Thinking about the Heiritsu-sei: A Structural-Learning Approach*

Steven R. Reed**

Political scientists know a great deal about how the single-member district electoral system (SMD) operates. We also understand quite a bit about proportional representation (PR). It would seem to follow, therefore, that we would be able to make some reasonable predictions about how the two systems should work when combined in parallel, as in the heiritsu-sei (parallel system) recently enacted in Japan. In fact, however, political scientists are not even certain about how to approach the problem.

A rational choice approach might try to calculate what a perfectly rational candidate would do to maximize his probability of winning a seat or what a perfectly rational party would do to maximize its number of seats. I do not think that this approach can get us very far. In fact, I would argue that the rational choice approach is most useful for explaining actions after the fact. The rational choice approach tends to observe behavior and

1) One critic of the rational choice approach suggests that economists spend a lot of energy demonstrating that “the actual is possible,” i.e., that what we observe happening in the real world can be fitted into a rational choice model. Alexander Rosenberg, Economics-Mathematical Politics or Science of Diminishing Returns? (The University of Chicago Press, 1992), p. 74 and p. 105.

To be fair, several scholars working within the rational choice framework have recognized some of the limitations of the approach and reached conclusions similar to mine. For example, in "Rational Choice and Social Theory," The Journal of Philosophy (1994), 71-87, Debra Satz and John Ferejohn offer a "nonpsychological" interpretation of rational choice and argue that rational choice explanations work best when individual action is severely constrained leading to a structural interpretation similar to the one I offer here. Similarly, in the second chapter of Nested Games: Rational Choice in Comparative Politics (University of California Press, 1990), George Tsebelis presents an interpretation of rational choice that overlaps a great deal with my own structural-learning model. Despite these similarities, however, I believe the approach illustrated here differs greatly from the rational choice perspective.

* I would like to thank Wada Junichiro, Hosono Sukehiro, Yokoyama Akira, Sone Yasunori, Kobayashi Yoshiaki, Leonard Schoppa, and Gary Cox for their helpful comments on an earlier draft of this article.

** 著者は中央大学教授。
※ 本稿は1994年8月受理。
then find a rational explanation for why someone would act that way, i.e., to rationalize actions already taken. In this paper I will present a structural-learning approach as an alternative to the rational choice approach and illustrate this approach by analyzing the *heiritsu-sei*. I will argue that the structural-learning approach solves several of the problems associated with rational choice and provides a dynamic analysis more effective at producing predictions before the fact.

**The Structural-Learning Approach**

Any given action can be explained by the interaction of the actor's preferences and the options available to him. Rational choice focuses on preferences, normally overlooking equally important factors in the actor's environment. In many cases, the situation virtually determines the choice because one option is clearly more attractive than the others. For example, if a robber gives you two choices, your money or your life, it is not necessary to explore your values very deeply to predict that you will prefer to give up your money and rather than your life. The social sciences have concentrated on cases in which choices are narrowly constrained and the options vary along only one or two dimensions: will the actor buy chocolate or vanilla ice cream, will he vote for the Democrat or the Republican. In such cases, the actor's preferences clearly dominate the situation because the options are so nearly equal and so easily comparable. The options presented by real life situations are often unbalanced and/or hard to compare. In many cases, preferences are less important in predicting behavior than the options produced by the situation

Rational choice analysts observe behavior and deduce preferences. An actor who regularly chooses chocolate over vanilla ice cream can safely be said to prefer chocolate because we can be certain that both options were equally available to him. In less constrained situations, however, such reasoning can produce absurd conclusions. For example, one often hears the unsophisticated argument that homeless people like living in the train station: if they did not like living there, they would move, probably to a mansion. If one considers the options actually available to homeless people, inside the station or on the street, a place with a roof or one outside exposed to the elements, choosing to live in the train station is not hard to understand. Variations in housing choices is better explained by the options available than by preferences. Do Japanese prefer to live in small (by international standards) apartments? Those Japanese who have the option of living in larger, less traditionally Japanese-style housing, tend to do so. It is more reasonable to assume that everyone has basically similar preferences for

---

2) Daniel Pound, who teaches political theory at the University of Alabama, regularly assigns students the task of keeping a diary of all the choices they make in a day. The students are always amazed at how few occasions they have for making decisions in the sense of weighing options and calculating costs and benefits. For a more extended discussion of the structural-learning approach, see Steven R. Reed, *Making Common Sense of Japan* (University of Pittsburgh Press, 1993), Chapter Three.

housing, that variation in tastes apply only to issues like the color of the walls, and that major choices are determined by the options available.

To take a more political example, one often hears the argument that the Japanese electorate liked the LDP governments between 1955 and 1993 despite the LDP's many corruption scandals: the voters must have liked the LDP or they would have voted for some other party. However, since at least 1976, with the Lockheed Scandal and the creation of the New Liberal Club, Japanese voters have consistently moved away from the LDP whenever they were offered an attractive alternative. I would argue that one can explain more about the electoral fortunes of the LDP by looking at the alternatives offered to the electorate than by voters' attitudes toward the LDP. In particular, it was a change in the alternatives available to the voters, not a change in their preferences, that resulted in the end of the Liberal Democrat's long reign in 1993.

Analysis of the alternatives available in the environment should be given equal weight with analysis of personal preferences. Moreover, since almost all social science analysis heretofore has focused on preferences, whether the rational choice or the cultural approach, progress will be faster if, for the time being, we focus on the relatively ignored aspects of the options available.

Structures

The most important type of situational variable is a structure. There has been a great deal of confusion over how to define structures but, starting from the distinction between preferences and choices available, we can define a structure as a stable and/or repeated situation which defines a set of options. Structures should not be defined as a set of written rules or a formal organization chart. Structural analysis must involve empirical research because informal rules are often as powerful as formal laws, and because formal rules are often little more than codifications of agreements worked out through informal political competition.

Repeated situations offer the best opportunities for scientific analysis in the social sciences because such structures produce large amounts of comparable data. Thus, in political science the greatest progress has been made in the analysis of budgeting and elections. In a stable or repeated situation actors learn how to get the rewards out of the system. Elsewhere I have argued that the connection between structure and behavior is not rationality but learning. The greatest weakness of rational choice theory is that it is based on poor psychology. Learning theory makes fewer and more realistic assump-

---


7) Similarly, Rosenberg, op. cit., argues that, within economics, it is agricultural economics which has made the greatest progress because of "the far greater completeness of the data base deemed relevant to agricultural policy. As much as anything, the existence of these systematic and reliable data is responsible for the success of the subject." (p.66).

Thinking about the Heiritsu-sei: A Structural-Learning Approach

tions about human psychology. Instead of many complex, debatable, and even clearly inaccurate psychological assumptions, we need only assume two things: (1) offered a choice between more or less of a good, actors will generally prefer more; and (2) actors are capable of adjusting their behavior in response to feedback from their environment. We thus move from a set of assumptions about rationality that could not be met by a supercomputer to a set of assumptions about learning which are easily met by an amoeba.

Structural-learning theory focuses on the actual options faced by actors, the feedback from the environment, not to a hypothetical range of possible options. I do not assume that actors maximize some function. For example, political parties do not maximize the number of seats they win, but if presented with the options of winning either zero seats or one seat, most parties will choose the latter. Japan’s middle-sized districts offer a clear illustration of the difference. Using a rational choice approach, it can be demonstrated that the optimal number of candidates equals one more than the number of seats in the district. However, parties regularly error by nominating too many candidates and these errors can cost the party seats. This phenomenon is called tomodaore (falling together). Tomodaore presents a clear choice to the party: continue the current strategy or reduce the number of candidates nominated and win one more seat with the same number of votes. Although a party might also try to increase their votes to elect all of their current nominees, in most actual cases of tomodaore such hopes are clearly unrealistic. Clearer choices seldom occur in politics, but Japanese political parties often continue running excess candidates for several elections before making the obviously rational decision. Japanese political parties are not rational but they do learn over time. Interestingly, it is the Communists who respond the quickest to a tomodaore. Very little of the JCP’s political behavior can be explained by assuming that the party maximizes its seats in the Diet, but when offered a choice between winning a seat or not, they choose to take the seat.

One need know very little about the preferences of political parties to predict their behavior in the face of tomodaore. More generally, if one understands how an electoral system rewards or punishes different strategies with more or fewer seats, one can calculate the advantages and disadvantages of different strategies and perhaps arrive at the optimal strategy. Much


of the analysis will rely on tools familiar to rational choice theorists, notably game theory, but because the assumptions of the model are simpler, most of the advanced math will prove superfluous. Moreover, once an optimum or equilibrium is known, structural-learning theory predicts a slow, trial-and-error process. Knowing the equilibrium of a structure is only the first step toward predicting behavior.

Learning

At the individual level, the connection between structure and behavior is learning, but people do not always learn. In structural-learning theory, learning is a hypothesis, not an assumption. Cases of people refusing to learn from what seems to be the most obvious situations are not hard to find. At the aggregate level, at least two other mechanisms produce evolution toward the equilibrium point. First, people who refuse to learn how the system works tend to be relatively powerless within the system. A Congressman who refuses to play the political game, for example by refusing to compromise, may be able to get elected and re-elected, but he will never be powerful inside the legislature. People who play the game properly, who act in the ways which are rewarded by the system, will tend to rise to the top of the organization. People who are good at the game, rise to leadership positions. People who neither learn nor adjust their values tend to retire, whether voluntarily or not. Given a stable structure and time, the system tends to evolve toward a predictable equilibrium. It is the goal of structural-learning analysis to identify this equilibrium, the mechanism which produce it, and the factors which upset it.

A major problem in the rational choice approach is determining who is the relevant actor. One reaches very different conclusions depending on whether one assumes that candidates are maximizing their re-election chances or that political parties are maximizing their number of seats. The structural-learning approach envisions many mechanisms, at both the individual and collective levels, all operating to produce movement towards equilibrium. Thus, one need not know precisely which mechanisms are operating before making predictions and the importance of individual and collective mechanisms becomes an empirical question.

Over time, people learn the rules of the game and adjust their preferences to value the goods produced by the system. People who are good at the game, rise to leadership positions. People who neither learn nor adjust their values tend to retire, whether voluntarily or not. Given a stable structure and time, the system tends to evolve toward a predictable equilibrium.
Since learning is based on feedback and not on rational calculation with full knowledge, one should not be surprised to find stable local equilibria. Even if a unique optimum does exist, there is no guarantee that the system will ever arrive at it. Moreover, the optimum may be rationally and/or scientifically incorrect. When analyzing their environment, people error in systematic, highly predictable ways\(^\text{14}\). Even if an actor makes reasonable deductions from the data provided by his environment, his deductions will be incorrect if the data is biased, and naturally occurring data is always biased. For example, if every foreigner you have ever met were a university professor, it would be scientifically incorrect but nevertheless common sensical to conclude that foreigners are better educated than we are. In fact, it proves to be extremely difficult psychologically to avoid coming to such incorrect conclusions.

Finally, the structural-learning approach expects many outcomes to be path-determined, i.e., where one winds up depends upon where one started. Many social scientific theories have tended to assume that if one knows the relevant independent variables at a given point in time, one should be able to predict the value of the dependent variable at that point in time. For example, if one knows a country’s current level of economic development, the percentage of the labor force engaged in secondary industry, the percentage of the labor force belonging to unions, and the current strength of the socialist parties in the government, one should be able to predict the percentage of the budget spent on social welfare. However, this is analogous to the question: “If you got in a helicopter and flew directly west at fifty miles per hours for two hours, where would you be?” The answer depends on where you started. One of the most consistent findings of studies which compare public policies across nations is the importance of the policy legacy\(^\text{15}\). More generally, in order to make predictions one must not only know how the world works, but also the initial conditions. Dynamic models that include the time dimension should be preferred over static models, especially in the social sciences\(^\text{16}\).

My basic claims for the structural-learning approach are: first, this approach provides a dynamic model that not only deals with equilibrium points but also with the paths and mechanisms by which systems arrive at equilibrium; second, it is based on simpler and more realistic psychological assumptions; and, third, by focusing on situations, the relevant variables are more available for measurement, analysis

\(^{14}\) James G. March and Johan P. Olsen provide a useful summary of these types of errors in Rediscovering Institutions: The Organizational Base of Politics (New York: The Free Press, 1989), p. 40. See also Daniel Kahneman et al. (eds.), Judgement Under Uncertainty (Cambridge University Press, 1982).

\(^{15}\) See, for example, the articles in R. Kent Weaver and Bert A. Rockman (eds.), Do Institutions Matter? (The Brookings Institution, 1993).

\(^{16}\) When a social scientist refer to the natural sciences, he normally has physics in mind and physicists find little use for history. However, if evolutionary biology were taken as the model, we would find more emphasis on reconstructing the history of particular adaptations, path-determined outcomes, and local optima. In “The Psychological Foundations of Culture” in Jerome H. Barkow, Leda Cosmides and John Tooby (eds.), The Adapted Mind (Oxford University Press, 1992) John Tooby and Leda Cosmides describe evolution in terms reminiscent of the way political scientists describe policy making: “In evolution, successive designs are always constructed out of modifications of whatever preexisting structures are there…” (p. 60).
and testing. I hope to illustrate these claims by analyzing the heiritsu-sei and making several predictions about how it will work.

Analyzing the Heiritsu-sei

Analyzing Electoral Systems

An electoral system allocates seats in the legislature based on the distribution of votes among parties and individual candidates and it does so in a mechanical fashion. One can thus analyze an electoral system as a mathematical relationship between a distribution of votes and the resulting distribution of seats. Theil has demonstrated that the relationship between a party’s proportion of the vote and its proportion of the seats must take the following form:

\[ s_i = \frac{v_i^p}{\sum v_i^p} \]

The proportion of seats allocated to the \( i^{th} \) party equals its vote share taken to the \( p^{th} \) power divided by the total votes of each of the parties also taken to the \( p^{th} \) power. The variable that determines the relationship between seats and votes is thus the exponent, \( p \).

When rho is equal to one, the result is perfect proportional representation: a party’s percentage of the seats equals its percentage of the vote. In the simplest two-party case, a \( p \) greater than one produces an S-curve, as illustrated in Figure 1. Similar multi-dimensional figures could be drawn for more complex situations. Figure 1 illustrates the difference between proportional representation (PR) and single-member districts (SMD). Single-member district systems tend to have relatively high rhos producing relatively sharp S-curves. PR, on the other hand, is highly reflective: the distribution of seats accurately reflects the distribution of votes. SMD is more responsive: near the fifty percent midpoint, a change in a party’s percentage of the vote produces a greater change in its percentage of seats. SMD magnifies shifts in public opinion into larger shifts in parliament, and this is one of the characteristics of SMD which makes alternation in power more likely. Thus, the first such alternation in Japan since 1955 occurred in the House of Councillors in 1989 and was due primarily to single-member districts. Taken by themselves, the single-member districts of the upper house have an extremely high rho of five. The LDP vote dropped from 53.8 percent in 1986 to 41.4 percent in 1989, but LDP seats dropped from 92 percent to 12 percent. A change of 12.4 percentage points in votes was magnified into a 80 percentage point change in seats.

Taagepera has derived a theoretical value for rho using a physics-style of analysis:

\[ p = \frac{1}{1 + \frac{1}{2} \log \left( \frac{E}{N} \right)} \]


18) This difference in electoral system is based on a deeper philosophical argument over the proper definition of democracy. See Arend Lijphart, *Democracies: Patterns of Majoritarian and Consensus Government in Twenty-One Countries* (New Haven: Yale University Press, 1984), especially the first two chapters.


where $V$ is the total number of votes, $D$ is the total number of districts, and $M$ is the district magnitude, i.e., the number of seats per district. Thus, the exponent $\rho$ should be the log of the total vote divided by the log of the total number of seats ($DM$) all taken to the $1/M$ power. Using this equation, we can estimate the relationship between seats and votes for any electoral system. For example, in Japan’s system of medium-sized districts, the three-member districts taken separately have a $\rho$ of 1.5, four-member districts 1.33, and five-member districts 1.25.

It is important to note that Theil’s equation is derived mathematically and Taagepera’s is based on a graphical analysis of actual distributions of seats and votes. Neither has the slightest behavioral basis. Both are generalizations about how electoral systems work, not about how people behave. Nevertheless, taken together, they explain the actual relationship between seats and votes remarkably well\[^{21}\].

![Diagram](image-url)

**The Heiritsu-sei**

Under the current proposal each voter will have two votes, one for a SMD election with 300 seats and the other for a PR election with 200 seats in eleven districts ranging in magnitude from 7 to 33, with an average of 18. The results in the two systems will be tabulated separately. Parties, candidates and voters will face some complex strategic choices, particularly because candidates may run in both elections simultaneously. The two systems will interact at the levels of party, candidate and voter strategy, but the mechanical aspects of the translation of votes into seats in the two systems will be independent. We can thus analyze the dynamics of the system using the tools described above. The first step is to calculate the value of $\rho$ for each system. The SMD system will have a rather steep $\rho$ of 3.14 while the PR system will have a $\rho$ of 1.07, deviating only slightly from pure proportionality.

The first prediction to be derived from the structural-learning approach is that the first election will have high entropy, by which I mean that it will be a mess. Although there may well be some prospective reduction in the number of parties through mergers, and though some candidates will retire because they do not expect to win, the number of candidates and parties will still be above equilibrium. Although everyone understands how important it is for smaller parties to merge, or at least cooperate, in the single-member districts, negotiations will break down over the all-important “detail” of whether...

\[^{21}\] Steven R. Reed, “Seats and Votes: Testing Taagepera in Japan manuscript under review, available from the author on request.
my party should support or join yours, or your party should support or join mine. Some parties and candidates will prefer to lose rather than abandon their principles. In merger talks, parties will argue over what principles the people really support and the only way to resolve these arguments will be to run separately and see who gets more votes.

Similarly, many candidates who do not receive their party’s nomination will switch parties, run as independents, or start their own parties. Remembering that the initial conditions for this new electoral system include candidates who campaigned primarily as representatives of well-organized local interests, we should expect incumbents to loathe to abandon their districts. They will therefore want to run in the nearest single-member district. There will thus be many single-member districts with three or four serious candidates and the results of those elections will be capricious. The media will carry stories of candidates winning with only a third of the vote and those opposed to single-member districts will say, “I told you so.”

The results of the first election under the new system will be subjected to intense analysis by scholars, journalists, and, most importantly, politicians. The learning process will begin immediately. Some of the debates between parties will have been resolved so we can expect some party mergers that rationally should have taken place before the election to be implemented only after the election. What will the parties learn? How will the heiritsu-sei work out in the long run?

**Where is the Equilibrium?**

My second prediction is that the SMD districts will dominate the PR districts. Because SMD is a more responsive electoral system, the feedback from any gains or losses will be magnified. A party gaining two percentage points of the vote will get two more percentage points worth of seats in the PR districts, but more than that in the SMD districts, how much more depending on how close the party is to the center of the graph. The “carrots and sticks” emanating from the SMD districts will be stronger than those emanating from the PR districts. Both the parties and the media will come to focus on the single-member districts when explaining who won or lost the election. Similarly, SMD over-represents larger parties and under-represents smaller parties. Whereas PR produces no particular incentives for voters to vote strategically or for parties to merge, SMD produces clear and strong incentives for both. SMD is a “strong” electoral system whereas PR is a “feeble” one. The PR districts will serve to moderate the effects of SMD, but the SMD districts will drive the system.

That the equilibrium under SMD is two equally-sized parties is the well-known Duverger’s Law. However, this is a rational choice equilibrium. In the real world, other factors keep the system from reaching the ideal equilibrium point. The electoral system is, after all, only one factor among many. After a long debate, political scientists have reached a consensus on the best way to measure the “effective number of parties” (ENP), the

---

Laakso-Taagepera index:

\[
\frac{1}{\Sigma (V^2)}
\]

The proportion of the vote \( (V) \) is squared and summed, then divided into one. Although this index has many attractive mathematical properties, it does tend to give higher numbers than simple common sense would suggest. For example, two parties with 40 percent of the vote and one with 20 percent gives us an effective number of parties equal to 2.78. We will use ENP throughout this paper, but the reader should take care in interpreting it.

Using this index, the ENP in parliamentary regimes with SMD averages 2.8 when calculated in votes and 2.2 when calculated in seats. These figures reflect the facts first, that most SMD systems do have significant third parties and, second, that such third parties do better in winning votes than in winning seats. The difference between the effective number of electoral parties (ENEP) and the effective number of parliamentary parties (ENPP) is an excellent measure of the degree to which the electoral system produces incentives reducing the number of parties, and I will use it in the analysis below. Assuming that the PR districts will make the equilibrium point higher than in pure SMD systems, we can estimate that, in the long run, the heiritsu-sei will produce about 3.0 parties when measured in votes and about 2.5 when measured in seats. The most likely configuration is two larger parties competing in the SMD districts with one to three smaller parties relying primarily on the PR districts.

### Table 1 Three Scenarios

<table>
<thead>
<tr>
<th>Party</th>
<th>Percent of the Vote</th>
<th>Percent of Seats, SMD</th>
<th>Percent of Seats, PR</th>
<th>Final Percent of the Seats</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>35</td>
<td>66.3</td>
<td>36.1</td>
<td>54.2</td>
</tr>
<tr>
<td>B</td>
<td>25</td>
<td>23.1</td>
<td>25.2</td>
<td>23.8</td>
</tr>
<tr>
<td>C</td>
<td>15</td>
<td>4.6</td>
<td>14.6</td>
<td>8.6</td>
</tr>
<tr>
<td>D</td>
<td>15</td>
<td>4.6</td>
<td>14.6</td>
<td>8.6</td>
</tr>
<tr>
<td>E</td>
<td>10</td>
<td>1.3</td>
<td>9.5</td>
<td>4.8</td>
</tr>
</tbody>
</table>

ENEP=4.16 ENPP=2.72 Difference=1.44

A Two Party Scenario

<table>
<thead>
<tr>
<th>Party</th>
<th>Percent of the Vote</th>
<th>Percent of Seats, SMD</th>
<th>Percent of Seats, PR</th>
<th>Final Percent of the Seats</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>35</td>
<td>48.6</td>
<td>35.9</td>
<td>43.6</td>
</tr>
<tr>
<td>B</td>
<td>35</td>
<td>48.6</td>
<td>35.9</td>
<td>43.6</td>
</tr>
<tr>
<td>C</td>
<td>10</td>
<td>1.0</td>
<td>9.4</td>
<td>4.3</td>
</tr>
<tr>
<td>D</td>
<td>10</td>
<td>1.0</td>
<td>9.4</td>
<td>4.3</td>
</tr>
<tr>
<td>E</td>
<td>10</td>
<td>1.0</td>
<td>9.4</td>
<td>4.3</td>
</tr>
</tbody>
</table>

ENEP=3.63 ENPP=2.59 Difference=1.04

A Balanced Multiparty Scenario

<table>
<thead>
<tr>
<th>Party</th>
<th>Percent of the Vote</th>
<th>Percent of Seats, SMD</th>
<th>Percent of Seats, PR</th>
<th>Final Percent of the Seats</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>22</td>
<td>26.5</td>
<td>22.1</td>
<td>24.8</td>
</tr>
<tr>
<td>B</td>
<td>21</td>
<td>22.9</td>
<td>21.1</td>
<td>22.2</td>
</tr>
<tr>
<td>C</td>
<td>20</td>
<td>19.7</td>
<td>20.0</td>
<td>19.8</td>
</tr>
<tr>
<td>D</td>
<td>19</td>
<td>16.7</td>
<td>18.9</td>
<td>17.6</td>
</tr>
<tr>
<td>E</td>
<td>18</td>
<td>14.1</td>
<td>17.9</td>
<td>15.6</td>
</tr>
</tbody>
</table>

ENEP=4.97 ENPP=4.87 Difference=0.10

(Note: These calculations predict the single most likely distribution of seats from the posited distribution of votes. Factors such as the geographical concentration of votes for specific parties and the luck of the draw in close races will produce some error variation around each point prediction.)

### The Path to Equilibrium

Critics of the social sciences tend to laugh at and ignore any prediction which contains the phrase, “in the long run.” One advantage of the structural-learning approach is that it allows us to answer these critics, at least in part. How fast will the parties and candidates learn? I
predict that it will depend on the results of the first election. In other words, the slope of the learning curve will be path determined. Table 1 presents forecasts based on the Taagepera equation for three different scenarios. The first scenario is a case of one large party competing with several smaller parties, the “dominant party” scenario. In second, “two party” scenario, two larger parties compete with three smaller ones. The final scenario is the “balanced multiparty” result, with five nearly equal parties.

Several basic points are illustrated in this table. First, the well-known point that larger parties are over-represented in the SMD districts is clear in each scenario. Similarly, the PR results illustrate the meaning of a rho equal to 1.07, proportional results marred only slightly by an extra seat or two gained or lost. An important but less widely understood point is the degree to which a party’s representation depends on the distribution of the votes among the other parties. Thus, a party with 35 percent of the vote can expect to win an absolute majority if the rest of the votes are divided evenly among its opponents as in the predominant party scenario. However, the same 35 percent of the vote will not produce a majority if there is a second party with nearly as many votes, as in the two-party scenario.

For present purposes, the most important lesson of Table 1 is that the incentives for party mergers vary greatly depending on electoral results. If the first election produces a predominant party, and especially if that party has an absolute majority, the opposition parties will face strong pressures to merge, as demonstrated by the very large difference between the effective number of electoral and parliamentary parties (ENEP−ENPP=1.44). Any small party which merged with Party B could expect to win an absolute majority and be running the government after the next election, assuming that no other mergers occurred. After a lot of complex maneuvering, we should expect major mergers before the second election and equilibrium by the third.

If the first election produces a two-party result, the incentives for merger will be less than the predominant party case, but still very strong. We should expect the two big parties to compete for the cooperation of the smaller parties. The potential reward for a successful merger is an absolute governing majority, but smaller parties can also imagine themselves as potential coalition partners, perhaps even controlling the swing vote in the Diet. Any successful merger could set off a rush toward equilibrium, but a slower process over three or four elections is more likely. Each party will try to use elections to increase its bargaining power and, after each election, will reassess its independent potential. Voters will tend to move to one of the larger parties forcing the smaller to choose sides before they get too small to make a difference.

The balanced multiparty scenario produces almost no incentives for merger, reflected in a difference between ENEP and ENPP nearly equal to zero. While any two parties that did merge could expect to become a dominant party, merger negotiations would be extremely difficult and might produce defections if some groups within a party felt that its leadership had
Thinking about the Heiritsu-sei: A Structural-Learning Approach

given up too much. Again, one successful merger could set off an avalanche of other mergers, the most likely result is extended negotiations putting off mergers until feedback from elections produces stronger incentives. We should not expect the system to reach equilibrium before six or seven elections had been experienced.

Note first that rational expectations theory leads to completely different predictions. The value of being a large party is constant; it does not depend on electoral results. If we assume that parties understand the system and act rationally, the incentives to merge into larger parties should be strong even in the balanced multi-party scenario. Thus, the first election in Japan could become a test of rational expectations theory versus structural-learning theory. The recent merger of the Liberal and Social Democratic parties in Great Britain gives me confidence in structural-learning theory25).

Secondly, note that we can perform these calculations as soon as the actual electoral results are in and get a precise estimate of the strength of the incentives for party mergers. Unfortunately, we still have not developed ways of translating the ENEP-ENPP index into units of time. It is meaningful to say that the incentives in the two-party case are ten times stronger than in the balance multiparty case, and I do predict that learning will be faster if the ENEP-ENPP index is higher, but my estimates for exactly how much faster are little more than educated guesses. Being able to say sooner and later is a small, but important, improvement over predicting that the system will reach equilibrium “in the long run.”

A Local Optimum?

The calculations in Table 1 are based on a false assumption: that the parties will get the same percentage of the vote in both the SMD and PR districts. The assumption was made for purposes of calculation and illustration only; none of my conclusions or predictions depend on it. However, the possibility of voters might prove willing to vote for a smaller party in the PR election means that some parties might find it possible to survive depending primarily on the PR districts. A party which does not depend on the SMD election at all and would be immune to the incentives for merging into one of two larger parties26). One can thus imagine a local optimum in which the effective number of parties is significantly higher than the ideal rational choice optimum but is nonetheless stable.

We have just entered the realm of speculation. I have no idea how voters will choose to use their two votes. In fact, this is just the kind of subtle choice for which structural-learning theory provides little guidance because there is no clear feedback from different ways of using the PR vote. Thus, I will make no predictions. I will instead hypothesize that such a local optimum exists and that it is much more likely to occur if the first election produces a balanced multiparty system than if it


26) Again rational expectations theory would predict that the party would calculate the advantages to merging with a larger party and winning some SMD seats. Thus, if the party understands the system and acts rationally, it is not immune to the incentives to merge. The structural-learning approach predicts that since there is no particular feedback from a party’s decision to depend only on the PR districts, it will be immune.
Figure 2 The Local Optimum Hypothesis

produces either a predominant or two-party system. This hypothesis is illustrated in Figure 2.

The vertical axis is entropy. The steeper the slope of the line, the greater the incentive for parties to merge. The left-hand side of the graph represents a balanced multiparty result in the first election. After this kind of electoral result, there is little incentive for party mergers, so the slope leading towards lower levels of entropy is relatively flat. The right-hand side of the graph represents a predominant party result and the slope is much steeper. A three-dimensional graph could represent other initial conditions, but I will compare just these two. The predominant party result leads directly and quickly to equilibrium, estimated to be 2.5 effective parties. The balanced multiparty result leads to a higher local equilibrium, perhaps 3.5 parties. The bump between the two equilibria is meant to indicate that the higher equilibrium is stable even though not ideal. Thus, after the system reaches one equilibrium or the other, I expect it to bounce up out of equilibrium in response to exogenous shocks such as the creation of new parties, major new issues, or shifts in party support. Under most circumstances, the system will simply return to the same equilibrium in the following election. Sometimes, however, a shock to the system could be large enough to cause it to shift from one equilibrium to the other. For example, if a small party in the ideal 2.5 equilibrium were to find a particularly attractive leader or issue, it could grow large enough to push the system into the higher equilibrium. Similarly, if the system were stable at the higher equilibrium, a major issue which divided the electorate into two camps could unbalance the system and push it over the hump into the two-party-plus equilibrium.

The bump is only the simplest example of a local optimum. Even if there were no bump, friction in the form of institutionalization could cause the long slow approach to equilibrium which started with a balanced multi-party result to stop well short of the optimum.

Making Predictions in the Social Sciences

There are two primary reasons why the social sciences have made little progress toward producing the kind of reliable generalizations which would permit us to make better predictions: the difficulty of doing experiments and the fact that our
subjects are not inanimate objects but human beings who react to being studied. The structural-learning approach attempts to overcome these problems by focusing on the easiest problems to solve.

Structures such as electoral systems produce lots of high quality data at regular intervals. Still, we cannot perform experiments. We cannot ask countries to run elections over again, controlling for some specific variable of scientific interest. Most of the truly impressive accomplishment of the natural sciences are based on experimentation. In particular, the certitude of natural science generalizations derive from repeated experiments under controlled conditions. To take a simple example, chemistry could not have developed very far if they had to conduct their experiments with water taken from a nearby stream. Precise generalizations were developed in and work with certainty only in artificial environments. Natural sciences which share the inability to experiment with the social sciences also share the inability to make precise predictions. When will the next earthquake rock California and how big will it be? Will it rain tomorrow? Certainty is hard to find in the world outside the laboratory. Prediction in the social sciences may well be doomed to a level of precision on a par with weather forecasting. Nevertheless, weather forecasting is worth doing and the social sciences should also be able to make worthwhile predictions.

Social sciences deal with people and people are less cooperative than inanimate objects. Because they deal with people, the social sciences have been based on some view of human psychology. No matter how common sensical this approach may seem, it is also focusing on the most problematic aspect of social science, mind reading. The structural-learning approach focuses on cases in which the situation is more important than preferences thus minimizing the need for mind reading. We focus on problems which can be solved using the simplest, least controversial assumptions about human psychology. Progress in the social sciences will probably depend ultimately upon advances in psychology but, while we are waiting, it makes sense to concentrate on areas where that dependence is minimized.

One problem with making predictions in the social sciences is that the objects of one’s predictions may read and understand them and change their behavior accordingly. If Japanese politicians read this article and believe it, they would be able to calculate the advantages of party mergers. Since the advantages are substantial, the ability to calculate them would be a powerful argument in favor of mergers. Mergers would become more likely than predicted by the theory because the actors would understand the theory. Fortunately, social science has such a bad reputation for being able to predict the future that I am not worried about anyone in a position of power paying any attention whatsoever to my research. I am confident that anyone who tried to use this research to promote a merger would find his opponents are unimpressed by such abstract academic research.

With the exception of the remote possibility that one’s predictions might influence political behavior, I find making predictions a no-lose proposition. If you turn out to be right, people think you are smart.
If you are wrong, you learn something. The important thing to remember is that social science predictions should be held to no higher standards than weather forecasts.