(L)

A DEFENCE OF INDEX-NUMBERS

[IN this article, which appeared in the ECONOMIC JOURNAL, March 1896, I ventured to differ from Mr. Pierson's "Further Considerations on Index-Numbers" (published in the same number of the Journal), on the ground that they ignore the character of Probabilities essential to the computation of index-numbers.]

It is justly observed by Adam Smith that the anxiety about public opinion is much greater among the candidates for excellence in some arts than it is in others. "The beauty of poetry is a matter of such nicety that a young beginner can scarce ever be certain that he has attained it. . . . Racine was so disgusted by the indifferent success of his *Phèdre*, that though in the vigour of his life and at the height of his abilities, he resolved to write no more for the stage. . . . Mathematicians, on the contrary, who may have the most perfect assurance both of the truth and of the importance of their discoveries, are frequently very indifferent about the reception which they may meet with from the public." [1] In the scale of susceptibility which is thus indicated, a high place must be assigned to the more refined parts of economic science. Even those investigations which at first sight appear to be wholly statistical—such as the calculation of index-numbers—may rest upon speculative assumptions, concerning which the consensus of authority is naturally desired. Accordingly, when the distinguished Dutch economist concludes in the immediately preceding paper that "all attempts to calculate and represent average movements of prices, either by index-numbers or otherwise, ought to be abandoned," those who have been making such attempts will anxiously reconsider the basis of their computation, and tremble for its safety. But the discouragement which such a condemnation coming from such an authority is calculated to produce may be mitigated by observing that the index-number which is the object of Mr. Pierson's crushing criticisms is one of a very peculiar character, differing in some essential attributes

[1] *Theory of Moral Sentiments*, Part III. ch. ii.

from the operation as ordinarily conceived and practised. Racine would not have been dejected by the indifferent success of his tragedy if the play, so badly received, had been a version of his masterpiece from which the characters of Phèdre and Hippolyte had been left out. Two equally serious omissions are presupposed by Mr. Pierson's animadversions.

There is, *first*, the character of probability. It is generally implied that the problem now before us, in its data, method, and result, is germane to the Calculus of Probabilities. The nature of the problem is happily indicated by Professor Nicholson when he compares the set of moving prices to a fleet of yachts which under the influence of a common cause—it may be rising wind or tide—are variously accelerated according to "the build of the various yachts or seamanship of the crews." The type of such problems is the investigation of what Mill calls a *residual phenomenon*,[1] illustrated by the discovery of the diurnal variation in the height of the barometer by comparing the averages of a great number of observations at different times of day. It is postulated in such reasoning that the error or deviation of one observation is independent of that which has been incurred by another observation;[2] just as, when a die is thrown a number of times, it may be assumed that the number of pips turned up at each throw is unaffected by the preceding throws. It is true that in concrete nature such ideal independence can hardly be expected. Thus in barometrical observations it is possibly not correct to treat the observation for each day as an independent sample. Probably the weather sometimes follows suit for two or three days together; but the deviation of the observations is doubtless sufficiently random to justify Laplace's application of the Calculus of Probabilities. So the grouping of human statures is perhaps not perfectly sporadic;[3] but it is sufficiently so to allow a Galton to infer with great probability that the conditions of a particular class—*e.g.*, boys in public schools, or men in the Royal Society—as compared with less favoured classes are particularly favourable to growth.[4] It is not necessary to

[1] Mill, *Logic*, Book III. ch. xvii. *Cp.* Laplace, *Probabilities*, Book II. ch. v.

[2] On this postulate see the present writer's "New Methods of Measuring Variations in General Prices," *Journal of the Royal Statistical Society*, 1888, p. 367, note. *Cp.* Laplace, *loc. cit.*: "Il faut avoir soin de varier les circonstances de chaque observation."

[3] As appears from the fact that in the group constituted by the measurements of a nation there will be sub-classes with different averages.

[4] *Cp.* "Methods of Statistics," *Journal of the Royal Statistical Society*, Jubilee volume.

discuss here whether the average would be of any scientific use if this condition of sporadic dispersion were not fulfilled—if all the observations were massed at two points, or collected into two sharply demarcated classes—*e.g.*, dwarfs and giants.[1] It is sufficient to observe that as a matter of fact the condition of sporadicity is very generally fulfilled both in physics and social phenomena : wherever there is at work a set of miscellaneous agencies, " a mass of fleeting causes " in Mill's phrase.[2]

It is by ignoring this character of sporadic dispersion that Mr. Pierson's criticisms acquire their plausibility. He begins : " Let us suppose ten commodities, all equally important. Five of them are doubled in price, and five of them fall to exactly one-half." But surely this is a very odd supposition, in view of the sporadic dispersion which very generally prevails in this world. It would have been more appropriate to suppose a number of figures representing variations of price (in one epoch as compared with another), not separately disposed in two heaps, but scattered about. Mr. Pierson's supposition would be appropriate if, for instance, Mr. Sauerbeck's percentages for the comparative prices of different commodities were massed at two points. But this is not so, as appears by considering his figures and diagrams representing annual or quarterly variations of price.[3] A common trend comes out in the average, but the particular movements are independent.

The recognition of this sporadic character is fatal to Mr. Pierson's principal objection, which is in effect, though perhaps not apparently, that if the particular observations be weighted differently the average will be seriously different. This objection recurs in different forms. In his first paragraph Mr. Pierson supposes ten observations : five commodities of which the price has been doubled, five of which it has been halved; in the second as compared with the first period the data may accordingly be regarded as consisting of ten ratios, or percentages, five of them each = 200 (: 100); five of them each = 50 (: 100). The simple arithmetical average of these may be written

[1] See the reference given in the preceding note. See also p. 279 in the Memorandum on *Methods of Ascertaining and Measuring Changes in the Value of the Monetary Standard*, by the present writer, published in the Report of the British Association for 1887. This Memorandum and the two supplementary ones, published in the Reports of the British Association for 1888 and 1889, should be referred to as containing justifications of statements made summarily in the present paper.

[2] Mill, *Logic*, loc. cit.

[3] A similar scrutiny of Laspeyres' statistics of relative prices is attempted in the Memorandum of 1887 [H, above, p. 245.]

$\frac{5 \times 200 + 5 \times 50}{10} = 125$. Now *weight*[1] each observed percentage with its own reciprocal, and you have $\frac{5 + 5}{5 \times \cdot 005 + 5 \times \cdot 02} = 100 \div 1 \cdot 25$. And the complaint is that these two results are not equal.

The complaint is virtually similar in the sixth paragraph (*loc. cit.* p. 128). There the simple observations are 75, 16·66, 25. And the simple arithmetical mean is $\frac{75 + 16 \cdot 66 + 25}{3} = 38 \cdot 88$. The other average which is contrasted with this one is obtained by weighting each observation with the value in money of a pound avoirdupois of the corresponding commodity at the initial period, that is 20, 12, 4, respectively. These weights being applied, the average becomes $\frac{1500 + 200 + 100}{60 + 12 + 36} = 16 \cdot 6$.

The same contrast is noticed in some other cases. "In Case I. there will be no change," "in Case II. there will be a rise of 25 per cent.," "in Case III. there will be a fall of 25 per cent."; the observations being weighted in the peculiar mode[2] which has just been described; whereas, according to the simple arithmetic mean, there is no change in any of the cases.

Such discrepancies seem very serious when we deal with artificially simplified examples; but they become insignificant when we deal with the concrete, sporadically dispersed, price-ratios. For it is a well-known proposition that a difference in the system of weights will not make much difference, provided that the number of independent observations is sufficiently great; provided also that the experiment is made in the spirit of Probabilities, with an *animus mensurandi*—in Herschel's phrase—not consciously selecting cases which will not work well. The reason and limits of the proposition are defined by theory,[3] and the theory is confirmed by experience.

As verifications of the theory *in aliâ materiâ* may be adduced the index-number constructed by Mr. Bowley to indicate the

[1] If $x_1, x_2, \ldots x_n$ are observations, the simple arithmetic mean is $\frac{x_1 + x_2 + \ldots + x_n}{n}$ the weighted arithmetic mean is $\frac{w_1 \times x_1 + w_2 \times x_2 + \ldots + w_n \times x}{w_1 + w_2 + \ldots + w_n}$, where $w_1, w_2, \ldots, w_n$ are the *weights*.

[2] The peculiarity of the mode being to assign as weight a pound or bushel, or, as in the passage before us, some unit, which is arbitrary and accidental with reference to the measurement of the depreciation of money. See below, p. 367.

[3] See I, above, p. 305 *et seq.*

increase of general wages. Weighting the percentages expressing the growth of wages in America according to the system which he thinks best, and according to the very different system employed by the American statisticians, Mr. Bowley obtains almost exactly the same result.[1] Another conspicuous example is afforded by the concurrence between the different methods which Sir R. Giffen in his census of wages has employed in order to determine the average wage. Using, in effect, different systems of weights, he obtains for the average weekly wage the values 29*s.* 5*d.*, 29*s.* 7*d.*, 29*s.* 7*d.*[2]

Experience more adjacent to the case in hand is afforded by the price-ratios which Mr. Sauerbeck has tabulated year by year. There is found to be a close agreement between the arithmetic mean and the averages which are obtained by taking account of quantity. The following figures are given by Mr. Sauerbeck in the ECONOMIC JOURNAL for June 1895 :—

SIMPLE AND WEIGHTED AVERAGE OF COMPARATIVE PRICES.

—	Arithmetical mean.	Making allowance for quantities.
1887	68	66·7
1888	70	68·8
1889	72	71·8
1890	72	72·1
1891	72	72·0
1892	68	67·7
1893	68	67·1
1894	63	62·0

Other comparisons of the two kinds of average are given by Mr. Sauerbeck in his well-known papers in the *Journal of the Statistical Society*. Further verifications will be found in the second of the Memoranda above referred to. It will be sufficient to make one extract. Of the percentages indicating the variations in price of nineteen commodities tabulated by Mr. Palgrave, the simple arithmetic mean and the mean weighted according to quantity are compared for sixteen successive years, and the sixteen differences between the two results for each year are as follows :[3] 4, 2, 2, 3·5, 1, 5, 0, 1, 2, 0, 4·5, 7, 2, 2·5, 1, 1.

[1] ECONOMIC JOURNAL, Vol V. p. 373.

[2] *Report on the Wages of the Manual Labour Classes.* [C. 6889—1893.]

[3] H, above. See also pp. 202 and 205. Attention may be called to the experiments with weights assigned arbitrarily : by forming the sum of a set of digits taken at random (p. 199, last paragraph), or—in the cognate case of the Median—tossing up a coin and assigning 1 or 2 as the weight, according as head or tail turned up.

But it is needless to labour this proposition further, as it is acknowledged by Mr. Pierson in a former paper [1] when he deals with real examples: in particular Mr. Palgrave's index-number, and Mr. Falkner's report on "wholesale prices," [2] in which the simple arithmetic mean of some hundreds of relative prices and the mean of the same weighted according to the importance of each commodity in the average household budget are found to agree. Here are some of the figures quoted by Mr. Pierson:—

—	Ordinary Average.	Corrected [weighted] Average.
1871–75	134·58	131·26
1876–80	106·78	108·14
1881–85	102·52	104·0
1886–90	93·04	95·20

"It is clear," comments Mr. Pierson, "that the relative weight may be left out of consideration without marked detriment when we extend our investigation to a great number of articles."

To sum up, several of Mr. Pierson's objections amount to this one: that the calculation of average relative prices is untrustworthy, because the result is seriously different according as different systems of weighting are employed. And this objection, though true in the abstract of artificially simplified index-numbers, is not true of the sets of figures with which we have actually to deal.

A similar reply may be made to the objection that the result of the calculation will be seriously different according as the arithmetic or the geometric mean is employed. This is true of the imaginary examples set up to be knocked down, but it is not true in the concrete. The arithmetic and geometric mean of the price-ratios for a large number of miscellaneous commodities are likely not to differ much from each other. This is a deduction from a more general proposition that, with certain reservations, *any* mean of a large group of observations is likely not to differ much from any other kind of mean.[3] Take, for example, the series of observations obtained by measuring the heights of different men. The arithmetic mean [4] of 1000 such observations obtained

[1] The paper described in the ECONOMIC JOURNAL, Vol. V. p. 109. See p. 8 of the German edition of *Goldmangel* (reprinted from the *Zeitschrift für Volkswirthschaft*, Band iv. Heft 1).

[2] Well summarised by Professor Taussig in the *Yale Review* for November 1893.

[3] For the evidence and limits of this proposition see the paper on the "Law of Error," by the present writer, in the *Philosophical Magazine* for November 1892. It is supposed that, as usual where miscellaneous agencies are at work, the *law of error* is approximately fulfilled by the observations; also that these are measured from a point outside the extreme value which an observation can possibly reach: for example, in the case of human statures or price-ratios, *zero*.

[4] The observations are given in the paper just referred to. Each of them is the mean height of twenty-five men.

by Mr. Elliott is 68·20 inches. Compare with this the mean value which is obtained by squaring all the observations, taking the arithmetic mean of the squares, and extracting the square root of that mean. The mean value so obtained is 68·25. The mean value obtained by cubing all the observations, taking the arithmetic mean of the cubes, and extracting the cube-root of that mean, is much the same, viz. 68·30. The geometric mean is 68·16.[1]

To adduce more specific experience, here are two rows of figures, of which one consists of the geometric means of thirty-nine percentages obtained by Jevons for several years, the other consists of the arithmetic means of the same percentages.[2]

	1851.	1853.	1855.	1857.	1859.
Geometric Mean	92·4	111·3	117·6	128·8	116
Arithmetic Mean	94·6	112·4	119	134	119

Mr. Sauerbeck has calculated the geometric mean of his forty-five percentages for two years and allows me to cite the results [3] :—

	1880.	1894.
Arithmetic Mean [4]	87·82	62·93
Geometric Mean	86·97	60·90

So much for the objection implied in the preceding paper that

[1] These calculations have been performed by Mrs. Bryant, D.Sc.

[2] From the Memorandum of 1888, p. 206. Alternate years were taken, the more to vary the circumstances of the experiments. There is no reason to suspect that successive years would have presented different results. For instance, for 1852 the geometric mean is 93·8, the arithmetic 94·6 (*loc. cit.*).

[3] That is, the simple arithmetic mean. The weighted (arithmetic) means were respectively 87·3 and 62·0.

[4] The geometric mean comes out a little less than the arithmetic, as might have been expected. This tendency may confer some advantage, but a very slight one (Memorandum, 1887, pp. 283–289), on the geometric mean. A more important prerogative of the geometric mean was noticed, as far as I know, first by Professor Harald Westergaard, and has not been sufficiently recognised by the *connoisseurs* of index-numbers. The geometric mean is the only one in which no alteration at all is produced by the change of basis. In the case of the arithmetic mean, if one year, x, be taken as basis, and the index-numbers for y and z, say I_y and I_z, be determined as percentages with reference to x, then the ratio of I_y to I_z will not in general be exactly the same when the index-numbers are calculated with reference to another basis, x', say I'_y and I'_z. The reason is that (as explained above with reference to a particular case where $x = y$ and $x' = z$) the two ratios $\frac{I_y}{I_z}$ and $\frac{I'_y}{I'}$ are to be regarded as *differently weighted* means of the same set of observations, viz. the set of ratios obtained by dividing the price of each commodity in y by its price in z. That the geometric mean follows in this respect the analogy of physical measurements is at least an elegance. The geometric mean is *pro tanto*—I do not say more accurate, but—more plausible than others. Unlike the arithmetic mean, it is not at all affected by the paradox pointed out by Mr. Sauerbeck in his article in the ECONOMIC JOURNAL (Vol. V. p. 163), that the extent of a fall (or rise) appears slightly different according as we start from a high or low basis

the index-number is the sport of the particular system of weights or species of mean which may be adopted. It is a more serious objection, expressed in former papers, that the result is materially affected when we take in additional data,[1] combining with Mr. Sauerbeck's forty-five prices the sixty-nine other prices treated by Soetbeer or his successors.[2] To reply that these commodities are unimportant in respect of quantity does not appear to me permissible *so long as* we treat the problem as simply statistical and purely objective.[3] From this point of view the *quæsitum* is such as the average barometric pressure at a certain time of day, to be ascertained, it might be, from observations with different barometers. For this scientific purpose there would be no propriety in attaching more importance to the observations made with barometers in which the column of liquid had a larger sectional area.[4]

The case may be as if it were required to find the average rise of the tide along an indented shore by observing the height of the water in several creeks. If the average of forty-five observations was materially altered by taking in sixty-nine additional ones we might conclude that we had not at first observed a sufficient number of samples. Perhaps we should have to content ourselves with a very rough figure, unless we took into account some practical purpose for the sake of which the measurement was undertaken. For instance, with reference to the purpose of using the reflux of the tide for the generation of energy, it might be desired to have a measure of the comparative number of foot-pounds available at different seasons. With reference to such a purpose no great error would be incurred by leaving out of account the smaller creeks. In such a case the Calculus of Probabilities by itself could tell us only the whereabouts of the

[1] This transition corresponds to division (3) of the analysis in the second Memorandum, p. 190 *et seq.* As observed there (p. 194), there is a greater inductive hazard involved in passing to new commodities than in allowing for inaccuracy in the weights of a constant set of commodities.

[2] See ECONOMIC JOURNAL, Vol. V. p. 110, and Mr. Sauerbeck's article in the same volume.

[3] As I understand Mr. Pierson to mean in the first paragraphs of the extract from his *Goldmangel* given in his article in the ECONOMIC JOURNAL (Vol. V. p. 331).

[4] Unless, indeed, there were some ground for believing that the smaller size was accompanied with some defect in the qualities of a good measurer: that the observations afforded by the thinner tube, or the commodity consumed in smaller quantities, were more liable to disturbance, or less independent, than other observations. Mr. Sauerbeck has suggested some reason for believing this in the case of commodities which are commercially unimportant (ECONOMIC JOURNAL, Vol. V. p. 171). Another reason has been suggested by the present writer (H, above, p. 247).

required average; the estimation of utility must be called in to render the result precise.

The direction to a practical purpose is the *second* attribute of an index-number which Mr. Pierson leaves out of account—not, indeed, ignoring this property, but deliberately omitting it, for reasons which he has given in a former paper.[1]

"One person consumes much bread and little meat; one person smokes tobacco, another drinks wine, a third neither smokes nor drinks, but makes a collection of books and etchings. In order to judge of the influence on the material condition of men exercised by the variations of prices it would be necessary to divide people into numerous groups, because the relative importance of commodities differs according to individual wants."[2]

There is, no doubt, much wisdom in these reflections; and I fully admit that the eminent author in his earlier and more temperate criticism of index-numbers has made important contributions to the determination of the probabilities and utilities that are pertinent to the subject. I submit, however, the following considerations as a counterpoise to his present scepticism:—

(1) Is it certain that the ground of weighting the variations in price according to their importance with reference to human welfare must be of the subjective kind just considered: taking account of individual wants? Is not a more objective criterion afforded by the increase in the amount of currency which would be required, in the case of appreciation, to raise a commodity to its original price, according to which criterion more weight should be assigned to those commodities which, being circulated in greater quantities, make greater demand on the currency?[3]

(2) With respect to more subjective determinations of importance, the mere diversity of tastes would not be fatal, I think, provided that the expenditure of different individuals is distributed among the different individuals in a normally sporadic fashion,[4] so that a particular system of quantities of commodities consumed tends to occur with maximum frequency, other systems

[1] Economic Journal, Vol. V. p. 331, quoting from *Goldmangel*, pp. 8–10.

[2] *Cp.* Professor Marshall in the *Contemporary Review* for 1887, p. 372.

[3] The allusion is here to the method described in the third of the Memoranda as Professor Foxwell's method (H, above, p. 261). "In averaging the respective price-variations he would assign to each an importance proportioned to the corresponding value." . . . "The question set to us is a pure currency question; and the answer to be sought primarily is not by how much are debts to be scaled up or down, but by how much the metallic currency is to be multiplied in order that the monetary *status in quo* may be restored."

[4] The variations in the quantities consumed with the price (Pierson, *loc. cit.*) might, I think, be treated as magnitudes of the second order.

with less and less frequency according to a well-known law.[1] It must be presumed also that the income of the individuals is about the same [2]—or, rather, distributed normally about an average. Under these circumstances it would be proper to take the average quantities consumed for the weights of the price ratios.

Where these conditions are not fulfilled the proper course would seem to be to construct index-numbers for the different strata of society each of which may have a *type* of expenditure and income in the sense above indicated. The various index-numbers thus constituted would almost certainly differ from each other less than Mr. Sauerbeck's and Soetbeer's (in recent times) have done; they would probably agree better with Mr. Sauerbeck's index-number, in which the component commodities are selected with some regard to their importance to the consumer, than with Soetbeer's, in which no such selection is made.

If practical exigencies require that some one measure of utility should be framed by combining the index-numbers pertaining to different strata of society, then presumably more importance should be assigned to that one which pertains to the masses.[3]

Upon some such principles may be justified the conclusion which Mr. Sauerbeck reaches in his discussion of this matter in the ECONOMIC JOURNAL, Vol. V. p. 171: "Small articles should not be taken acount of in an index-number constructed like Soetbeer's" (that is, a simple arithmetic average of relative prices).[4]

Let it be freely admitted that this measurement of utility has

[1] The prevalence of the *Compound Law of Error*, or probability function of several variables, is proved for the attributes of organisms by the researches of Messrs. Galton and Weldon. With respect to its prevalence and significance in social phenomena see *Statistical Correlation between Social Phenomena*, by the present writer, in the *Journal of the Statistical Society* for December 1893.

[2] On the conditions postulated for the measurement of utility see Professor Marshall's *Principles of Economics*, 3rd edition, Book I. ch. iv.

[3] Because they are more in number and the final utility of money to them is greater.

[4] Analogous remarks apply to the construction of an index-number for measuring the appreciation or depreciation of money, not by the variation in the utility, which is procured by the unit of money, but by the variation in the disutility of labour, by which a unit of money is procured. This is the Labour Standard discussed in the third Memorandum (1889). This method of measuring appreciation has been adopted by Professor Simon Newcomb and some other eminent writers (see J, above, p. 345). It has been unfavourably criticised by Professor Foxwell in the *National Review* for January 1895. No doubt the measurement of appreciation in terms either of disutility or of utility becomes a delicate matter when the production and the consumption of goods per head vary. The subject has been recently discussed by several able writers in the American Academy for Political Science.

not quite the objective character of physical science. It may nevertheless be a postulate of practical economics.[1]

It sometimes happens that an original thinker who rebels against unscientific assumptions himself assumes first principles which are not more demonstrable than the received ones. Of this character, if my interpretation is right, is Mr. Pierson's tacit assumption that the *primâ facie* proper method of dealing with observed variations in price is, in his own words: "If a pound of sugar, a pound of wheat, a yard of cotton yarn, and whatever else is purchasable could be bought in the period 1847–1850 for a sum of money which we call 100, and this sum of money has risen in the period 1851–1860 to 116, we are fully entitled to conclude that the purchasing power of money in those years has fallen in the proportion of 116 to 100." [2]

And, again: "Let us suppose three commodities, costing (A) 20*d.*, (B) 12*d.*, and (C) 4*d.* a pound, and falling respectively to 15*d.*, 2*d.*, and 1*d.* a pound. This will be an average fall at the rate of 100 to 50, for—

$$20 + 12 + 4 = 36d.$$
$$15 + 2 + 1 = 18d.$$

In other terms, twice the quantities of these commodities will be purchasable for the same amount of money as before." But "index-numbers" (that is, the ordinary arithmetic mean of the price-ratios expressed as percentages) will show a fall from 100 to 38·88. "Which is manifestly wrong," says Mr. Pierson.

And, again, of the ordinary arithmetic and geometric mean of price-ratios, "both methods are wrong": as disagreeing with a method which in its essential feature resembles that which has just been described. If we consider the ratio between the prices at different epochs to constitute the datum of observation, Mr. Pierson's method of combining these data is to weight each observation with the money-value of the unit of avoirdupois or volume measure.[3]

Where is the peculiar propriety of this system of weighting, according to which a variation in the price of, say, *argon* or *iridium* should count for more than a variation in the price of *coals* or *cotton*, because each pound-weight of the former articles is dearer than a pound-weight of the latter? I do not now so much complain that the system has no reference to any useful purpose.

[1] The practical validity of index-numbers is well shown in Mr. L. L. Price's excellent *Money and its Relation to Prices.*

[2] Economic Journal, Vol. V. p. 331.

[3] Cp. H, above, p. 258.

The statistician is within his rights in making abstraction of human welfare; but, viewing the problem as purely objective and merely statistical, why should we employ this principle of preference? [1]

Let us, after Mill [2] and Hume, represent the phenomenon under consideration, depreciation,[3] by supposing that, *ceteris paribus*, every piece of money and instrument of credit has been on an average increased in a certain ratio. With reference to the measurement of that ratio [4] it is surely an accidental circumstance whether the unit of mass or volume of one commodity as compared with another exchanges (prior to the depreciation) for more or fewer units of money. The following would be a fair analogy in physical measurement of the proposed system of weighting. Let it be required to determine the expansion due to a rise in temperature for the diamond, from observations made on several portions of the substance. Lay out several units of money—say pounds sterling or ten-pound notes—in purchasing so many parcels of diamonds. Make an observation with each of these portions, and *weight* each observation with the mass or the volume of the diamond which is obtained in exchange for the unit of money. According to this arrangement an observation on a compact and glittering diamond shall count for less than one made upon a mass of less commercial value. This system of weighting the observations is on a par with Mr. Pierson's system. The number of units of mass or volume exchanged for the unit of money is not more irrelevant in the physical measurement than is the number of units of money exchanged for the unit of mass or volume in the monetary measurement.

Yet Mr. Pierson treats this system as *primâ facie* reasonable, and abandons it only because its two modes—mass and volume—lead, in imaginary examples, to inconsistent results. And he deduces from this inconsistency the futility of the whole measurement: that "all attempts to calculate and represent average movements of prices either by index-numbers or otherwise ought to be abandoned."

Let us see how this sort of objection would apply to the typical physical problem above instanced, the determination of the diurnal

[1] The principle does not seem to have found much favour among the constructors of index-numbers. It is mentioned by Geyer in his *Theorie und Praxis der Zettelbankkursens*, appendix vi. But he at once introduces a modification which makes his system practically identical with the ordinary arithmetic mean.

[2] Mill, *Political Economy*, Book III. ch. viii. § 2.

[3] Or, *mutatis mutandis*, appreciation.

[4] Another appropriate conception of the *quæsitum* might be, I think, the change in the quantity of final utility which is the equivalent of the unit of money, assuming the marginal utility of goods not to have altered.

variation of the average atmospheric pressure. Suppose that the observations have been made with barometers consisting of different liquids—mercury, water, etc. Weight each observation *first* inversely as the money-value of the unit-weight of the corresponding liquid, and *secondly* inversely as the money-value of the unit-volume of the liquid. Then, *if* the observations are not sporadically dispersed, but collected at two or three points, it will make all the difference whether the first or the second system of weighting be employed. Therefore the calculation of average variations in barometric pressure—performed by Laplace and approved by Mill—is to be "abandoned altogether" as "faulty in principle."