Chapter 4

The Comparative Statics of Collective Action Problems

The time has come to put together all the theoretical elements developed in the previous chapters. In this chapter I will apply the method of stability sets to the games presented in Chapter 2 to see what new insight results.

This chapter will prove one conclusion: in general, in a collective action problem, the likelihood of cooperation is a continuous function of the game’s payoffs. The exception to this occurs when the model is a single-equilibrium game because then outcome is deterministic and does not respond to changes in the payoffs. In light of the preceding analysis, this may not come as a surprise, but without a rigorous mathematical proof, this statement cannot be used to develop applications such as the ones in this book’s second part. To prove this central claim, I will analyze the different models discussed in Chapter 2 with the aid of the method of stability sets.

4.1 Single-equilibrium Models

Since the goal of the method of stability sets is to calculate the relative likelihood of the different equilibria in a game, when
that game has only one equilibrium the method is largely redundant: this unique equilibrium will be the only outcome with a positive probability. Applying the method to these models will not tell us anything we do not already know. But it is worth spending some time on these cases given their prominence in the scholarship on collective action. Additionally, when using unfamiliar methods, it is a good idea to test them first in the familiar cases where we already know what the solution should be.

4.1.1 The Public Goods Model

Let’s start by reviewing the payoffs of the standard Olsonian model. This is the same model already studied in Section 2.4.1:

\[
\begin{align*}
    w_1 &= B - c, \\
    w_2 &= B, \\
    w_3 &= -c, \\
    w_4 &= 0.
\end{align*}
\]

With these payoffs we obtain \( W = \infty \). If we fix \( \lambda = 1 \) in Inequalities 3.7 and 3.8, we can retrieve the game’s Nash equilibrium. It is clear that the only value \( \gamma_{\mu} \) that can satisfy these conditions is \( \gamma_{\mu} = 0 \). (This makes Inequality 3.7 hold because it becomes \( 0 \geq 0 \).) This is Olson’s free-riding result: the only solution is a turnout of zero.

Now let’s turn to the computation of the stability set. In Section 3.2.1 I showed with the aid of a \( 2 \times 2 \) example that, when a game has dominant strategies, the entire strategy space is the stability set of the game’s unique equilibrium. The framework developed here comes to the same. To see why, consider first the case \( \lambda = 0 \). Just as when \( \lambda = 1 \), the only solution is \( \gamma_{\mu} = 0 \). In other words, regardless of what players initially believe about each other, they will always free-ride.

Under these conditions, checking for the entire tracing path is rather superfluous, but the reader can see that, regardless of the value of \( \lambda \), the only solution for both inequali-
ties is always the same: $\bar{\gamma}_\mu = 0$. This proves that, just as we had already established, in an Olsonian public goods problem, the only possible outcome, which occurs with probability 1, is universal defection.

4.1.1.1 The Public Goods Model with Selective Incentives

This is also a case we have already studied exhaustively (Section 2.4.2) but that can help illustrate the tracing procedure. The payoff structure of this model is:

\[
\begin{align*}
  w_1 &= B - c + s, \\
  w_2 &= B, \\
  w_3 &= s - c, \\
  w_4 &= 0.
\end{align*}
\]

If, as Olson claims it must, the selective incentive takes values $s > c$, then $W = -\infty$. This becomes the mirror image of the previous case. Now the only solution to Inequalities 3.7 and 3.8 is $\gamma_\mu = 1$, regardless of the value of $\lambda$ or the value of $\gamma_\eta$. This game is, just as the previous one, dominance-solvable. The only difference is that now the dominant strategy is to cooperate. As a result, the only equilibrium is for every player to cooperate ($\gamma_\mu = 1$) and the stability set of this equilibrium is the entire strategy space, so that this result happens with probability 1.

4.1.1.2 Generalizing Strategic Dominance

The analysis of these first two cases would have achieved the same results for values of $W$ different from $\infty$ and $-\infty$. In fact, the results would have been the same had $W$ been $> 1$ in the first case and $< 0$ in the second. This suggests that strategic dominance can occur in payoff structures more general than the one of the original Olsonian model.

In any model in which $W < 0$ cooperation will be a dominant strategy. Likewise, in any model in which $W > 1$ defection will be the dominant strategy. This latter case gives us
the implicit assumptions of the public goods model and, by contrast, allows us to see also the implicit assumptions of the multiple-equilibrium approach. If \( w_4 > w_3 \), that is, if participating in a failed attempt at collective action entails a cost, then \( W > 1 \) is true if and only if \( w_2 > w_1 \). In Section 2.6.5.2 (pg. 53) I stated that any model of collective action that relies on focal points, tipping, or any other concept tied to the existence of multiple equilibria is, wittingly or unwittingly, assuming that players who participate in a successful instance of collective action are compensated, at least ex ante. Here we see a more formal proof of that same statement. If there is no such compensation, that is, if \( w_2 > w_1 \), the collective action problem does not have multiple equilibria and invoking tipping mechanisms or focal points becomes plainly nonsensical.

While the method of stability sets does not add anything we did not already know to the study of single-equilibrium models, it does not do any harm either. The same standard conclusions about universal free-riding that result from the typical equilibrium analysis are also true here. This was to be expected: the method of stability sets is more general, not less, than equilibrium analysis.

4.2 A Basic Model with Multiple Equilibria

In Chapter 2 I presented several variants of the collective action problem that had multiple equilibria (e.g., the model with differential costs or the “stock-option” model). Here I will subsume most of them under one general structure and will analyze it thoroughly.

I will simplify the analysis by assuming that all players have identical payoffs. This is the assumption made in Sections 2.6.1 - 2.6.4. What these models have in common, from the point of view of their payoff structure, is that \( 0 < W < 1 \). As we have just seen, if this condition is violated, the game has only one equilibrium. It is now time to study the multiple-equilibria case.
The procedure I will follow here is the same as the one already shown for $2 \times 2$ games although, of course, the notation and the mathematical arguments involved may be a bit intimidating for some readers. So, let’s describe it in essence:

**Step 1: Compute the Equilibria of the Original Game.**
This step is to some extent redundant because we already computed these equilibria in Chapter 2. But the analysis will be tighter if we have all the results together and coming from the same place. Inequalities 3.7 and 3.8 give us the equilibria of any game along the tracing procedure, including the original one. All we have to do is to adapt these inequalities to the game at hand and let $\lambda = 1$. Once this is done, whatever value of $\gamma_\mu$ that satisfies both inequalities simultaneously is an equilibrium.

Not surprisingly, we will find the same equilibria: one equilibrium where nobody cooperates, another one where everyone does so and a third equilibrium with an intermediate level of turnout $W$ ($\gamma_\mu = 0$, $\gamma_\mu = 1$ and $\gamma_\mu$ such that $F(\gamma_\mu) = W$). It may be useful for the reader to verify that this equilibrium coincides with the mixed strategy equilibria computed in Sections 2.6.1 - 2.6.4 and that it is also the same as the unstable equilibrium of the tipping game in Section 2.6.5 if we assume identical payoffs. This is an instructive way of verifying that the analysis in this chapter truly generalizes the conventional equilibrium arguments. In light of what has already been discussed, it should come as no surprise that the first two equilibria will have sizeable stability sets, while the latter one will be a razor-edge equilibrium.

**Step 2: Compute the Equilibrium of Game $\Gamma^0$.** In this step we compute the equilibrium of the game when players base their decisions on their initial beliefs only. This is the equilibrium with which the tracing path starts. Just as in the examples already studied, the solution here will differ in two important ways from the solution obtained in the first step: there will be only one equilibrium and its strategies will depend on the initial belief conditions chosen.
Step 3: Compute the Tracing Path. This step generates the tracing path for any set of initial belief conditions. It will tell us all the equilibria of the game as $\lambda$ varies from 0 to 1. The single most important information we need from this step is which of the equilibria at $\lambda = 1$ is continuously connected to the equilibrium at $\lambda = 0$. This will give us, for each initial condition, the stability set to which it belongs.

Step 4: Characterize the Stability Sets. Once we know how to assign each possible vector of initial belief conditions to a stability set, all we need to do is to put together the information. In the cases studied below, the tracing procedure will result in a rule that tells us what kind of initial belief conditions are mapped onto what equilibrium. In fact, the final result is not entirely surprising.

4.2.1 The Main Result

This section presents the main result that culminates the tracing analysis of the preceding chapters. Let $W$ be defined, following pg. 127, as:

$$W = \frac{w_4 - w_3}{(w_4 - w_3) + (w_1 - w_2)}.$$  

If the mutual expectations players have are summarized by initial belief conditions with expected aggregate turnout $\gamma_\eta < W$, those expectations belong to the stability set of noncooperation ($\gamma_\mu = 0$). If, instead, the initial belief conditions are such that $\gamma_\eta > W$, they belong to the stability set of cooperation ($\gamma_\mu = 1$).

Intuitively, levels of expected turnout below $W$ are not enough to persuade players that it is worthwhile to cooperate. Just as successful collective action results in benefits for the participants, failed collective action results in a cost. If the expected turnout is too low, viz. below $W$, the risk of failure outweighs the prospects of success. Thus deterred, players opt out of cooperation and collective action does not take off. In contrast, if the expected turnout level is above $W$, players find
the risk worth taking and, as these expectations are validated, they reach the cooperative equilibrium.

This is the cornerstone of the comparative statics of collective action. With this result, if we want to compute the probability of cooperation, all we need is to specify our assumptions about the distribution of initial belief conditions.

The method of stability sets is just that, a method. It does not legislate what conclusions we must obtain from the study of a game with multiple equilibria but simply gives us a rigorous language in which to state our conclusions and sheds light on the connection between beliefs and payoffs in a game. The method of stability sets does not debunk any theory of collective action but clarifies its assumptions so that we can decide on its merits.

For example, if we want to cling to the notion that collective action problems are inherently unpredictable, that their outcome depends so much on human agency that no a priori statement makes sense, that knowing their structure tells us nothing about their result, we can simply disregard the information contained in the stability sets. Such extreme agnosticism implicitly assumes that any knowledge we may have about initial belief conditions is of no use in inferring anything about the prospects for coordination in a group. This is an entirely consistent stance but I do not see its rationale. Stability sets are mathematical objects, just the way equilibria are. In both cases it is up to the model’s user to interpret their meaning in ways relevant to understand human interactions. I do not see how we can regard equilibria as meaningful constructs for the study of social coordination while at the same time denying that status to their stability sets.

The above type of agnosticism has some further difficulties. In the analysis of a game with multiple equilibria, just as in the analysis of any system that may arrive at different outcomes, there comes a point when we need to spell out our conjectures about its behavior. Given that country A’s currency entered a free fall, should we expect the risk of regime collapse to go up or down? Given that country B granted more autonomy
to province B1, should we expect the risk of civil war to go up or down? Given that country C signed a free-trade agreement with D, should we expect union membership to go up or down? Given that country F is now producing good G at a lower price, should we expect the membership of the “Protect Domestic G Producers” lobby in country H to go up or down?

Without knowing the details, it is pointless to try to answer these questions. In all these instances the answer may well be “up,” “down” or “stay roughly the same.” But whatever answer we give, it must be preceded by a statement of how likely we believe these outcomes are. Those of us who believe that game theory offers a good way to represent some aspects of human interactions, will be inclined to answer these questions with the aid of a model and if this model turns out to have multiple equilibria, our final assessment will depend on the probability we assign to each of them. At that point we face a choice: either we assign probabilities without micro-foundations to support our judgment, or we try to come up with probabilities that in some way respond to the underlying game-theoretic logic that we adopted from the start. The first choice strikes me as inconsistent: if, when push comes to shove, we are willing to give up the search for micro-foundations, maybe we should have not used game theory in the first place. The method of stability sets formalizes the second choice.

Another option is to cling to the standard arguments of focal points. Suppose that, in analyzing a collective action problem, we are convinced that cooperation is focal. Then we simply postulate that the distribution of initial belief conditions is concentrated on $\gamma = 1$.

The analysis just developed is entirely compatible with such an argument. But its wisdom seems dubious. Unless we have overwhelming evidence, it is hard to justify the assumption that initial belief conditions are fixed at any single level. The advocate of focal points claims to know that under no circumstance could there be initial belief conditions in a stability set different from the one he is defending. In making that claim, he is invoking some added insight beyond game
theory because the analysis of a game does not tell us anything about focality. Those of us who take a more skeptical stance are entitled to ask where that added insight comes from.

It seems more reasonable to hedge, allowing that initial belief conditions may, in fact, be at different levels. This is what tipping games do. But we do not have to stop there and simply conclude that many equilibria are possible. If we have some knowledge about initial belief conditions, which can be very vague and does not have to be as absolute as that claimed by defenders of focal points, we can translate our knowledge about stability sets into probabilities of the different equilibria.

Suppose, for example, that in this particular case we want to remain absolutely agnostic about the exact location of the initial belief conditions. Then we can adopt the Laplacian stance and represent that agnosticism by assuming that $\gamma_\eta$ is uniformly distributed over $[0, 1]$. Now, the border between the two stability sets of this collective action problem is given by $F(\gamma_\eta) = W$. So, we can put together our results about the stability sets and our probabilistic assessment of the initial belief conditions to say that the probability of cooperation $(P(\gamma_\mu = 1))$ is such that:

$$P(\gamma_\mu = 1) = F^{-1}(W) = F^{-1}\left(\frac{w_4 - w_3}{(w_1 - w_3) - (w_2 - w_4)}\right).$$

For illustrative purposes, let’s assume that $F(x) = x$. That is, let’s assume that the probability of success of collective action given any level of turnout is equal to that level of turnout. Then, the probability of cooperation would be:

$$P(\gamma_\mu = 1) = \frac{w_4 - w_3}{(w_1 - w_3) - (w_2 - w_4)}.$$

Although the exact algebraic expression for the probability depends on our assessment of the initial belief conditions, the qualitative properties will not. The value $W$ is the border between the stability sets of the game’s equilibria regardless of the probability distribution we want to use to represent $\gamma_\eta$. We are free to assume any other distribution over those initial belief conditions. But, unless we take an extreme assumption
such as the one of the focal points model, changes in $W$ will change the resulting probability of cooperation.

At long last, we are in a position to enter the water fees debate with which this book opens. The method of stability sets allows us to pronounce General B’s analysis as essentially correct but lacking nuance. Here is a way to formalize the intuition he had in mind. Suppose that the citizens of country $X$ understand that, were they to engage in collective action, they could overthrow this odious regime. If there is some insurgent organization waiting on the wings, it faces the task of offering prospective rewards to the would-be participants. In the terminology adopted here, this organization, if it is to be taken seriously, must see to it that, at least ex ante the potential members believe that $w_1 > w_2$. In other words, they believe that, if the organization were to take control, it would arrange the state of affairs so that direct participants receive $w_1$ and those who remained on the fence receive a smaller benefit $w_2$. The regime, instead, controls the variables $w_3$ and $w_4$: it determines the status quo payoff for the citizens as long as the regime survives and the payoff that awaits those who challenge the regime but fail to bring it down ($w_3$). By his own reckoning, General B, who is an expert in $w_3$, feels that, as much as he can raise it with his ruthless secret police, there is only so much he can do. In formal terms, he worries about the effect that the Finance Minister’s proposal will have over $w_4$. A gratuitous increase in water fees, without any visible benefit for the citizenry will depress $w_4$ and, as a result, will reduce the value of $W$. We know that as $W$ decreases, this means that the stability set of collective quiescence decreases. In more pedestrian terms, the ones that General B would understand, a decrease in $W$ means that the population is more restive, more likely to coordinate and revolt.

Of course, he cannot be sure that the insurrection will happen. He simply expressed his gut feeling that the water fees plan would be the straw that would break the camel’s back. There have been several straws heaped on this particular camel in the past so it is tricky to say that this is the decisive one. He may be wrong and the only way he could prove
his argument is if he knew exactly the values of $W$ before and after the water fees increase and the level $\gamma_\eta$ representing the aggregate initial belief conditions of this collective action problem. If he could, say, prove that $\gamma_\eta = 0.6$ and that currently $W = 0.7$ but the water fees increase would bring it down to 0.5, he would be entitled to conclude that this scheme will spell the regime’s demise. But he does not know these values and it is dubious that he could ever know them.

From a conceptual point of view, however, it is irrelevant whether he is exactly right or wrong. What matters is that his entirely untrained insight into the citizens’ collective action problem is compatible with what we know about rational decision making, both individual and collective, as codified by game theory with the help of the method of stability sets. It would be wrong to rebuke him for being inconsistent with rational-choice theory because he is not. To claim, as the advisors in that story do, that water fees are irrelevant for the regime’s survival, one would have to be ready to defend the restrictive assumptions of the Olsonian framework, or to claim that the initial belief conditions among the populace happen to be just so that an increase in water fees will not increase the likelihood of insurgent coordination.

The lengthy theoretical and mathematical exercise that has brought us here would be pointless if all it accomplished was to bring support to General B’s uncouth views. But with a solid formalism we can go beyond that. For instance, suppose that we now abandon the stance of the security chief of an odious regime (a stance many of us might find uncomfortable) and instead adopt the viewpoint of social scientists who want to understand how the economic transformations ravaging Country X shape its political process.

At first glance, we would surmise that a five-fold increase in water fees has a larger impact over the likelihood of massive revolt than, say, a tax on jewelry. The preceding analysis of collective action confirms this because, on average, the effect of a tax on jewelry over the $w_4$ of this country’s citizens is smaller than the effect of water fees. But we do not need to stop there. We can use this theoretical framework to bring
together deeper and more extensive findings about Country X’s economic and social structure. The bare-bones formula of page 136 is simply a tool; it is up to us to put it to use in fruitful ways by developing plausible accounts of the forces shaping the value $W$. Changes in Country X’s export performance will affect the status quo and may be mapped into effects over $w_4$. Increases in the share of public-sector employment, unaccompanied by modern civil-service legislation, might increase the power of the government to threat dissenters, something that will impact $w_3$. (I present a more detailed example of this line of thought in Chapter 5.) This is not the place to describe in detail the possible uses of this chapter’s main result. Instead I just want to claim, much more modestly, that this result offers promise in our search for comparative statics results about politico-economic processes, given how much those processes are shaped by collective action, be it actual or potential.

Not only does the method of stability sets give us probabilities of the game’s outcomes; it also gives us probabilities that depend functionally on its payoff structure. In other words, it gives us comparative statics of the game. With the previous result in hand we can study how changes in the payoffs (the $w$’s) affect the probability of cooperation. I submit that this type of study cannot be conducted with the conventional methods. At the end of the day, focal points and tipping games are not tools to investigate the effect of structural changes upon outcomes of collective action.

In essence, this is the analysis I will carry out in the second part of this book. I will make explicit the collective action problem that agents in a polity face, be this in wage bargaining (Chapter 5) or in a clientelistic regime (Chapter 6). Then, I will show that the severity of those collective action problems depend on the politico-economic structure in which they are embedded. To prove this, I will appeal to the results I have obtained in this part of the book from using the method of stability sets. In other words, the method will help us generate predictions about the comparative statics of the collective action problems analyzed in these chapters.

In this sense the method of stability sets provides the game-theoretic foundations of a structural analysis of collec-
A Basic Model with Multiple Equilibria

Rational-choice approaches have shown that mutual beliefs and expectations shape the outcomes of collective action and that those beliefs can be, in turn, shaped by organizations, institutions and leaders. But although beliefs are fundamental, so are payoffs. The method of stability sets shows that there is no analytic reason to reject this point and that, on the contrary, game theory has the right tools to transform it into a source of testable hypotheses.

4.2.2 Computing Stability Sets in a Generalized Collective Action Game∗

Now I will offer the mathematical proof behind the results just discussed. To that end, let \( \lambda = 0 \) to compute the unique equilibrium of \( \Gamma^0 \) as a function of the initial belief conditions \( \eta \). In this case, Inequalities 3.7 and 3.8 become:

\[
\begin{align*}
F(\gamma_{\eta})\gamma_{\mu} &\geq \gamma_{\mu}W, \\
F(\gamma_{\eta})(1 - \gamma_{\mu}) &\leq (1 - \gamma_{\mu})W.
\end{align*}
\]

So, if \( F(\gamma_{\eta}) < W \), the only solution is \( \gamma_{\mu} = 0 \). Likewise, if \( F(\gamma_{\eta}) > W \), then the equilibrium is \( \gamma_{\mu} = 1 \). If \( F(\gamma_{\eta}) = W \), then any value \( \gamma_{\mu} \) is an equilibrium.

Now we need to know which equilibria will exist for each value of \( \lambda \). This will depend on the initial belief conditions.

Consider first the aggregate turnout level \( \gamma_{\mu} = 0 \). It can only be an equilibrium if:

\[
\begin{align*}
\lambda F(\gamma_{\eta}) + (1 - \lambda)F(0) &\leq W; \\
\lambda F(0) + (1 - \lambda)F(\gamma_{\eta}) &\leq W; \\
(1 - \lambda)F(\gamma_{\eta}) &\leq W; \\
F(\gamma_{\eta}) &\leq \frac{W}{1 - \lambda}.
\end{align*}
\]

If \( F(\gamma_{\eta}) \leq W \), this inequality holds for every value of \( \lambda \). This already means that for initial belief conditions that

∗Technical section.
satisfy this inequality there is a continuous tracing path to equilibrium \( \gamma_\mu = 0 \).

If, instead, \( F(\gamma_\eta) \geq W \), this inequality holds only for values such that:

\[
\lambda \geq \frac{F(\gamma_\eta) - W}{F(\gamma_\eta)}.
\]

If the level of turnout \( \gamma_\mu = 1 \) is to be an equilibrium of an auxiliary game, it must be that:

\[
\begin{align*}
\lambda F(\gamma_\mu) + (1 - \lambda) F(\gamma_\eta) & \geq W, \\
\lambda F(1) + (1 - \lambda) F(\gamma_\eta) & \geq W, \\
\lambda + (1 - \lambda) F(\gamma_\eta) & \geq W, \\
F(\gamma_\eta) & \geq \frac{W - \lambda}{1 - \lambda}.
\end{align*}
\]

In a mirror image of the previous case, if the initial belief conditions are such that \( F(\gamma_\eta) \geq W \), then the equilibrium \( \gamma_\mu = 1 \) exists for every value of \( \lambda \). Otherwise, it will exist only if:

\[
\lambda \geq \frac{W - F(\gamma_\eta)}{1 - F(\gamma_\eta)}.
\]

What about the equilibrium with randomized play \( 0 < \gamma_\mu < 1 \)? This equilibrium only exists if:

\[
\lambda F(\gamma_\mu) + (1 - \lambda) F(\gamma_\eta) = W.
\]

If \( F(\gamma_\eta) < W \), there is no possible solution to this equation unless \( \lambda \geq (W - F(\gamma_\eta))/(1 - F(\gamma_\eta)) \). In that interval, the solution will be a level of turnout \( \gamma_\mu \) that is a decreasing function of \( \lambda \). Analogously, if \( F(\gamma_\eta) > W \), any solution to this equation can only exist if \( \lambda \geq (F(\gamma_\eta) - W)/F(\gamma_\eta) \). Then the equilibrium \( \gamma_\mu \) will be an increasing function of \( \lambda \).

### 4.2.3 A Diagrammatic Presentation

The graphs in Figures 4.1 and 4.2 summarize all the results thus far:
Figure 4.1: Tracing Path for Initial Belief Conditions $\gamma_\eta > W$. The only continuous path corresponds to the equilibrium with cooperation.

Conceptually, this means that initial belief conditions with an expected turnout such that $F(\gamma_\eta) < W$ belong to the stability set of noncooperation ($\gamma_\mu = 0$) and that, conversely, those initial belief conditions for which $F(\gamma_\eta) > W$ belong to the stability set of cooperation ($\gamma_\mu = 1$). Just as we expected, the randomized equilibrium has a measure-zero stability set.

As already explained, once we have described the stability sets of a game’s equilibria, it is up to us to translate them into probabilities. The final result will depend on the assumptions we want to make about the distribution of initial belief conditions.

### 4.2.4 A “Simple” Tipping Game∗

In Chapter 2 I showed that, starting from Model 0, we could obtain the structure of a tipping game if we allow payoffs to differ across agents. There are very good reasons to do so. Most collective action problems differently affect people in different stations of life. We may want to analyze the prospects

---

*Technical section.
of collective action aimed at, say, income redistribution in a group that includes blue-collar and white-collar workers, or the prospects of collective action aimed at trade reform in a group that includes producers of tradeables and nontradeables, or of labor-intensive goods and capital-intensive goods, and so on.

Proving a general result for any conceivable tipping game is beyond the scope of this book. But for the purposes of a full-blown agenda of structural analysis of collective action, differences in payoffs are crucial, as the examples in the previous paragraph attest. So, I will offer here an illustration of how to handle this issue. To that end, I will study a “simple” tipping game in which there are only two possible payoff vectors.

Consider a version of Model 0 in which there are two groups of players: Group 1 with size $\rho$ and Group 2 with size $1 - \rho$. The payoffs for any member of Group 1 are $w_{11}, w_{21}, w_{31}, w_{41}$ and, likewise, the payoffs for any member of Group 2 are $w_{12}, w_{22}, w_{32}, w_{42}$. In keeping with the previous notation, we shall introduce values:

Figure 4.2: Tracing Path for Initial Belief Conditions $\gamma_0 < W$. The only continuous path corresponds to the equilibrium with no cooperation.
A Basic Model with Multiple Equilibria

\[ W_1 \equiv \frac{w_{41} - w_{31}}{(w_{11} - w_{31}) - (w_{21} - w_{41})}; \]
\[ W_2 \equiv \frac{w_{42} - w_{32}}{(w_{12} - w_{32}) - (w_{22} - w_{42})}. \]

Without loss of generality, I will assume that \( W_1 < W_2 \). Notice that this can come about if, for instance, Group 1 stands to gain more from success or to lose more from failure. All else being equal, members of Group 1 will be more prone to cooperate. From our previous discussion of tipping games we know that each player’s threshold value, that is, the expected aggregate turnout that will lead each to cooperate, depends on the payoffs faced. In this case, we have two such threshold values. Not coincidentally, the threshold values computed in Chapter 2 are exactly the same as the \( W \) values introduced here.

In this game with two possible threshold values, the distribution \( G \) is a discrete distribution with the following form:

\[
G(x) = \begin{cases} 
0 & \text{if } x < W_1, \\
\rho & \text{if } W_1 \leq x < W_2, \\
1 & \text{if } W_2 \leq x.
\end{cases}
\]

Depending on the values \( \rho, W_1 \) and \( W_2 \) we will have either two or three equilibria in pure strategies. This is somewhat anomalous for a tipping game: it is an artifact of having only two threshold values.

But the standard tipping diagrams focus only on pure strategies while, as we have shown, there is much to be learned about the game by looking at the entire set of equilibria. The preceding analysis retrieves not only the pure-strategy equilibria that the tipping diagram shows but also those that involve mixed strategies.

In this case, we need to satisfy Inequalities 3.7 and 3.8 for both sets of payoffs. Once we introduce the assumption \( F(x) = x \) we conclude that an equilibrium of this game will be described by turnout levels \( \gamma_1^1 \) and \( \gamma_2^2 \) for Group 1 and Group 2 respectively that satisfy the following conditions:
These inequalities show that there are five possible equilibria of the original game (when $\lambda = 1$):

- $\gamma^1_{\mu} = \gamma^2_{\mu} = 0$. This equilibrium exists for any values $W_1, W_2$ and $\rho$. 
- $\gamma^1_{\mu} = \gamma^2_{\mu} = 1$. This equilibrium also exists for any set of parameters.
- $\gamma^1_{\mu} = 1, \gamma^2_{\mu} = 0$. This equilibrium exists only if $W_1 < \rho < W_2$.
- $0 < \gamma^1_{\mu} < 1, \gamma^2_{\mu} = 0$. This equilibrium exists only if $\rho < W_2$.
- $\gamma^1_{\mu} = 1, 0 < \gamma^2_{\mu} < 1$. This equilibrium also requires $\rho < W_2$.

Let’s now calculate the starting points of the tracing paths by calculating the equilibrium of each auxiliary game $\Gamma^0$ for a given set of initial belief conditions. To that end, let’s fix $\lambda = 0$ and we obtain the following solutions to Inequalities 4.1 - 4.4.

- $\gamma^1_{\mu} = \gamma^2_{\mu} = 0$. This equilibrium exists only if $\rho \gamma^1_{\eta} + (1 - \rho) \gamma^2_{\eta} < W_1$.
- $\gamma^1_{\mu} = \gamma^2_{\mu} = 1$. This equilibrium exists only if $\rho \gamma^1_{\eta} + (1 - \rho) \gamma^2_{\eta} > W_2$. 

A Basic Model with Multiple Equilibria

- \( \gamma_1^\mu = 1, \gamma_2^\mu = 0 \). This equilibrium exists if \( W_1 < \rho \gamma_1^\mu + (1 - \rho) \gamma_2^\mu < W_2 \).

To map out the tracing paths it will be useful to find the smallest value of \( \lambda \) for which a given equilibrium exists. This gives us the following results:

- Equilibrium \( \gamma_1^\mu = \gamma_2^\mu = 0 \) exists for values of \( \lambda \geq (\rho \gamma_1^\mu + (1 - \rho) \gamma_2^\mu - W_1)/(\rho \gamma_1^\mu + (1 - \rho) \gamma_2^\mu) \).

- Equilibrium \( \gamma_1^\mu = \gamma_2^\mu = 1 \) exists for values of \( \lambda \geq (W_2 - \rho \gamma_1^\mu - (1 - \rho) \gamma_2^\mu)/(1 - \rho \gamma_1^\mu - (1 - \rho) \gamma_2^\mu) \).

- Equilibrium \( \gamma_1^\mu = 1, \gamma_2^\mu = 0 \) exists for values of \( W_1 < \rho \gamma_1^\mu + (1 - \rho) \gamma_2^\mu < W_2 \).

The same reasoning proves that the paths that lead to the equilibria where members of one of the two groups randomize branch out of the paths with pure strategies. More exactly, the equilibrium \( \gamma_1^\mu = 1, 0 < \gamma_2^\mu < 1 \) becomes feasible only when \( \gamma_1^\mu = 1, \gamma_2^\mu = 0 \) stops being feasible and the equilibrium \( 0 < \gamma_1^\mu < 1, \gamma_2^\mu = 0 \) becomes feasible only when the equilibrium \( \gamma_1^\mu = 1, \gamma_2^\mu = 1 \) is also feasible.

Summing up, the stability sets of this tipping game’s equilibria are the following:

<table>
<thead>
<tr>
<th>For Equilibrium</th>
<th>the stability set is formed by priors such that</th>
<th>for size parameter ( \rho ) such that</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \gamma_1^\mu = \gamma_2^\mu = 0 )</td>
<td>( \rho \gamma_1^\mu + (1 - \rho) \gamma_2^\mu &lt; W_1 )</td>
<td>( 0 &lt; \rho &lt; 1 )</td>
</tr>
<tr>
<td>( \gamma_1^\mu = \gamma_2^\mu = 1 )</td>
<td>( \rho \gamma_1^\mu + \gamma_2^\mu &gt; W_2 )</td>
<td>( 0 &lt; \rho &lt; 1 )</td>
</tr>
<tr>
<td>( \gamma_1^\mu = 1, \gamma_2^\mu = 0 )</td>
<td>( W_1 &lt; \rho \gamma_1^\mu + (1 - \rho) \gamma_2^\mu &lt; W_2 )</td>
<td>( W_1 &lt; \rho &lt; W_2 )</td>
</tr>
<tr>
<td>( \gamma_1^\mu = 1, 0 &lt; \gamma_2^\mu &lt; 1 )</td>
<td>( W_1 &lt; \rho \gamma_1^\mu + (1 - \rho) \gamma_2^\mu &lt; W_2 )</td>
<td>( \rho &lt; W_2 )</td>
</tr>
<tr>
<td>( 0 &lt; \gamma_1^\mu &lt; 1, \gamma_2^\mu = 0 )</td>
<td>( W_1 &lt; \rho \gamma_1^\mu + (1 - \rho) \gamma_2^\mu &lt; W_2 )</td>
<td>( \rho &lt; W_2 )</td>
</tr>
</tbody>
</table>
In two of these equilibria, some players randomize their choice but they nevertheless do have stability sets. This does not contradict the conclusions obtained in our discussion of $2 \times 2$ games: the fact that some players use pure strategies is enough to precipitate the existence of a stability set with positive measure.

Since those equilibria have stability sets with positive measure they will form part of any final forecast. To sharpen the result, then, we can compute the aggregate turnout that each of them generates. This can be done by solving the following equations that describe them when $\lambda = 1$:

\[
0 < \gamma_1 \mu < 1, \gamma_2 \mu = 0 \iff \gamma_1 \mu = \frac{W_1}{\rho}; \\
\gamma_1 \mu = 1, 0 < \gamma_2 \mu < 1 \iff \gamma_2 \mu = \frac{W_2 - \rho}{1 - \rho}
\]

Just as in the examples above, once we have computed the stability sets, we can complement them with our knowledge (or ignorance) about the players’ initial belief conditions to obtain a probabilistic assessment of the different equilibria which, in turn, will give us the probability of success as a function of the game’s structure. The method of stability sets subsumes the focal point arguments because, if we want, we can postulate a distribution of the initial belief conditions such that, in the end, all the probability is assigned to the equilibrium we believe is focal. Under the approach taken in this book, such procedure would be entirely legitimate because the method of stability sets does not legislate the distribution of initial belief conditions. But, as the method makes clear, if we say that one equilibrium is focal we are implicitly restricting the distribution of initial belief conditions. However legitimate from a purely theoretical standpoint, this reasoning is questionable. Unless a theorist has powerful arguments in favor of a distribution of initial belief conditions restricted to justify a focal point, he is liable to the charge of simply assuming what he is supposed to prove.

If we maintain our assumption of a uniform distribution over initial belief conditions, we can arrive at a precise prob-
A Basic Model with Multiple Equilibria

A probabilistic assessment of this collective action problem: the stability sets map into probabilities of each equilibrium. So, we have to consider several cases depending on the groups’ relative size, determined by the parameter $\rho$.

If $\rho < W_1$:

\[
\begin{align*}
P(\gamma_\mu^1 = 0, \gamma_\mu^2 = 0) &= W_1, \\
P(\gamma_\mu^1 = 1, \gamma_\mu^2 = (W_2 - \rho)/(1 - \rho)) &= W_2 - W_1, \\
P(\gamma_\mu^1 = 1, \gamma_\mu^2 = 1) &= 1 - W_2.
\end{align*}
\]

This translates into a probability of success:

\[
P(S) = \frac{W_2 - \rho}{1 - \rho} (W_2 - W_1) + (1 - W_2).
\]

If $W_1 < \rho < W_2$:

\[
\begin{align*}
P(\gamma_\mu^1 = 0, \gamma_\mu^2 = 0) &= W_1, \\
P(\gamma_\mu^1 = 1, \gamma_\mu^2 = 0) &= W_2 - W_1, \\
P(\gamma_\mu^1 = 1, \gamma_\mu^2 = 1) &= 1 - W_2.
\end{align*}
\]

Then, the probability of success is:

\[
P(S) = \rho(W_2 - W_1) + (1 - W_2).
\]

If $W_2 < \rho$:

\[
\begin{align*}
P(\gamma_\mu^1 = 0, \gamma_\mu^2 = 0) &= W_1, \\
P(\gamma_\mu^1 = W_1/\rho, \gamma_\mu^2 = 0) &= W_2 - W_3, \\
P(\gamma_\mu^1 = 1, \gamma_\mu^2 = 1) &= 1 - W_2.
\end{align*}
\]

In this case, the probability of success is:

\[
P(S) = \frac{W_1}{\rho} (W_2 - W_1) + (1 - W_2).
\]

Compared to the standard way of dealing with tipping games, this approach is much more laborious, especially when
we consider that this was the “simple” case with only two types of payoffs. Given its immature state, the method of stability sets still needs to generate more powerful algorithms for more general cases. But I believe that the extra effort is well worth it. Had we settled for the conventional analysis of tipping, all we could have said about this collective action problem was that several outcomes were possible, depending on the location of the initial belief conditions. We could have even refined this idea and arrived at conclusions about how changes in those beliefs would map into changes in the outcome. But we could have not made any claim about the effect of the game’s structure, viz. its payoffs and the size of the groups, upon such outcome.

With the method of stability sets, in contrast, we have expressions for the probability of cooperation that depend in a systematic way on the structural parameters. We can formulate deductive hypotheses about the game’s comparative statics without eroding the theoretical foundations on which our analysis is based.

4.3 Toward an Analysis of Stability Sets in Repeated Games

In Chapter 2 (Section 2.8) we saw that infinitely repeated PDs do not differ much from the general structure we have been analyzing from the start. In fact, they have many more equilibria than the other games in this book and, as such, their study offers some possibilities for the method of stability sets. In this section I will illustrate this point by continuing the study of the example of a repeated game introduced above.

This same trading game plays a crucial role in the vast literature on trust so it is instructive to see, however tentatively, what the method of stability sets can add to this research.\(^1\) Much of the research on trust games has focused on the endogenous emergence of trust. Considering a game such as the

\(^1\)For a recent treatment of the problem similar in spirit to the one adopted here, see Yamagishi et al. (2005).
one studied here, this literature asks whether repetition itself will lead players to be more willing to trade, in other words, if cooperation leads to trust or if it is the other way around.

Cooperation and trust, one could argue, always depend on the context in which they occur. Just because two people trust each other enough to lend and borrow a pen while waiting in a line does not mean that they will also serve as each other’s co-signer for a home loan. If we want to investigate the effects of trust as a facilitator of interactions in a society, one of the pillars of the research on social capital (in the tradition of Putnam, Leonardi and Nanetti (1993)), we cannot remain oblivious to the external circumstances, to the stakes involved. If trust is a scarce resource, like any material resource, we want to use it wisely and not require or expect people to trust each other in situations where their social capital is not up to the task.

Let’s see what the method of stability sets can contribute to this topic. Not much can be expected at this stage because I will not present a complete analysis of the game but will instead focus on one of its parameters: the discount rate. For purposes of illustration this is at once easy and fruitful to study. One standard interpretation of the discount rate links it to a structural aspect of a repeated interaction: the probability with which it continues over time. A low discount rate in this situation represents the trade between two agents that know that there is a high likelihood of the game being cut short for some exogenous reason. The higher the discount rate, the more the agents can rely on there being another round of interaction.

Intuitively, we would expect that cooperation is more likely in a stable environment, where the players are confident that the opportunities for mutual benefit that cooperation generates will not vanish on short notice, than in a volatile context where today’s partner may tomorrow be gone for good. Inhabitants of small villages often complain of how modernization and trade erode trust in their communities for precisely that reason: with new opportunities available, agents that were bound before for the long haul have now to take into account
that their would-be partner may tomorrow move to the big city. A systematic study of trust and its social effects should be able to articulate this intuition and turn it into a solid basis for comparative analysis. With the method of stability sets we can do just that.

For convenience, I will reproduce here the game that captures the decision of both players to enter in a contract that is self-enforced through iteration. (This is the same payoff matrix as on page 78.) In the game in Figure 4.3, we can compute the mixed strategy equilibrium with the standard method. Let \( \alpha_i \) denote the probability that Player \( i \) will choose \( C \). Then, the expected payoffs become:

\[
\begin{array}{cc}
C & N \\
1 & \frac{1}{1-\delta}, \frac{1}{1-\delta} \\
N & 2, -1 \\
\end{array}
\]

Figure 4.3: Will They Trade?

\[
v_1(\alpha_1, \alpha_2) = \alpha_1 \left( \frac{\alpha_2}{1-\delta} - (1 - \alpha_2) \right) + (1 - \alpha_1)2\alpha_2 \\
= \alpha_1 \left[ \alpha_2 \left( \frac{\delta}{1-\delta} \right) - 1 \right] + 2\alpha_2;
\]

\[
v_2(\alpha_1, \alpha_2) = \alpha_2 \left[ \alpha_1 \left( \frac{\delta}{1-\delta} \right) - 1 \right] + 2\alpha_1.
\]

These equations give us the following best-response correspondences:

\[
\alpha_1^*(\alpha_2) = \begin{cases} 
1 & \text{if } \alpha_2 > \frac{1-\delta}{\delta}, \\
[0, 1] & \text{if } \alpha_2 = \frac{1-\delta}{\delta}, \\
0 & \text{if } \alpha_2 < \frac{1-\delta}{\delta},
\end{cases}
\]

with an analogous best-response correspondence for \( \alpha_2^*(\alpha_1) \). If, just as before, we plot these best-response correspondences, we will see graphically the stability sets of the equilibria (Figure 4.4).

The stability sets of this game articulate formally the intuition discussed above: the higher the discount rate, that is, the
greater the likelihood of the interaction continuing into the future, the larger the stability set of the equilibrium where both parties enter and honor the contract, and so the more likely it is that the players will take the risks involved in the contract and attain the mutual benefits from trade. By the same token, the lower the discount rate, the more likely it is that the players will simply abstain from this self-enforcing contract, lest they misplace their trust.

The preceding analysis shows that the method of stability sets, far from being a competitor, could become a useful complement of the folk theorems often invoked in the study of repeated interactions. By their very nature, the folk theorems of repeated games leave many crucial questions open. All they tell us is that mutually beneficial exchange is possible in a repeated PD, but they cannot guarantee that such exchange will occur nor can they tell the exact terms under which it will. Several strands of literature have worked on this topic,
notably the theory of bargaining solutions, and each of them has made worthy contributions to it.

As regards the overarching problem of trust and cooperation in mutually beneficial exchanges, the method of stability sets could conceivably clarify the role external circumstances play in facilitating these types of interactions. After all, as social scientists we are interested in how different societies, facing different conditions, arrive at different answers to similar predicaments. The method of stability sets can help us isolate the role of material conditions so that we can understand better how all the other ingredients (e.g., social beliefs, traditions, institutions) combine.

Much work remains to be done. All I have shown here is how to calculate the stability sets of two narrowly defined equilibria, one with symmetric exchange and one with no exchange whatsoever. But we know from the folk theorems that a repeated game has a continuum of equilibria: the result presented in this section is no substitute for a general theorem and can only count as an indication, a clue that the method’s reach is much more general than what its current version suggests.

### 4.4 Conclusions

This chapter culminates the theoretical analysis of Chapters 2 and 3 and sets the stage for the politico-economic models of Chapters 6 and 5. It is, then, appropriate that we take stock of what has been accomplished thus far.

We began with a general model of collective action that unifies the Olsonian and the Schellingean paradigms. According to that model’s fundamental result, the difference between these approaches turns on their assumptions regarding the cost of successful collective action: if we believe that players can recoup the cost of participating whenever their collective endeavor succeeds, we are implicitly using a model with multiple equilibria. If we deny this, then our model is, of necessity, a single-equilibrium model.
After showing that these two main paradigms are special cases of a general model, we studied a method, the method of stability sets, that can be applied to the general framework to calculate the relative probabilities of all the equilibria in a game. That method is not incompatible with standard equilibrium analysis and in fact generalizes it. It is also closely related to other approaches in game theory, such as evolutionary games and quantal-response equilibria.

Although the formal calculations of stability sets are not well-known, the underlying intuition is not entirely unfamiliar. Tipping games offer a good but informal template. When studying a tipping game, we begin by postulating different possible levels of expected turnout as an input for the model. Unlike standard analysis, tipping games do consider expectations out of equilibrium and study how the game itself sets in motion an adjustment process that pushes them toward an equilibrium. In a typical tipping game there are multiple equilibria and so different levels of expected turnout may be mapped into different equilibria.

The method of stability sets formalizes this reasoning and goes beyond it. Rather than simply determining that some initial beliefs lead to coordination and others to defection, the method transforms this insight into a basis for probabilities of coordination that depend on the game’s payoffs. In short, the method gives us the key element for comparative statics of collective action.

The third step, the one taken in this chapter, was to apply the method of stability sets to collective action problems, something that results in crisp, analytical expressions for the probabilities of cooperation and defection. In a collective action problem, if we assume that all players face the same payoffs, the size of the stability set of the noncooperative equilibrium is $W = (w_4 - w_3)/(w_1 - w_3 - (w_2 - w_4))$. This expression covers both the single-equilibrium and the multiple-equilibria versions and tells us that in games with strict dominance, players will either cooperate with probability 1 (with selective incentives $W > 1$), or will defect with probability 1 (without selective incentives $W < 0$) and that in games without dom-
inant strategies, that is, Schellingian games where \( w_1 > w_2 \), the probability of collective action varies as the payoffs of the game vary.

This line of reasoning can be extended to other games and I have illustrated how to do this in two simple cases: an iterated PD and a collective action game with nonidentical payoffs. In the first case, if we fix the terms of trade, by virtue of, say, the Nash bargaining solution, the method of stability sets tells us how changes in the external environment (e.g., the probability of iteration) affect the likelihood of trade. In the second case, the method arrives at an expression more complex than the one above, where now the probability of the cooperative equilibrium depends not only on the payoffs but also on the relative size of the groups.

From a technical point of view, this is all there is to these three chapters. A set of long-winded mathematical steps, covering pages and pages of definitions and graphs, boils down to a one-line equation. But these technical details mask important issues about the research agenda I want to launch here and how it relates to other lines of inquiry.

The method of stability sets offers us a way to deal with the lack of predictive power of game theory when it comes to games with multiple equilibria. Instead of leaving the game’s outcome entirely indeterminate, and instead of imposing on it a set of arbitrary restrictions, the method obtains the relative likelihood of different outcomes, a relative likelihood entirely derived from the game’s microfoundations.

I believe this technical solution has far-reaching implications for the way we study collective action problems. In its early stages, rational-choice theory’s involvement with collective action was perceived by friends and foes alike as a calculated attack on what, for lack of a better word, I will call “structuralism,” that is, the belief that instances of collective action, or the lack thereof, are responses to external, objective circumstances.

Game theory, the argument would go, had taught us that the real triggers of collective action are internal to the agents themselves, be it their organization responsible for offering
selective incentives or the mutual beliefs they cultivate about each other and that determine their ability to coordinate in the pursuit of a common goal. With these variables doing all the explanatory work, it would seem that there is no room in a coherent theory of collective action for the kind of structural, especially socioeconomic, variables that studies of an earlier generation had privileged.

The method of stability sets throws this received wisdom into question. Once we generalize the standard methods and look not only at a game’s equilibria but at their stability conditions, we see that exogenous parameters, the payoffs, do play a role in shaping the game’s outcome. They certainly do not determine the result, but they make some results more likely than others.

History is replete with instances where agents have been able to overcome the most formidable barriers imposed by their external circumstances. Countless tales of treason and courage are based on this. After all, structural conditions do not exert any deterministic causal impact over collective action: they are just enablers or impediments that cannot substitute for human agency. But as social scientists, we care about enablers and impediments. Quite often those are the ones we can affect. We may never find a way to prevent a decent democratic regime from collapsing under massive protests. After all, focal points of coordination may emerge in the most unexpected places. But if it turns out that helping such regime avoid the worst effects of a drought forestalls such instability, social scientists should take notice of this whenever their analysis and advice is required.

Institutions, ideas and often sheer individual resolution drive collective action but always in a context given by objective circumstances. We cannot claim to understand a particular instance of collective action if we ignore this context.

This suggests that research about collective action should try to integrate the study of external, structural factors with the study of institutions, organizations and mutual expectations, rather than belittling the impact of either of these. The method of stability sets offers a way to do this. In a game with
multiple equilibria, the likelihood of the different outcomes depends on two main types of factors: the payoffs and the initial belief conditions. The payoffs are exogenous parameters that constitute the context where collective action occurs, whereas the belief conditions are shaped by the shared ideas and interactions of the agents concerned.

If we knew exactly what the belief conditions in a group are, we would not need any extra inputs, we could make a deterministic and accurate prognosis in any collective action problem. But as social scientists, most of the time we cannot know exactly what the belief conditions are. At best, we can surmise what they could be. In the technical framework developed above, this is what the distribution of initial belief conditions does: it captures the analyst’s information, or ignorance, about the expectations agents have about each other.

Stability sets capture the other side of the problem: payoffs. The payoffs inform us which initial belief conditions are conducive to coordination and which are not. Combining the stability sets and the distribution of beliefs brings together the structural context where agents operate and the expectations they have when they act. This avoids two types of one-sidedness: that of attributing all the explanatory role to structural conditions and that of ignoring them altogether.

Structural accounts of collective action have been criticized within the tradition of rational-choice theory so often that there is hardly any need to rehearse the criticisms in a book that relies so much on it. Perhaps it is more salutary to spend some time discussing the drawbacks of the other extreme and what the method of stability sets can do to overcome them.

If we focus entirely on belief conditions, on the work of the organizations, institutions and ideas that shape them, we forgo the possibility of analyzing them from a systematic comparative perspective. Let’s call this the “exceptional-people bias.” Collective action requires mobilizing strategies, but the strategy that succeeds famously in one context may fail in another. It is tempting to conclude that this is a matter of talent: some people are simply more gifted than others when it comes to mobilizing large groups. It would be obtuse, however, to deny
that part of this erratic pattern of success and failure has to do with exogenous circumstances, some of which affect the payoffs of collective action.

Part of the job of a good political entrepreneur consists of finding the right time and place, the right combination of external factors, in a word, the right context, to operate. Many activists that would otherwise be considered masterful mobilizers have seen their enterprise blow up in their face, often costing them their life, for having been cavalier about the context, for choosing as a base a town where their opponents had co-opted vast parts of the population, for beginning their mobilizing efforts when other alternatives were already gathering momentum, for overestimating the impact of the upcoming economic crisis, and so on. Context is not everything, but politicians ignore it at their peril. In fact, most of the time, they do not. Any organization serious about collective action spends time and effort getting a good reading of its context. The social scientists that study them must do the same.

For those of us who use the language of game theory to analyze social phenomena, the method of stability sets offers a tool with which to keep in mind the objective conditions that surround collective action without reifying them into an iron law of determinacy, recognizing the scope of human agency. Changes in payoffs change likelihood; this is a piece of knowledge that is at once useful and modest.

This lack of a systematic basis for comparative analysis biases the results by directing the researcher only to look at success stories. For every social movement that grows into an overwhelming political force, there are many that never get off the ground, often with similar goals but in different times or places. Once we know that a collective endeavor is successful, it is easy to marvel at the political genius of its participants forgetting that equally talented agents failed in similar, but ultimately different, circumstances.

But if this first drawback vitiates results of ongoing research, there is a second drawback that simply makes other inquiries impossible: the current approaches are not adequate to study counterfactuals, the instances of collective action that
do not even get started but that have a powerful effect on the course followed by a political process. In parallel to the previous bias, let’s call this the “exceptional-times bias.” A complete theory of collective action cannot be a theory tailored only to study periods of unrest. Coordination is a powerful social force even when it operates as an undercurrent. Conflicting goals among social groups are ubiquitous, but they rarely lead to open confrontation. Most of the time, all sides involved in such conflicts see the status quo as a better alternative than any other. If, as social scientists, we want to understand how this happens, we must understand what are those other alternatives, many of which involve collective action. We need to have a theoretically informed account of instances of collective action that we will never be able to observe because they simply lurk behind as threats that keep the current situation in place.

If our conceptual apparatus is entirely concentrated on mobilization strategies, we are at a loss here. Potential, counterfactual collective action is in the minds of all rulers, through all their waking hours, whether they rule the most open democracies or the harshest tyrannies. In deciding on an economic package, or a piece of legislation, or a constitutional change, all of them, or at least the really able among them, are mindful of the possible surges of collective action that may destroy their plans. In their current state, rational-choice theories cannot even start addressing this question because they rely too heavily on the role of organizations. In instances like this, however, the organization does not exist: it is simply a prospect. But even as a prospect, it is a force to be reckoned with.

The method of stability sets can study such counterfactuals. In fact, in the second part of this book I have deliberately stayed away from the study of ostensible instances of collective action, such as revolutions or mass protests. In Chapters 5 and 6 I illustrate the method by using it where its advantages are clearer: in the study of situations where collective action, even if it does not occur, contributes to shaping the polity’s outcomes. For all their differences, the reader of these chapters will be able to discern a pattern. First, I will spell
out a politico-economic model where some agents, clients of a political machine in one case, workers with no capital in the other, could potentially benefit from collective action (voting against the patron or creating a strong union with full bargaining powers). The likelihood of them doing so depends on the problem’s structure, the type of economy they inhabit, just as the method of stability sets would have it. Collective action is neither a necessity, dictated by the shrewdness of some entrepreneur, nor an impossibility always trammeled by free-riding: it is a counterfactual. Although this counterfactual may never materialize, it exerts a potent effect because it is the benchmark against which agents evaluate their current status quo and decide on it. By calculating the likelihoods of the different counterfactual scenarios, we can understand better the sources of stability (or disruption) of the status quo.

The preceding statement may seem to exaggerate the situation because most of the times organizations exist, even if not yet mature and in full swing. That is correct, but for the same reason, our knowledge of the belief conditions of collective action is so hazy that we should not put too much stock in it. It is wiser to acknowledge that we can only surmise what those conditions are and combine this uncertainty with the knowledge obtained from the stability sets. In the applications I will supply, I arrive at the likelihood of collective action, leaving room for improvement if we incorporate our knowledge about belief conditions. The fact remains, though, that this likelihood allows us to study counterfactual instances of collective action. More knowledge about beliefs would not make counterfactuals useless but would instead allow us to understand them better.

Rational-choice theorists often consider contemporary economics the most successful social science, an example to be imitated by the rest. Whatever truth one may find in this perception of economic science, rational-choice approaches to collective action have thus far failed to live up to that standard in one critical aspect: comparative statics. Economics as a discipline would not be where it is now without comparative statics. The exercises of formalization typical in economics
and that so baffle observers foreign to the discipline are, ultimately, procedures to enable comparative statics involving long chains of causal reasoning.

An example might help. The law of supply and demand can be expressed easily without any formalization. If we want to find out the effect of a drought on the price of wheat in a country, we do not need formal models of any kind: this is a chain of reasoning that involves only one link. But if we want to find out the effect of said drought on that country’s interest rate, more links are involved. We need to know whether it is a net importer or net exporter of wheat, whether it has a fixed or floating exchange rate, whether the central bank targets inflation or growth or some hybrid of both, and so on. Any estimate of the final impact will depend on our estimates of each of the stages of reasoning involved. Economic models are the tools to negotiate each of these steps, keeping all the pieces in sight at the same time. Much of economic analysis is about this type of comparative statics, about finding the effect of exogenous shocks over endogenous variables that are not immediately or transparently related to them. If it did not contribute to comparative statics, most of economic formalization would be an exercise in futility.

Ultimately, the methodological purpose of the analysis of stability sets is to put the rational-choice approach to collective action in the service of comparative statics. Coordination is the key to most political processes and understanding how structural conditions shape coordination is the key to a rigorous treatment of political economy. This type of political economy cannot dispense with a comparative statics of collective action because its polar star is the study of how exogenous changes in the economic structure affect the realm of political alternatives that the actors in a polity face.

The chapters that follow illustrate the research agenda I have in mind. Simplified as they are, these models help us transform our intuitions about the interactions between economics and politics into testable hypotheses, derived from the analysis of rational decision making. In that sense, they show that the method of stability sets does not see the macrostruc-
tural dimensions of collective action as in conflict with their microlevel logic. It combines both to arrive at a picture of social interactions that is both textured and informative.

A theory of collective action compatible with rational-choice theory, attuned to the context where agents act, general enough to encompass successes and failures, to encompass times of calm and times of agitation, able to illuminate broader questions of political economy: such is the theory we need to develop. Such is the theory the method of stability sets is meant to facilitate.