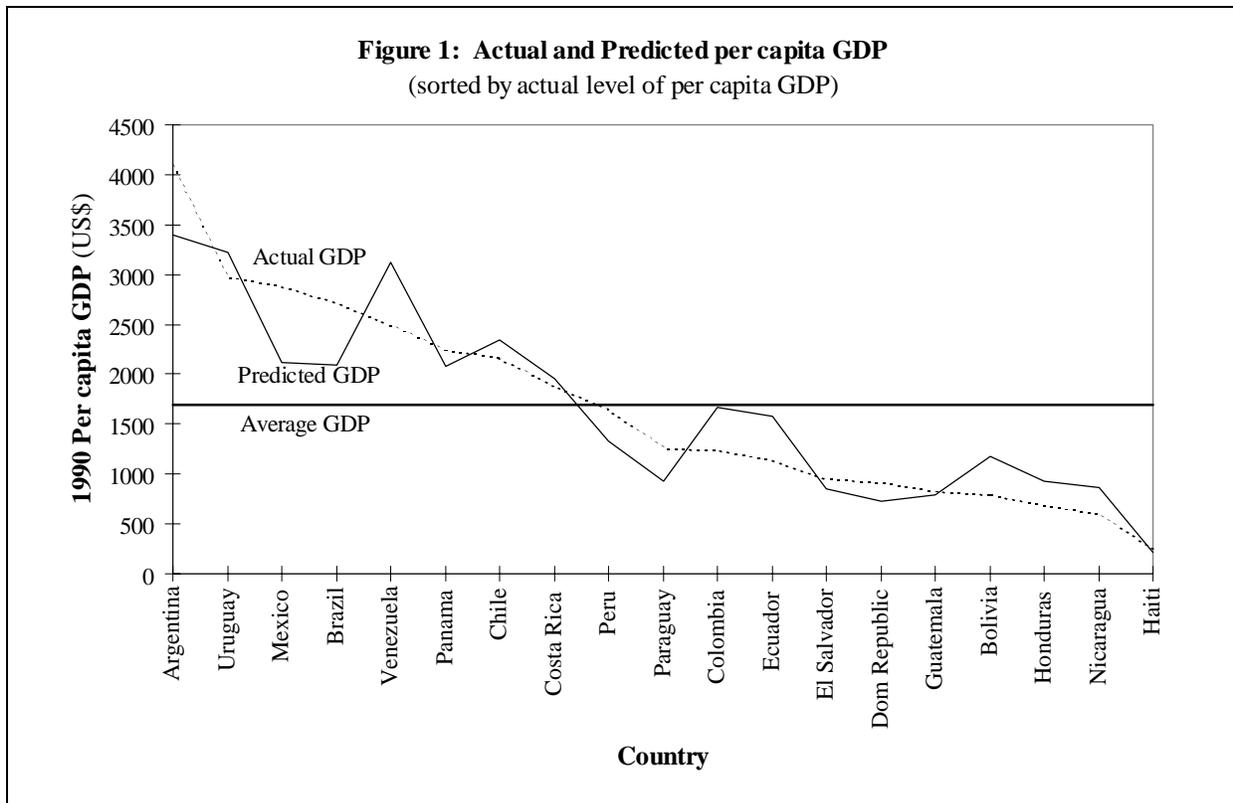
ESTIMATING CUBAN PER CAPITA GDP

This section gives an explanation of the factor analysis procedure used in this paper, and specifically how

---

5. The countries include Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, the Dominican Republic, Ecuador, El Salvador, Guatemala, Haiti, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, Uruguay, and Venezuela.

6. The list of variables is presented in the Appendix, which also includes a brief explanation of their computation.

7. Cuban data, excluding GDP estimates by other organizations, were used to extract the original seven factors.

Figure 1: Actual and Predicted per capita GDP
(sorted by actual level of per capita GDP)

Actual GDP

Predicted GDP

Average GDP

Cuban per capita GDP may be estimated using this method. Readers who are more interested in the final results, rather than in the details of the work, should move on to Table 3 and the relevant text at the end of this section.

Factor analysis was used, first, to reduce the 1990 data for 37 variables (not including per capita GDP) to seven factors in the sample of all 20 Latin American economies (including Cuba).[8] Principal components was used to extract these factors. The number of factors was chosen using a commonly accepted rule: use each factor with an eigenvalue of one or greater (for otherwise individual variables would have greater predictive power than the estimated factor). The seventh factor had an eigenvalue of 1.096, and the eighth of 0.968. A screen plot, also traditional in the literature of this procedure, confirmed that seven factors were an appropriate number. Corroboration of the effectiveness of the factor analysis was given by

the fact that the seven factors reproduced the correlation between the 37 variables well. In particular, only 18% of the residuals had absolute values greater than 0.05.

The extraction phase of factor analysis produced a component matrix, which essentially displayed the correlation coefficients of each of the seven factors for each of the 37 variables. These correlation coefficients are equivalent to regression coefficients for each of the factors regressed on the 37 variables, and are called factor loadings.

The component matrix was rotated to allow for an interpretation of the factors. Factor rotation recomputes the factor loadings by shifting the amount of variance accounted for by each individual factor. It does not change the total variance accounted for by all of the factors combined. The Varimax technique with Kaiser Normalization was used for the rotation (Norušis, 1994). This method minimized the num-

---

8. As a side note, the data was converted to standardized (z-score) form. This was done by subtracting each data point by the mean of the sample and dividing the remainder by the sample standard deviation.

Table 1.   Total Variance Explained by Rotated Factors

Component (factor)	Eigenvalue	% of Variance explained by factor	Cumulative % explained by factors
1	11.12	30.06	30.06
2	6.64	17.93	47.99
3	5.45	14.73	62.72
4	2.94	7.95	70.67
5	2.43	6.56	77.22
6	2.33	6.30	83.53
7	1.40	3.78	87.30

ber of variables that have high loadings for each factor. The rotated component matrix appears in a separate appendix which is not reported here, but is available directly from the authors upon request.

The factors appear to measure different aspects of modernity. In particular, the values of factors 1, 3, 5 and 7 are negatively associated with modernity, while the rest of the factors are positively associated with it. The variables most strongly correlated with factor 1 are demographic and health variables; factor 2 relates mainly to energy consumption; factor 3 (except for the death rate) accounts both directly and indirectly for certain inputs; factor 4 addresses population density considerations; factor 5 is correlated mainly with the construction sector; and factors 6 and 7 closely track education. Table 1 shows the percent of the variance of the sample that is accounted for by each of the factors after rotation.

Next, values for the seven factors were estimated for each of the 20 Latin American countries using a linear regression technique. The component score coefficient matrix (available in the separate appendix, not included here) displays a vector of coefficients for each factor that, when multiplied against the vector of data (for all 37 variables) for each country, produces factor scores for each country. These factor scores essentially constitute seven new variables, which are estimates of the underlying factors (or components) that explain 87.3% of the variance of the 37 original variables.

To estimate an equation that relates per capita GDP to these seven factors, a linear regression was per-

formed using data from the 19 market economies (but not Cuba). Specifically, 1990 per capita GDP was regressed on a constant and the seven series of factor scores. Since the Varimax method is an orthogonal rotation method, the factor scores are truly independent and hence not correlated with one another.

The regression results show how per capita GDP is correlated with the seven factors in market economies. The results of the regression are presented below, in Table 2. Table 2 also reports results for GDP adjusted for terms of trade (GDPTT) (from the Penn World Tables). This latter measure uses international prices for domestic absorption but current prices for exports and imports.

The predicted values derived from these regression results were those plotted in Figure 1 of the previous section. It was also from this regression that the estimate of per capita GDP for Cuba was obtained. Specifically, the regression coefficients in Table 2 were used in conjunction with the factor scores for Cuba (obtained from the original factor analysis). It should be noted that because the Penn World Tables does not report data for Haiti, those estimates were obtained without Haitian data.

Table 3, below, reports the predicted values for each of the 19 original countries and Cuba. Actual values for the 19 other Latin American countries are also provided for reference. All data has been converted to dollar values from the standardized (z-score) form.

The predicted level of 1990 per capita GDP for Cuba, nearly $2,600, is fourth highest, between Venezuela and Chile. It is almost $900 higher than the average of the countries (almost $1,700). This estimate is consistent with other estimates, including those of the CIA ($2,000 for per capita GNP) and the original estimate reported in the Human Development Report ($2,200 for per capita GDP). When GDPTT data from the Penn World Tables are used, the estimates for Cuban per capita GDPTT are even higher. Estimates using 1990 real per capita GDP in 1985 prices adjusted for changes in terms of trade (GDPTT) is $5,420, second only to Venezuela. These higher estimates are consistent with the revised

Table 2.   Regression results

Dependent variable	Per capita GDP (z-score)			Per capita GDPTT (z-score)		
Independent Variable	Regression Coefficient	t-statistic	Significance level	Regression Coefficient	t-statistic	Significance level
Constant	0.000	0.000	1.00	0.002	0.154	0.88
Factor 1	-0.593	-5.170	0.00	-0.546	-3.620	0.01
Factor 2	0.563	4.907	0.00	0.739	7.017	0.00
Factor 3	-0.220	-1.915	0.08	-0.134	-0.486	0.64
Factor 4	-0.246	-2.141	0.05	0.004	-0.239	0.82
Factor 5	0.076	0.660	0.52	-0.144	-0.953	0.36
Factor 6	0.233	2.033	0.06	0.250	1.682	0.12
Factor 7	-0.063	-0.549	0.59	-0.221	-2.239	0.05
R^2	0.842			0.893		
Adjusted R^2	0.750			0.818		
Std. Error of estimate	0.500			0.426		
F	9.135			11.928		

Table 3.   Actual and Predicted per capita GDP

Country	Predicted GDP	Actual GDP	Predicted GDPTT	Actual GDPTT
Argentina	3403	4081	5076	4663
Bolivia	1176	796	1485	1591
Brazil	2094	2715	3529	4029
Chile	2338	2170	4778	4399
Colombia	1662	1236	3122	3287
Costa Rica	1957	1881	3536	3447
Cuba	2578	—	5420	—
Dom Republic	726	912	2233	2233
Ecuador	1583	1137	2600	2669
El Salvador	847	954	2101	1760
Guatemala	787	832	1363	2138
Haiti	212	254	N/A	N/A
Honduras	933	686	1691	1401
Mexico	2125	2888	4492	5793
Nicaragua	864	599	1593	1207
Panama	2079	2249	3371	2785
Paraguay	926	1248	2315	2093
Peru	1334	1656	2448	2185
Uruguay	3222	2975	4304	4663
Venezuela	3120	2495	6172	5867

figures found in Human Development Report when the Penn World Tables were used. The use of chain-weight real GDP results in similar findings.[9]

THE WEAKNESSES OF FACTOR ANALYSIS

The GDP estimates of the previous section are remarkable in that they match the previous *relative rankings* obtained by Joglekar and Zimbalist (1989) in their study of per capita GDP for Cuba in 1980, although in their study Venezuela had come out on top among Latin American countries, at $3,621.

---

9. Specifically, estimates for Penn World Tables 1990 chain weight real per capita GDP (at 1990 prices) is $5,233 (second in the region only to Venezuela). These estimates are consistent with the revised figures found in *Human Development Report* when the Penn World Tables were used.

This country was followed by Argentina, Uruguay, Cuba, Chile, Brazil and Mexico (at $3,265, $2,815, $2,691, $2469, $2,427 and $2,224, respectively).[10] Hence, the Joglekar and Zimbalist study, together with the results of the previous section, lead to the conclusion that Cuba, at least for the decade of the 1980s, had a per capita GDP that approximated that of Chile. Or else, one has to draw the conclusion that the methodology of factor analysis is somehow in error. It should be noted that the estimates of economic output are even higher when GDPTT statistics are used, an issue that will be addressed below.

Using other "rudimentary" methods (according to their own terminology), Joglekar and Zimbalist (1989) estimated that the per capita GDP of the Cuban economy in 1980 was between the Venezuelan and Chilean figures, matching the estimates obtained in this paper. These results, then, must be confronted head on by those who have been critical of the economic performance of the Castro regime. What could possibly be wrong with consistent estimates across independent studies?

Critics of the regime have argued that Cuba did relatively well prior to the 1990s economic collapse due to massive Soviet subsidies—and these subsidies cannot be denied. They have also argued that the regime has been lax or deceitful in calculating economic and social data—and again there is some evidence to support this contention (Aguirre and Vichot, 1996). Yet even most critics of the Cuban regime accept the notion that it achieved significant goals in some fields, such as health, education and basic income security (Romeu, 1995).

With the collapse of the Cuban economy, the focus of attention among Cuban scholars has shifted to *what should be done under present circumstances.*[11] But an important question remains: is it worthwhile to save some of the characteristics of the pre-1990 economic regime, or should Cuba discard its command structure and become a market economy? To answer this question it is necessary to understand as well as possible what was *really* happening in the previous decades.

This study will not dispute the important role of subsidies in propping up the old Cuban economy or the general weakness of Cuban statistics. Rather, it will focus on the fact that (a) the GDP estimates from regression or factor analysis will generally overestimate economic performance; and, even if they do not, (b) the GDP figure is an ambiguous measure of economic welfare.

As noted earlier, the procedure used to estimate per capita GDP essentially found the correlation between the factors and per capita GDP. The seven factors did not contain output data; instead, they consisted of data relating the structure of demographics, health care, education, infrastructure, agriculture, energy, and the relative importance of different economic sectors. The implicit assumption is that countries with a given level of GDP tend to have the same demographic characteristics, level of health care, education, etc. In market economies, it is generally understood that greater levels of income lead to the demand for more public goods or goods with important public goods characteristics, higher energy use, lower birth rates, and certain shifts in economic structure (from agricultural-based to manufacturing and service-based economies).

Public goods have two major characteristics: (a) the consumption of the good by one individual does not subtract from that of others, and (b) it is costly to exclude any individual from enjoying its benefits. National defense is a pure public good, as is the legal framework within a country. A public good may be thought of as providing significant positive externalities or indirect benefits to consumers. A sewer or water system, for example, will tend to improve the overall health of the population, even when some

---

10. Given that Joglekar and Zimbalist had placed more emphasis on output measures (including steel and cement production) and that Venezuela had a strong economy in 1980 as a result of its oil production, the top ranking for Venezuela among the Latin American countries is not surprising.

11. Although many examples could be given, the article by Castañeda and Montalván (1997) is typical of this shift in interest.

people may not have direct access to it. Measures to decrease inequality may tend to reduce social conflict, and potentially make everyone better off. The success of national sports teams makes every citizen proud and hence better off. An educated citizenship will hopefully make better collective choices, which is a positive benefit even to those with limited education.

In contrast, private goods benefit only those consuming them, and it is relatively cheap to acquire exclusive rights to them. Some goods and services produced by governments are purely private in nature, such as mail delivery, but others, like education, have some significant public-goods characteristics. Although in practice it is impossible to find an exact method of dividing goods into two mutually exclusive categories, such categories are commonly used as theoretical constructs within the economics literature; for an example, see Stiglitz (1988, p. 11).

In command economies, many of the socioeconomic behaviors of people (which are associated with modernity in market economies) are not necessarily linked to the income level that the economies have achieved. For example, higher per capita GDP in market economies are associated with lower population growth and fertility rates, as families do not have to rely on children for household labor and elder care. These elements of development are contained in the regression equation of Table 2; factor 1, which is negatively correlated with these health and demographic factors, has a negative regression coefficient when used to estimate per capita GDP. In command economies, elder care and household income is determined and provided by the state, reducing the incentive to have children. Thus, the link between these demographic variables and per capita GDP in the regression equation should not apply—The regression equation is not an appropriate tool for estimating the GDP of command economies. Because command economies with low incomes generally have demographic characteristics similar to countries with high incomes, the factor analysis model is likely to overestimate the per capita GDP of command economies.

In general, in command economies, the quantities of many public goods are not necessarily correlated with

per capita GDP in the same way that they are in market economies. At any output level, the government may choose to limit or restrict the provision of private goods to emphasize the provision of either public goods or goods with important public-goods characteristics. Consider the following example. Marshall (1998, p. 287) notes that in Cuba, "one is struck by the lack of automobiles…[and] a massive six lane super-highway…is normally empty." Apparently, in Cuba, the number of road miles is not related to the number of vehicles. Generally, however, road mileage is related to GDP directly (as commerce increases, so do the number of vehicles and consequently the demand for roads) and indirectly (as the number of roads increases, commerce becomes less costly, raising GDP). In the regression results for the Latin American market economies, GDP indeed increases as road mileage increases. The large number of roads (per capita) in Cuba, then, leads to a spuriously high prediction of per capita GDP. Similar arguments can be made about health care, education, and many other variables whose proportion to GDP in Cuba is greater than in market economies.

The conclusion that must be drawn is that regression or factor analyses, based on the characteristics of market economies, cannot properly estimate the per capita GDP of a command economy. The proportion of public goods—or goods with important public goods characteristics—to private goods, and the relationship between demographic characteristics and GDP will differ significantly between market and command economies. Yet, the factor analysis technique estimated the relationship between the quantities of public goods and GDP, and demographic characteristics and GDP (among other correlations) in market economies to estimate the per capita GDP of Cuba. Since command economies provide public goods at high rates (regardless of GDP) resulting in demographic characteristics similar to higher income market countries, it is likely that GDP is overestimated in command economies.

However, even if that was not the case, one other question remains: does per capita GDP in a command economy reflect the true level of economic welfare? Economic theory is quite clear that to calcu-

late consumer surplus, a measure of economic welfare, one needs to take into account the subjective valuation that consumers place on goods. This is done by considering utility maps and the compensation that consumers must receive to freely move from one consumption bundle to another. The value of commodities to consumers depends on how well the commodities satisfy the consumers' tastes; in fact, it is impossible to derive the demand for goods without acknowledging that these functions depend on the utility or subjective values that consumers derive from commodities.

Suppose that a country produces only two types of goods: private and public goods. In a market economy of a developing nation, or at least in those that are lightly regulated, one is likely to observe a preponderance of private goods benefiting the citizens directly. This is so because the production of public goods requires some degree of centralization, which is costly to achieve, and because it is difficult to prevent anyone from consuming the public goods, giving the citizens an incentive to avoid paying for them (thereby creating the need for a coercive tax mechanism which is resented by the population). It has been observed, however, that as a society becomes more productive and wealthier, market economies produce more public goods, or at least goods with important public-goods characteristics.[12] This is not surprising, for the relative abundance of private goods that accompany development, and the relative dearth of public goods, makes the former less valuable than the latter at the margin, and the citizens are willing to trade off private for public goods. Also, the process of development itself lowers the cost of organizing a central authority. Thus, as countries develop, consumers demand and receive more public goods.

It is now possible to represent the normal growth process of a developing economy with the assistance of Graph 1. It will be assumed that consumers (but

Graph 1.
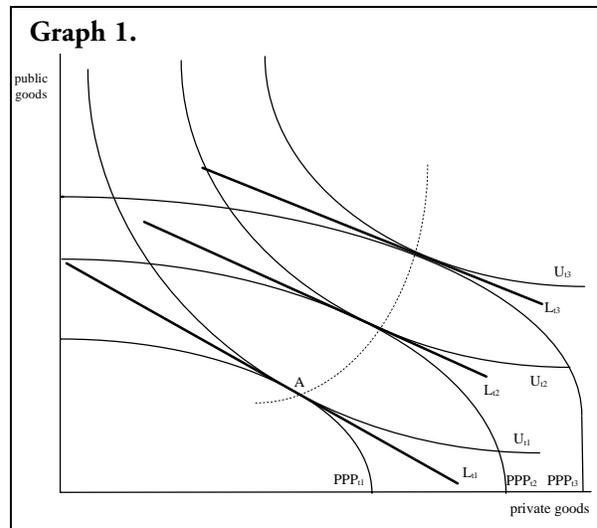
not the central planner, oligarchy or "ruling class"[13]) have identical and homothetic indifference maps, which are then aggregated in the form of community indifference curves.

The first production possibility frontier PPP_{t1} represents the productive capacity of the economy at the initial period t1, while the community indifference curve U_{t1} represents the aggregate preferences of consumers in that economy. The tangent line L_{t1} represents the price ratio between private and public goods. Initially, consumers desire a large number of private goods, making them relatively expensive.

As the growth process takes place in real time, the productive capacity of the economy is enhanced, and consumers reach higher levels of satisfaction as they move to the tangency of U_{t3} and PPP_{t3}. The broken line shows the expected relative growth of public goods, and tangent line L_{t3} (which is flatter than L_{t1}) demonstrates that private goods are becoming cheaper to obtain while public goods are becoming more expensive.

It is now possible to hypothesize a set of events that will characterize the development process of a command economy, such as has occurred in Cuba since

---

12. This has been the experience throughout most of this century, although there is now evidence that the growth of the government sector is slowing down. See Saunders (1993) and Lin (1993).

13. More on this point below.
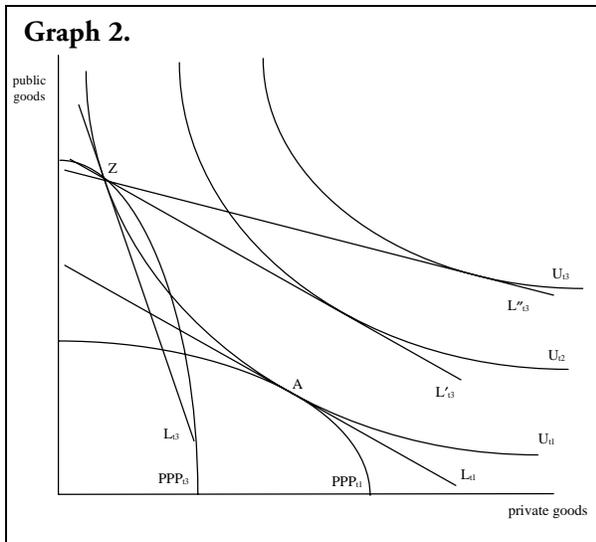
Graph 2.

1960. Starting out from the same set of initial conditions depicted in Graph 1 (with PPP_{t1} and U_{t1}), the adoption of a command economy will bring forth a dramatic restructuring of the economy in real time, leading to greater production of public goods.[14] This will happen both for ideological reasons (since the leaders will prefer public to private goods) and for economic reasons (since it will be easier to maintain the production efficiency of public goods, which are normally produced by national or regional monopolies even in market economies, than the production efficiency of private goods, which normally thrive under market competition).

This restructuring can now be observed in Graph 2, where it is represented by the shift of the production frontier to PPP_{t3}. For analytical purposes, the utility map remains the same, although it is recognized that those in control of the economy may try to influence the preferences of consumers to get them to desire a greater consumption of public goods.

Although prices are not determined in the market in a command economy, that does not mean that it is not possible to calculate the valuation that consumers place on goods when the utility map is known. Assume that A represents the initial consumption level of consumers, and that Z represents the final bundle of goods available to them (as prescribed by the au-

thorities). In this case line L_{t3} is tangent to U_{t1}, representing the subjective value to consumers of the goods that they purchase; but note that consumer welfare, by construction, has not improved at all, since consumers remain on the initial utility line U_{t1}. However, if bundle Z is evaluated at the old market prices, using L'_{t3}, or at market prices that prevail at other economies where public goods have naturally become more expensive, using line L''_{t3}, an outside evaluator will reach the conclusion that the level of output of this economy as measured by GDP has increased or kept up with that of other countries, and then infer that consumer welfare has increased. This conclusion, however, is an illusion.

It is not surprising, then, that when GDPTT statistics are used (which make corrections for the terms of trade—equivalent in developing countries to using line L''_{t3}) one will obtain even higher estimates for the value of economic output.

The typical economic and social variables that are used to compare economies at the macro level contribute to this illusion. Better education, better health and basic income security are perceived by most political leaders as goals that, if attained, provide many benefits to all citizens; but when those economic activities that lead to the achievement of these goals are emphasized, the provision of private goods will have to be neglected, for the simple reason that any society has limited resources. The absence of private markets, including those that contribute to a vigorous financial sector that helps direct the allocation of resources, will destroy the efficient production of private goods.

Neither production statistics, nor variables normally associated with development, allow the researcher to draw unambiguous inferences about consumer welfare in a command economy. It could be argued, of course, that *no one* is in the condition of making welfare comparisons: using identical and homothetic indifference maps for consumers, as is done in this paper, is a very restrictive assumption indeed. While

---

14. With one significant exception: the maintenance of environmental quality. See Sáez (1998).

there is some validity to this argument, it should be noted that (a) the use of this assumption is widespread in the literature; (b) introducing interpersonal comparability without cardinality resolves the theoretical problems associated with social welfare functions and aggregation, making social welfare comparisons possible (Sen, 1987); and (c) dividing the Cuban population between the vast majority of consumers, who may have, if not identical, at least very similar preferences, from a central planner, dictator, oligarchy or a "ruling class" that imposes its own preferences on the rest of the population, is not unrealistic.

Consumers in Cuba are not free to make their choices known, and those in control of the economy do not know what the consumers want. A quote from Moreno Fraginals (1997, pp. 213-214) goes to the heart of this observation: "But when a regime is created in the image of a few people, these few believe it to be their duty to direct others' lives from the cradle to the grave… When the capacity to direct one's own life is reduced, people either reach a point of desperation or become utterly ambivalent about their own lives." This problem becomes particularly acute when the citizens are not even given the opportunity to choose their leaders through a democratic process. The construction industry, for instance, may produce public buildings rather than private houses highly desired by consumers. Hospitals may increase their ability to admit patients, but the consumers will not have access to over-the-counter medicines that would have been sold in private markets. Schools may engage more people in the educational process, but the citizens will be limited as to the choice of books that they can purchase and the career choices open to them. New highways are built which consumers cannot use, for lack of cars and gasoline.

The emphasis on the provision of public goods can be sustained for a long time under two sets of circumstances: if the economy is subsidized from abroad, or if the capital which used to sustain the production of private goods is neglected—and there is no doubt that the capital stock for private good production has been neglected in Cuba (Pastor and Zimbalist, 1998). This has been best observed in Cuba's most important industry, sugar (Sanguinetty, 1997), but it has also affected other industries that are very important for consumer welfare: private transportation and housing (Pérez, Jr., 1988).

When individuals are free to choose their trading opportunities, including those in the labor market, they will choose those occupations that will maximize their utility. It is for this reason that few people will engage in "voluntary" (unpaid) labor for significant amounts of time, unless they are coerced to do so. When coercion is applied, the individual suffers great disutility, even if the outcome of the effort is some measured output. Measures of GDP for command economies fail to account for this disutility, whereas one can assume that labor in a market economy will be engaged in production only if the effort leads to a net gain in personal satisfaction.

The Cuban population has been restricted in its human capital investment plans. By destroying, or at the very least minimizing the size of the private goods economy, the population has failed to develop the skills that allow most populations to engage in foreign transactions, which normally involve trade in private goods. Trading itself will help increase the human capital of the population (Pissarides, 1997). This may explain the inability of the Cuban economy to adapt to producing for the global market. The one exception to this is the tourist trade, which occurs within national borders; not surprisingly, the Cuban government has turned to this industry for its economic survival.[15]

CONCLUSIONS

This study has updated estimates of per capita GDP for Cuba, for the year 1990. In relative terms, the per capita figure for Cuba is similar to that of Chile, confirming the estimates of the Central Intelligence

---

15. Another possible exception might be the biotechnology sector. This industry, however, received support from the government because of its symbolic significance (Feinselver, 1992). It is difficult to believe that these firms will be able to compete in global markets over the long run, since they disregard the patents of other companies and are bound to find legal difficulties in various nations.

Agency for that year. In absolute terms, the estimated $2,578 figure of this study is close to the original United Nations estimate of $2,200 for 1990 and $2,500 for 1989. The consistency of all three estimates cannot be ignored.

In 1960, Cuba was one of the most economically advanced countries of Latin America, matching the per capita consumption of electricity that was observed in Chile for that year (UN, 1961). Thirty years later, if one is to believe the estimates of this study, the CIA, and the United Nations, the Cuban and Chilean economies had similar levels of economic performance; their income growth would presumably indicate an improvement in consumer welfare.

This optimistic picture has been changed with the collapse of the Soviet Union. Chile's per capita GDP continues to be among the highest in Latin America, while the Cuban economic performance is beginning to match that of the poorest country of the hemisphere. This paper has argued, however, that the Cuban performance in the first thirty years of the revolution cannot be interpreted unambiguously as an improvement in economic welfare, regardless of what the GDP estimates say. This is so because once consumer preferences are taken into account, the restructuring of the economy prevents us from concluding that an improvement in economic welfare took place; Cuba was producing more public goods *but* fewer private goods.

There are also reasons to believe that indirect estimates of per capita GDP for a command economy, including Cuba, tend to overestimate economic activity because these estimates are biased by the greater availability of public goods. Market economies produce more public goods as they become wealthier, within a natural development process; command economies, on the other hand, artificially *force* the production of public goods.

It has also been noted that the restructuring of the Cuban economy has made it dramatically vulnerable to global competition: the Cuban people simply do not produce enough private goods for the international market. This may explain why, once Soviet subsidies ended, Cuba's economic performance began to approach that of Haiti.

To repeat, since Cuban authorities do not have access to their countrymen's preference maps, which are normally revealed and acted upon in market economies, it is not possible to conclude that the Cuban people were better off after the first thirty years of the revolution: relatively high GDP figures for 1990 may mean little in terms of economic welfare. Although one cannot strictly argue that the Cuban people are worse off, either, since no one else has access to those preference maps, one must reach the conclusion that any argument about economic welfare based on per capita GDP for a command economy is not compelling.

BIBLIOGRAPHY

Aguirre, Benigno E. and Vichot, Roberto. "Are Cuba's Educational Statistics Reliable?" in *Cuba in Transition* 6 (1996), pp. 371-386.

Betancourt, Roger R. "Growth Capabilities and Development; Implications for Transition Process in Cuba, " in *Economic Development and Cultural Change* 44 (1996), pp.315-31.

Carranza Valdés, Julio. "Cuba: Los Retos de la Economía," in *Cuadernos de Nuestra América* 19 (1993), pp. 131-159.

Castañeda, R. H. and Montalván, G. P. "Cuba: Una Plena Reinserción Internacional, Comercial y Financiera con Estabilización Rápida, sin Recesión y con Crecimiento Alto, Estable y Sustentable," in *Cuba in Transition* 7 (1997), pp. 99-122.

Central Intelligence Agency. *World Fact Book*. Washington: Government Printing Office, (1990-present).

Feinsilver, Julie. "Will Cuba's Wonder Drugs Lead to Political and Economic Wonders? Capitalizing on Biotechnology and Medical Exports," in *Cuban Studies* 22 (1992), pp. 79-111.

Green, William H. *Econometric Analysis*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 1997.

Joglekar, Gitanjali and Zimbalist, Andrew. "Dollar GDP per Capita in Cuba," in *Journal of Comparative Economics* 13 (1989), pp.85-114.

Kline, Paul. *An Easy Guide to Factor Analysis*. New York: Routledge, 1994.

Lin, David. "Recent Trends in the Size and Growth of Government in Developing Countries, in *The Growth of the Public Sector*, ed. By Norman Gemmell. Brookfield, VT.: Edward Elgar, 1993.

Marshall, Jeffery H. "The Political Viability of Free Market Experimentation in Cuba: Evidence from *Los Mercados Agropecuarios*," in *World Development* 26 (1998), pp. 277-288.

Mesa-Lago, Carmelo. "The State of the Cuban Economy: 1995-96" in *Cuba in Transition* 6 (1996), pp. 4-7.

Moreno Fraginals, Manuel. "Transition to What?" in *Toward a New Cuba? Legacies of a* Revolution, ed. by M. A. Centeno and M. Font. Boulder, CO: Lynne Reinner Publishers, 1997.

Mulaik, Stanley A. The *Foundations of Factor Analysis*. New York: McGraw Hill, 1972.

National Bureau of Economic Research, *Penn World Table*, <http://www.nber.org/pwt56.doc.html> 1998.

Norušis, Marija J. *SPSS, SPSS Professional Statistics 6.1*. Chicago, IL: SPSS Inc., 1994, pp. 47-82.

Pastor, Jr., Manuel and Zimbalist, Andrew. "Has Cuba Turned the Corner? Macroeconomic Stabilization and Reform in Contemporary Cuba," in *Cuban Studies* 27 (1997), pp. 1-20.

Pérez, Jr., Louis. *Cuba, Between Reform and Revolution*. New York: Oxford University Press, 1988.

Pérez-López, Jorge F. "The Cuban Economy in the Age of Hemispheric Integration," in *Journal of Interamerican Studies and World Affairs* 39 (1997), pp. 3-47.

Pissarides, C. A. "Learning by Trading and the Returns to Human Capital in Developing Countries," in *The World Bank Economic Review* 11 (1997), pp.17-32.

Ritter, Archibald R. "The Cuban Economy in the Mid-1990s: Structural/Monetary Pathology and Public Policy, in *Toward a New Cuba? Legacies of a Revolution*, ed. by M. A. Centeno and M. Font. Boulder, CO: Lynne Reinner Publishers, 1997.

Romeu, Jorge Luís. "More on the Statistical Comparison of Cuban Socioeconomic Development," in *Cuba in Transition* 5 (1995), pp.293-301.

Sáez, Hector. "Resource Degradation, Agricultural Policies, and Conservation in Cuba," in *Cuban Studies* 27 (1997), pp.40-67.

Sanguinetty, Jorge A. "The Structural Transformation of the Cuban Economy: A Report of the Last Twelve Months," in *Cuba in Transition* 7, 1997, pp. 8-12.

Saunders, Peter. "Recent Trends in the Size and Growth of Government in OECD Countries," in *The Growth of the Public Sector*, ed. by Norman Gemmell. Brookfield, VT: Edward Elgar, 1993.

Sen, Amartya. "Social Choice," in *The New Palgrave Dictionary of Economics* IV, ed. by John Eatwell *et. al.* London: Macmillan Press, 1987, pp.382-393.

Stiglitz, Joseph E. *Economics of the Public Sector*. New York: W. W. Norton and Company, 1988.

Szilágyi, G. Y. "Factor-Analytical Comparison of Economic Level and Structure," in *Acta Oeconomica* 21 (1978), pp.379-403.

United Nations. Economic Commission for Latin America and the Caribbean. *Anuario Estadístico de América Latina y el Caribe*. Chile (1990-present).

United Nations Development Programme, *Human Development Report*. New York: Oxford University Press (1991-present).

United Nations. *Statistical Yearbook*. New York: United Nations, 1961.

Wilkie, James W. (Ed). *Statistical Abstract of Latin America*. Los Angeles: UCLA Latin American Center Publications (1995-present).

World Bank, *Social Indicators of Development*. <http://www.ciesin.org/charlotte> 1998.

Yotopoulos, Pan A. and Nugent, Jeffrey B. *Economics of Development; Empirical Investigations*. New York: Harper and Row, 1976.

Appendix
DESCRIPTION OF DATA

AgeDependency—age dependence ratio:(pop. age 0-14 + pop. age 65-)/(pop. age 15-64)

Autos—number of automobiles per 1000 population (1991)

Births—births per 1000 persons, average for 1985-1990

Deaths—deaths per 1000 persons, average for 1985-1990

Dentists—number of dentists per 1000 persons (1993)

LaborParticipationFemale—percent of females age 10+ in labor force (employed or unemployed)

LaborParticipation—percent of total persons age 10+ in labor force

Electricity—electrical energy consumption in kilowatt hours per person

ConstructionEmploy—number of workers in construction sector as a % of total labor force (1994)

FinanceEmploy—number of workers in financial sector as a % of total labor force (1994)

AgEmploy—number of workers in agricultural sector as a % of total labor force (1994)

TradeEmploy—number of workers in trade sector as a % of total labor force (1994)

Energy—total energy consumption in kg of oil equivalent per person

AgSector—% of output produced by agricultural sector

ConstructionSector—% of output produced by construction sector

ManufactureSector—% of output produced by manufacturing sector

TradeSector—% of output produced by trade sector

HospitalBeds—number of hospital beds per 100,000 persons (1993)

Hydrocarbon—hydrocarbon (oil and natural gas) consumption in kg of oil equivalent per person

InfantMort—infant mortality rate: (children age 0-1 deaths/1,000 live births) (1985-90)

Irrigate—% of lands irrigated (1989-94)

Tractors—hectares of land per tractor

Newspapers—newspaper circulation per 1000 persons

Physicians—physicians per 1000 persons (1993)

PopAgeUnder15—population 0-14 as a percent of total population

ChildFemaleRatio—population 0-4 as a percent of female population aged 15-44

PopAgeOver65—population 65- as a percent of total population

PopDensity—population density (persons per square mile)

PopGrowth—average population growth rate (1985-90)

Railroads—1000 persons per mile of rail (1991)

Telephones—persons per telephone lines

Vehicles—persons per vehicle (autos, trucks, busses) (1991)

Roads—road miles per 1000 persons (1991)

St/TeacherPrimary—student teacher ratio at primary level of schooling (1991)

St/TeacherSecondary—student teacher ratio at secondary level of schooling (1991)

CollegeStudents—students in tertiary level of education per 100,000 persons (1992)

Fertility—total fertility rate

Source: NBER, 1998; UN, 1990-present; Wilkie, 1990-present; World Bank, 1998. Unless otherwise noted, all data is for 1990.