(9 January 2004, revised version; comments are welcome)

Some comments on cross-sectional price-value deviation measures

Dr. Emilio Díaz
Departamento de Economía Aplicada II, University of Seville
Avenida Ramón y Cajal 1
28018 Seville
Voice +34 954 55 44 84
emidiaz@us.es

Dr. Rubén Osuna
Departamento de Análisis Económico I, UNED
Paseo Senda del Rey 11
28040 Madrid
Voice +34 91 398 78 19
rosuna@cee.uned.es
http://www.uned.es/dpto-analisis-economico1/fichprof/rosuna/rosuna(es).htm

Key words:

JEL classifications:

Abstract

Recent debates on empirical price-value relations have posed some important objections to the so-called *strong correlation hypothesis*, first proposed by Shaikh (1984). This concept implies the existence of a high degree of correlation between different relative price systems for both time-series and cross-sectional data. Particularly in the last case, Kliman's (2002 and 2003) critique based on the spurious correlation affecting empirical tests and Steedman and Tomkins' (1998) critical discussion on limitations of vectorial deviation measures imply a revision of the literature on the subject. In this paper we show that traditional formal procedures in analyzing cross-sectional price-value correlations are extremely dubious due to severe specification problems. We also show that frequently used geometric deviation measures are not capable of overcoming the problems raised by the arbitrary selection of both normalisation criteria and physical units of measurement.

1. Introduction

Since the seminal work of Anwar Shaikh (1975), a growing body of literature on empirical relations among prices and labour values has proliferated. Two of the most important developments in this extensive line of research were a detailed revision of the theoretical foundations of the National Accounts by incorporating several classical insights (see Shaikh and Tonak, 1994), and the implementation of formal procedures aimed at testing the empirical validity of Ricardian and Marxian theories of value, from Shaikh (1984) and Ochoa (1984) to Tsoulfidis and Maniatis (2002)[1].

In the last case, the strategy has been to test the existence of a substantial amount of correlation among sectoral market prices and labour values –measured in monetary terms, using both cross-sectional and time-series data. This is known as the test of *strong correlation hypothesis*. The testing procedure usually deals with a set of linear or log-linear simple regressions, the null hypothesis being a slope coefficient equal to one and an intercept equal to zero, in the context of supposedly well-behaved and small-sized error terms.

Although the data seemed to support this hypothesis, from the outset there were several objections to the employment of cross-sectional log-lineal regression models in measuring price-value deviations (see Petrovic 1987 and Ochoa 1989). In fact, as we will show, there are even more fundamental reasons for a complete rejection of testing procedures based on regression models. As a consequence, the only cross-sectional exercise that appears to remain meaningful in the analysis of relations among different prices systems is that of the vectorial deviation measures. However, Steedman and Tomkins (1998) critical paper on those measures pointed to two unavoidable problems: the normalisation criterion selection, and the arbitrary adoption of a concrete set of physical units of measurement. On this ground, only *dispersion* measures of price-value ratios –like the coefficient of variation— appear to have an unambiguous interpretation.

In any case, the results obtained (strong correlation among sectoral aggregates and moderate relative vectorial price-value deviations) have been interpreted either as a corroboration of a Ricardian flavoured theory of labour value –with labour values as mere proxies to prices- or as evidence in favour of the Marxian theory of value –with labour values "explaining" or "causing" prices in some sense.

In a recent paper Kliman (2002) has presented an interesting critique to these results. The core argument is based on the idea that traditional procedures inevitably fails in correctly measuring cross-sectional correlations between prices and values due to a *spurious correlation problem*: big sectors will show both high sectoral values and high sectoral prices for its output, and the contrary will be the case for the small ones. Following Kliman (2002, p. 302), the problem may only be overcame by focusing on the very notion of unitary prices and values and that requires accounting for industry size influences, a point largely absent in the literature (for a more detailed discussion, see Osuna 2003). However, a careful analysis of Kliman's procedure led us to the conclusion that the results obtained from his proposed cost-deflation, aimed at controlling for industry size, are not meaningful because it shares theoretical and empirical flaws with the traditional approach.

In this paper we will analyse *four* related problems that affect the bulk of the literature: a problem of misspecification of the error terms, a problem of parameter instability, a problem of spurious correlation, all of them aggravated by a problem of aggregation. In the first section we adopt a restrictive assumption in order to make a separate analysis of the three specification problems. Therefore, we will start dealing with a theoretical model of fully disaggregated industries, where the aggregation problem (the aggregation of different, non-homogeneous commodities in the same 'sector') does not appear. Throughout the second section that assumption will be relaxed, and the aggregation problem will make its appearance, although in the context of the spurious correlation problem. The effect of some of these problems on the deviation measures widely employed by the literature will be analysed briefly in the last section.

2. The log-linear regression model

Anwar Shaikh has suggested that «a natural form for cross-sectional … hypotheses is that empirical correlations between relative prices and relative values is log linear» (1984, p. 70). Let us suppose that each «industry» produces only one physically homogeneous commodity, and that quantities and prices of such commodities are observable magnitudes. Thus, the Shaikh proposal can be expressed as

$$\log p_{ij} = \log d_{ij} + \log (w_{ij}\, z_{ij}) \qquad (1)$$

where p_{ij} is the relative (market or production) prices of commodity «i», and d_{ij} is the relative *direct prices* (or labour-values in monetary terms)[2] of commodity «i», both of them expressed in terms of some commodity «j». On the other hand, the variable z_{ij} is

$$z_{ij} \equiv (1+z_i)/(1+z_j) \equiv [(1+ ð_i^T/w_i^T)/(1+ ð_j^T/w_j^T)]$$

where $z_i = (ð_i^T/w_i^T)$ and $z_j = (ð_j^T/w_j^T)$ are the *vertically integrated profit-wages ratios* for commodities «i» and «j» (for 'vertical integration' notion, see Pasinetti, 1973). Finally, w_{ij} is the *vertically integrated relative wage rate* (wage per hour of work) for commodities «i» and «j», which can be defined as

$$w_{ij} \equiv (w_i/w_j) \equiv (w_i^T/l_i^T)/(w_j^T/l_j^T)$$

where l^T symbolizes the *vertically integrated hours of work per unit of output*, and w^T is the *vertically integrated wage per unit of output*, being $\lambda_{ij} = (l_i^T/l_j^T)$ the relative labour-value of the *i*th commodity in term of the *j*th commodity[3]. Notice that d_{ij} is the relative *direct price* of commodity «i» with respect to commodity «j», and it is equal to λ_{ij} –the relative labour-values— because direct prices are simply the labour-values magnitudes times a constant, the so-called *value of money*[4].

Notwithstanding, the *parametric* version of equation (1) is written as (Shaikh, 1984, p. 71; Ochoa, 1989, note 1 of page 420; and Petrovic, 1987, equation 17 of page 202 and note 2 of page 207):

$$\mathbf{\log(p_{ij}) = \alpha + \beta \log(d_{ij}) + u_{ij}} \qquad \textbf{(2)}$$

In this section we deal with three different but related specification problems found in model (2). The first is a problem of *error terms misspecification*, due to a possible correlation between the error terms and the independent variable. The second is a problem of *parameter instability*, due to the arbitrary selection of the physical measurement units for commodities. The third is a problem of *spurious correlation*, which can arise due to the omission of a relevant variable –i.e., the physical production of each industry.

Error terms misspecification

It must be stressed that equation (1) represents the *strong correlation hypothesis* not only because α=0 and β=1, but also because its error terms are supposedly well-behaved and small-sized. However, the error terms incorporated into model (2) are written as

$$u_{ij} \equiv \log(w_{ij}\ z_{ij})$$

This definition is due to the fact that equation (1) represents an stochastic process only because the term $\log(w_{ij}\ z_{ij})$ can play the part of a "disturbance" term (Shaikh 1984, p. 70). But the properties of these *random* terms are extremely important, because all the statistical properties of a regression model depend on them[5].

Thus, the proposed error terms specification implies that ‘distributional factors’ such as wage rates and profit-wage ratios are random influences which *only* affect the dependent variable. This is not plausible at all, because the labour-values variable $\log(d_{is})$ can keep a strong relationship with the distributional factors included into the error terms $\log(w_{ij}\ z_{ij})$, because not only market prices formation but also the choice of techniques depend on them, and so do the Input-Output and NIPA data as well.

The consequence is that standard techniques like *ordinary least squared* (OLS) –which typically requires independent error terms identically distributed with respect to (zero) mean and variance; *IID property* thereafter– are inadequate for the estimation of model (2), because its error terms are possibly correlated with the independent variable, and parameter estimates will be therefore *biased* and *inconsistent*[6].

Parameter instability problem

Nevertheless, the problems of model (2) are not reduced to error terms lacking for the IID property. There is also, and above all, a problem of *parametric instability*. This problem is caused by the fact that we can always select an arbitrary measurement unit for every commodity considered. For instance, if we measure steel in tons the unit price (of each ton) will be different from the unit price if the measurement unit selected is instead kilograms. An unconventional criterion for this choice does not exist, and different choices will yield different ratios of unit prices and therefore different statistical results.

We can try to avoid the problem working with the *total output* of every industry –the total physical production of each commodity multiplied by its price— because the changes in the units of measurement do not affect these magnitudes. In order to do so, we can add to both sides of the model (2) the term $\log(q_i/q_j)$, and then add and subtract $\beta \log(q_i/q_j)$ in the right hand side. Writing explicitly relative prices and values, we have

$$\log(p_i/p_j) + \log(q_i/q_j) = \alpha + \beta \log(d_i/d_j) + \log(q_i/q_j) + \beta \log(q_i/q_j) - \beta \log(q_i/q_j) + u_{ij}$$

and rearranging

$$\mathbf{\log(p_i q_i / p_j q_j) = \alpha + \beta \log(d_i q_i / d_j q_j) + (1-\beta) \log(q_i / q_j) + u_{ij}} \qquad \textbf{(3)}$$

Notice that the OLS estimator is invariant to any non-singular linear transformation of the columns of the data matrix, so under such a condition we can always reparametrise any regression, not affecting in any way the ability of the regressors to explain the regressand (Davidson and MacKinnon, 1993, p. 16). As a consequence, OLS estimation of both models (2) and (3) yields exactly the same results.

However, the term we have just introduced to transform unitary magnitudes into sectoral ones, $\log(q_i/q_j)$ represents the *relative physical production* of each sector with respect to the 'standard' sector, a magnitude which is sensitive to changes in the physical measurement units actually adopted for commodities[7].

Therefore, working with total output prices and quantities requires accounting for the relative sizes of the industries, and it reintroduces the problem we were trying to avoid. This means that the model (3) also *lacks for parametric stability*, and therefore it cannot be estimated by means of standard techniques. As a matter of fact, correlation measures as r-squared can be arbitrarily manipulated to give any value between 0 and 1, as Ochoa first realized (1984, p. 130).

Spurious correlation problem

A possible 'solution' to the problem of the units of measurement is to omit the problematic variable $\log(q_i / q_j)$ and proceed to test a model like

$$\log(p_i q_i / p_j q_j) = \alpha + \beta \log(d_i q_i / d_j q_j) + \varepsilon_{ij} \qquad (4)$$

where the 'industry size' factor term $\log(q_i/q_j)$ does not explicitly appears in the right-hand side of the equation. However, omitting that variable implies to incorporate it into the new error terms

$$\varepsilon_{ij} \equiv (1-\beta) \log(q_i / q_j) + u_{ij},$$

and therefore the estimation of model (4) presents the well-know *omitted-variable misspecification error*, a problem that invalidates *again* all the statistical results. When the omitted variable has a *high correlation* not only with the dependent, but also with the independent variable, this problem can appear as one of *spurious correlation* between relative price and relative labour-value vectors.

3. Spurious correlation in aggregated models

Kliman (2002) seems to be aware of some of the problems discussed above, although he did not made explicit the analysis of the previous section. He correctly believes that even under the conditions discussed in the first section (and the assumption of full disaggregation) the omitted variable problem would not be easily avoided by returning from model (4) to model (3), due to a problem with the units of measurement (Kliman, 2002, p. 202, quoting Freeman, 1998, and p. 203). Using aggregated data, the omitted variable problem typically appears as a *spurious correlation* problem (first detected by Ochoa, 1984, p.129), because the omitted variable frequently has got a *high* influence on the dependent as well as on the independent variable.

Kliman's proposal for a correction

Consequently, Kliman attempts to find a testable version of model (2) that overcomes the spurious correlation problem, starting from a version of model (4) adapted to a context of conventionally defined, and empirically operative, 'sector' or 'branch' of activity, for which there is available data. To do so, the *strong correlation hypothesis* is first redefined as follows (Kliman, 2002, p. xxx):

$$\log(P_i) = \log(D_i) + u_i \qquad (5)$$

where P_i is the total output at market prices of sector «i», and D_i is the total output at direct prices of the same sector, and the error terms are supposed to verify the *IID property*[8]. We have not expressed the aggregated magnitudes as the product of two terms, price and quantity, because these observable magnitudes are not decomposable in such a way, which is due to the fact that every 'sector' contains different (physically heterogeneous) commodities.

The procedure followed by Kliman is rather intuitive. Starting with equation (5), he identifies, but does not define, the "sector size" as the *hidden* influence causing the spurious correlation problem. Then he looks for an index of this undefined variable, and he chooses the total cost (C_i) as a natural 'deflator' of the aggregated output of every sector; finally, he proceeds to 'deflate' sectoral outputs by means of that index, which is a way of taking account of the omitted variable problem. Equation (6) is the result:

$$\log(P_i/C_i) = \log(D_i/C_i) + u_i \qquad (6)$$

where C_i is the total cost of sector «i». As Kliman claims, this deflation does not affect the size of the errors (ibid., p. 303), as can be seen by comparing equations (5) and (6). But this does not means these two equations are the same thing.

Notice that what Kliman's procedure gives us is a set of *proxies* to real unit prices, and therefore equation (6) can rather be seen as a 'proxy' of equation (1). But we have not a proxy for each existing commodity, but only one for each observable ‘sector’. We can think on this set as a *sample* of the ‘true’ population –the fully disaggregated original industries. However, we cannot be sure about the representativeness of this *sample*, because the population is unknown. This problem ought to be added to the ones found in the previous section, which hold equally true when the assumption of full disaggregating is relaxed, as we point in the introductory section[9].

But the worse problem is that Kliman’s coarse procedure is even unable to account for the omitted variable problem, and therefore for the spurious correlation problem that entails to start from equation (5). Because of equations (5) and (6) are not parametric equations, we need to parametrise both equations, but in such a way that the *original parameter space* remains unchanged. The proposed parametric form of equation (5) is simply (Kliman, 2002, p. xxx):

$$\log(P_i) = \alpha + \beta \log(D_i) + u_i \qquad (7)$$

In order to obtain the parametric form of equation (6), but without changing the parameter space defined by model (7), we subtract $\log(C_i)$ from the right hand and the left hand side of equation (7), and then add and subtract the term $\beta \log(C_i)$ from the right hand side. After rearranging, we get

$$\log(P_i/C_i) = \alpha + \beta \log(D_i/C_i) + (\beta-1) \log(C_i) + u_i \qquad (8)$$

This reparametrisation of model (7) implies only a linear non-singular transformation of the logarithms of the original variables, and therefore the vector of coefficients estimated by means of OLS will be exactly the same for models (7) and (8). Therefore, the deflation *per se* is unable to solve the problem Kliman was interested in –i.e., the problem of spurious correlation that affects model (7), *if the size of the error terms is to be the same in both models*.

However, Kliman tests a model like the following one

$$\mathbf{\log(P_i/C_i) = \alpha' + \beta' \log(D_i/C_i) + \varepsilon_i} \qquad \mathbf{(9)}$$

where $\alpha \neq \alpha'$ and $\beta \neq \beta'$. This is because the deflation procedure has clearly changed the original parameter space. Thus, the new error terms are given by

$$\mathbf{\varepsilon_i = (\beta-1) \log(C_i) + u_i}$$

Kliman compares the estimated coefficients of models (7) and (9) and, obviously, he find that they are not the same. Kliman interprets that the difference is due to the hidden influence of the omitted variable problem affecting model (7). However, the cause of the difference is rather in the misspecification of the ‘deflated’ model (9), which is related to a parameter space other than the original one, and whose error terms are not therefore the same that those incorporated into model (7).

Some consequences of the problems discussed

These problems let us to explain why anomalous results could easily appear. For example, the observed correlation between prices and values can be very high although the wage rate is zero (Petrovic, 1987, p. 207), in strong contrast to expected results. Petrovic findings are counterintuitive because «workers are assumed to receive no wages, so that labour does not enter at all into the cost of production, the calculated prices… bear no relation to labour-times» (Shaikh, 1984, p. 71-72).

In other context, following now Kliman's procedure to overcome spurious correlation, but taking instead several indexes of industry size other than total costs –for it is clear that there is not any unconventional criterion in the selection of such indexes, we could observe very high correlations as well (see Osuna, 2002, pp. xx-xx). Clearly, there are not unambiguous results arising from this type of testing procedures. These are only two crude examples of the oddities we can find in cross-sectional studies of correlation.

4. Other price-value deviation measures

Steedman and Tomkins (1998) point to a double problem affecting deviation measures based on the *geometric* distance between price and value vectors, widely employed in the literature: they depend on the selection of physical measurement units or, alternatively, they depend on the normalization criterion choice. This is the case of the Mean Absolute Deviation (MAD) proposed by Ochoa (1989), or the Root-Mean-Square-per-cent-Error (RMS%E) suggested by Petrovic (1987), both of them affected by the normalization criterion choice.

However, Steedman and Tomkins claim it is possible to obtain measures not affected by the choice of the normalization criterion, being one of such measures, which has an intuitive descriptive-statistical character, the *coefficient of variation* (CV) of the price-value ratios (ibid, p.381). But, as we will show, the CV measure could be easily interpreted as a special case of the RMS%E measure.

To begin with, consider

$$n^{-1/2}\left\{\left(\sum \frac{p_{is}-\lambda_{is}}{\lambda_{is}}\right)^{2}\right\}^{1/2}$$

which is the RMS%E measure proposed by Petrovic (1987, p. 202, n. 2), being here p_{is} the relative prices of the *i*th sector, and λ_{is} the relative value of the *i*th sector, in both cases with respect to the selected standard sector «s». Note that the weight factor $n^{-1/2}$ is the (inverse of the) length of the n-ones vector, where n is the number of sectors. As Steedman and Tomkins assert, this measure is independent from the units of measurement, but depends on the arbitrary choice of a *numeraire*, as was recognized by Petrovic (1987, p. 204). Therefore it is not suitable as a measure of price-value deviations; this is also the case for the MAD measure proposed by Ochoa (Steedman and Tomkins 1998, p. 380)[10].

Note that we can explicitly incorporate the problem of the choice of the *numeraire* in the last expression by substituting $p_{is} = p_i / p_s$ and $\lambda_{is} = \lambda_i / \lambda_s$, and rearranging

$$n^{-1/2}\left\{\sum\left(\frac{p_i}{\lambda_i}\left(\frac{p_s}{\lambda_s}\right)^{-1}-1\right)^2\right\}^{1/2}$$

Denoting now the price-value ratios as $h_i = p_i/\lambda_i$ and the 'factor of scale' of the standard commodity as $\theta_s = p_s/\lambda_s$ –a factor that varies with the *numeraire*–, we can re-write the last expression as follows,

$$n^{-1/2}\left\{\sum\left(\theta_s^{-1}h_i-1\right)^2\right\}^{1/2}$$

Thus, when the chosen *numeraire* is the value of total output, we have $\theta_s = \eta$, where η is the so-called *value of money*, which is a case frequently considered in the literature.

Consider now the coefficient of variation (CV) of the h_i terms (Steedman and Tomkins 1998, p.381)

$$\frac{1}{h_m}\left\{\frac{\sum\left(h_i-h_m\right)^2}{n}\right\}^{1/2}$$

where $h_m = n^{-1}\sum h_i$ is simply the mean of the price-value ratios, and n is the number of sectors considered. We can rewrite this expression as follows

$$n^{-1/2}\left(\sum\left(h_m^{-1}h_i-1\right)^2\right)^{1/2}$$

being clear that the coefficient of variation (CV) of the price-value ratios has a direct relation to Petrovic's measure: it is *exactly* the RMS%E value when the standard commodity is the mean-price-value-ratio's sector and then $\theta_s = h_m$. Therefore, the coefficient of variation is not a measure which is independent from the choice of *numeraire*, but rather a measure which incorporates a special normalization criterion.

5. Some concluding remarks

Throughout this paper, we have demonstrated the existence of severe problems of specification in the log-linear models employed in testing the cross-sectional *strong correlation hypothesis*. The verification of that hypothesis was supposed to prove a causal connection among labour values and prices.

The critique posed by Kliman is based on a problem of spurious correlation related to an unadvertised hidden influence of an omitted variable in widely used regression models. Such an influence is identified as the sector size. The proposed correction consists in the deflation of the sectoral variables using sectoral costs. A detailed analysis of the problem of aggregation and the problem of choice of units of measurement is absent in Kliman's analysis. Even worse, the specification problems present in the original model criticized are not clearly detected, and hence those specification errors are reproduced.

In addition, measures of geometric deviations among relative prices and relative values, that are aimed to support the Ricardian hypothesis of *proximity*, can overcome the measurement's units problem, but only by incorporating the *numeraire* choice problem. As a result, they cannot be considered as adequate measures of "proximity". However, statistical measures of price-value ratios *dispersion*, as it is the case for the variation coefficient measure proposed by Steedman and Tomkins, can be interpreted as geometric measures of deviation of the weighted price-value ratios vector, but it depends on the *numeraire* selection as well.

Several conclusions arise from these results. First, testing procedures in analysing cross-sectional price-value relations based on regressions models must be clearly rejected. There is not "deflation" procedure able to overcome the problems raised by the sectoral data aggregation and related misspecification errors. Second, geometric deviation measures based on the observed price-value ratios only can be univocally interpreted as statistical *dispersion* measures, because the calculated degree of "proximity" arbitrarily varies with the *numeraire* choice. To sum-up, neither measures of "correlation" among prices and values nor "proximity" among relative prices and relative values, can overcome the problems of physical measurement units and normalisation criteria selection in cross-sectional data analysis.

Bibliography

Bienenfeld, M. 1988. Regularity in price changes as an effect of changes in distribution, *Cambridge Journal of Economics*, vol. 12, pp. 247-255.

Chilcote, E. 1997. *Interindustry structure, relative prices and productivity: an input-output study of the U.S. and O.C.D.E. countries*, Doctoral Dissertation, New School for Social Research, New York.

Cockshott, P. and Cottrell, A. 1997. Labour Time versus Alternative Value Bases: A research note, *Cambridge Journal of Economics*, vol. 21, pp. 545-549.

Cockshott, P. and Cottrell, A. 1998. Does Marx need to transform? pp. 75-80 in Bellofiore, R. (ed.), *Marxian Economics: A Reappraisal*, Vol. 2, Basingstoke, Macmillan.

Cockshott, W. P. and Cottrell, A. 2003. Robust correlations between sectoral prices and labour values: a comment, *Cambridge Journal of Economics*, forthcoming

Cockshott, P., Cottrell, A. and Michaelson, G. 1995. Testing Marx: some new results from UK data, *Capital and Class*, 55, pp. 103-129.

Davidson, R. and MacKinnon, J.G. 1993. *Estimation and Inference in Econometrics*, Oxford University Press, New York, 1993.

Febrero, E. 1998. *Valor trabajo: un indicador de productividad y competitividad. Una aplicación al caso español: 1970-1992*, Doctoral Dissertation, Universidad de Castilla-La Mancha.

Freeman, A. and Carchedi, G. 1996 (Eds). *Marx and non-equilibrium economics*, Edward Elgar, London.

Giussani, P. 1991-92. The Determination of Prices of Production, *International Journal of Political Economy*, vol. 21, pp. 67-87.

Griliches, Z. 1971. Hedonic Price Indexes Revisited, pp. 3-15 in Griliches, Z. (Ed.) *Prices Indexes and Quality Change*, Harvard University Press, Cambridge, Massachusetts, 1971.

Guerrero, D. 2000. *La teoría del valor y el análisis insumo-producto*, mimeo, New School for Social Research, New York.

Kliman, A. 2002. The Law of Value and Laws of Statistics: Sectoral Values and Prices in the U.S. Economy, 1977-1997, *Cambridge Journal of Economics*, 26, pp. 299-311.

Kliman, A. 2003. Spurious value-price correlations: some additional evidence and arguments, *Research in Political Economy*, forthcoming

Kliman, A. and McGlone, T. 1999. A Temporal Single-System Interpretation of Marx's Value Theory, *Review of Political Economy*, vol. 11, pp. 33-59.

Kmenta, J. 1997. *Elements of Econometrics*, University of Michigan Press.

Ochoa, E. M. 1984. *Labour Values and Prices of Production: an Interindustry Study of the U.S. Economy, 1947-1972*, Doctoral Dissertation, Ann Arbor, Mich (University Microfilms International).

Ochoa, E. M. 1989. Values, prices, and wage-profit curves in the US economy, *Cambridge Journal of Economics*, vol. 13, pp. 413-429.

Osuna, R. 2003. *Un modelo secuencial para el cálculo de precios. El caso español, 1986-1994*, Doctoral Dissertation, UNED, Madrid.

Panethimitakis, A. (1993): Direct *versus* total labour productivity in Greek manufacturing: 1958-1980, *Economic Systems Research*, vol 5, pp. 79-93.

Pasinetti, L. (1973): The Notion of Vertical Integration in Economic Analysis, *Metroeconomica*, vol. 25, pp. 1-29.

Petrovic, P. 1987. The deviation of production prices from labour values: some methodology and empirical evidence, *Cambridge Journal of Economics*, vol. 11, pp. 197-210.

Shaikh, A. 1975. *Input-Output Accounts and Marxian Categories*, mimeo, New School for Social Research, New York.

Shaikh, A. 1984. The Transformation from Marx to Sraffa, pp. 43-84 in Mandel, E. and Freeman, A. (Eds.). *Ricardo, Marx, Sraffa*, Verso, London.

Shaikh, A. 1998. The Empirical Strenght of the Labour Theory of Value, pp. 225-251 in Bellofiore, R. (ed.) 1998. *Marxian Economics. A reappraisal. Essays on Volume III of Capital. Profits, Prices and Dynamics*, St. Martin Press, Nueva York.

Shaikh, A. and Tonak, E. A. 1994. *Measuring the Wealth of Nations. The Political Economy of National Accounts*, Cambridge University Press, Cambridge.

Steedman, I. and Tomkins, J. 1998. On measuring the deviation of prices from values, *Cambridge Journal of Economics*, vol. 22, pp. 379-385.

Tsoulfidis, L. and Maniatis, T. 2002. Values, prices of production and market prices: some more evidence from the Greek economy, *Cambridge Journal of Economics*, vol. 26, pp. 359-369.

Valle Baeza, A. 1994. Correspondence between labor values and prices: a new approach, *Review of Radical Political Economics*, vol. 26, pp. 57-66.

---

[1] See also Shaikh (1998), Ochoa (1989), Chilcote (1997), Cockshott, Cottrell and Michaelson (1995), Cockshott and Cottrell (1997, 1998), Febrero (1998) and Guerrero (2000). There are others who were not directly inspired by Shaikh's work, although they used a similar approach, as Bienenfeld (1988), Panethimitakis (1993), Petrovic (1987) or Valle (1994).

[2] Shaikh (1984) often calls them 'values', whereas Petrovic (1987, p. 198 and note 1 of page 207) uses the term 'value price'.

[3] When a uniform wage rate is assumed $w_{ij} = 1$ and therefore $p_{ij} = \lambda_{ij} z_{ij}$ (Shaikh 1984, p. 67), but in general wage rates will differ across industries, as equation (1) explicitly shows (then $w_{ij} \neq 1$).

[4] It will be needed, for empirical purposes, to choose a normalisation criterion. The so-called *first Marxian invariance* is usually adopted

$$\Sigma(p_i q_i) = \eta^{-1} \Sigma(\lambda_i q_i)$$

where λ_i is the labour value of one unit of commodity «i», expressed in hours of work, $\Sigma(p_i q_i)$ represents total output at market prices, and $\Sigma(\lambda_i q_i)$ represents total output at labour values, and where $\eta \equiv \Sigma(\lambda_i q_i)/\Sigma(p_i q_i)$ represents the *value of money* (see, for instance, Ochoa 1989, p. 416 or Petrovic, 1987, p. 202). Now, calling $d_i = \eta^{-1} \lambda_i$ *direct prices* (prices proportional to labour values), we are able to re-write the former equality as

$$\Sigma(p_i q_i) = \Sigma(d_i q_i) \; (=1)$$

[5] The literature implicitly establishes that $E[\log(w_{is} z_{is})]=0$ –which implies that both w_{is} and z_{is} tend to unity–, and that error term's variances are bounded, $Var[\log(w_{is} z_{is})]<\sigma^2$.

[6] Thus, the general condition for the *consistency* of the OLS parameter estimates is

$$p\lim_{n\to\infty}\left(n^{-1}[1 \;\; \log(\lambda)]^t u\right) = 0$$

This means that error terms vector *u* must be asymptotically uncorrelated with data matrix **[1 log(λ)]**, being **1** a n-ones column vector and **log(λ)** the nx1 sectoral log-value vector. Furthermore, if parameter estimates are to be *unbiased*, the stronger condition

$$E([1 \; \log(\lambda)]^t u)=0$$

must hold. These necessary conditions are not directly verifiable, since the orthogonality property of OLS ensures that regardless of whether *u* is correlated with data matrix or not, the residuals of the corresponding regression are always orthogonal to that matrix (see Davidson and MacKinnon 1993, pp. 209-210).

[7] This term could be easily interpreted as an *index of industrial size*. In such a framework, the problem of the change in the physical measurement units can be viewed as a particular (cross-sectional) form of a more general problem: the sensitivity of the indexes to changes in the structure of base-weights. For a case related to *qualitative change* in commodities, see Griliches (1971, p. 6-7).

[8] Notice that the relative variables are gone (there is no *j* sector here). It is due to the methodology followed by Kliman, based on a set of difference equations that, unlike Shaikh's simultaneous equation based approach, does not need for normalization. For a formulation of Kliman's referential model, see Giussani (1991) or Kliman and McGlone (1999). See, for a broader outlook, Freeman and Carchedi (1996).

[9] Kliman (2003a, fourth section) presents an exercise of simulation: he creates 7000 artificial disaggregated industries, and then he proceeds to aggregate them in 71 sectors. In this experiment, the ‘population’ is known, and we could check if a deflation using some index of sector size produces a representative *sample* of this population. However, Kliman’s exercise is aimed at demonstrating that the r-squared for aggregated sectors is higher than for disaggregated sectors, an already known result (see, for example, Kmenta, 1997, pp. 367-373).

[10] It must be stressed that these measures «makes the apparent extent of price-value deviation depend on a choice (the choice of *numeraire*) which has no relevance to that deviation» (Steedman and Tomkins 1998, p. 380).