Preface

The following pages constitute a book with an unusual structure, a structure that may be hard to negotiate without an initial summary. Part of the difficulty is due to the book's unusual genesis: although this is ostensibly a volume on the theory of collective action, my original purpose was not to write on this topic. In fact, I did not set out to write a book at all. My goal was rather to write a short paper on political clientelism in developing countries, a paper that ultimately became Chapter 5 in this book.[1] As I progressed in this project, I began to realize that it was built on shaky ground and that I could only salvage its initial insights, which I still found valuable, by going back to its theoretical foundations. This book is my response to this original impasse and much about it can be clarified by retracing the steps I had to follow.

I began working on the problem of clientelism hoping to develop a model that would extend to the study of patronage regimes the rigor of standard electoral models. Since my preparation consisted solely of some familiarity with modeling and some personal acquaintance with clientelism, I was fortunate to collaborate in the early stages with Susan Stokes, a scholar with ample research experience with this problem. From the outset I was convinced that whatever model I developed should capture two basic properties of clientelism: the fact that, at least as a first approximation, clientelism tends to be weakened by economic development and the fact that, judging from the narrative of clients in a political machine, politi-

---

[1] Some of its main ideas also form part of another joint piece: Medina and Stokes (2007).

cians in a clientelistic regime wield with respect to their voters
a power unmatched by anything in a fully mature democracy.
While patrons can lavish their voters with all kinds of favors,
they can also punish disloyal behavior with the loss of jobs,
goods or services.

Since the literature in industrial organization is full of
game-theoretic models of such power asymmetries, at first
I thought I would adapt any of the models on offer to the
specifics of the problem at hand. But soon I realized that
there was a significant difference between the threats issued
by a patron and the threats with which, say, a firm can disci-
pline its workers, or an investor can discipline a manager: in
an electoral context, the power of the patron depends crucially
on having the voters confirm his position via the ballot box.
Bosses exist independently of the wishes of their workers just
as investors exist independently of the wishes of managers.
But patrons of a political machine owe their position as such
patrons to the very people they try to control: the voters.

This meant that, instead of the usual models of industrial
organization, I would need to rely on the theory of collective
action. After all, voters in this situation are trapped in a
coordination problem: while they can get rid of an undesirable
patron if enough of them vote against him, dire consequences
await those who try if they end up going it alone.

There were, then, two paths open at this juncture. I could
conceptualize this collective action problem as an Olsonian
Prisoners' Dilemma or I could, instead, adopt the framework
of tipping games proposed by Thomas Schelling, where coor-
dination problems have multiple equilibria. I quickly ruled
out the first option for reasons that will become apparent in
Chapter 2. The Olsonian model is very fragile and even tiny
changes in its specification are enough to destroy its main re-
sult, viz. universal free-riding. But, for my purposes, models
with multiple equilibria were not much better because it is
not clear how they respond to exogenous shocks. Within that
framework I would not be able to study the connection be-
tween exogenous variables, especially economic ones, and the
survival or demise of clientelism; I would have had to give up
my original goal.

So I probed a different approach. Once we go beyond the mere reporting of equilibria in a game and inquire about their robustness, or lack thereof, we realize that the multiple solutions of a game are not entirely capricious and instead, their relative likelihood depends in intuitive and illuminating ways on the exogenous parameters that govern the payoffs. This means that, going back to the problem of clientelism, in a precise and formal sense, the probability that the voters coordinate against the patron depends on economic variables such as levels of development. Such approach would allow us to formulate rigorously the connections between economic transformations and the putative demise of clientelistic machines. Having produced a fully working "toy example" of an economy with two voters and one patron, I was soon convinced that, even at the price of departing from the canon, it was worth putting together these ideas in a more systematic framework.

The method of stability sets is just such framework. It may not be the only possible one, or even the best one. It has its limitations, which will be discussed later on. But it has several advantages that convinced me to put it into print. First, it is not entirely original. Although we all place a premium on scientific innovation, there is something to be said for gradualism and cumulative progress as opposed to sweeping changes in outlook. The method of stability sets builds on ideas already developed in game theory, especially by John Harsanyi and Reinhardt Selten in their joint work *A General Theory of Equilibrium Selection*. In developing the method of stability sets I have borrowed liberally from their classic work but I disagree with its ultimate goal. The method of stability sets is, deliberately, *not* a method of equilibrium selection. Another advantage is its comparative simplicity, which makes it a good tool for obtaining results when we do not have much information or when we are simply in the initial stages of model building. Although the method of stability sets requires mathematical tools that may not form part of the average political scientist's toolkit, it is relatively simple. Nowhere is this more apparent than in the study of large-scale coordination games.

As I will show in Chapter 4, with the method of stability sets we can compute the relative likelihood of equilibria in games involving large numbers of players. Using this result, it is possible to extend the intuition of the two-voter clientelistic election to a true electoral model involving an arbitrary number of voters and two candidates. In fact, thanks to the method of stability sets it is straightforward that some types of economic growth can alleviate the coordination problem of voters in a clientelistic polity and, thus, erode the advantages enjoyed by their patron. The intuition is simple: changes in the economic environment that make the voters less dependent on the patron's resources reduce the weight of his threats and, therefore, make it more likely that the voters will coordinate against him if a better alternative comes around. But it should be noted that this intuition cannot be represented formally with the tools of standard game theory because, with their exclusive focus on the equilibria themselves rather than on their stability conditions, these tools cannot establish a link between the predictions of the game and its exogenous parameters.

The results pertaining to the method of stability sets and its extension to large coordination games, presented in Chapters 3 and 4, form the backbone of the present book. Chapter 3 introduces the method and Chapter 4 shows the results of applying it to the standard models of collective action already discussed in Chapter 2. These parts are the most mathematically demanding of all the book and many readers will prefer not to spend much time on them. Aware of this, I have written these chapters trying to keep separate the purely formal foundations, of interest only for readers familiar with game theory, and the substantive results that will be necessary later on in the book.

For readers uninterested in following all the technical arguments, down to the tiniest detail, there are several ways of navigating this part of the project. The easiest is to look at the central result of the method of stability sets applied to collective action games (pg. 136) and then move on to the chapters where that formula is used. For some purposes this

might be an efficient way to use this book but, of course, readers who follow this path will have to accept as a given central conclusions of this book that are at odds with some received wisdom on the matter. A second option is to read the verbal discussions of the method. I suspect this is what most readers will do and, undoubtedly, for many it will be the wisest choice. But I think that some readers in that group may benefit from a third course of action: looking occasionally at the numerical examples of 2×2 games. I am convinced that the key elements of the method of stability sets are accessible to people without a state-of-the-art technical training. In fact, these examples are written at a level of mathematics not too different from the one that students often attain in their first course on game theory. Furthermore, attentive readers will realize that some topics that are a constant source of confusion in elementary game theory, especially the concept of mixed strategy Nash equilibria, come under a new, simpler and more intuitive light once connected to the notions of stability. In sum, although there is no guarantee that these examples will work for readers already familiar with basic game theory, there is no harm in trying them. They can clarify much of the method's fundamental principles and even some notions of game theory in general.

Beyond the interest that the method may have for game theory, however, I concluded that it had several substantive ramifications, sufficient to support a research program on collective action. If the method of stability sets could bring the precision of formal analysis to the idea that economic development lowers the barriers to collective action among clients of a political machine, this suggested that similar progress could be made in other areas. Collective action is a potent mechanism of social change but it always occurs within a particular set of structural conditions. As social scientists, we should try to make sense of how changes in said structure lead to changes in the possibilities of collective action that individuals face. This would be a crucial step toward a rigorous theory of social transformations. I am convinced that the method developed in this book offers promise in such an endeavor and

have tried to defend this view at length throughout the text. To that end, I offer two lines of argument, each one somewhat independent of the other: one, comprising Chapters 1 and 2, that pertains to the general structure of the theory of collective action and another, developed in Chapters 5 and 6 where I present examples, admittedly crude, of the research program I envision.

The first such line of argument pertains to the general structure of the theory of collective action. Its main arguments are contained in Chapters 1 and 2. For all their accomplishments, the existing approaches to collective action from the perspective of rational-choice theory are unable to answer some fundamental questions. If we want to develop a systematic treatment of collective action we cannot just look at the mechanisms that individuals use to coordinate. We also need to investigate how the circumstances in which they operate help or hinder their efforts. Such is, precisely, the central objective of the method of stability sets.

Thanks to its ability to handle this problem the method is a useful tool in areas where other methods more elaborate and with a more solid foundation have not worked thus far. I give special importance to one such area: collective action as a counterfactual. Although real episodes of large-scale coordination are rare in societies, latent coordination is a potent undercurrent running through most social affairs. An example from the game of chess is entirely appropriate here, given the extensive use of game theory in this book.

Checkmate is the ultimate goal in chess; whoever checkmates the opponent wins. In professional chess, however, only a handful of games end with checkmate. In the vast majority of tournament matches one player resigns when the situation seems hopeless, even if the prospect of a checkmate is still several moves away. In fact, there is something of a convention: a player who waits until the bitter end of a checkmate is regarded as either too incompetent to realize that a checkmate was in the making or too arrogant to believe that the opponent would be able to pull it off. While observationally, checkmates seem negligible, it would be absurd to conclude that they are.

It is not possible to understand chess, to make any sense of the moves of the players at any juncture of the game, without knowing what a checkmate is and how to produce it.

Collective action is the checkmate of social sciences. The members of a society can destroy all its existing institutions, no matter how old and revered, through collective action and yet such action surfaces only in exceptional times. Beneath the surface, its possibility guides the choices of political actors at any given moment, no matter how quiet things seem.

This poses a technical problem for any game-theoretic enterprise. If the multiplicity of equilibria is one of the defining features of collective action, how can we study threats of collective action as opposed to actual instances? What threat assessment can we attribute to the players involved if we do not even know which outcome is the right one? A moment's reflection shows that this question, although entirely pertinent, should not be blown out of proportion. In real life we are constantly making decisions that we regard as good because they avoid other sequences of events. But those sequences themselves may be contingent and we can nevertheless see that we are better off avoiding them. Many drunk drivers, perhaps even a majority, make it safely to their destination. But, although it is not a sure thing that taking a ride from an ostensibly drunken friend will result in disaster, we often think it is wise not to: the risks entailed can be too high, and are higher the more inebriated he seems or the more difficult the road. When we make our decisions, we rely on risk assessments of events that do not occur and, in making such assessments, we draw from our knowledge of how those events could come to happen.

With its emphasis on relative likelihood rather than point-predictions, the method of stability sets can help us develop assessments of the likelihood of coordination in a collective action game even if we never observe it. Those assessments capture in an intuitive and rigorous way our knowledge about the objective circumstances where said collective action could eventually occur. I doubt that something similar can be formulated with the existing methods.

Chapter 1 makes these reflections explicit and spells out the goals of my research program. After clarifying the program the way I envision it, in Chapter 2 I lay down its foundations. In its current state, the rational-choice theory of collective action is not entirely unified and there are some loose ends in some of its formulations. In fact, although having several schools of thought can be healthy, the coexistence of the Olsonian and the Schellingean paradigms is on occasions unfortunate. Quite often one scholar considers some phenomenon an instance of a Prisoners' Dilemma and another scholar sees it instead as a tipping game and it is hard to see the underlying reasons for each choice. The difference is far from trivial: Olsonian models have one equilibrium, Schellingean models, many. In this chapter, then, I offer a unified framework for the study of collective action problems, a framework that covers both the Olsonian and the Schellingean cases. Although models of iterated games are not strictly speaking a particular case of the framework I propose here, I show that they are connected to the other models in ways that can be studied with the method of stability sets. This chapter may be of interest independent of the rest of the book since its results do not require the method of stability sets. Expressing both models as cases of the same mathematical structure allows us to bring out their implicit assumptions so that, when we analyze a collective action problem with the help of game theory, we can know which model is better suited for our purposes.

The analysis developed in that chapter is not kind to the Olsonian model of collective action. As it turns out, the central prediction of the public goods model, that individuals will always free-ride unless they are offered selective incentives, is an artifact of some parameter choices that are hard to justify. Once we relax the restrictions imposed by this model, we realize that collective action problems rather resemble games with multiple equilibria, in the spirit of Schelling's tipping games. This is not to say that the public goods model is useless, but simply that its scope is much more limited than what is often asserted.

The method of stability sets is not simply an alternative to the Olsonian model. It is a method that can be applied to any game, Olsonian, Schellingean or otherwise, even games that do not involve collective action. To prove that point, in Chapter 4 I apply it to the models of collective action already discussed, showing the results it allows us to obtain. Since the method of stability sets is designed to deal with multiple equilibria, when applied to a model with a unique equilibrium, such as Olson's, it gives results that are true but trivial. Its true relevance becomes apparent only when studying games with multiple equilibria, such as the Schellingean models or the models of repeated interactions.

This concludes the first part of this book, the part where I develop the technical foundations for my research program. In the second part I exemplify the work that such a program can generate. Since this part ventures into substantive issues, it is likely to be more controversial than the first one. To mature, a research program must leave behind its speculative stage and produce results that can be compared with already existing knowledge. Beyond a certain point, it becomes futile to argue in the abstract about the connection between collective action and the socioeconomic structures in which it occurs.

As already mentioned, Chapter 5 presents my discussion of clientelism in light of the preceding study on collective action and its game-theoretic underpinnings. Confirming the intuition with which I began this project, the model shows that, under very general circumstances and holding constant other elements, economic growth undermines the grip of a clientelistic machine. But the real test for models is not whether they ratify our previous intuitions but whether they allow us to develop other new intuitions that would have been harder to articulate without a precise language. Thus, I use the same model to study the connection between clientelism and policies of universalistic redistribution, a connection that has not been explored as extensively as the link between clientelism and development. For all the model's limitations, I think these results constitute a first success for the method of stability sets: thanks to it we can formulate new hypotheses about the re-

lationship between clientelism and its wider politico-economic environment.

The path that led me to Chapter 5 also convinced me that with the method of stability sets we can expand the range of applicability of our ideas about collective action. Whereas with the standard tools we can only deal with explicit manifestations of collective action, with this method we can also study its more hidden instances and their far-reaching politico-economic effects. I would not have wanted to put into print a book where all the methodological weight lifting comes down to developing one model of one particular characteristic of some electoral regimes. If, as I believe, the work in this book can serve to launch a research program, it should be possible to show more than one instance of it. In Chapter 6 I do so. Its starting point is simple enough: modern democracies are more than just electoral systems; democratic rights include not just the right to vote, but also the right, subject to varying constraints, of engaging in collective action with other fellow citizens.

Whereas an outstanding literature in political economy has studied the consequences of expanding the right to vote, much less is known about the consequences of expanding the right to associate in collective action. This is a regrettable omission because a society's fundamental institutions owe their stability to the acceptance, or at least acquiescence, they command from the citizens. If that acceptance or that acquiescence disappears, the said institutions go with it as the floodgates to collective action open up. This book is not the right place to elaborate this principle in full detail so I decided instead to illustrate it with an example.

A vast literature in economics has studied how different institutional settings to regulate labor markets can lead to different patterns of economic performance. As some of that literature has already recognized, these regulations emerge from a process of confrontation between capital and labor where collective action is the ultimate source of power, especially for the latter.

The labor market illustrates the role of collective action as a threat and, hence, the need for a method to understand it. Strikes, lockouts and other forms of industrial conflict are, even at worst, a rarity. In most countries, most of the time, the factories and the labor market run smoothly without being disrupted by class struggle. But this normalcy represents a consensus between all the parties involved about many aspects, e.g., wages, paid leaves, pensions. Ultimately, the specifics of this consensus are determined by collective action, understood not as an overt explosion but as a constant threat that both parties wield to make sure that the terms agreed upon are honored.

In their current form, techniques such as those of tipping games and focal points are at a loss in studying games where collective action is only a possibility, so to speak, off the equilibrium path. Instead, with the method of stability sets it is possible to analyze this problem and establish what structural conditions inhibit the possibility of labor's collective action. Such latent collective action plays a decisive role in determining an economy's performance and distribution and, hence, the political coalitions that such economy will engender. Since the method of stability sets shows the connections between the potential for collective action and the economic structure that supports it, the result is a model that explains how the asset distribution and the technology of an economy shape the prospects for different types of interclass political coalitions.

In what I regard as a very fortunate development, the recent literature has already produced deductive arguments of this type, the best of them being real exemplars of rigor and relevance in political economy. Chapter 6 can be read as an attempt to supplement some of that literature. While most of the models in this tradition consider elections as the link through which economic structures affect political outcomes, the model I present here, while recognizing the centrality of elections, goes on to show that more nuanced patterns of coalition building emerge once we make clear the importance of extra-electoral collective action.

By now it must be clear that my original project became something entirely different. Such transformation came at a

price. I did not obtain all the results I wanted about political
clientelism. For example, I wanted to study the effects of trade
on these political systems, in particular, whether it is possible
to discern differences in those effects depending on the sector
where trade occurs. In principle, one would expect that what-
ever effect trade has on the stability of a clientelistic regime
must depend on its relative factor intensity, on the volatility
of the income it generates, on the exchange rate regime that
harnesses it, on the infrastructure it needs and so on. But
I have no regrets for not having gotten there. Such studies
are still possible and, moreover, I think that the method of
stability sets is the most appropriate tool to conduct them.
I did not arrive at the port I intended but I believe that, in
the process of trying, I learned how to build a sturdier vessel
that will take me there next time. The vessel can use several
improvements and there are many more places to explore with
it. This led me to a second conclusion: the next journeys will
be much more interesting if accompanied. That is why I have
decided to publish this book.