REVUE BELGE
DE
NUMISMATIQUE
ET DE SIGILLOGRAPHIE

BELGISCH TIJDSCHRIFT
VOOR NUMISMATIEK EN ZEGELKUNDE

PUBLIÉE
SOUS LE HAUT PATRONAGE
DE S. M. LE ROI
PAR LA
SOCIÉTÉ ROYALE
DE NUMISMATIQUE DE BELGIQUE

UITGEGEVEN
ONDER DE HOGE BESCHERMING
VAN Z. M. DE KONING
DOOR HET
KONINKLIJK BELGISCH
GENOOTSCHAP VOOR NUMISMATIEK

Directeurs
GHISLAIN MOUCHARTE, PIERRE COCKSHAW,
FRANÇOIS DE CALLATAY et JOHAN VAN HEESCH

CXLVII - 2001

BRUXELLES  BRUSSEL
D. CAIRNS and T.P. HUTCHINSON (*)

DID THE GOLD CONTENT OF CYZICENE ELECTRUM COINS DECLINE OVER TIME?
A STUDY USING ELABORATION AS A STATISTICAL STRATEGY

Was there really a steady decline in the gold content of electrum coins from Cyzicus, as argued by Smekalova and Djukov (1)? We shall reexamine their dataset (which was based upon coins in the State Hermitage collection, St. Petersburg), and will find that gold content did change, but the changes were irregular, rather than being a steady decline. Smekalova and Djukov did not take into account the following features of their data: that the average gold content of the lower denominations is higher than for staters, and that the lower denominations tend to be from earlier periods than the staters do. Consequently, the straightforward calculation of average percentage gold in the different periods reflects the coin denomination as well as the period, and can be misleading.

In Smekalova and Djukov's Table 1, they report that the average gold content of the coins fell from 61% for period I coins to 54% for period IV coins. The change is statistically significant. Fortunately, Smekalova and Djukov presented their original data in the Supplement to their paper: for each of their 54 coins, they listed several of its characteristics, including its period, its denomination, and the results from X-ray fluorescent spectroscopy (percentages of gold, silver, and copper). Both they and the journal editor deserve praise for this publication of data, as any interested readers, such as ourselves, can critically reexamine the conclusions drawn. We shall express doubt about the validity of the change found in the average gold content.

Elaboration as a Statistical Strategy

Is there any information within this dataset that can explain the apparent association of percentage gold with time? Procter (2) describes this

(*) T.P. HUTCHINSON, Department of Psychology, Macquarie University, Sydney, N.S.W. 2109, Australia.
e-mail: phutchin@bunyip.bhs.mq.edu.au


step in an analysis very well. « So to find a relationship between two variables is not the end but little more than the beginning of the analysis. It is the researcher's job to be sceptical about the causal interpretations that might be made from such a relationship, to think carefully, in a theoretically sensitive way, about what nuisance factors might be involved in the system of variables, and to find ways of testing alternative explanatory hypotheses ». The approach to accounting for alternative explanations concerns mainly holding constant the suspected nuisance factors. It is quite easy to do, at least when there are only few variables — if one has discovered a relationship between two variables, one then extends this by examining whether this holds whatever value a third variable takes. This approach has a long history, and variants of it are called by a number of names — « the logic of causal order », « elaboration », « statistical control » and « multiple regression », as well as others. Vogt (3) describes elaboration as « A process of studying correlations between variables by observing how they are affected when controlling for the effects of other, intervening, variables ». (The word « correlations » here carries the broad meaning of « associations » or « relationships », not some highly specific formula-based meaning.) Watts (4) describes elaboration as « That mode of data analysis where the systematic examination of the relationship between two variables proceeds by way of the introduction of an antecedent variable or test factor into the analysis ». Some other authors appear to restrict elaboration to analyses where the dependent variable is categorical rather than metric (i.e., is expressed as percentages rather than means), but this seems pointless to us. The central idea is that elaboration envisages starting with the simple situation of two variables and their relationship, and then introducing other variables into the scope of the research; the contrast is with starting with lots of independent variables contributing to the prediction of the dependent variable, and trying to omit some of them because they improve the accuracy of prediction very little. Very readable references to these methods include Bowen and Weisberg (5) and Neuman (6). A more advanced discussion of statistical control can be found in Darlington (7).

We now look for other measured variables related to gold content. There is one very obvious candidate to be an additional factor. The coin

---


characteristics included in the Supplement to Smekalova and Djukov's paper are: denomination, weight, diameter, period, gold percentage, silver percentage, copper percentage. The weight and diameter merely reflect its denomination, period is the independent variable already being discussed, and the composition percentages are the dependent variables to be analysed. Thus denomination is the only extra independent variable that we can examine. A further reason for including denomination as a factor is that there is clear evidence in Figure 1 of Guepin (8) that the gold content of staters may be lower than that of the lower denomination coins. (We do not go further with Guepin’s dataset than merely noting that it warns us that denomination may be an important factor, as the reported gold content of the coins is very different from that in Smekalova and Djukov's dataset: the mean is lower, and the coin-to-coin variability is greater. Whether this was due to the different method of measuring the alloy composition, or to the different set of coins examined, or to something else, we do not know).

Results

We will concentrate on percentages of gold. Variations in the percentage of silver are largely the reverse of variations in the percentage of gold.

Our first contribution is to observe that the 30 staters averaged 54% gold, the 13 hectai averaged 56% gold, and the 10 hemihectai averaged 59% gold. Smekalova and Djukov did not tabulate composition using denomination as the explanatory variable, and so did not notice this. As the differences between denominations are nearly as large as the differences between periods, it seems unsafe to ignore them. Figure 1 illustrates the difference between denominations. The coins included are only from Period II. Both gold content and silver content are shown; the impression gained from comparing the three plots is certainly that gold and silver percentages do depend upon denomination. (Smekalova and Djukov mention the possibility that there is a small systematic error in their figures, arising from the surface of a coin having a higher gold content than the bulk of the coin. They do not discuss whether this systematic error might be greater for small coins than for large. If this were the case, it would destroy the validity of any apparent relationship between denomination and gold content. But in the present context, it does not matter whether gold content does genuinely differ between denominations or whether a bias is present: in either case, it is necessary to take denomination into account when attempting to perceive the effect of period. We ourselves do not have expert knowledge of metallurgy or X-ray

fluorescent spectroscopy, so we cannot say whether it is plausible that there is a bias, we merely raise the possibility).

Our second contribution, in the spirit of « elaboration », is to examine how gold percentage is related simultaneously to period and denomination. Table 1 shows the results. Between 7 and 14 coins contribute to each of the means shown there. Omitted from the table, simply because there are so few of them, are two Period III hectai (average percentage gold = 55), two Period I hemihectai (average percentage gold = 61), and one Period III mishemihecta (percentage gold = 55). Almost the only information about the effect of period on percentage gold comes from the staters. That being so, we might ask whether percentage gold is significantly related to period, when considering only staters. The answer is yes, it is. But the nature of the relationship is not a steady decline. Instead, the mean for Period IV is intermediate between those for Periods II and III (and is actually rather closer to the former).
Table 1. Average percentage gold, coins being classified by period and denomination

<table>
<thead>
<tr>
<th>Period</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staters</td>
<td>55.7</td>
<td>51.1</td>
<td>54.2</td>
</tr>
<tr>
<td>Hectai</td>
<td>56.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hemihectai</td>
<td>58.1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Our opinion is that further statistical significance tests would only distract from the point we are making. The reason is that there would need to be detailed discussion of which statistical test to choose and how to implement it. For example, coin 27 has an exceptionally high gold content and an exceptionally low silver content, coin 42 has the reverse, and so possibly these coins should be treated as outliers and excluded from calculations (as suggested in Smekalova and Djukov's footnote 11). But if these coins are to be excluded, should not the same apply to coin 2 (exceptionally low in silver and high in both copper and other elements) and to coins 4, 15, 44 (which are exceptionally high in other elements)? And perhaps a judgment as to whether or not a coin is an outlier should be based on comparing it not with all the other coins, but only with coins of the same denomination that are from the same period? Or if there is concern but no proof that various coins might be outliers, should not we retain all of them in the dataset, but choose a statistical test that is robust to the presence of outliers? Another issue is should Period II be split into IIa and IIb, c? And again, why choose percentage gold as the dependent variable of interest, rather than percentage silver or something else? Unfortunately, it turns out that the results of significance tests are affected by the choices made on these issues. (Other fields of study have similar problems. It may be of interest that in some fields, such as clinical trials of drug effectiveness, it is common practice to specify in advance precise details of how the data will be processed and how the statistical analysis will be performed — it is felt that the conclusions will not be convincing to outsiders if the investigators have discretion to do one thing or do another).

In the absence of a compelling reason for preferring one statistical analysis over the others, we simply make the weaker point that, on the basis of this dataset, it seems that the gold content of staters did vary over time, but the variation was irregular, not a steady decline. The means are consistent with a hypothesis of a decline from Period II to Period III, followed by a return to the previous level in Period IV. (We say this merely as a description of the pattern — we do not know whether there is any evidence for deliberate changes).