Reprinted from Public Opinion Quarterly, Vol. 18, No. 3

Bandwagon and Underdog Effects and the Possibility of Election Predictions

By HERBERT A. SIMON

Social research has often been attacked on the grounds that the research itself so altered the original situations as to make accurate predictions impossible. In this article, the author deals particularly with the effects of published predictions and the adjustments necessary to account for reactions to those predictions.

Herbert A. Simon is Professor of Administration in the Graduate School of Industrial Administration at the Carnegie Institute of Technology.

THERE HAS been a considerable amount of discussion, and some empirical investigation, of the possibility that the publication of an election prediction (particularly one based on poll data) might influence voting behavior, and hence—among other effects—falsify the prediction. Practically, we might be more interested in the influence of an election prediction from the standpoint of its significance for the working of democratic government than from the standpoint of its significance for the methodology of social science. Nevertheless, the latter question—involving as it does such issues as the "self-confirming" and "self-falsifying" prophecies, "pluralistic ignorance," and, indeed the entire possibility of public prediction in the social sciences—is of considerable importance in its own right. It is with the latter issue that we shall be chiefly concerned: Under what conditions will a public prediction, although it influences behavior, still be confirmed?[1]

Before we analyse this point, it will be helpful to define what we mean by a "bandwagon" and by an "underdog" effect. It is supposed that the voting behavior of at least some persons is a function of their expectations of the election outcome; published poll data are assumed to influence

[1]A diligent scholar could, no doubt, trace the history of this problem back to Aristotle. The author's first encounter with it came through the teaching and writings of Professors Knight and von Hayek (the latter in his *Economic* articles on "Scientism and the Study of Society"), and through discussions with Professor Milton Singer. It is considered briefly and inconclusively in *Administrative Behavior*, pp. 251–2. More recently, Emile Grunberg and Franco Modigliani, solved the problem for certain cases of economic prediction, and their solution suggested a generalization by means of the fixed-point theorem of topology. For the present exposition, I have drawn heavily on their paper reporting these results in the *Journal of Political Economy*, December, 1954. I am grateful to the Ford Foundation for a grant that made this work possible.

these expectations, hence to affect the voting behavior of these persons. If persons are more likely to vote for a candidate when they expect him to win than when they expect him to lose, we have a "bandwagon" effect; if the opposite holds, we have an "underdog" effect. Notice that we are not concerned with the converse mechanism: the effect of an individual's own voting preference upon his expectations of the election outcome.

THE CONFIRMATION OF A PUBLISHED PREDICTION

Of course, the question of the confirmation of a prediction is of interest only if the publication of the prediction is supposed to affect the behavior of at least some people (e.g., a bandwagon or underdog effect). But in this case we must carefully distinguish between: (a) what the outcome would have been *in the absence of a published prediction;* and (b) what the outcome actually was *after a prediction had been published.*

We take as a specific example the percentage of voters who will vote for candidate A in a two-candidate election. We let:

I = the percentage of voters who would have voted for A in the absence of the published prediction.

V = the percentage of the voters who in fact voted for A after publication of the prediction.

The difference between these two percentages, (I–V), measures the effect upon the voting behavior of publishing the prediction. Now it is reasonable to assume that, V, the percentage of persons who actually voted for A, depended upon two factors: (1) the percentage who intended to vote for him prior to publication of the prediction—that is, I; and (2) the prediction itself—that is, the percentage who, according to the prediction, intended to vote for him. This latter percentage we shall call P. The assumption then amounts to asserting that V is some function of I and P, or symbolically:

(1) $V = f(I, P)$

Now let us regard I as a "given" quantity—it is whatever it is, although we may or may not actually know its value. If I is fixed, then V may be regarded as a function of P alone. That is, given I, the percentage, V, of voters who will vote for A still depends on the published prediction, P.

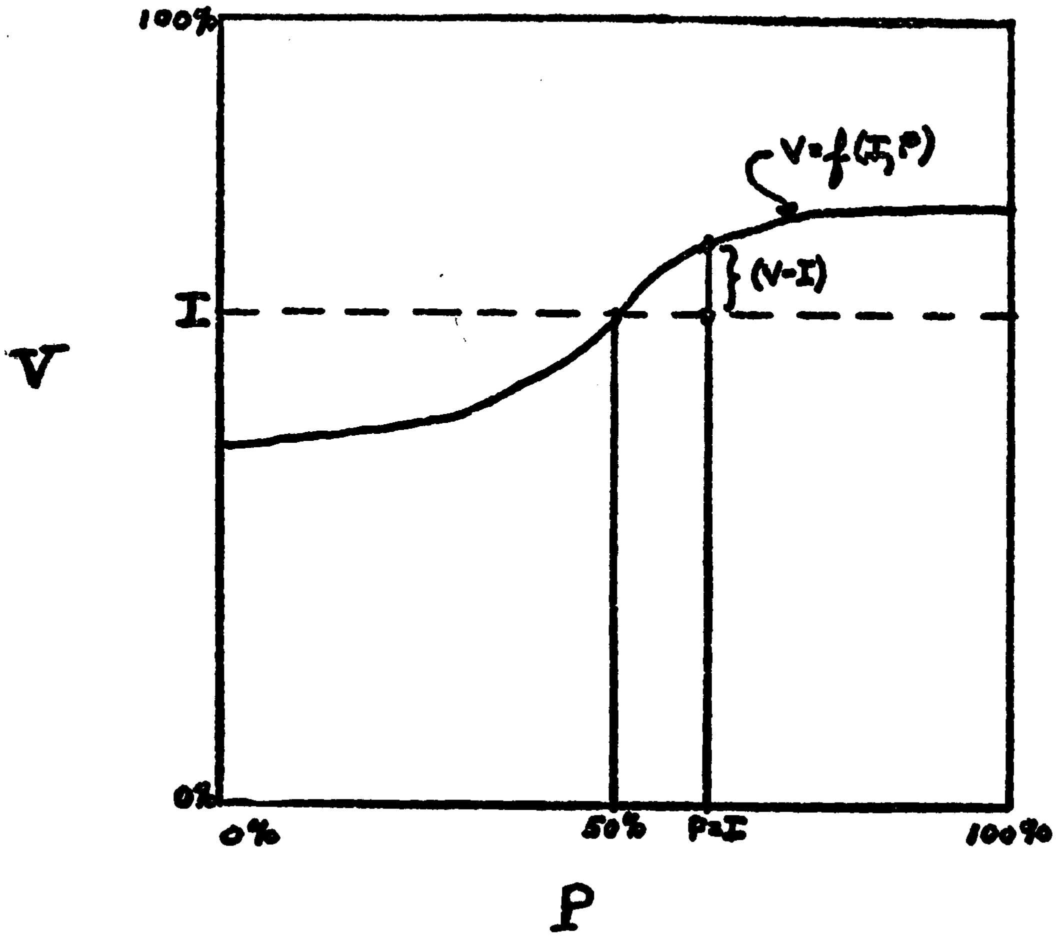
The situation is illustrated in Figure 1. The x-axis measures P, the published prediction, which may range from 0 per cent to 100 per cent. The y-axis measures V, the actual vote, which may also range from 0 per cent to 100 per cent. Two hypothetical curves have been drawn in the figure. The broken horizontal line, intersecting the y-axis at I, shows what the vote would be if the prediction were made privately, but not published. In this case, the vote is exactly the same, no matter what value we assume for P, for the private prediction cannot affect the vote. Thus, we have:

(2) $V = I$,

which is exactly what the horizontal line shows.

The solid curve shows the assumed relation between V and P derived from equation (1). Figure 1 illustrates the particular case of a bandwagon effect. The solid curve is drawn on the assumption that if a prediction of a victory for A is published (P>50%), then V>I—that is, some people will switch their votes to A; while if victory is predicted for A's opponent (P<50%), then V<I—that is, some people will switch their votes to A's opponent. The vertical distance between the two lines, (I–V), is a measure of the effect of the prediction. This vertical distance is of course not constant, but as we have seen, depends upon P.

FIGURE 1



Under what circumstances will we say that the prediction, P, is "confirmed"? The pollster generally proceeds, in the design of his sample, as if he were trying to make an accurate estimate of I. But in fact, the only way in which he can assess his accuracy is to compare his prediction, after the election, with V. If the poll is accurate in the first sense, if $P = I$, and if publication of the poll does in fact, have an effect on voting behavior, then we will have:

(3) $P = I = V + (I-V) \neq V$,

for under the assumed conditions, (I–V) will not be zero.

ACCOUNTING FOR PUBLICATION EFFECTS

We see that the only way in which the pollster can arrive at a prediction that will coincide with the election result is by privately adjusting his poll results (which we assume for the moment to be an accurate estimate of I) for the effect that their publication will have upon the voters' behavior. But is even this possible? If he makes such an adjustment, will not the adjustment itself alter the effect of the prediction and again lead to its own falsification? Is there not involved here a vicious circle, whereby any attempt to anticipate the reactions of the voters alters those reactions and hence invalidates the prediction?

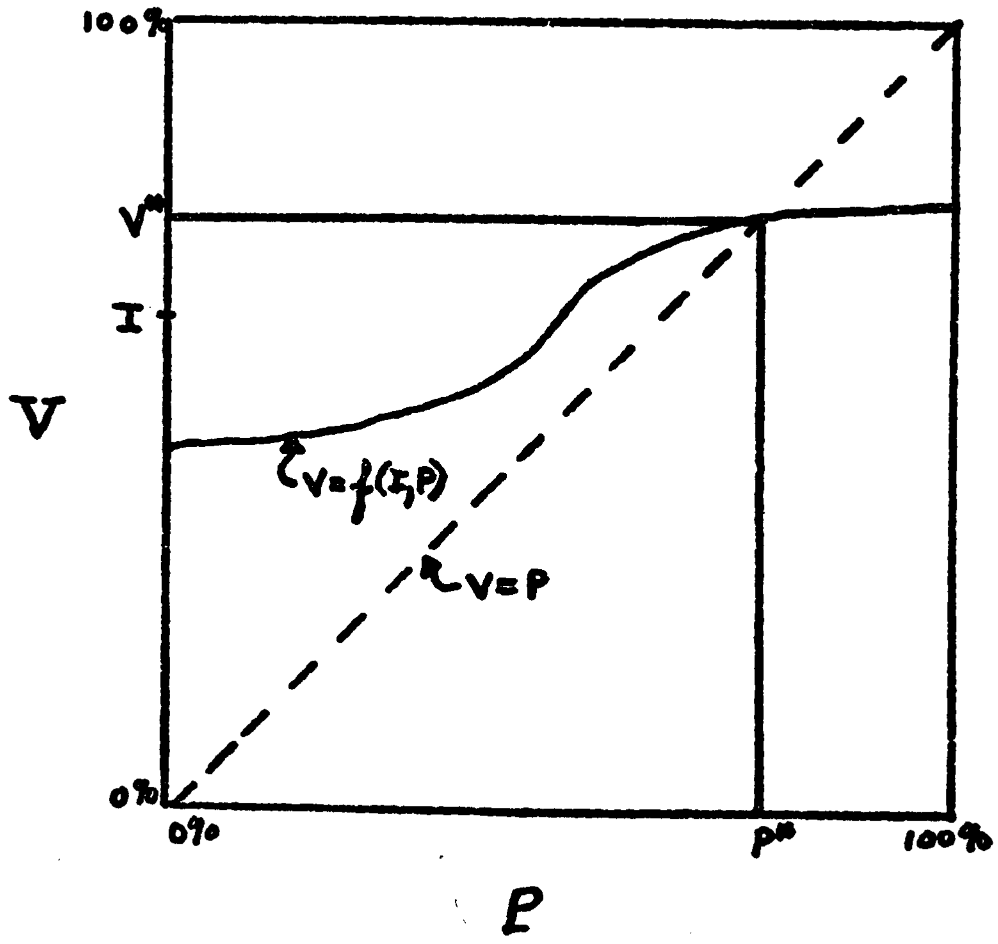
In principle, the last question can be answered in the negative: there is no vicious circle. Whether accurate predictions can be made *in practice* will be discussed later.

The "in principle" situation is illustrated in Figure 2. We have taken as our criterion of confirmation of a published prediction that the actual behavior should coincide with the prediction. The axes in Figure 2 represent P and V, respectively, as before. We have drawn again the solid curve that represents V as a function of P (from equation (1), assuming I to be fixed and given). A broken straight line through the origin with a slope of 45° has also been entered in the figure. For any point on this broken line, $V = P$, the actual voting percentage is equal to the predicted percentage.

Now consider the point of intersection of the two curves—the solid curve and the broken line. Let us call the specific value of P at the point where this intersection occurs P^*, and the value of V at this point V^*. Because this point lies on the solid curve, it is true that if the pollster published the prediction that A will receive P^* per cent of the vote, he will in fact receive V^*. But because this point also lies on the broken line, it is true that $P^* = V^*$—that the actual vote will be equal to the predicted vote. (On the other hand there is no reason to suppose that $P^* = I$, and in general this will not be the case.)

We see, therefore, that if the curve given by equation (1) intersects with the line $P = V$, it will be possible, if the point of intersection is known, to make a prediction that will be confirmed. But will such a point of intersection always exist? It can be shown that it will, under the sole condition that V in equation (1) is a *continuous* function of P (roughly, that the function not have any finite "jumps").

FIGURE 2



A formal proof of the theorem will not be given here. It is a classical theorem of topology due to Brouwer (the "fixed-point" theorem), and a non-technical exposition may be found in *What is Mathematics?*.[2] The reader who does not demand a rigorous proof may satisfy himself of the correctness of the theorem by graphical means. Construct a figure like Figure 2, but omit the solid curve. Mark any point on the y-axis between V = 0 per cent and V = 100 per cent; and a second arbitrary point on the vertical line, P = 100 per cent, within the same limits. Now try to connect these two points, without lifting the pencil from the paper, without going outside the limits 0 per cent to 100 per cent for V and P (that is, without going outside the square), and without intersecting the broken line. Since this is impossible, *any* continuous curve relating V and P for the whole range of values $0\% \leqq P \leqq 100\%$ must intersect the line V = P in at least one point.

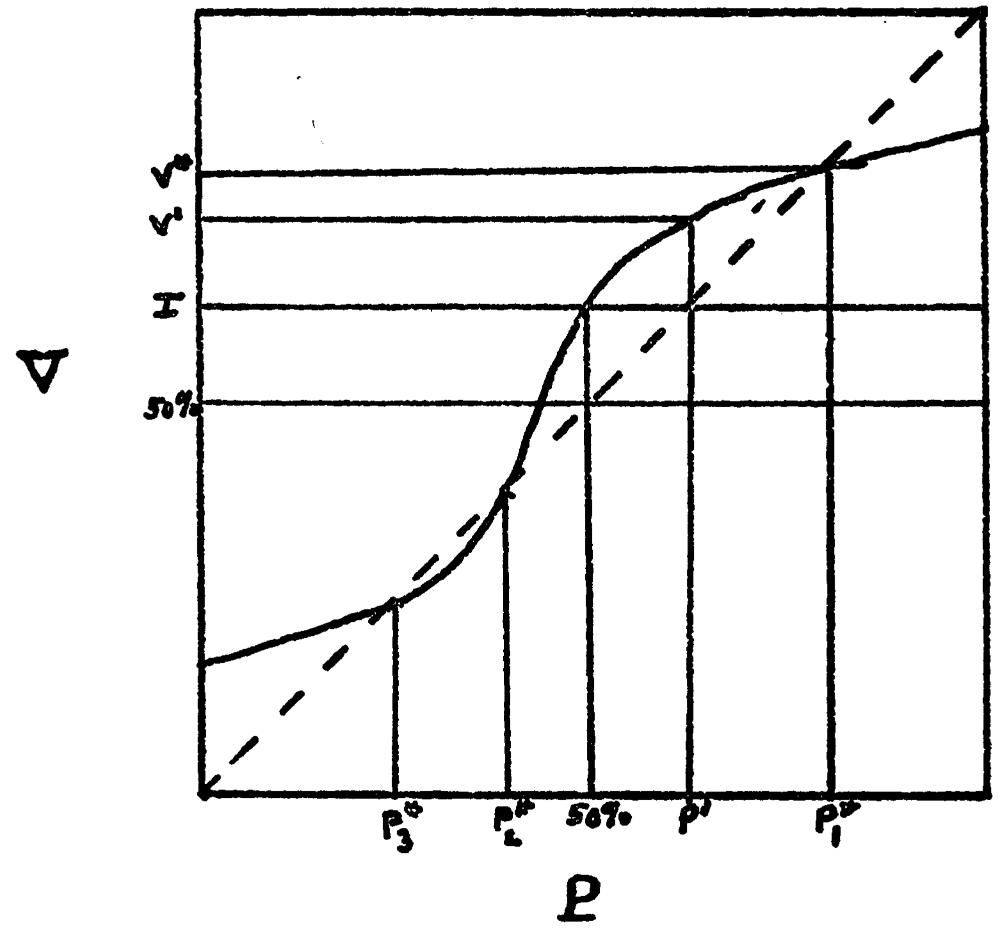
[2]Courant and Robbins, *What is Mathematics?*, pp. 251–255.

PREDICTION PROBLEMS

We have proved that it is always possible in principle to take account of reactions to a published prediction in such a way that the prediction will be confirmed by the event. But can this procedure be carried out in practice by a pollster? Stated otherwise, what information would the pollster have to possess in order to adjust his prediction for the anticipated reaction? The answer is that he would have to know the function (1), at least in the neighborhood of the actual value of I, and that he would have to have an accurate estimate of I. It is the aim of his poll to give him the latter; it is less obvious where he can obtain the former.

Even if the adjustment factor, as set forth in equation (1), is not known precisely, it may be possible to improve a prediction on the basis of knowledge of the direction of the reaction (the sign of (I–V)). In Figure 3, we illustrate the case where we have an accurate estimate of I (from a poll), and where we know that there is a "bandwagon" effect. Again, we draw a solid curve to represent the relation V = f(I,P). If there is a bandwagon effect, then this curve must lie above the horizontal straight line,

FIGURE 3



$V = I$, when $P>50$ per cent, and below that line when $P<50$ per cent. For in the former case, when the prediction is published some voters will switch to candidate A, while in the latter case some voters will switch to his opponent.

A prediction, P^*, will be confirmed if it lies on an intersection of the solid curve with the 45° diagonal. It can be seen from the figure (and can be shown rigorously by another application of the fixed-point theorem) that there always exists at least one prediction, P_1^*, with the following two properties: (a) the prediction, if published, will be confirmed, and (b) publication of the prediction will not change the outcome of the election (i.e., $P_1^*>50\%$ only if $I>50\%$). However, examination of the figure will show that there may also exist other values of P^* possessing the first property but not the second. If one of these latter predictions is published, it will be confirmed by the election result, but the candidate who would have won if no prediction had been published will be defeated. In the figure as drawn, two such values of P^* exist, corresponding to the two intersections of the solid curve with the 45° line in the lower left-hand quarter of the figure. It is intuitively obvious that such points will exist only if the bandwagon effect is "very strong." The exact conditions can be stated analytically, but will be omitted here.

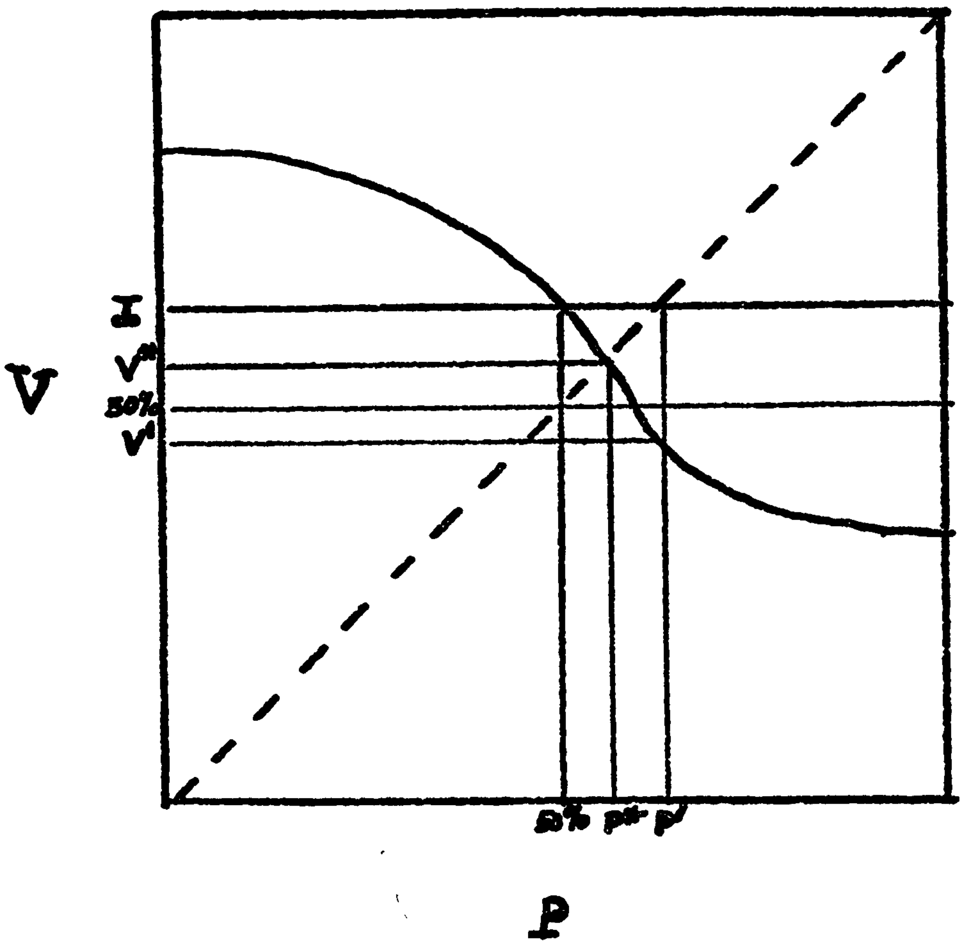
On the other hand, it can be seen from the figure that if the unadjusted poll result, P', is an accurate estimate of I, then, in the case of a bandwagon effect, the publication of P' cannot change the outcome of the election.[3]

The case of an "underdog" effect is illustrated in the same way in Figure 4. In this case we see by examination of the figure (and may prove analytically) that the publication of a prediction, P^*, correctly adjusted for the reaction to its publication, *cannot* reverse the outcome of a two-candidate election; while the publication of a prediction, P', that is an accurate estimate of I, *may* reverse the outcome. Again, the latter possibility will occur only when the underdog effect is "very strong."

From the results of this section, we see that there is no simple relationship between the "adjustment" of poll results prior to publication and the "manipulation" of an election. If we assume the original poll to be accurate, in the usual sense; and if by "adjustment" we mean taking account (accurately) of the reaction to publication; then: (1) adjustment can have a manipulatory effect in the bandwagon case, but not in the

[3]The bandwagon effect may, of course, change the outcome of a contest among three or more candidates by diverting votes from the weakest candidate to one of the stronger. Also, the published prediction may affect the number of voters preferring a given candidate who actually go to the polls. The model can be generalized to permit discussion of such effects.

FIGURE 4



underdog case, and need not in either case; (2) failure to adjust can have a manipulatory effect in the underdog case, but not in the bandwagon case. By "manipulatory effect" we mean, of course, not merely an effect on the voting percentages, but an actual reversal of the outcome. I hasten to add that pollsters might experience some difficulty in explaining even "non-manipulatory" adjustments to members of Congressional investigating committees.

SUMMARY

In this paper we have raised and answered the question of whether, and under what circumstances, a published prediction will be confirmed, even if there is reaction to the prediction. The problem was stated, and a graphical method for analyzing it was set forth in Section I. In Section II, it was shown that it is always possible *in principle* to make a public prediction that will be confirmed by the event. This proof refutes allegations commonly made about the impossibility, in principle, of correct prediction of social behavior. In Section III, certain practical problems were examined that arise in actually making predictions. It was shown that correct prediction requires at least some knowledge of the reaction function; that whether publication of a prediction (adjusted for expected reactions or unadjusted) will affect the outcome of an election between two candidates depends on the shape of the reaction function; and that publication of an adjusted prediction will sometimes have more, and sometimes less effect on behavior than publication of an unadjusted prediction. These results were applied in particular to the reaction functions corresponding to "bandwagon" and "underdog" effects.