ECONOMIA

Volume XIX □ **Número 3** □ **Outubro 1995**

Carlos Robalo Marques
Paulo Soares Esteves

Portuguese GDP and its deflator before 1947: a revision of the data produced by Nunes, Mata and Valério (1989)

Separata

Portuguese GDP and its deflator before 1947: a revision of the data produced by Nunes, Mata and Valério(1989)

CARLOS ROBALO MARQUES
PAULO SOARES ESTEVES

Abstract: Thit paper presents new estimates for the Portuguese GDP at current prices and its deflator, that are immune to the criticisms raised in Esteves(1993). The new figures suggest that the data produced by Nunes, Mata and Valério(1989) overestimate both the GDP series and its deflator.

1 - Introduction

Nunes, Mata and Valério(1989) (hereafter, NMV(1989)) have constructed long series for the Portuguese economy that go back to 1833, for GDP at product market current prices and its deflator. Soon after their publication, these data started to be used in many different empirical studies (see, for instance, Correia, Neves and Rebelo(1992), Crato(1992) and Oxley(1993)). However, recently, Esteves(1993) argued that GDP deflator, as estimated by NMV(1989), is inconsistent with the so called "surrogate index of the cost of living", which has been used by the authors to generate the former series. This inconsistency stems from the fact that the estimated relationship between the two variables lacks some important properties. Let PY and PC be the GDP deflator and the cost of living index, respectively. In order to compute the GDP deflator for the period 1833-1946, the authors, using the generalized least-squares (GLS), estimated the following regression

$$\ln PY = -1.417 + 0.816 \ln PC \qquad (1)$$
$$R^2=.994 \quad \sigma=3\ o\% \quad DW=1.43$$

for the period 1947-1981, and assumed that this regression was also valid for the period 1833-1946. Esteves(1993) raised the point that the estimated elasticity between the two series is smaller than unity (0. 816). This fact implies that the relative price, In(PC/PY), is always increasing, so that the two series diverge in the long-run and, consequently, the GDP deflator is seriously overestimated. For the series to be consistent, the long-run elasticity between them must be equal to unity, that is, the two series must exhibit a parallel evolution in the long run, even though there may be temporary deviations in the short-run.

Furthermore, Esteves(l 993) argues that there are some reasons to suspect that the constructed GDP series at current prices may also be overestimated. This is because when GDP at current prices is deflated using the cost of living index (instead of the GDP deflator), one gets the unacceptable result that the real GDP per capita, in Portugal, during the second half of the XIX century, does not differ much from the one that has been estimated for England, in the prosperous Victorian years.

For these reasons, Esteves(1993) concludes that the construction of data series for GDP and its dellator, ror the Portuguese economy, remains an issue that deserves further investigation. This is exactly the subject of the present paper. Resorting to econometric models, more elaborated than the ones utilized by NMV(1989), we compute new estimates for GDP at current prices and its deflator. The GDP deflator is obtained using a model that exhibits long-run unitary elasticity, so that the inconsistency mentioned above is eliminated. The GDP series at current prices, which in NMV(1989) was computed using a simple static model based on exports, fiscal receipts and public expenditure, is now computed through a dynamic regression that includes exports, imports and public expenditure, as explanatory variables. The fiscal receipts were excluded because they were not significant in the estimated regression.

The new series computed in this paper start only m 1867 (instead of 1833), because data on exports, imports and public expenditures are available, on a continuous basis, only after 1865 . We think that the simplifying assumptions made by NMV(1989), in order to prolong the series back to 1833, are so strong that the data they obtained suffers from lack of content. That is why we decided not to carry out a similar exercise.

This paper is organized as follows. Section 2 discusses the estimation of GDP at current prices and its deflator and presents the models used to build the new series. Section 3 compares the series computed in section 2 with the ones estimated by NMV(1989). Section 4 reports the main conclusions.

2 - Computation of new estimates for GDP and its deflator

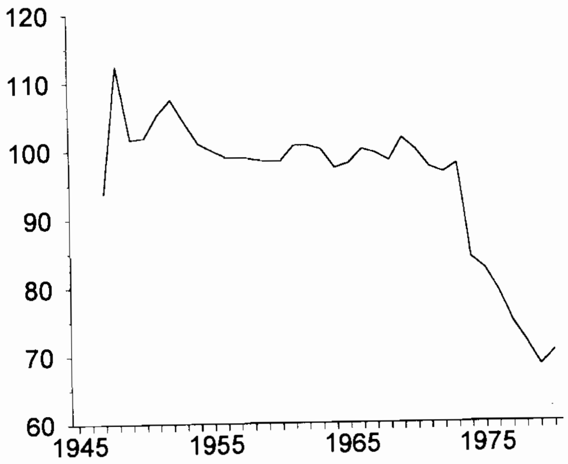
It is important to mention that the data used in this section to estimate the models shown below are of NMV(1989) and the approach underlying the construction of the two series is also similar to the one of NMV(1989). Let us first begin with the estimation of the GDP deflator.

2.1 - Estimation of GDP deflator

As mentioned above, NMV(1989) computed the GDP deflator series using model (1). One difficulty with model (1) is that the classical results for the properties of the estimators are not applicable, because InPY and InPC are not stationary variables. Another problem has to do with the estimation period used by NMV(1989): 1947-1981. The period after 1972, for reasons that are very well known (the oil price shock in 1973 and the Portuguese revolution in 1974) is characterized by important structural changes. For this reason, it seems wise to exclude the data after 1972 from the estimation period. Otherwise, the parameters of the estimated models would be distorted and, so would be, the figures of the series we want to compute.

Figure I depicts the ratio PY/PC. It is apparent that this ratio is basically constant wntil 1973, exhibits an mcreasing trend from that date to 1979 and starts to decrease in 1980. Thus, one should not be swprised to get distorted estimated coefficients if the observations after 1972 are included in the estimation period. For example, if we use the period 1947-1972 to estimate model (1), either by ordinary least-squares(OLS) or by GLS, the elasticity increases from 0.82 to 0.88 and to 0.87, respectively. These results show how dependent, on the estimation period, is the GDP deflator computed by NMV(1989).

Figure 1 - PY/PC (1963=100)



To overcome the inconsistency pointed out in Esteves(1993), one needs to specify a dynamic model for PY. However, this model must be able to make predictions for the past, that is, to make backcasts, and so, it cannot include the lagged endogenous variable as a regressor. To see why, let us take the following simple dynamic model

$$\ln PY_t = \alpha_0 + \alpha_1 \ln PY_{t-1} + \alpha_2 \ln PC_t + \varepsilon_t \qquad (2)$$

After estimating model (2), in order to generate figures for InPY prior to the estimation period, we have to resort to the following inverted model

$$\ln PY_{t-1} = -(\alpha_0/\alpha_1) + (1/\alpha_1) \ln PY_t - (\alpha_2/\alpha_1) \ln PC_t \qquad (3)$$

in which the figures for lnPY are computed recursively. The difficulty with model (3) is that it is not stable. Under the usual asswmption that $O<\alpha_1<l$ in model (2), we must have $(1/\alpha_1>1$ in (3). So, model (3) is unstable and thus not appropriate to backcast.

This difficulty may however be overcome if we resort to a finite approximation of the infinite distributed lag model, implicit in model (2). Let us assume, as the starting point, the general case of the so called autoregressive distributed-lag model

$$A(L) \ln PY_t = c + B(L) \ln PC_t + \varepsilon_t \qquad (4)$$

where A(L) and B(L) are polynomials in the lag operator L, of finite order. Asswme also that the long-run multiplier is equal to one, that is, [B(1)/A(1)]=1. As it stands, model (4) cannot be used to backcast, as we have seen. However, it may be written in the so called transfer function form

$$\ln PY_t = c/A(1) + [B(L)/A(L)] \ln PC_t + u_t \qquad (5)$$

where B(L)/A(L) is a potentially infinite polynomial, m the lag operator. In practice, however, this polynomial can always be approximated by a polynomial of finite order, giving rise to a model of the following type

$$\ln PY_t = d + C(L) \ln PC_t + u_t \qquad (6)$$

where the order of C(L) is determined by the data and by the condition that $C(1) \approx B(1)/A(1) = 1$. This has been the strategy utilized in the paper.

For the GDP deflator we started with the following general distributed lag model

$$\ln PY_t = d_0 + c_0\Delta^2\ln PC_t + c_1\Delta^2\ln PC_{t-1} + c_2\Delta^2\ln PC_{t-2} + c_3\Delta^2\ln PC_{t-3} +$$
$$+ c_4\Delta^2\ln PC_{t-4} + c_5\Delta^2\ln PC_{t-5} + c_6\Delta\ln PC_{t-1} + c_7\ln PC_{t-1} \qquad (7)$$

which is equivalent to a model in levels with seven lags in InPC. Model (7) is reparameterized in order to reduce multicollinearity among the regressors. Fwther, it has the advantage of producing a direct estimate of the long-run multiplier, which is given by C7, the coefficient of the variable in levels.

When we estimate model (7), using data for the period 1954-1972 (to allow for the lags in lnPC,), we obtain an estimated long-run multiplier equal to 1.035, which is statistically not different from one. Imposing this restriction on the model and dropping the nonsignificant coefficients, one ends up with the following model (estimated by OLS)

$$\ln PY_t = 0.841\Delta^2 \ln PC_t + 0.605\Delta \ln PC_{t-1} + \ln PC_{t-1} \quad (8)$$
$$(6.45) \qquad (7.76)$$

R^2=.855[1] σ=1.07% DW=1.4 Q(6)=4.3 (P(6)=0.64)

Model (8) exhibits long-run wlitary elasticity, and thus, does not suffer from the inconsistency problem raised in Esteves(1993). Furthermore, the fit of model (8) is significantly better than that of model (1), the standarderror of the regression being three times smaller than that of model (1). In model (8) the two series, InPY and InPC are expressed in the same base year: 1963=100 (this explains why the intercept term twned out to be non significant). Of course, changing the base year in a log-log model only the estimate of the intercept changes, but in this case, we decided to use both series in the same base year to make the computation of the GDP deflator easier. Using directly equation (8) we computed the GDP deflator up to 1946.

2.2 - Estimation of GDP at current prices

To estimate GDP at current prices up to 1946, NMV(1989) started by estimating a static regression of the following type

$$\ln Y_t = \alpha_0 + \alpha_1 \ln X_t + \alpha_2 \ln M_t + \alpha_3 \ln F_t + \alpha_4 \ln G_t + u_t \quad (9)$$

where Y, X, M, F e G stand for GDP at cwrent prices, exports, imports, fiscal receipts and public expenditure, respectively. All variables were taken at current values. After estimating model (9) for 19471985, the authors concluded that imports were not relevant, and so they selected the following model

$$\ln Y_t = 3.455 + 0.136 \ln X_t + 0.447 \ln F_t + 0.273 \ln G_t \quad (10)$$

R^2=0.999 σ=2.3% DW=1.62

1 Notice that the reported R^2 is relative to the variable ($\ln PY_t$-$\ln PC_{t-1}$) and, so, it does not compare with the R^2 of equation (1). Besides, the vanables of the regression being non stationary, the R^2 is not appropriate to measure the model's ability to fit the data (see, forinstance, Wooldridge (1991)). Q(6) represents the Ljung-Boxstatistic for 6 lags and P(6) the marginal significance level of Q(6).

estimated by GLS, which was utilized to compute the estimates of GDP at current prices, up to 1946.

When one tries to replicate equation (10) using the same data set and GLS, one obtains the following regression

$$\ln Y_t = 3.458 + 0.147 \ln X_t + 0.427 \ln F_t + 0.281 \ln G_t \quad (11)$$

whose coefficients are very close to the ones exhibited by equation (10).

Similarly to GDP deflator, the utilization of model (10) presents several problems. First, it is a simple static regression with non stationary variables, so that classical inference does not apply. In particular, the t-statistics in this regression are not valid and, consequently, is not a good strategy to decide wether or not to include a variable in a regression based on that information. Secondly, it is well known that in this kind of static regressions, even though superconsistent (if the variables are cointegrated) the OLS estimators are biased in finite samples and this bias is, sometimes, very significant[2]. Finally, but even more important in the present situation, specifying a dynamic model instead of a static one, increases the probability of getting a model that fits the data much better and so to produce more reliable figures for GDP. For these reasons, we have decided to start with a dynamic model including all the potential relevant variables (except, of course, the lagged endogenous variable). For the same reason presented in the previous section, we have decided to exclude from the estimation period the observations after 1972. This model is a generalization of model (7) with four explanatory variables, but with only two lags in each, in order to preserve degrees of freedom. The final selected specification is in equation (12). It does not include the fiscal receipts because they turned out to be non significant. Note also, that contrary to NMV(1989), the model includes imports among the regressors.

$$\ln Y_t = 2.590 + 0.104\Delta^2 \ln X_t + 0.214\Delta^2 \ln M_t + 0.247 \ln X_{t-1} +$$
$$(28.3) \qquad (3.25) \qquad (5.95) \qquad (7.31)$$
$$+ 0.371 \ln M_{t-1} + 0.301 \ln G_t \quad (12)$$
$$(7.18) \qquad (8.21)$$

2 Using GLS does not solve the problem. It is possible to demonstrate (see Park and Phillips (1988)), that OLS and GLS are asymptotically equivalent for cointegrating regressions with variables integrated of order one and where the residuals follow an autoregressive process and the explanatory variables are strictly exogenous.

$R^2 = 0.9993 \quad \sigma = 1.61\% \quad DW = 1.65 \quad Q(6) = 5.1\ (P(6)=0.53)$

Estimation period: 1949-1972

Model (12) exhibits a much better fit (the standard-error of the regression is significantly smaller) than model (10), which was selected by NMV (1989). Equation (12) was utilized to estimate GDP at current prices up to 1946. Both series - GDP deflator and GDP at current prices - are presented in the appendix.

3 - Comparing the results with the NMV(1989) series

Our estimations for the period 1867-1946 confirm our initial conjecture that NMV (1989) overestimate both the GDP and its deflator. As can be seen in figure 2, the NMV(1989) GDP price is always above our estimation, which is an immediate consequence of the elasticity between the GDP deflator and the cost of living index that they estimated in equation (1). In fact, given that their elasticity is lower than one and that the index of cost of living has a positive trend, the ratio PY/PC is increasing when we go back in time. In contrast, our estimation is compatible with a common long-run evolution between the two price measures.

Figure 2 - PY/PC (1963 = 100)

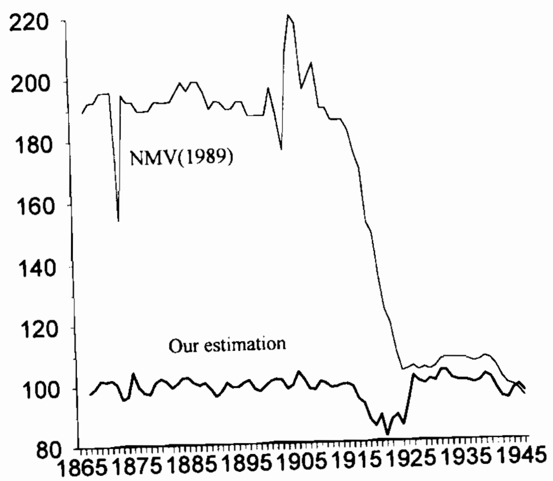


Figure 3 - Nominal GDP (NMV (1989) / Our estimation)

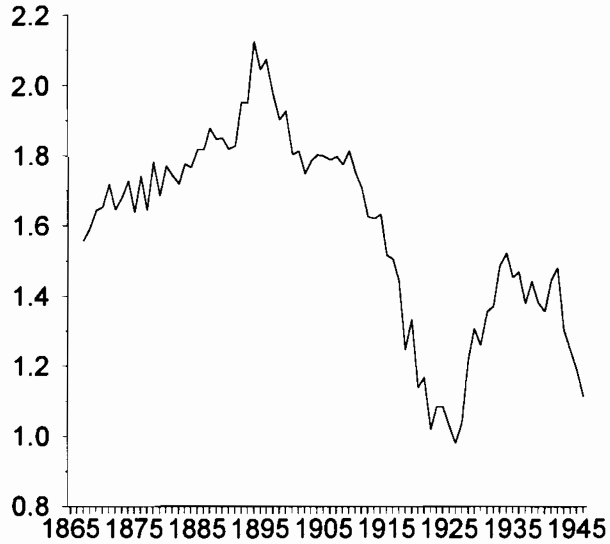


Figure 3 compares the two estimates of the nominal GDP. With the exception of the middle twenties, NMV(1989) estimates are clearly higher than ours. This deviation seems to confirm the suspicion that their nominal GDP could be overestimated, which is justified by the unrealistic results reached when the index of cost of living is used as deflator of GDP (see table 1). With our estimates, given that production and consumption prices have the same long-run evolution, this inconsistency does not arises.

Despite these differences in estimated nominal GDP and its deflator, our results for the real GDP are not very distant from the ones of NMV (1989), suggesting that the deviations are offset when we consider the ratio between the two variables. Table 1 and Figure 4 compares several estimations of the real per capita GDP. Ourresults are close to those of NMV(1989) and of Bairoch (1976). The main difference appears in the period after 1920, where our values point to a slower economic growth.

Figure 4 - Real per capita GDP (contos, 1960 prices)

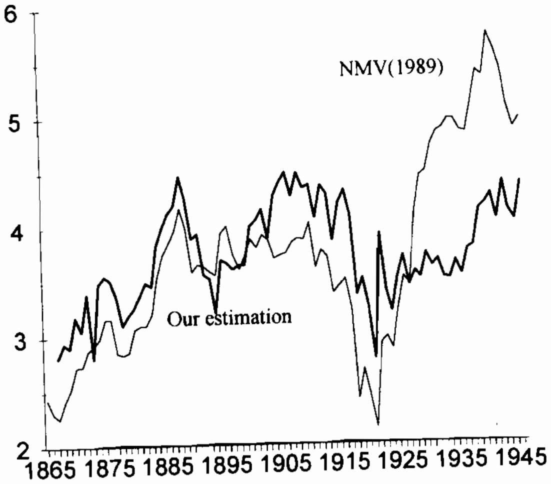


Table 1 - Relative per capita GDP between Portugal and the UK (UK=100) (*)

	1880([a])	1910	1938
Bairoch (1976)	39.7	32.1	29.7
NMV(1989)	33.8	34	33 1
Using PC as deflator	65.8	65.5	35 9
Ourestimation	36.8	36.9	24 5
Using PC as deflator	36.8	36.9	24.5

(*) in 1960 PPP exchange rates and prices.
([a]) three-year annual averages.

4 - Concluding remarks

Among other results, NMV (1989) presented estimations of the nominal GDP and its deflator for the period before 1947. These estimations have

been used in many empirical research that required long-data beyond published official figures. However, as pointed out in Esteves (1993), there are reasons to believe that both nominal GDP and its deflator could be overestimated. First, these results produce a divergence between the observed index of cost of living and the estimated production price. Second, when we consider the cost of living index as the deflator of nominal GDP, we find the unrealistic result that the Portuguese per capita GDP was not very distant from that of the United Kingdom in second half of the XIX century.

In this paper, using the same approach of NMV(1989), we compute alternative data through a more consistent estimation process and we conclude that our inital suspicions seem to be confirmed. Both nominal GDP and its deflator computed by NMV(1989) are likely to be overestimated. But, despite these differences in each individual series, the two estimates for the real GDP do not seem very distant, with the exception of the 1920-1946 period where our estimates point to a slower economic growth.

Statistically, our estimates are more reliable than those of NMV(1989). However, we leave to the experts the critical appraisal of our data from an historical point of view.

References

Bairoch P., 1976, "Europe's Gross National Product: 1800-1975", The Journal of European Economic History, Vol. 5, 2, 273-340.

Correia I.H., Neves J.C. e Rebelo S., 1992, "Business cycles in Portugal: theory and evidence", The Portuguese Economy towards 1992, ed. J.F.Amaral, D.Lucena, A. S.Melo, Kluwer Academic Publishers.

Crato N., 1992, "Persistence in Portuguese economic activity", Estudos de Economia, Vol. XII, n° 4.

Esteves P. S., 1993, "Portuguese prices before 1947: inconsistency between the observed cost of living index and the GDP price estimation of Nunes, Mata e Valério (1989)", Banco de Portugal, 1993, Working Paper 19/93.

Nunes A., Mata E., Valério N., 1989, "Portuguese economic growth 1833-1985", The Journal of European Economic History, Vol. 18, 2, 291-230.

Oxley, L., 1993, "Cointegration, causality and export-led growth in Portugal", 1865-1985, Economics Letters, 43, 163-166.

Park J.Y., Phillips P.C.B., 1988, "Asymptotic equivalence of ordinary least squares and generalized least squares in regressions with integrated regressors", Journal of The American Statistical Association, Vol. 83, n° 401, 111-115.

Wooldridge J. M.,1991, "A note on computing r-squared and adjusted r-squared for trendig and seasonal data", Economics Letters, 36, 49--54;

Appendix

	Nominal GDP 103 contos	GDP price (1914=100)	Real GDP		Nominal GDP 103 contos	GDP price (1914=100)	Real GDP
1867	231.2	84.7	272.9	1907	502.4	89.7	559.9
1868	229.7	80.0	287.1	1908	523.1	88.5	591.8
1869	231.1	81.0	285.4	1909	518.8	89.9	577.2
1870	237.1	75.4	314.4	1910	540.8	92.3	586.2
1871	224.2	73.8	303.6	1911	530.2	96.0	552.5
1872	247.3	72.9	339.4	1912	575.6	97.2	592.4
1873	257.7	91.4	282.0	1913	585.9	100.8	581.1
1874	267.6	75.8	352.9	1914	527.7	100.0	527.7
1875	300.5	83.0	361.8	1915	643.3	110.9	580.2
1876	287.9	80.0	359.7	1916	779.1	129.9	599.8
1877	309.5	89.4	346.2	1917	864.5	151.3	571.5
1878	283.6	87.9	322.4	1918	1194.3	257.7	463.4
1879	292.0	87.3	334.3	1919	1384.4	286.8	482.8
1880	283.6	82.3	344.8	1920	2306.7	517.8	445.5
1881	300.6	83.5	359.9	1921	2884.3	747.7	385.8
1882	311.2	82.8	375.8	1922	5303.2	964.5	549.9
1883	299.6	79.8	375.2	1923	7427.5	1544.5	480.9
1884	315.5	75.2	419.5	1924	9455.4	2058.1	459.4
1885	318.1	72.2	440.7	1925	11037.9	2159.3	511.2
1886	334.7	72.8	459.8	1926	12320.9	2256.6	546.0
1887	331.6	70.2	472.2	1927	12192.0	2381.5	512.0
1888	358.4	70.9	505.9	1928	12131.0	2259.4	636.9
1889	367.9	75.4	487.8	1929	12726.6	2383.6	533.9
1890	373.4	83.2	449.0	1930	12942.4	2252.7	574.5
1891	372.5	81.5	456.9	1931	11651.4	2064.3	564.4
1892	358.1	85.9	417.0	1932	11756.7	2022.5	581.3
1893	360.3	86.7	415.7	1933	11063.0	1968.6	562.0
1894	337.5	88.7	380.6	1934	11225.2	1975.9	568.1
1895	373.1	84.8	440.o	1935	11902.8	1977.5	601.9
1896	379.4	86.5	438.8	1936	11931.6	2024.1	589.5
1897	396.0	91.1	434.7	1937	13272.4	2088.2	635.6
1898	416.3	94.1	442.3	1938	13253.9	2039.8	649.8
1899	418.3	93.9	445.6	1939	14054.5	1961.7	716.5
1900	459.9	93.3	492.9	1940	14916.4	2033.9	733.4
1901	447.0	88.7	503.9	1941	16579.2	2204.4	752.1
1902	466.7	89.5	521.4	1942	18751.0	2613.3	717.5
1903	472.0	95.8	492.8	1943	23030.9	2933.0	785.2
1904	476.4	87.0	547.7	1944	23275.8	3119.7	746.1
1905	480.5	84.5	568.5	1945	25389.5	3445.2	737.0
1906	488.9	83.7	584.2	1946	30991.1	3841.5	806.7