THESIS / THÈSE

DOCTEUR EN SCIENCES ÉCONOMIQUES ET DE GESTION

Essays in information aggregation and political economics

Righi, Simone

Award date:
2012

Awarding institution:
Universite de Namur

FACULTES UNIVERSITAIRES NOTRE DAME DE LA PAIX

Membre de l'Académie universitaire "Louvain"

Faculté des Sciences Economiques, Sociales et de Gestion
Département des Sciences Economiques

Dissertation pour l'obtention du titre de

DOCTEUR EN SCIENCES ECONOMIQUES ET DE GESTION

Essays in Information Aggregation and Political Economics

Simone Righi

14 June 2012

Composition of the Committee

Supervisor:	Prof. Gani Aldashev (secretary)
Co-Supervisor:	Prof. Timoteo Carletti (president)
Other members of the committee:	Prof. Micael Castanheira
	Prof. Guillaume Deffuant
	Prof. Enrico Diecidue
	Prof. Marco LiCalzi

Contents

Acknowledgements

It is a pleasure and an honor to thank the many people who made this thesis possible.

It is difficult to overstate my gratitude to my Ph.D. supervisor, Prof. Gani Aldashev. With his inspiration and his enormous knowledge of every field of economics he has been the perfect match for someone like me. Throughout the Ph.D. he provided encouragement, sound advices and lots of good insights. Gani was always ready to see the merits of my ideas even the most crazy ones. I would have been lost without him.

I have an immense debt of gratitude with Prof. Timoteo (Teo) Carletti . Timoteo is the person who brought me to Namur five years ago to collaborate with him. Without his kindness, extreme intelligence, infinite patience, scientific curiosity and versatility I would not have managed to arrive here. He has been a wonderful co-author, a superb co-supervisor and a very good friend.

I would like to thank the many people who taught me and encouraged my scientific curiosity especially David Lane, Roberto Serra and Marco Villani. I acknowledge the contribution of my Ph.D. thesis committee members: Professors Micael Castanheira, Guillaume Deffuant, Enrico Diecidue and Marco LiCalzi; with their comments they helped me improving my thesis. For writing many letters of reference on my behalf and providing helpful advices for the job search I own an additional thanks to Prof. Marco LiCalzi.

I am indebted to all my colleagues, in both the departments of economics and mathematics for providing a stimulating and fun environment in which to learn and grow. I am especially grateful to Elias Carroni, Elena Serfilippi, Helene Laurent, Giulia Camilotti and Wouter Gelade in the department of economics and to my former colleagues Jehan Boreux, Audrey Compere and Jojo Barthelemy. Without their help, the discussions and especially the *happy hours* my life would have been much less fun these years.

I would like to thank my best friend Anne-Sophie Valschaerts and my close graduate student friends Margherita Negri, Francesco Violante, Riccardo Costantini and Mauro

Francesconi, for helping me get through the difficult times and for all the emotional support, entertainment, discussion time and fun they provided.

I am grateful to University of Namur for supporting financially and materially my work across all these years.

I wish to thank my family, especially my parents, Olivio Righi and Barbara Reggiani and my grandparents Ave, Leo e Maria which together raised me, supported me and loved me. A special thanks, finally goes to my brother Fabio, which provided an infinite source of discussion topics that greatly improved my knowledge of the world.

Introduction

This thesis covers, in its four chapters, a diverse set of topics linked to the current research on complex systems. For their nature, these systems cannot be studied only with the paradigms and the toolboxes of a single science. They require a multi-disciplinary approach able to gather the insights and techniques from different domains. Therefore, in order to provide the reader with a general overview of our work, we will first introduce the logical connections between the different chapters and then we will explain the rationales, the intuitions and the methodologies that characterize each contribution.

The main focus of this thesis is the analysis of socio-economic systems characterized by the presence of many heterogeneous agents, each endowed with a piece of information (an opinion, a belief or a preference), which can be shared with some of the other members of the population. Since agents interact mainly with their first order social "neighborhood", the setups that we study as characterized as distributed systems.

Creating micro founded dynamic models involving localized interactions among many heterogeneous agents is a new trend in social sciences and economics that has the objective of developing better tools to reproduce and to predict the behaviour of these, inherently complex, systems. Such models overcome the simplifying, but sometimes inaccurate, assumption that the diversity in a society can be reduced assuming the presence of a representative agent. In this context, it is important to model parsimoniously the features of the agents and the local interaction rules, avoiding unnecessary complications. However, to obtain a global understanding of the problem studied, it is necessary to aggregate a large number of agents and interactions and to observe the consequent *emerging properties* at a macro level. The operation is sometimes challenging, indeed, here the total is rarely simply the sum of its parts and small changes in the micro-setup can have extended influences on the macro characteristics of the system observed.

In this thesis we put great care in trying to keep the inherent complexity of the systems we study analytically tractable. Nevertheless, some of the most interesting phenomena can be only be observed using different methods such as network theory, fractals and numerical methods.

The first two chapters of this manuscript share the same methodological approach: they both answer their research questions analyzing the topological characteristics of networks. These constructions, where the individuals are represented as nodes and the interactions between them as edges that connect pairs of nodes, naturally represent localized relationships at the most micro level while providing efficient tools for studying the macro level in statistical terms. Indeed, while their use is relatively recent in economics, the theory of graphs and the social network analysis developed across several centuries (the birth of graph theory is usually attributed to L. Euler with his seminal paper concerning the "Konigsberg bridges problem" of 1736), so the set of analytical and computational tools available to researchers is quite ample. Network theory is today mainly devoted to construct and characterize "complex networks" exhibiting some of the remarkable and universal properties that are empirically observed when a large group of individuals interact. Among the most intensely studied features of complex networks are: the small world property, the presence of weak ties and the scale-free distribution of the node degrees. We hereby briefly introduce the reader to these concepts, as they are useful to understand our findings.

A network is *scale-free* when the degrees of its nodes (i.e., the number of connections for each node) are distributed according to a power-law. This implies that such network is composed by a limited number of highly connected individuals (hubs) and a large number of nodes with few connections. Empirical socio-economic networks of this kind are, for example, those describing the sexual contacts in a group, the scientific collaborations, the international trade between countries and the financial relationships between banks. All these networks share the common characteristic of being very resilient to the failure of random nodes, but highly sensible to targeted attacks that aim at the elimination of the hubs.

In simple terms, a network is defined *small-world* when it is such that most of the nodes are not directly neighbors of each other but they can be reached from every other in a small number of steps. This feature is typical of graphs where the clustering coefficient (the number of triangles) is relatively large and the average distance among nodes relatively small compared to the one of a comparable network where connections are random. Examples of small-world networks in the real world have been observed in the field of job research, in the bilateral trade between countries and in the political discussions among voters.

Weak-ties are empirically observed features of social networks that emerge when the graph is partitioned in hierarchies of well tied groups of acquaintances weakly connected with each other. To clarify this definition an example is useful. Typically the strong ties of an individual are those with his family and his close friends (most these people normally also know each other), while the weak ties are those with people that our individual sees, and speak with, only rarely (such as the ex-schoolmates or the ex-coworkers). The concept of weak ties is strictly related to the one of information diffusion on a network. In his seminal paper, Granovetter (1983), shows that individuals with weak ties are much more likely to receive new information than peers with only strong ties. This is typically the case of the job-market, where it has been observed that news about new vacancies frequently reach unemployed workers through people outside their own close social neighborhood.

With these definitions in mind, the fundamental questions shared by both chapter one and two are: how ubiquitous are these characteristics? Can they be reproduced with very simple network-based models? In two different setups, both of them relatively simple with respect to the realities that they intend to model, we observe the emergence of some of these properties. On one side this underlines the universality of these features and, on the other side, it indicates that they can be successfully captured even with minimalistic models.

Networks are not the only point of contact between the chapters of this thesis. Both the second and the third chapter address the role of information exchange in large groups of individuals. While micro-economic theory studies (strategic) interactions between agents, it is often assumed that everybody share the same information set (or alternatively the same beliefs). Our approach is different. We start with a group of agents endowed with an information set and we allow them to interact, in one-to-one relationships, and to dynamically change their future behaviour as a consequence of these interactions. This allows us to study both the process of formation and the stable states of opinions and expectations.

The last logical connection between the chapters of this thesis is that, in each of the last three papers, we construct models with microscopic rules and then we study their influence on the aggregated features of the system. The fundamental question is then: which are the emerging properties of these systems? The answer we give depends on the aggregation mechanism used. Specifically, in chapter two the process is completely driven by local (one-to-one) interactions; in chapter three instead the information is integrated by the agents partially through a social mechanism and partially via the signals of a centralized market (prices); finally, in chapter four, we embed heterogeneity among voters preferences in a model a political economics model, where the aggregation of preferences happens through elections.

Let us now briefly introduce, one by one, the papers of this manuscript. For each of them we will explain the setup, we will introduce the reader to those concepts that some could find complicate, and we will integrate the results of the chapter providing and overview of related works and extensions that we realized.

In the first chapter[1] we introduce a particular class of weighted networks inspired by fractal sets. We then proceed to analytically study the topological characteristics of such constructions, finding some surprising similarity between them and the social networks observed empirically.

In order to understand the results of this paper, it is important have an intuition of the main properties of fractal sets. The term fractal comes from the latin *frāctus* which translates as *broken* or *interrupted*. It was first introduced by Mandelbrot (1975) to describe natural objects that are so *fractured* that it is impossible to accurately describe them with the shapes of classical geometry. Classical geometrical objects can be classified on the base

[1] Written with **Timoteo Carletti** and published on *Physica A*

of their dimensionality, which in turn can be defined as the number of distinct parameters required to univocally identify a point in the set (a line has one dimension since a single parameter can univocally define a point on that set; a triangle has two dimensions for analogous reasons). However, it is difficult to apply this definition to fractals, as they would turn out to have a dimension and to behave like sets with a different one. The Hausdorff dimension provides an alternative definition of dimensionality that allows us to say how well a set covers the space in which it is embedded. With this definition it is possible to characterize fractals as sets whose Hausdorff (fractal) dimension exceeds its topological one. Specifically, their dimensionality can be computed using the two parameters of a particular type of contraction mapping called Iterated Function System (IFS): the *number of copies* of the original set (that are made each time the IFS is iterated) and the *scale* of each of them with respect to the original set. This particular construction mechanism produces self-similar sets: if we progressively zoom to finer and finer levels of detail we always observe conformal copies of the whole set.

As we construct weighted networks that share some features of the fractal sets, we name the outcome Weighted Fractal Networks (WFN). The construction mechanism is simple: we start from an initial network (seed) and, iteratively applying an IFS, we construct a family of weighted networks, in which each iteration builds a certain number of copies of the original network, each scaled of a certain factor. Using an explicitly defined algorithm and the self-similarity of fractals, we are able to completely characterize the topology of these networks as a function of the parameters involved in their construction. In this way we prove that WFNs exhibit the small world property. Moreover we show that the probability distribution of the node strength follows a power law whose exponent is the fractal dimension of the underlying fractal, thus proving that WFN are scale-free.

Many of our results are made more simple by the symmetry of our construction algorithm and by its completely deterministic nature. At the end of this chapter we therefore relax the first assumption, generalizing the previous results to non-homogeneous scaling factors for the different subnetworks (so that our graphs grow more irregularly). We prove that the same properties hold true in this case and that the homogeneous WFN are a special case of the non-homogeneous one. Moreover Carletti (2010), further extends our setup (obtaining similar results) to stochastic constructions, where also the number of copies of the original network (made at each step of iteration of the construction) is drawn from a distribution. These two extensions show that fractal-inspired networks always share some characteristics with the original fractal sets and that the properties studied in this thesis can be reproduced using very general construction mechanisms.

While this is a theoretical paper, the construction algorithm for the class of networks we study can be applied to problems interesting for social scientists. This setup is useful to model the diffusion of information, with losses of flow, in hierarchical social networks. One interesting example is the structure of political parties in communist countries. There, the need to control the society and the means of production, requires a government with many layers (from the central government down to the industry management) each self-similar to the initial one (e.g., secretary, politburo, central committee and general assembly). It is well known (see Hayek, 1944) that these political systems deal with extreme problems of information loss due to the massive bureaucracy. WFN can be of help in analyzing the extension (the number of levels of command), and thus the degree of control on the society, that can be sustained.

The second chapter of this thesis [2] examines the dynamical evolution of a social network emerging from a model of opinion exchanges between a closed group of agents. The network analysis done in this chapter contributes to the studies of continuous opinion dynamics models. Many different models have been proposed by this literature with the aim of uncovering the mechanisms through which people change opinions due to contacts with others and to understand under which conditions the process of opinion formation leads to

[2] Written with **Duccio Fanelli** and **Timoteo Carletti**, and published on *Advances in Complex Systems*

polarization (fragmentation of the opinions in different groups) or to consensus (all agents share the same opinion) among the agents. The typical setup is one with a group of agents, each endowed with an opinion (or a preference) on some subject of discussion. These agents are generally allowed to interact with each other through some local rule, but they cannot broadcast directly their points of view to all their peers at the same time.

Many of the contributions to this literature assume that agents meet randomly or impose a fixed network topology over which interactions are then studied. However, in reality, exchanges of opinions influence people's chances to remain in contact, thus modifying the social network. For instance, an exchange between two persons, that discover to have very different opinions on some subject, can lead to the end of a well established relationship. On the other hand, two people with no previous interactions with each other, can discover that they have very similar opinions and thus develop a friendship.

Bagnoli et al. (2007), elaborating on the previous literature, introduces a model in which the social network of affinities between agents is assumed to evolve, coupled with opinions exchanges, following a specific set of rules. When two agents interact they may reach a compromise and, if so, they increase their mutual degree of acquaintance thus strengthening their reciprocal affinity. The specificity of this model is that it allows the co-evolution of opinions and social structures (emblematized by the affinity network), thus generating a genuine *adaptive network*. Our aim, in chapter two, is to capture some of the properties of the networks that emerge from this model. More specifically, the key question that we ask ourself is whether starting from a random group (both in terms of initial opinions and social ties), this can evolve toward a final state where the social structure shares some of the features observed in real world social networks.

We numerically find that, under some quite general conditions, this minimalistic setup evolves networks that are characterized by the presence of weak ties and in which the network topology is small-world. Moreover, we are able to obtain analytically a very good approximation of the time evolution of the average degree in the network. The emergence of these properties in our model is a direct consequence of the co-evolution of opinions and affinities (which evolve self-consistently in a two dimensional space). Indeed, repeating the same experiments on other models of opinion dynamics on non-trivial fixed network topologies (which evolve on a one-dimensional space), these characteristics disappear. Our results underline the fact that, departing from unidimensional models, it is possible to capture some of the properties of real social networks even with very simple scenarios and without imposing many restrictive assumptions a priory.

The paper presented in this thesis is the result of a wider endeavor, on our side, to uncover the properties of these kind of models. Indeed in Righi and Carletti (2009a, 2009b) we extend the analysis initializing the social network to non trivial topologies. The aim is to understand if the initial structural properties of the social network are preserved by the dynamic evolution. We show that, imposing a Barabási-Albert scale-free initialization (where the strength of the affinities among agents is distributed as a power-law), the system progressively - but slowly - loses memory of the initial state and the degree distribution becomes progressively log-normal. With this initialization it is initially more difficult to realize successful interactions (those that change opinions) as most of them must involve at least one hub to succeed. However, using the Watts and Strogatz (1998) algorithm to construct an initial network of acquaintanceship with the small-world property, we observe that this initial distribution is partially preserved as the opinions converge. These papers thus confirm that this model tends to generate and maintain small-world networks.

In Carletti and Righi (2008) we extend Bagnoli et al. (2007) in a different direction. We allow contemporaneous meetings between more than two agents (i.e., group interactions) applying a more general dynamic rule of interaction, in which not only the network and the opinions evolve, but also the size of the groups that interact. In this, more realistic, setup we study the phase transition between the consensus and the polarization state. We show that there exists a polarization-consensus phase transition (in the steady state of the system) that depends only on how open minded are the agents and not on the parameters relative to the social network. Moreover, we find that this three-dimensional space allows for more complex relationships among the agents. The simulated people first interact in affinities, essentially selecting their friend, and only eventually they begin to modify their opinions.

Finally in Righi and Carletti (2012), we make one last step forward. Using real world data on the lengths of social interactions among a group of scientists (from Cattuto, 2010) participating to a conference, we show evidence that the observed supra linear growth of total discussion time with respect to the agent connectivity is compatible with two different hypotheses - about the nature of their interaction - based on the group structure. These can be summarized in two micro-founded models of social interaction. In the first one we assume discussions among the agents to be performed about a neutral topic (the opinion exchanges do not influence the discussion times nor the selection of the partner), i.e. the goal of the discussions is to introduce himself to the partner (networking). In the second model, we instead assume that agents encounters are aimed to find a compromise among them and thus that there could be an opinion update, as consequence of their meeting. We find that when the group exhibits clearly identified well reputed people, then the outcomes of our model are consistent with the assumption that discussions are engaged mainly to share time with the most reputed person (networking). On the other hand, when the reputation hierarchy is not strong, our findings are consistent with the hypothesis that discussions are finalized to opinions exchange.

For future works we envision several directions of research, mainly linked to applications of this modeling technique to economics. First, the modeling needs to be improved with the introduction of some degree of rationality in the agents' behaviour and of a well defined utility function (while preserving the heterogeneity in the objectives: for instance some people enjoy diversity of opinions, while some other prefer to deal with peers with similar ideas). This kind of implementation could then be used to study economic problems, prominently within models of asset pricing (see chapter three of this thesis) and political economics. Specifically, in political economics, the model discussed here could be useful to account for how the social interactions among voters influence their political positions.

In the third chapter of the thesis [3], we study the informational efficiency of a financial market where traders are subject to two types of behavioral bias. It has been long known that markets can be studied not only for their characteristics in terms of allocative efficiency (the proportion of profitable exchanges that are actually realized), but also as aggregators of information. A market is informationally efficient when the price converges, over time, to the value that would obtain if all market participants had full information about the fundamental value of the asset exchanged. The study of the informational properties of financial markets dates back to Hayek (1945) and concentrates on setups where it is assumed that the traders are all fully rational. Under this assumption it has been shown that both centralized markets, where an *auctioneer* centralizes all the trade proposals and there is a unique price (most of the financial markets belong to this category) and decentralized ones, where buyers and sellers meet directly and make exchanges so that a plurality of prices can co-exist (goods markets are frequently of this category), are efficient in aggregating information.

However, financial markets are ultimately composed of human beings whose decisions can be subject to a variety of behavioral bias. Specifically, there is consistent evidence from experimental economics and behavioral finance which indicates that traders do not behave in a fashion consistent with the full rationality assumption. Indeed, traders seem to have *adaptive expectations*: they give more importance to past prices than what a fully-rational agent would (under full rationality past prices cannot serve as predictors for future ones). Moreover, there is a social dimension in human behaviour that extends to trading. Market participants do not act on markets solely upon their perceptions or data, they also entertain discussions with each other that can modify their expectations. On this regard it has been noted that there exists some degree of *confirmatory bias* in the behaviour of traders: they tend to discard the information that differs too much from their priors. In this context

[3]Written with **Gani Aldashev** and **Timoteo Carletti**, and published on *Journal of Economic Behaviour and Organization*

models of opinion dynamics, similar to the one studied in the second chapter of this thesis, can be effectively applied to reproduce this kind of behaviour.

The relevant question here is whether the financial markets can still be informationally efficient when agents do not behave fully rationally. We answer this question with a simple model, in which traders are initially endowed with some noisy private information (correct on average) about the fundamental value of an asset and they are subject to the two behavioral bias discussed. A certain fraction of the agents revises his expectation each period giving some weight to past prices and also exchanging opinions about future prices with some other agent. Not all the social information is however integrated by agents, which are subject to some degree of confirmatory bias. We show that, taken separately, each of the deviations from rationality worsen the informational efficiency of the market (as expected). However, when the two bias are combined, the degree of informational efficiency of the market can be non-monotonic both in the weight of the adaptive component and in the degree of the confirmatory bias. For some ranges of parameters, we find the unexpected result that the two biases mutually dampen their effects on price dynamics, thus increasing the informational efficiency. From the policy perspective, our results contribute to the debate on the nature of speculative bubbles, suggesting that it is the irrationality of the agents (and not the incomplete information) to allow the asset prices to deviate - in the long run - from their fundamental values.

In order to extend and generalized the results of this chapter, in Righi et al. (2011) we relax some of the technical assumptions made here. We confirm that, with different mechanisms of entrance of new agents and with different hypotheses on which agents do update their expectations, most of the results still hold.

The work of this chapter opens the doors for a considerable amount of future research. First, the model itself should be further generalized in several directions. Indeed, our setup makes two strong simplifying assumptions: that agents are homogeneous in everything except their initial expectations and that the strength of their bias is fixed. The introduction of some degree of heterogeneity in the behavioral bias and the possibility for them to evolve toward rationality would make the model more realistic and able to capture learning effects. Complementarily, here we study only the interaction between two well known behavioral bias, but many others are studied in the empirical literature (e.g. loss-aversion, anchoring and systematic errors). In the future we will try to integrate some of these other deviations from rationality in our model.

Finally, while our model does not allow for the emergence of speculative bubbles, we suggest that behavioral bias (that prevent the market from discovering the fundamental price), could explain their emergence. A current (2012) research proposal of Biondi and Righi goes in this direction. We propose to study the effects of social interactions in a context characterized by a realistic modeling of the market micro-structure. Our intuition is that social interactions (together with behavioral bias) amplify the market oscillations due to the flux of new information (whose characteristics have been shown to depend on the accounting rules, imposed by regulators, by Biondi, Giannoccolo, and Galam, 2011), increasing the chances of appearance of significative deviations of the markets from the fundamental value of the assets traded. This proposal, encompass both the creation of an enhanced theoretical model and the realization of, much needed, experimental tests of the specification presented here. The experimental setup that we envision is one with an experimental market where the subjects are allowed to trade, to retain informations about the past prices and in which they are given the opportunity to share information with each other in a limited fashion. This would allow to study in the laboratory (without the spurious influence of new exogenous information reaching the traders) the realism of the evolution of both market prices and expectations.

Finally, in the fourth paper of the thesis[4], we deal with an issue that emerges when studying electoral campaigns: the relationship between campaign spending of candidates

[4] Which is also our Job Market Paper

and corruption of the political system. Representative democracy, where policy choices are delegated to elected representatives, is considered as an efficient mechanism to aggregate the preferences of large social groups. However, when citizens delegate economic choices to professional politicians, the latter can use their power (and the imperfect control that the voters have on their actions) in order to obtain personal advantages (hereby denominated *rents*). In the terminology of political economics, this is the problem of the agency costs of political delegation. This literature - that deals with the conflicts of interest among voters with different policy preferences and between voters (that would like to be governed by honest politicians) and candidates - predicts the degrees of efficiency of electoral competitions as well as the measure to which candidates can obtain personal advantages at the expenses of voters.

In political economics, the electoral competition is modeled similarly to any other type of competition with rational players (such as the one between firms). Therefore, when candidates are perfectly informed about voters' preferences, and when voters are rational (i.e. in perfect competition), the force of the electoral competition is enough to avoid candidates' misbehavior. However, when voters are ideological or candidates are uncertain about voters' evaluation of their platforms (imperfect competition), the competition relaxes and the candidates can manage to obtain positive rents. Our main contribution is the introduction of an analysis of candidates' campaign choices this literature.

We are interested, in particular, in studying countries where corruption is endemic and where the electoral law limits the possibility for lobbies to bias the political outcome with their donations to candidates. Our key questions are therefore: How are the levels of corruption and embezzlement of a political system influenced by electoral campaigns? How rent extraction can be reduced with anti-corruption policies? We answer these questions in the context of a probabilistic voting model characterized by the presence of ideological voters whose preferences can be manipulated, by political candidates, through campaign spending. Moreover, we assume the existence of a direct link between campaign expenditures and utility of voters. With our modeling choice we are able to capture two opposite empirical observations. On one side, voters participate to political activities (e.g., political rallies) even without the perspective of obtaining a direct monetary advantage from this. On the other side, in advanced democracies the large availability of information about candidates seems to be connected with a progressive hostility of the population towards politics.

Our main finding and testable prediction is that campaigning choices are orthogonal to decisions about rent extractions (one does not influence the other directly). Moreover, when given the opportunity, candidates always invest a significative amount of their resources in advertisements, in the attempt of being elected. However, the measure in which they do so is directly correlated with the amount of (dis-)utility obtained by the voters as a consequence of their campaigns.

In most of the cases positive net rents (i.e. gross rents minus campaign expenditures) are still possible in equilibrium. We therefore turn our attention to the study of which welfare policies can reduce the inefficiency of the electoral competition. In absence of lobbies, reducing the amount of resources that can be invested in electoral campaign by the candidates always reduces the welfare of the voters. Therefore the typical mechanism studied by researchers through which the voters' welfare can be increased is ineffective here. Our main policy suggestion is to introduce an anti-corruption policy, i.e. a policy that reduces the ability of candidates to extract rents by abating the incentives to rent accumulation. In practice, we propose to introduce a new tax on voters that, beyond the control of the politicians, finances the strengthening of the police, the justice system and the other institutions of control on the political system, so to make more difficult (costly) for the candidate to embezzle. We show that the introduction of such tax can make the citizens better off. Surprisingly, it may also make the candidates better off if the policy is not sufficiently efficient (i.e. when it is insufficiently financed). Finally, we establish the conditions under which a policy of this kind can achieve the popular support required for an effective implementation and we show that these conditions are difficult to achieve in countries with large income inequalities. This results allow us to give a theoretical explanation to the observation that, while most of the counties where corruption is endemic have policies to fight the phenomenon, these are usually ineffective. The politicians use the cost of these policies, and the consequent unwillingness of part of the population to pay for them, as an excuse to implement only partial reforms, and to allow the prosecution of embezzlement behind a façade of honesty.

Chapter 1

Weighted Fractal Networks

1.1 Introduction

Complex networks have recently attracted a growing interest of scientists from different fields of research, mainly because they define a powerful framework, among others developed in the complex systems sciences, for describing, analyzing and modeling real systems that can be found in Nature and/or society. This framework allows to conjugate the micro to the macro abstraction levels: nodes can be endowed with local dynamical rules, while the whole network can be thought to be composed by hierarchies of clusters of nodes, that thus exhibits aggregated behavior.

The birth of graph theory is usually attributed to L. Euler with his seminal paper concerning the "Königsberg bridge problem"(1736), but it is only in the 50's that network theory started to develop autonomously with the pioneering works of Erdős and Rényi (1959). Nowadays network theory defines a research field in its own (see Albert and Barabási, 2002 and Boccaletti et al., 2006) and the scientific activity is mainly devoted to construct and characterize complex networks exhibiting some of the remarkable properties of real networks, scale–free (see Barabási and Albert, 1999), small–world (see Watts and Strogatz, 1998) , communities (Fortunato, 2009) weighted links (for instance Yook et al., 2001; Barrat et al., 2004 and Barrat, Bartelemy and Vespignani, 2004), just to mention few of them.

In recent years we observed an increasing production of papers (for instance, Barabasi et al., 2001; Jung et al., 2002; Dorogovtsev et al., 2002; Ravasz and Barabási, 2003; Dorogovtsev et al., 2004; Barrat et al., 2004; Zhang, 2008; Zhang, Zhou et al., 2008 and Guan et al., 2009) where authors proposed a new point of view by constructing networks exhibiting scale-free and hierarchical structures by adapting ideas taken from fractal construction, e.g. Koch curve or Sierpinski gasket. The aim of the present paper is to generalize these latter constructions and to define a general framework, hereby named *Weighted Fractal Networks*, WFN for short, whose networks share with fractal sets several interesting properties, for instance the self-similarity and the hierarchical structures.

The WFN are constructed via an explicit algorithm and we are able to completely analytically characterize their topology as a function of the parameters involved in the construction. We are thus able to prove that WFN exhibit the "small–world" property, i.e. slow (logarithmic) increase of the average shortest path with the network size, and large average clustering coefficient. Moreover the probability distribution of node strength follows a power law whose exponent is the Hausdorff (fractal) dimension of the "underlying" fractal, hence the WFN are scale–free.

WFN also represent an explicitely computable model for the renormalization procedure recently applied to complex networks (see Song et al. 2005, 2006; Radicchi et al. 2008, 2009).

The paper is organized as follows. In the next section we will introduce the model, we outline the similarities with fractal sets and the differences with respect to the above mentioned earlier results. In Section 1.3 we present the analytical characterization of such networks also supported by dedicated numerical simulations. We then introduce in Section 1.4 a straightforward generalization of the previous theory, and thus we conclude by showing a possible application of WFN to the study of fractal structures emerging in Nature.

1.2 The model

According to Mandelbrot (1982) "a fractal is by definition a set for which the Hausdorff dimension strictly exceeds the topological dimension". One of the most amazing and interesting feature of fractals is their *self-similarity*, namely looking at all scales we can find conformal copies of the whole set. Starting from this property one can provide rules to build up fractals as fixed point of *Iterated Function Systems* (see for details, Barnsley, 1988 and Edgar, 1990), IFS for short, whose Hausdorff dimension is completely characterized by two main parameters, the number of copies $s > 1$ and the scaling factor $0 < f < 1$ of the IFS. Let us observe that in this case this dimension coincides with the so called similarity dimension

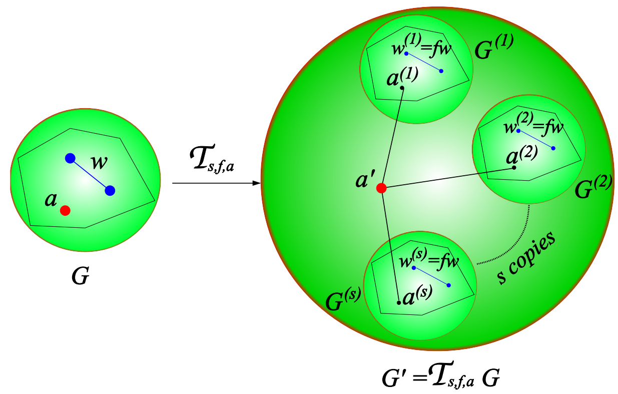

Figure 1.1: The definition of the map $\mathcal{T}_{s,f,a}$. On the left a generic initial graph G with its attaching node a (red on-line) and a generic weighted edge $w \in G$ (blue on-line). On the right the new graph G' obtained as follows: Let $G^{(1)}, \dots, G^{(s)}$ be s copies of G, whose weighted edges (blue on-line) have been scaled by a factor f. For $i = 1, \dots, s$ let us denote by $a^{(i)}$ the node in $G^{(i)}$ image of the labeled node $a \in G$, then link all those labeled nodes to a new node a' (red on-line) through edges of unitary weight. The connected network obtained linking the s copies $G^{(i)}$ to the node a' will be by definition the image of G through the map: $G' = \mathcal{T}_{s,f,a}(G)$.

(Edgar, 1990), $d_{fract} = -\log s/\log f$.

The main goal of this paper is to generalize such ideas to networks, aimed at constructing weighted complex networks [1] with some a priori prescribed topology, that will be described in terms of node strength distribution, average (weighted) shortest path and average (weighted) clustering coefficient, depending on the two main parameters: the number of copies and the scaling factor [2]. Moreover taking advantage of the similarity with the IFS fractals, some topological properties of the networks will depend on the fractal dimension of the IFS fractal.

Let us fix a positive real number $f < 1$ and a positive integer $s > 1$ and let us consider a (possibly) weighted network G composed by N nodes, one of which has been labeled *attaching node* and denoted by a. We then define a map, $\mathcal{T}_{s,f,a}$, depending on the two parameters s, f and on the labeled node a, whose action on networks is described in Fig. 1.1.

[1]We hereby present the construction for undirected networks, but it can be straightforwardly generalized to directed graphs as well.

[2]A straightforward generalization will be presented in the next Section 1.4. See also Carletti (2009) where the WFN theory will be generalized as to include a stochastic iteration process.

So starting with a given initial network G_0 we can construct a family of weighted networks $(G_k)_{k\geq 0}$ iteratively applying the previously defined map: $G_k := \mathcal{T}_{s,f,a}(G_{k-1})$.

Because of its general definition, the map $\mathcal{T}_{s,f,a}$ could be used to cast the various models of fractal or hierarchical networks recently proposed (Barabasi et al., 2001; Jung et al., 2002; Dorogovtsev et al., 2002; Ravasz and Barabási, 2003; Dorogovtsev et al., 2004; Barrat et al., 2004; Zhang, 2008; Zhang, Zhou et al., 2008 and Guan et al., 2009) into a unified scheme, instead of using "ad hoc"constructions and computations, and so to obtain informations about some relevant topological quantities, such as average shortest path and nodes strength, in a straightforward way using the proposed framework and thus prove, for instance, the small–world character of WFN. In this framework we are also able to construct scale–free networks with any desired power–law just by setting s and f, thus overcoming some limitations of Dorogovtsev et al. (2002).

Let us observe that the main differences between WFN and the above mentioned models reside in the tree–like structure, instead of a star–like (such as those of Barabási et al., 2001; Jung et al., 2002; and Ravasz and Barabási, 2003), that the former acquire because of the growth process, although WFN will exhibit high clustering coefficient, and the absence of a preferential attachment mechanism weight dependent (such as those of Yook, 2001; Barrat et al., 2004 and Dorogovtsev and Mendes, 2004).

For the sake of completeness we present numerical results for two WFN: the *Sierpinski* one (see Fig. 1.2) and the *Cantor dust* (see Fig. 1.3).

Given G_0 and the map $\mathcal{T}_{s,f,a}$ we are able to completely characterize the topology of each G_k and also of the limit network G_∞, defined as the fixed point of the map: $G_\infty = \mathcal{T}_{s,f,a}(G_\infty)$. Let us observe that the topology of G_∞ will not depend on the initial "seed"G_0 as it happens for the IFS fractal sets.

Thus the WFN undergo through a growth process strictly related to the inverse of the renormalization procedure (see Song et al. 2005, 2006); at the same time G_∞ will be infinitely renormalizable.

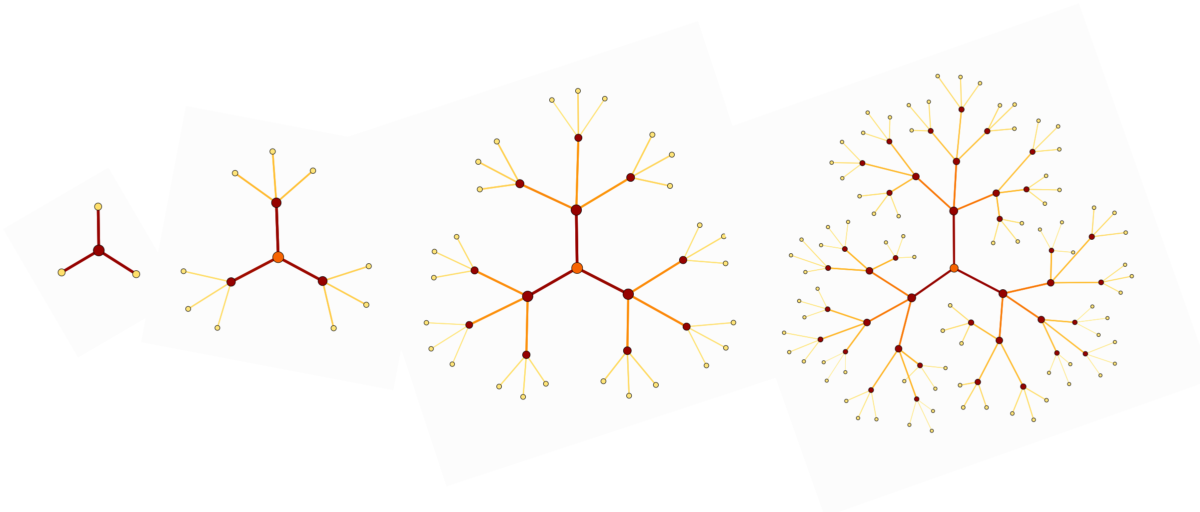

Figure 1.2: The “Sierpinski” WFN, $s = 3$, $f = 1/2$ and G_0 is composed by a single node. From the left to the right G_1, G_2, G_3 and G_4. Gray scale (color on-line) reproduces edges weights: the darker the color the larger the weight. The dimension of the fractal is $\log 3/\log 2 \sim 1.5850$. Visualization was done using Himmeli software [38] .

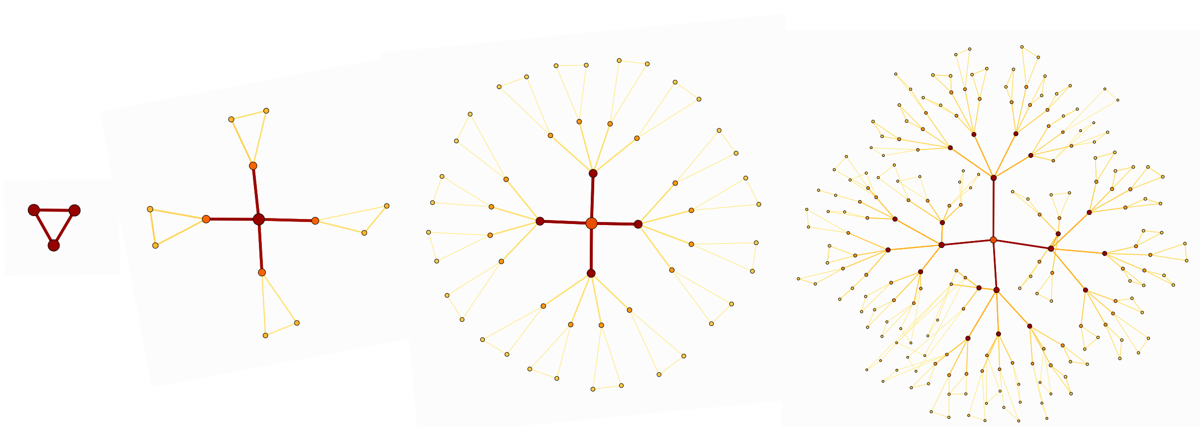

Figure 1.3: The “Cantor dust” WFN, $s = 4$, $f = 1/5$ and G_0 is a triangle. From the left to the right G_0, G_1, G_2 and G_3. Gray scale (color on-line) reproduces edges weights: the darker the color the larger the weight. The dimension of the fractal is $\log 4/\log 5 \sim 0.8614$. Visualization was done using Himmeli software [38].

1.3 Results

The aim of this section is to characterize the topology of the graphs G_k for all $k \geq 1$ and G_∞, by analytically studying their properties such as the average degree, the node strength distribution, the average (weighted) shortest path and the average (weighted) clustering coefficient.

At each iteration step the graph G_k grows as the number of its nodes increases according to

$$N_k = s^k N_0 + (s^k - 1)/(s-1)\,, \tag{1.1}$$

being N_0 the number of nodes in the initial graph, while the number of edges satisfies

$$E_k = s^k E_0 + s(s^k - 1)/(s-1)\,, \tag{1.2}$$

being E_0 the number of edges in the graph G_0. Hence in the limit of large k the *average degree* is finite and it is asymptotically given by

$$\frac{E_k}{N_k} \underset{k\to\infty}{\longrightarrow} \frac{s + E_0(s-1)}{1 + (s-1)N_0}\,. \tag{1.3}$$

Let us denote the weighted degree of node $i \in G_k$, also called *node strength* (see Barrat et al, 2004), by $\omega_i^{(k)} = \sum_j w_{ij}^{(k)}$, being $w_{ij}^{(k)}$ the weight of the edge $(ij) \in G_k$; then using the recursive construction, we can explicitly compute the total node strength, $W_k = \sum_i \omega_i^{(k)}$, and, provided $sf \neq 1$, easily show that

$$W_k = 2s\frac{(sf)^k - 1}{sf - 1} + (sf)^k W_0\,.$$

Because $f < 1$, we trivially find that the *average node strength* goes to zero [3] as k increases: $W_k/N_k \underset{k\to\infty}{\longrightarrow} 0$.

[3] Let us observe that the same is true if $sf = 1$; in this case, in fact W_k grows linearly with k, thus slower than N_k.

Node strength distribution.

Let $g_k(x)$ denote the number of nodes in G_k that have strength $\omega_i^{(k)} = x$ and let us assume g_0 to have values in some finite discrete subset of the positive reals, namely:

$$g_0(x) > 0 \text{ if and only if } x \in \{x_1, \dots, x_m\} \,,$$

otherwise $g_0(x) = 0$. Using the property of the map $\mathcal{T}_{s,f,a}$ we straightforwardly get $g_k(x) = sg_{k-1}(x/f)$ provided [4] $x \neq s$ and $x \neq fs+1$, from which we can conclude that for all k:

$$g_k(x) = s^k g_0(x/f^k)\,, \quad g_k(fs+1) = s \quad \text{and} \quad g_k(s) = 1\,. \tag{1.4}$$

This implies than the node strengths are distributed according to a power law with exponent $d_{fract} = -\log s/\log f$, that equals the fractal dimension of the fractal obtained as fixed point of the IFS with the same parameters s and f. In fact defining $x_{ik} = f^k x_i$ we get:

$$\begin{aligned}\log g_k(x_{ik}) &= k\log s + \log g_0(x_i)\\ &= \frac{\log s}{\log f}\log x_{ik} + \log g_0(x_i) - \frac{\log s}{\log f}\log x_i\,,\end{aligned}$$

namely for large k (see Fig. 1.4)

$$g_k(x) \sim C/x^{d_{frac}}\,. \tag{1.5}$$

Average weighted shortest path.

By definition the average *weighted shortest path* (Boccaletti et al., 2006) of the graph G_k is given by

$$\lambda_k = \frac{\Lambda_k}{N_k(N_k-1)}\,, \tag{1.6}$$

where

$$\Lambda_k = \sum_{ij\in G_k} p_{ij}^{(k)}\,, \tag{1.7}$$

[4] Without loose of generality we can assume that for all integers $i,j \in \{1,\dots,m\}$ and $k>0$ we have $f^k x_j \neq x_i$ and $f^k(fx_j+1) \neq x_i$.

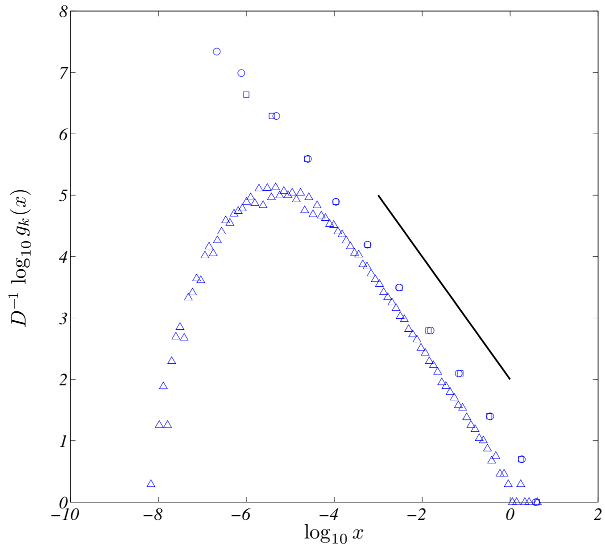

Figure 1.4: Node Strengths Distribution. Plot of the renormalized node strengths distribution $D^{-1}\log_{10} g_k(x)$, where $D = d_{frac}$ in the homogeneous case, while $D = -\frac{s \log s}{\log(f_1 \ldots f_s)}$ in the non–homogeneous one. Symbols refer to : $\square$ the finite approximation G_{14} with 2391484 nodes of the "Sierpinski" WFN, $s = 3$, $f = 1/2$ and G_0 is formed by one initial node; $\bigcirc$ the finite approximation G_{11} composed by 3495253 nodes of the "Cantor dust" WFN, $s = 4$, $f = 1/5$ and G_0 is made by a triangle; $\triangle$ the finite approximation G_{11} composed by 3495253 nodes of the non–homogeneous "Cantor dust" WFN, $s = 4$, $f_1 = 1/2$, $f_2 = 1/3$, $f_3 = 1/5$, $f_4 = 1/7$ and G_0 is formed by a triangle. The reference line has slope -1; linear best fits (data not shown) provide a slope -0.9964 ± 0.034 with $R^2 = 0.9993$ for the Sierpinski WFN, a slope -1.002 ± 0.064 with $R^2 = 0.9996$ for the Cantor dust WFN and a slope -1.006 ± 0.024 with $R^2 = 0.9976$ for the non–homogeneous Cantor dust WFN.

being $p_{ij}^{(k)}$ the weighted shortest path linking nodes i and j in G_k.

To simplify the remaining part of the proof it is useful to introduce $\Lambda_k^{(a_k)} = \sum_{i \in G_k} p_{ia_k}^{(k)}$, i.e. the sum of all weighted shortest paths ending at the attaching node, $a_k \in G_k$. One can prove (see Appendix 1.6) that for large k the asymptotic behavior of $\Lambda_k^{(a_k)}$ is given by

$$\Lambda_k^{(a_k)} \underset{k \to \infty}{\sim} \frac{N_0(s-1)+1}{(1-f)(s-1)} s^{k-1} \,. \tag{1.8}$$

Using the construction algorithm and its symmetry one can prove (see Appendix 1.6)

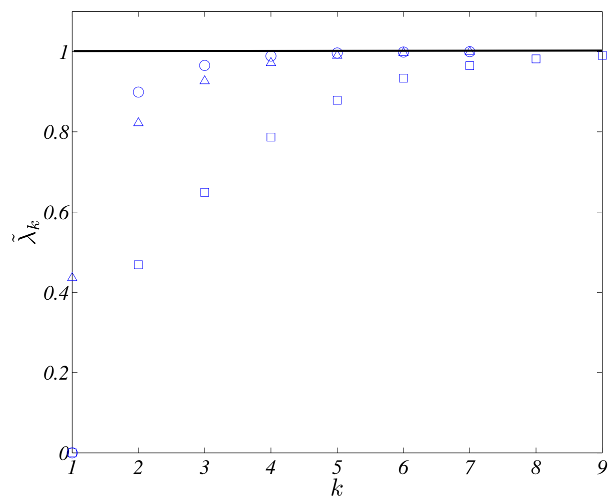

Figure 1.5: The average weighted shortest path. Plot of the renormalized average weighted shortest path $\tilde{\lambda}_k$ versus the iteration number k, where $\tilde{\lambda}_k = \lambda_k \frac{(s-F)(s^2-F)}{2s^2(s-1)}$ and $F = f_1 + \cdots + f_s$ for the non–homogeneous case, while $F = sf$ for the homogeneous one. Symbols refer to: □ the "Sierpinski" WFN, $s = 3$, $f = 1/2$ and G_0 is formed by one initial node; ◯ the "Cantor dust" WFN, $s = 4$, $f = 1/5$ and G_0 is made by a triangle; △ the non–homogeneous "Cantor dust" WFN, $s = 4$, $f_1 = 1/2$, $f_2 = 1/3$, $f_3 = 1/5$, $f_4 = 1/7$ and G_0 is formed by a triangle.

that Λ_k satisfies the recursive relation

$$\Lambda_k = sf\Lambda_{k-1} + 2s[(s-1)N_{k-1} + 1][N_{k-1} + f\Lambda_{k-1}^{(a_{k-1})}] , \tag{1.9}$$

that provides the following asymptotic behavior in the limit of large k (see Fig. 1.5)

$$\lambda_k = \frac{\Lambda_k}{N_k(N_k-1)} \underset{k\to\infty}{\longrightarrow} \frac{2(s-1)}{(1-f)(s-f)} . \tag{1.10}$$

We can also compute the *average shortest path*, ℓ_k, formally obtained by setting $f = 1$ in the previous formulas (1.6) and (1.7). Hence slightly modifying the results previously presented we can prove that asymptotically we have

$$\ell_k \underset{k\to\infty}{\sim} 2\left(k - \frac{s}{s-1}\right) \underset{k\to\infty}{\sim} \frac{2}{\log s} \log N_k , \tag{1.11}$$

where the last relation has been obtained using the growth law of N_k given by equation (1.1) (see Fig. 1.6). Let us remark that the average shortest path is a topological quantity and

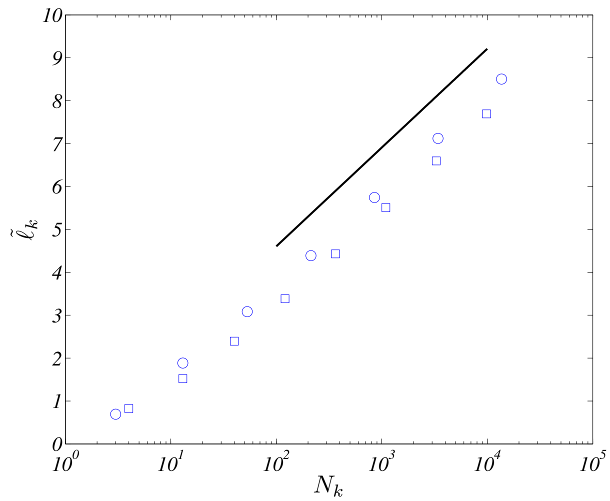

Figure 1.6: The average shortest path ℓ_k as a function of the network size (semilog graph). Plot of the renormalized average shortest path $\tilde{\ell}_k$ versus the network size N_k, where $\tilde{\ell}_k = \ell_k \frac{\log s}{2}$. Symbols refer to : □ the "Sierpinski" WFN, $s = 3$, $f = 1/2$ and G_0 is formed by one initial node; ◯ the "Cantor dust" WFN, $s = 4$, $f = 1/5$ and G_0 is made by a triangle. The reference line has slope 1. Linear best fits (data not shown) provides a slope 0.9942 ± 0.019 and $R^2 = 1$ for the Sierpinski WFN and a slope 0.9952 ± 0.019 and $R^2 = 1$ for the Cantor dust WFN.

thus it doesn't depend on the scaling factor, that's why we don't report in Fig. 1.6 the case of the non-homogeneous WFN.

Thus, as previously stated, the network grows unbounded but with the logarithm of the network size, while the weighted shortest distances stay bounded.

Average clustering coefficient.

The average clustering coefficient (Watts and Strogatz, 1998 and Boccaletti et al. 2006) of the graph G_k is defined as the average over the whole set of nodes of the local clustering coefficient $c_i^{(k)}$, namely $< c_k >= C_k/N_k$, where $C_k = \sum_{i \in G_k} c_i^{(k)}$. Because of the construction algorithm the network inherits a tree–like structure preventing the inner core from acquiring new triangles, in such a way that the number of possible triangles, hence the local clustering coefficient, at each step increases just by a factor s. Thus after k iterations we will have

$C_k = s^k C_0$, being C_0 the sum of local clustering coefficients in the initial graph. We can thus conclude that the clustering coefficient of the graph is asymptotically given by:

$$< c_k > \underset{k\to\infty}{\longrightarrow} \frac{s-1}{s} \frac{< c_0 > N_0}{(s-1)N_0+1} \,. \tag{1.12}$$

On the other hand, one can use edges' values to weigh the clustering coefficient (Saramäki et al. 2007); hence generalizing the previous relation, we can easily prove that the average *weighted clustering coefficient* of the graph is asymptotically given by:

$$< \gamma_k > \underset{k\to\infty}{\sim} \frac{s-1}{fs} \frac{< \gamma_0 > N_0}{(s-1)N_0+1} f^k \underset{k\to\infty}{\sim} \frac{1}{N_k^{1/d_{fract}}} \,, \tag{1.13}$$

where once again, the fractal dimension d_{fract} of the IFS fractal play a relevant role.

1.4 Non–homogeneous Weighted Fractal Networks

The aim of this section is to slightly generalize the previous construction to the case of *non–homogeneous* scaling factors for each subnetwork $G^{(i)}$. So given an integer $s > 1$ and s real numbers $f_1, \dots, f_s \in (0,1)$, we modify the map $\mathcal{T}_{s,f,a}$ by allowing a different scaling for each edge weight according to which subgraph it belongs to: if the edge $w^{(j)}$, image of $w \in G$, belongs to $G^{(j)}$, then $w^{(j)} = f_j w$.

Let us remark that the construction presented in the former Section 1.2 is a particular case of the latter, once we take $f_1 = \dots = f_s = f$; we nevertheless decided for a sake of clarity, to present it before, because the computations involved in this latter general construction could have hidden the simplicity of the underlying idea. We hereby present some results for the non–homogeneous "Cantor dust"WFN (see Fig. 1.7).

Using the recursiveness of the algorithm we can, once again, completely characterize the topology of the non-homogeneous WFN, moreover only the weighted quantities will vary with respect to the homogeneous case. For instance, a straightforward, but cumbersome, generalization of the computations presented in the previous Sections allows us to prove that

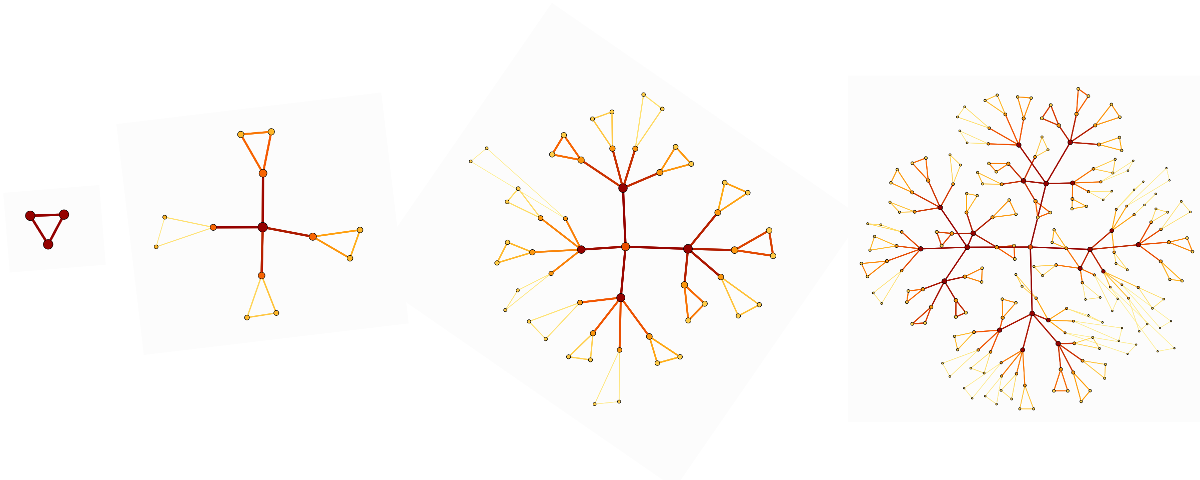

Figure 1.7: The non–homogeneous "Cantor dust" WFN, $s = 4$, $f_1 = 1/2$, $f_2 = 1/3$, $f_3 = 1/5$, $f_4 = 1/7$ and G_0 is formed by a triangle. From the left to the right G_0, G_1, G_2 and G_3. Gray scale (color on-line) reproduces edges weights: the darker the color the larger the weight. Visualization was done using Himmeli software [38].

the average weighted shortest path exhibits the following asymptotic behavior (see Fig. 1.5)

$$\lambda_k \underset{k\to\infty}{\longrightarrow} \frac{2s^2(s-1)}{(s-F)(s^2-F)}\,, \tag{1.14}$$

where $F = f_1 + \cdots + f_s$. Let us observe that Eq. (1.14) reduces to Eq. (1.10) once we set $f_1 = \cdots = f_s = f$ and thus $F = sf$.

Let $g_0(x)$ denote the number of nodes with node strength equal to x in the initial network G_0; then after k steps of the algorithm, all nodes strengths will be rescaled by a factor $f_1^{k_1} \ldots f_s^{k_s}$, where the non-negative integers k_i do satisfy $k_1 + \cdots + k_s = k$. Because this can be done in $k!/(k_1! \ldots k_s!)$ possible different ways, we get the following relation for the node strength distribution for the network G_k:

$$g_k(f_1^{k_1} \ldots f_s^{k_s} x) = \frac{k!}{k_1! \ldots k_s!} g_0(x) \quad \text{with } k_1 + \cdots + k_s = k\,. \tag{1.15}$$

After sufficiently many steps and assuming that the main contribution arises from the choice $k_1 \sim \cdots \sim k_s \sim k/s$, we can use Stirling formula to get the approximate distribution (see

Fig. 1.4)

$$\log g_k(x) \sim \frac{s \log s}{\log(f_1 \dots f_s)} \log x \,, \tag{1.16}$$

so once again the nodes strength distribution follows a power law.

1.5 Conclusions

In this paper we introduced a unifying framework for complex networks sharing several properties with fractal sets, hereby named *Weighted Fractal Networks*. This theory, that generalizes to graphs the construction of IFS fractals, allows us to build complex networks with a prescribed topology, whose main quantities can be analytically predicted and have been shown to depend on the fractal dimension of the IFS fractal; for instance the networks are scale–free with exponent the fractal dimension. Moreover the weighted fractal networks share with IFS fractals, the self-similarity structure, and are explicitly computable examples of renormalizable complex networks.

These networks exhibit the *small–world* property. In fact the average shortest path increases logarithmically with the system size (1.11), hence it is small as the average shortest path of a random network with the same number of nodes and same average degree. On the other hand the clustering coefficient is asymptotically constant (1.12), thus larger than the clustering coefficient of a random network that shrinks to zero as the system size increases.

The self-similarity property of the weighted fractal networks makes them suitable to model real problems involving generic diffusion over the network coupled with local looses of flow, here modeled via the parameter $f < 1$. For instance electrical grids in the case of mankind artifacts, metabolic networks of living organisms (West et al., 1997; Banavar et al., 1999) or air flows in mammalian lungs (Suki et al., 1994; Barabási et al., 1996; Kitaoka and Suki, 1997; Andrade et al., 1998; Kamiya et al., 2007; and Almeida et al. 1999) in the case of natural networks. In all these studies, assuming some hierarchical fractal structure, scientists could explain some natural laws, such as the allometric scaling or the avalanches in the mammalian respiratory system, although the actual branching numbers could not be explained. Using our framework we suggest a possible analysis for these values; for instance in mammalian lungs where air flows through bronchi–bronchioles, submitted to air vessels' section reduction. Assuming that the recorded avalanches and power–laws, hence the fractal structure, have some functional–biological reason, e.g. a typical time needed to fulfill alveoli with oxygenated air from the primary bronchi, we can relate this time to the average shortest weighted path, which in turn depends of the number of copies s and the scaling factor f. In other words the topology of such networks could have been shaped by evolution in such a way any two nodes can be connected in a finite optimal time, whatever their physical distance is.

Let us finally observe that in this case the fragility of the network with respect to failure of the edges, see Moreira (2009) and references therein, is a main biological question, that can be restated as follows: how many air vessels should fail before the lung stops to correctly operate ?

1.6 Appendix - Complementary material

Computation of $\Lambda_k^{(a_k)}$

Let a_k be the attaching node of the graph G_k. Let us define $\Lambda_k^{(a_k)} = \sum_{i \in G_k} p_{i a_k}^{(k)}$, i.e. the sum of all weighted shortest paths to a_k. Then using the recursive property and the symmetry of the map $\mathcal{T}_{s,f,a_k}$ we can easily obtain a recursive relation for $\Lambda_k^{(a_k)}$:

$$\Lambda_k^{(a_k)} = sf\Lambda_{k-1}^{(a_{k-1})} + sN_{k-1}\,,$$

where N_{k-1} is the number of nodes in G_{k-1}. This recursion can be easily solved to get for all $k \geq 1$

$$\Lambda_k^{(a_k)} = (sf)^{k-1}\Lambda_1^{(a_1)} + \frac{1-f^{k-1}}{1-f}\frac{(s-1)N_0+1}{s-1}s^{k-1} - \frac{s}{s-1}\frac{(sf)^{k-1}-1}{sf-1}\,, \tag{1.17}$$

from which we can conclude, because $f < 1$, that $\Lambda_k^{(a_k)}$ exhibits the asymptotic behavior given by equation (1.8).

Computation of Λ_k

Starting from the definition of the sum of all weighted shortest paths (1.7), the recursive construction and its symmetry we can decompose the sum Λ_k into three terms:

$$\Lambda_k = s \sum_{ij \in G_k^{(1)}} p_{ij}^{(k)} + s(s-1) \sum_{i \in G_k^{(1)}, j \in G_k^{(2)}} p_{ij}^{(k)} + 2s \sum_{i \in G_k^{(1)}} p_{ia_k}^{(k)} \tag{1.18}$$

where the first contribution takes into account all paths starting from and arriving to nodes belonging to the same subgraph, that using the symmetry can be chosen to be $G_k^{(1)}$. The second term takes into account all the possible paths where the initial point and the final one belong to two different subgraphs, and still using the symmetry we can set them to $G_k^{(1)}$ and $G_k^{(2)}$ and multiply the contribution by a combinatorial factor $s(s-1)$. Finally the last term is the sum of all paths arriving to the attaching node a_k; once again the symmetry allows us to reduce the sum to only one subgraph, say $G_k^{(1)}$, and multiply the contribution by $2s$.

Using the scaling mechanism for the edges, the first term in the right hand side of equation (1.18) can be easily identified with

$$\sum_{ij \in G_k^{(1)}} p_{ij}^{(k)} = f \Lambda_{k-1} \, .$$

By construction, each shortest path connecting two nodes belonging to two different subgraphs, must pass through the attaching node, hence using $p_{ij}^{(k)} = p_{ia_k}^{(k)} + p_{a_k j}^{(k)}$ the second term of equation (1.18) can be split into two parts:

$$\sum_{i \in G_k^{(1)}, j \in G_k^{(2)}} p_{ij}^{(k)} = \sum_{i \in G_k^{(1)}} p_{ia_k}^{(k)} N_k^{(2)} + \sum_{j \in G_k^{(2)}} p_{a_k j}^{(k)} N_k^{(1)} \, ,$$

where $N_k^{(i)}$ denotes the number of nodes in the subgraph $G_k^{(i)}$. Using the symmetry of the construction, the previous relation can be rewritten as

$$\sum_{i \in G_k^{(1)}, j \in G_k^{(2)}} p_{ij}^{(k)} = 2N_k^{(1)} \sum_{i \in G_k^{(1)}} p_{ia_k}^{(k)} .$$

The last term of equation (1.18) can be related to $\Lambda_{k-1}^{(a_{k-1})}$ by observing that each path arriving at a_k must pass through $a_k^{(i)}$ for some $i \in \{1, \ldots, s\}$, thus

$$\begin{aligned} \sum_{i \in G_k^{(1)}} p_{ia_k}^{(k)} &= \sum_{i \in G_k^{(1)}} (p_{ia_k^{(1)}}^{(k)} + p_{a_k^{(1)} a_k}^{(k)}) = N_k^{(1)} + \sum_{i \in G_k^{(1)}} p_{ia_k^{(1)}}^{(k)} \\ &= N_k^{(1)} + f\Lambda_{k-1}^{(a_{k-1})} , \end{aligned} \tag{1.19}$$

Observing that $G_k^{(1)}$ has as many nodes as G_{k-1} we can conclude that $N_k^{(1)} = N_{k-1}$ and finally to rewrite equation (1.18) as:

$$\Lambda_k = sf\Lambda_{k-1} + 2s[(s-1)N_{k-1} + 1][N_{k-1} + f\Lambda_{k-1}^{(a_{k-1})}] .$$

Chapter 2

Endogenous Social Networks under Opinion Exchange

2.1 Introduction

Modeling social phenomena represents a major challenge that has in recent years attracted a growing interest. Insight into the problem can be gained by resorting, among others, to the so called *Agent Based Models*, an approach that is well suited to bridge the gap between hypotheses concerning the microscopic behavior of individual agents and the emergence of collective phenomena in a population composed of many interacting heterogeneous entities.

Constructing sound models deputed to return a reasonable approximation of the scrutinized dynamics is a delicate operation, given the degree of arbitrariness in assigning the rules that govern mutual interactions. In the vast majority of cases, data are scarce and do not sufficiently constrain the model, hence the provided answers can be questionable. Despite this intrinsic limitation, it is however important to inspect the emerging dynamical properties of abstract models, formulated so to incorporate the main distinctive traits of a social interaction scheme. In this paper we aim at discussing one of such models, by combining analytical and numerical techniques. In particular, we will focus on characterizing the evolution of the underlying social network in terms of dynamical indicators.

It is nowadays well accepted that several social groups display two main features: the *small world property* (Watts and Strogatz, 1998) and the presence of *weak ties* (Granovetter, 1983). The first property implies that the network exhibits clear tendency to organize in densely connected clusters. As an example, the probability that two friends of mine are also, and independently, friends to each other is large. Moreover, the shortest path between two generic individuals is small as compared to the analogous distance computed for a random network made of the same number of individuals and inter-links connections. This observation signals the existence of short cuts in the social tissue. The second property is related to the cohesion of the group which is mediated by small groups of well tied elements, that are conversely weakly connected to other groups. The skeleton of a social community is hence a hierarchy of subgroups.

A natural question arise on the ubiquity of the aforementioned peculiar aspects, distinctive traits of a real social networks: can they eventually emerge, starting from a finite group of initially randomly connected actors? We here provide an answer to this question in the framework of a minimalistic opinion dynamics model, which exploit an underlying substrate where opinions can flow. More specifically, the network that defines the topological structure is imagined to evolve, coupled to the opinions and following a specific set of rules: once two agents reach a compromise and share a common opinion, they also increase their mutual degree of acquaintance, so strengthening the reciprocal link. In this respect, the model that we are shortly going to introduce hypothesize a co-evolution of opinions and social structure, in the spirit of a genuine adaptive network (Gross and Blasius, 2008 and Zimmermann et al., 2004).

Working within this framework, we will show that an initially generated random group, with respect to both opinion and social ties, can evolve towards a final state where small worlds and weak ties effects are indeed present. The results of this paper constitute the natural follow up of a series of papers (Bagnoli et al., 2007; Carletti et al., 2008a, 2008b), where the time evolution of the opinions and affinity, together with the fragmentation vs. polarization phenomena, have been discussed.

Different continuous opinion dynamics models have been presented in literature, see for instance Deffuant et al. (2000), Galam (2008) and Castellano et al. (2009), dealing with the general consensus problem. The aim is to shed light onto the assumptions that can eventually yield to fixation, a final mono-clustered configuration where all agents share the same belief, starting from an initial condition where the inspected population is instead fragmented into several groups. In doing so, and in most cases, a fixed network of interactions is a priori imposed (as in Amblard and Deffuant, 2004), and the polarization dynamics studied under the constraint of the imposed topology. At variance, and as previously remarked, we will instead allow the underlying network to dynamically adjust in time, so modifying its initially imposed characteristics. Let us start by revisiting the main ingredients of the model. A more detailed account can be found in Bagnoli et al. (2007).

Consider a closed group of N agents, each one possessing its own opinion on a given subject. We here represent the opinion of element i as a continuous real variable $O_i \in [0,1]$. Each agent is also characterized by its affinity score with respect to the remaining $N-1$ agents, namely a vector α_{ij}, whose entries are real number defined in the interval $[0,1]$: the larger the value of the affinity α_{ij}, the more reliable the relation of i with the end node j.

Both opinion and affinity evolve in time because of binary encounters between agents. It is likely that more interactions can potentially occur among individuals that are more affine, as defined by the preceding indicator, or that share a close opinion on a debated subject. Mathematically, these requirements can be accommodated for by favoring the encounters between agents that minimizes a *social metric* , as defined below. More concretely, select at random, with uniform probability, the agent i and quantify its *social distance* with the other members of the community: this is the $N-1$ vector $d_{ij} = |\Delta O^t_{ij}|(1-\alpha^t_{ij})$, where $\Delta O^t_{ij} = O^t_i - O^t_j$ is the opinions' difference of agents i and j at time t. The smaller the value of d^t_{ij} the closer the agent j to i, both in term of affinity and opinion. Mutual affinity can in fact mitigate the difference in opinion, thus determining the degree of social similarity of two individuals, an observation that inspires the proposed definition of d^t_{ij}. A Gaussian random perturbation $\mathcal{N}_j(0,\sigma)$ (mean zero and variance σ) is added to d^t_{ij} so to mimic the impact of a social mixing effect, the obtained vector $D^t_{ij} = d^t_{ij} + \mathcal{N}_j(0,\sigma)$ being the social metric. The second agent j for the paired interaction is the one closer to i with respect to D^t_{ij}. For a more detailed analysis on the interpretation of σ as a *social temperature* responsible of a increased mixing ability of the population, we refer to Bagnoli et al. (2007), Carletti et al. 2008a, 2008b). Let us observe that other models, see for instance Schweitzer and Hołys (2000), Bordogna and Albano (2007) and Klimek et al. (2008), make use of the social temperature concept: beyond the specificity of each the formulation, the social temperature is always invoked to control the degree of mixing in the population.

Once two agents are selected for interaction they possibly update their opinions (if they are affine enough) and/or change their affinities (if they have close enough opinions), following:

$$\begin{cases} O_i^{t+1} & = O_i^t - \frac{1}{2}\,\Delta O_{ij}^t\,\Gamma_1\left(\alpha_{ij}^t\right) \\ \alpha_{ij}^{t+1} & = \alpha_{ij}^t + \alpha_{ij}^t(1-\alpha_{ij}^t)\,\Gamma_2\left(\Delta O_{ij}^t\right)\,, \end{cases} \tag{2.1}$$

being:

$$\Gamma_1\left(x\right) = \frac{\tanh(\beta_1(x-\alpha_c))+1}{2} \quad \text{and} \quad \Gamma_2\left(x\right) = -\tanh(\beta_2(|x|-\Delta O_c))\,, \tag{2.2}$$

two *activating functions* which formally reduce to step functions for large enough values of the parameters β_1 and β_2, as it is the case in the numerical simulations reported below.

Let us briefly comment on the mathematical construction of the model. Suppose two subjects meet and imagine they challenge their respective opinions, assumed to be divergent, i.e. $|\Delta O_{ij}| \simeq 1$. According to the bounded confidence assumption, see for instance Deffuant et al. (2000), when the disagreement falls beyond a given threshold, the agents stick to their positions. As opposed to this simplistic view, in the present case, and as follows a punctual interaction, the agents can still modify each other beliefs, provided the mutual affinity α_{ij}^t is larger than the reference value α_c. This scenario accounts for a plausible strategy that individual can adopt when processing a contradictory information: if $\alpha_{ij}^t < \alpha_c$, the agent ignores the dissonating input, which is therefore not assimilated; Conversely, when the opinion comes from a trustable source ($\alpha_{ij}^t > \alpha_c$) the agent is naturally inclined to restore the consistency among the cognitions, and thus adjust its belief. The scalar quantity α_{ij} schematically accounts for a large number of hidden variables (personality, attitudes, behaviors,..), all here integrated in the abstract affinity concept. Similarly each affinity entry evolves in a self-consistent fashion, as guided by the individual dynamics. When two subjects gather together and discover to share common interests, $|\Delta O_{ij}^t| < \Delta O_c$, they increase their mutual affinity score, $\alpha_{ij}^t \to 1$. The opposite ($\alpha_{ij}^t \to 0$) holds if $|\Delta O_{ij}^t| > \Delta O_c$. The logistic contribution in Eqs. (2.1) confines α_{ij}^t in the interval $[0, 1]$, while it maximizes the change in affinity for pairs with $\alpha_{ij}^t \simeq 0.5$. Pairs of individuals with $\alpha_{ij}^t \simeq 1$ (resp. 0) have already formed their mind and so can expected to behave more conservatively.

Despite its simplicity the model exhibits an highly non linear dependence on the involved parameters, α_c, ΔO_c and σ, with a phase transition between a polarized and fragmented dynamics (Bagnoli et al., 2007).

A typical run for $N = 100$ agents is reported in the main panel of Fig. 2.1, for a choice of the parameters which yields to a consensus state. The insets represent three successive time snapshots of the underlying social network: The N nodes are the individuals, while the links are assigned based on the associated values of the affinity. The figures respectively refer to a relatively early stage of the evolution $t = 1000$, to an intermediate time $t = 5000$ and to the convergence time $T_c = 10763$. Time is here calculated as the number of iterations (not normalized with respect to N). The corresponding three networks can be characterized using standard topological indicators [1]. (see Table 2.1), e.g. the mean degree $\langle k \rangle$, the network clustering coefficient C and the average shortest path $\langle \ell \rangle$. An explicit definition of those quantities will be given below.

In the forthcoming discussion we will focus on the evolution of the network topology, limited to a choice of the parameters that yield to a final mono cluster.

	$\langle k \rangle(t)$	$C(t)$	$\langle \ell \rangle(t)$
$t = 1000$	0.073	0.120	3.292
$t = 5000$	0.244	0.337	2.013
$t = T_c$	0.772	0.594	1.228

Table 2.1: Topological indicators of the social networks presented in Fig. 2.1. The mean degree $\langle k \rangle$, the network clustering C and the average shortest path $\langle \ell \rangle$ are reported for the three time configurations depicted in the figure.

Before proceeding further let us anticipate the main results of this paper so to mark the differences with previous analysis. On the one hand, we will provide an analytical solution for the dynamical evolution of the average network properties (mean and variance). In

[1]See Albert and Barabási (2002) and Boccaletti et al. (2006)

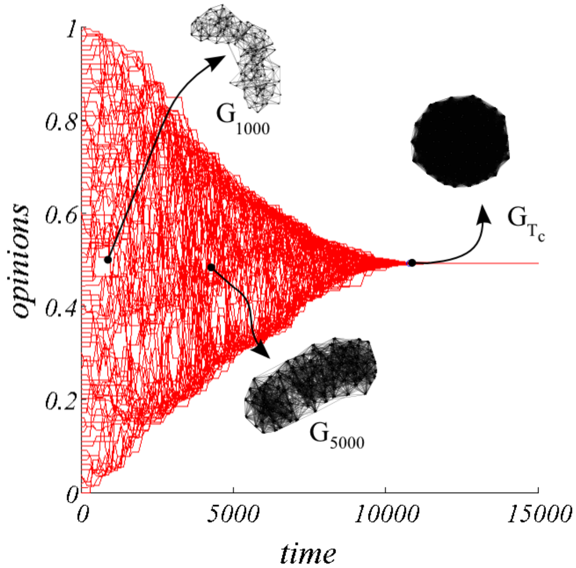

Figure 2.1: Opinions as function of time. The run refers to $\alpha_c = 0.5$, $\Delta O_c = 0.5$ and $\sigma = 0.01$. The underlying network is displayed at different times, testifying on its natural tendency to evolve towards a single cluster of affine individuals. Initial opinions are uniformly distributed in the interval $[0, 1]$, while α_{ij}^0 are randomly assigned in $[0, 1/2]$ with uniform distribution.

this respect we shall clearly expand over previous investigation (see Carletti et al., 2008) where a closure for the equations of the moments was imposed by neglecting the variance contribution. On the other hand, we will characterize in depth the network properties by computing and monitoring numerically a large set of topological indicators.

2.2 The social network

The affinity matrix drives the interaction via the selection mechanism. It hence can be interpreted as the *adjacency* matrix of the underlying *social network*, i.e. the network of social ties that influences the exchange of opinions between acquaintances, as mediated by the encounters. Because the affinity is a dynamical variable of the model, we are actually focusing on an *adaptive* social network (Gross and Blasius, 2008; Zimmermann et al., 2004)

: The network topology influences in turn the dynamics of opinions, this latter providing a feedback on the network itself and so modifying its topology. In other words, the evolution of the topology is inherent to the dynamics of the model because of the proposed self-consistent formulation and not imposed as an additional stochastic ingredient, as e.g. rewire and/or add/remove links according to a given probability (see Holme and Newman, 2006 and Kozma and Barrat, 2008) once the state variables have been updated. It is the inherent dynamics of the system (which includes the noise source that we accommodated for) which governs the network evolution, the links being not assigned on the basis of a pure stochastic mechanisms. Moreover, in our model, at a given time t all possible pairs have a finite chance of interaction, as opposed to Holme and Newman (2006) and Kozma and Barrat (2008), where the interaction is instead dictated by the existing links.

Remark 1 (Weighted network)**.** *Let us observe that the affinity assumes positive real values, hence we can consider a weighted social networks, where agents weigh the relationships. Alternatively, one can introduce a cut-off parameter, α_f: agents i and j are socially linked if and only if the recorded relative affinity is large enough, meaning $\alpha_{ij} > \alpha_f$. Roughly, the agent chooses its closest friends among all his neighbors.*

The first approach avoids the introduction of non–smooth functions and it is suitable to carry on the analytical calculations. The latter results more straightforward for numerical oriented applications.

As anticipated, we are thus interested in analyzing the model, for a specific choice of the parameters, α_c, σ and ΔO_c, yielding to consensus, and studying the evolution of the network topology, here analyzed via standard network indicators: the average value of *weighted degree*, the *cluster coefficient* and the *averaged shortest path*. These quantities will be quantified for (i) a fixed population, monitoring their time dependence; (ii) as a function of the population size, photographing the dynamics at convergence, namely when consensus has been reached.

Time evolution of weighted degree

The simplest and the most intensively studied one–vertex (i.e. local) characteristic is the node *degree* [2]: the *total number of its connections* or its nearest neighbors. Because we are dealing with a weighted network we can also introduce the *normalized weighted node degree*, also called *node strength* (Barrat et al., 2004), namely $s_i(t) = \sum_j \alpha^t_{ij}/(N-1)$. Its mean value averaged over the whole network reads:

$$\langle s \rangle(t) = \frac{1}{N}\sum_{i=1}^{N} s_i(t)\,. \tag{2.3}$$

Let us observe that the normalization factor $N-1$ holds for a population of N agents, self-interaction being disregarded, $\langle s \rangle$ belongs hence to the interval $[0,1]$ and having eliminated the relic of the population size, one can properly compare quantities calculated for networks made of different number of agents.

All these quantities evolve in time because of the dynamics of the opinions and/or affinities. Passing to continuous time and using the second relation of (2.1), we obtain:

$$\frac{d}{dt}\langle s \rangle = \frac{1}{N(N-1)}\sum_{i,j=1}^{N} \frac{d}{dt}\alpha^t_{ij}\,. \tag{2.4}$$

Let us observe that the evolution of affinity and opinion can be decoupled when $\Delta O_c = 1$. For $\Delta O_c < 1$, this is not formally true. However on can argue for an approximated strategy (as in Carletti et al., 2008), by replacing the step function Γ_2 by its time average counterpart γ_2, where the dependence in ΔO^t_{ij} has been silenced. In this way, we obtain form (2.4)

$$\frac{d}{dt}\langle s \rangle = \frac{\gamma_2}{N(N-1)}\sum_{i,j=1}^{N} \alpha^t_{ij}(1-\alpha^t_{ij}) = \gamma_2(\langle s \rangle - \langle s^2 \rangle)\,, \tag{2.5}$$

where $\langle s^2 \rangle = \sum \alpha^2_{ij}/(N(N-1))$. Let us observe that γ_2 is of the order of $1/N^2$ times, a factor taking care of the asynchronous dynamics (see Carletti et al., 2008).

[2] Let us observe that the affinity may not be symmetric and thus the inspected social network will be directed. One has thus to distinguish between *In–degree*, k_{in}, being the number of *incoming edges* of a vertex and *Out–degree*, k_{out}, being the number of its *outgoing edges*. In the following we will be interested only in the outgoing degree, from here on simply referred to as to degree.

In Carletti and Fanelli (2008) authors proved that (2.5) can be analytically solved once we provide the initial distribution of node strengths (see 2.4 for a short discussion of the involved methods). Assuming $s_i(0)$ to be uniformly distributed in $[0, 1/2]$, we get the following exact solution (see Fig. 2.2):

$$\langle s\rangle(t) = \frac{e^{\gamma_2 t}}{e^{\gamma_2 t}-1} - \frac{2e^{\gamma_2 t}}{(e^{\gamma_2 t}-1)^2}\log\left(\frac{e^{\gamma_2 t}+1}{2}\right) , \tag{2.6}$$

Using similar ideas we can prove (see Carletti and Fanelli, 2008) that the variance $\sigma_s^2(t) = \langle s^2\rangle - \langle s\rangle^2$ is analytically given by

$$\sigma^2(t) = \frac{2e^{2\gamma_2 t}}{(e^{\gamma_2 t}-1)^2(e^{\gamma_2 t}+1)} - \frac{4e^{2\gamma_2 t}}{(e^{\gamma_2 t}-1)^4}\left[\log\left(\frac{e^{\gamma_2 t}+1}{2}\right)\right]^2 . \tag{2.7}$$

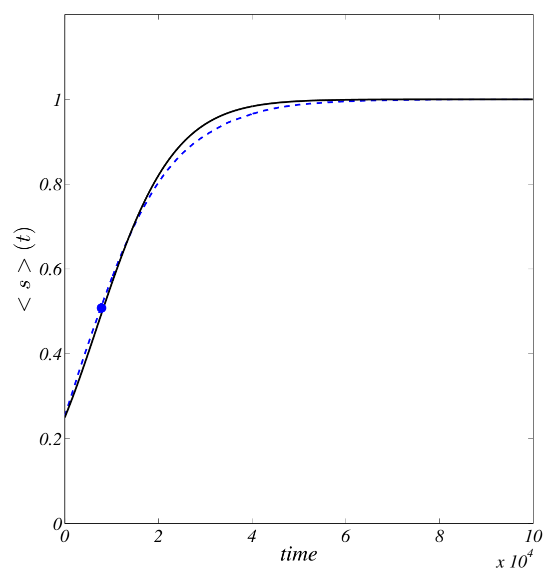

Figure 2.2: Evolution of $\langle s\rangle(t)$. Dashed line (blue on-line) refers to numerical simulations with parameters $\alpha_c = 0.5$, $\Delta O_c = 0.5$ and $\sigma = 0.3$. The full line (black on-line) is the analytical solution (2.6) with a best fitted parameter $\gamma_2 = 1.6\,10^{-4}$. The dot denotes the convergence time in the opinion space to the consensus state, for the used parameters affinities did not yet converge. Let us observe in fact that affinities and opinions do converge on different time scale (see Carletti et al, 2008).

The comparison between analytical and numerical profiles is enclosed in Fig. 2.2, where the evolution of $\langle s\rangle(t)$ is traced. Let us observe that here γ_2 will serve as a fitting parameter, when testing the adequacy of the proposed analytical curves versus direct simulations, instead of using its computed numerical value (see Carletti et al, 2008). The qualitative correspondence is rather satisfying, in accordance with the analytical results.

Assume T_c to label the time needed for the consensus in opinion space to be reached. Clearly, T_c depends on the size of the simulated system [3]. From the above relation (2.6), the average node strength at convergence as an implicit function of the population size N reads:

$$\langle s\rangle(T_c(N)) = \frac{e^{\gamma_2(N)T_c(N)}}{e^{\gamma_2(N)T_c(N)}-1} - \frac{2e^{\gamma_2(N)T_c(N)}}{(e^{\gamma_2(N)T_c(N)}-1)^2}\log\left(\frac{e^{\gamma_2(N)T_c(N)}+1}{2}\right), \qquad (2.8)$$

where we emphasized the dependence of γ_2 and of T_c on N. However, as already observed $\gamma_2(N) = \mathcal{O}\left(N^{-2}\right)$ and $T_c(N) = \mathcal{O}\left(N^a\right)$, with $a \in (1,2)$. Hence $\gamma_2(N)T_c(N) \to 0$ when $N \to \infty$ and thus $\langle s\rangle(T_c(N))$ is predicted to be a decreasing function of the population size N, which converges to the asymptotic value $1/4$, the initial average node strength (see Fig. 2.3), given the selected initial condition. In sociological terms this means that even when consensus is achieved the larger the group the smaller, on average, the number of local acquaintances. This is a second conclusion that one can reach on the basis of the above analytical developments.

Small world

Several social networks exhibit the remarkable property that one can reach an arbitrary far member of the community, via a relatively small number of intermediate acquaintances. This holds true irrespectively of the size of the underlying network. Experiments (see Milgram and Traver, 1969) have been devised to quantify the “degree of separation” in real system,

[3] In Bagnoli et al. (2007) and Carletti et al. (2006) it was shown that T_c scales faster than linearly but slower than quadratically with respect to the population size N.

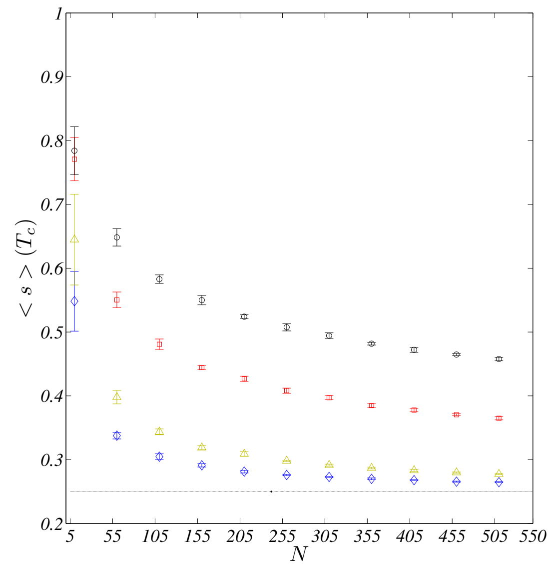

Figure 2.3: Average node strength at convergence as a function of the population size. Parameters are $\Delta O_c = 0.5$, $\sigma = 0.5$ and four values of α_c have been used : ($\diamond$) $\alpha_c = 0$, ($\triangle$) $\alpha_c = 0.25$, ($\square$) $\alpha_c = 0.5$ and ($\bigcirc$) $\alpha_c = 0.75$. Vertical bars are standard deviations computed over 10 replicas of the numerical simulation using the same initial conditions.

and such phenomenon is nowadays termed the “small world”effect, also referred to as the “six degree of separation”.

On the other hand several, models have been proposed (prominently Watts and Strogatz, 1998 and Newman and Watts, 1999) to construct complex networks with the small world property. Mathematically, one requires that the average shortest path grows at most logarithmic with respect to the network size, while the network still displays a large clustering coefficient. Namely, the network has an average shortest path comparable to that of a random network, with the same number of nodes and links, while the clustering coefficient is instead significantly larger.

In this section we present numerical results aimed at describing the time evolution of both the average shortest path and the clustering coefficient of the social network emerging from the model. As before, the parameters are set so to induce the convergence to a consensus state in the opinion space.

We will be particularly interested in their values at consensus, terming the associated values respectively $\langle \ell \rangle (T_c)$ and $C(T_c)$ once the consensus state has been achieved.

In Fig. 2.4 we report these quantities (normalized to the homologous values estimated for a random network with identical number of nodes and links) versus the system size, computed using the adjacency matrix obtained by binarizing the affinity matrix as prescribed in Remark 1 using a value of α_f equal to 0.5 . The (normalized) clustering coefficient is sensibly larger than one, this effect being more pronounced the smaller the value of α_c. On the other hand the (normalized) average shortest path is always very close to 1.

Based on the above we are hence brought to conclude that the social network emerging from the opinion exchanges, has the small world property. This is a remarkable feature because the social network evolves guided by the opinions and not result from an artificially imposed recipe. The implications of these findings on real social networks deserve to be further and deeply analyzed.

Weak ties

Social networks are characterized by the presence of hierarchies of well tied small groups of acquaintances, that are possibly linked to other such groups via "weak ties"(Csermely, 2009). According to Granovetter (1983) these weak links are fundamental for the cohesion of the society, being at the basis of the social tissue, so motivating the statement "the strength of weak ties". Such phenomenon has been already shown to be relevant in social technological networks by Onnela et al. (2007).

In general any structured social tissue, can be generically decomposed into communities (internally highly connected subgroups of affine individuals) weakly linked together, generically called "small and closed social circles"or "acquaintances"versus "close friends"by Granovetter (1983). To quantify these concepts, in the following, we invoke a strong working

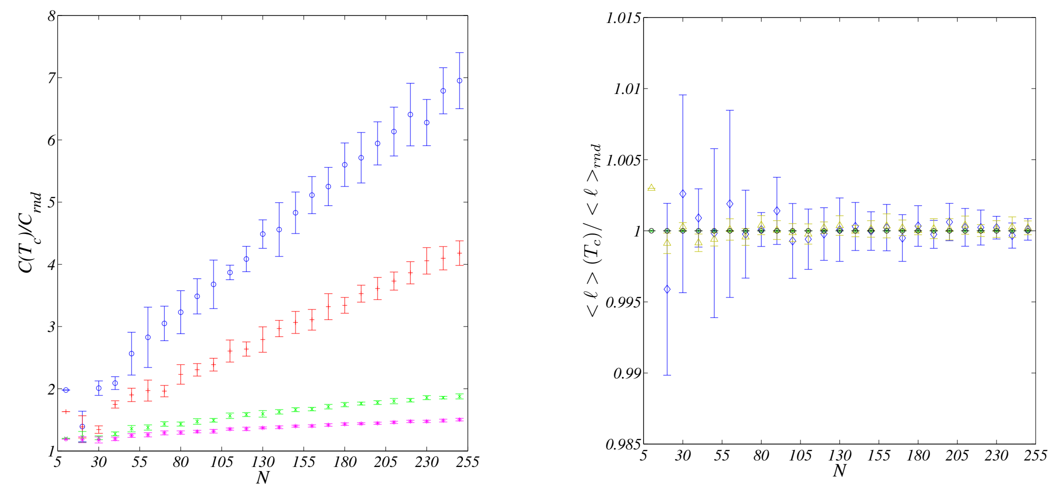

Figure 2.4: Normalized clustering coefficient (left panel) and normalized average mean path (right panel) as functions of the network size at the convergence time. Parameters are $\Delta O_c = 0.5$, $\sigma = 0.5$ and four values of α_c have been used : ($\diamond$) $\alpha_c = 0$, ($\triangle$) $\alpha_c = 0.25$, ($\square$) $\alpha_c = 0.5$ and ($\bigcirc$) $\alpha_c = 0.75$. Vertical bars are standard deviations computed over 10 repetitions. The adjacency matrix has been obtained from the affinity matrix using $\alpha_f = 0.5$.

hypothesis by requiring that agents belonging to each small and closed subgroup of size m are indeed all linked together, thus defining a clique (see Albert and Barabási, 2002) of size m, hereafter simply m-clique. A weaker request consists in assuming that just a subset of all possible links are active. This alternative choice implies dealing with communities (see Fortunato, 2010) rather than m–cliques, as it is instead the case in the following.

The degree of cliqueness of a social network is hence a measure of its cohesion/fragmentation: the presence of a large number of m-cliques together with very few, m'-cliques, for $m' > m$, means that the population is actually fragmented into small pieces, of size m weakly interacting with each other (Berlow,1999).

We are interested in studying such phenomenon within the social network emerging from the opinion dynamics model here considered, still operating in consensus regime. To this end we proceed as follows. We introduce a cut–off parameter α_f used to binarize the affinity matrix, which hence transforms into an adjacency matrix a. More precisely, agents i and j will be connected, i.e. $a_{ij} = 1$, if and only if $\alpha_{ij} \geq \alpha_f$. Once the adjacency matrix is being constructed, we compute the number of m–cliques in the network; more precisely, because a m'–clique contains several m–cliques, with $m' > m$, to have a precise information about the network topology, we count only maximal m–cliques, i.e. those not contained in any m'–cliques with $m' > m$. Let us observe that this last step is highly time consuming, the clique problem being NP-complete. We thus restrict our analysis to the cases $m \in \{3, 4, 5\}$.

For small values of α_f the network is almost complete, while for large ones it can in principle fragment into a vast number of finite small groups of agents. As reported in the inset of Fig. 2.5, for $\alpha_f \sim 1$ only 3–cliques are present. Their number rapidly increases as α_f is lowered. On the other hand for $\alpha_f \sim 0.98$ few 4–cliques emerge while 5–cliques appear around $\alpha_f \sim 0.73$. This means that the social networks is mainly composed by 3–cliques, i.e. agents sharing high mutual affinities, that are connected together to form larger cliques, for instance 4 and 5–cliques by weaker links, i.e. whose mutual affinities are lower than the above ones. The network has thus acquired some non-trivial topology, starting from a random one.

To critically examine our conclusions we compare our findings to that obtained for a random network. This latter is made of a number of nodes and links identical to that of the binarized social network. The social network displays many more cliques as compared to the random reference case: the relative number of cliques goes from tens, for small α_f amounts to, hundreds for larger α_f values (data not shown). On the other hand 4 and 5–cliques are almost absent in the random network.

If we increase the value of σ, i.e. the social mixing effect, then we can show (results not reported) that the weak ties phenomenon is prevented to occur. This is an interesting point that will deserve future investigations, benchmarked to the relevant sociological literature.

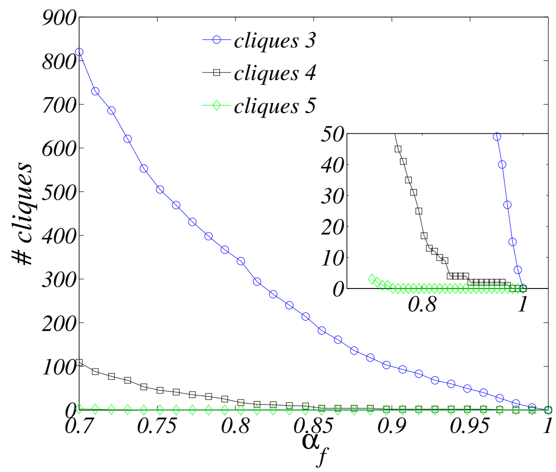

Figure 2.5: Number of maximal 3, 4 and 5–cliques in the social network once consensus has been achieved. Parameters are $N = 100$, $\Delta O_c = 0.5$, $\alpha_c = 0.5$ and $\sigma = 0.5$.

2.3 Conclusion

Social system and opinion dynamics models are intensively investigated within simplified mathematical schemes. One of such model is here revisited and analyzed. The evolution of the underlying network of connections, here emblematized by the mutual affinity score, is in particular studied. This is a dynamical quantity which adjusts all along the system evolution, as follows a complex coupling with the opinion variables. In other words, the embedding social structure is adaptively created and not a priori assigned, as it is customarily done. Starting from this setting, the model is solved analytically, under specific approximations. The functional dependence on time of the networks mean characteristics are consequently elucidated. The obtained solutions correlate with direct simulations, returning a satisfying agreement. Moreover, the structure of the social network is numerically monitored, via a set of classical indicators. Small world effect, as well weak ties connections, are found as an emerging property of the model, contrary to other opinion dynamics models on non-trivial network topology, for instance Kozma and Barrat (2008) and Holme and Newman (2006).

We in fact repeated the present analysis for the case of the bounded confidence model of Deffuant et al. (2000) and could not detected emerging topological structures as seen in the affinity model, namely small world and weak ties phenomena. It could be speculated that the richness of the model here inspected stems from its embedding dimensionality: opinion and affinity evolve self–consistently in a two dimensional space, while the classical formulation á la Deffuant et al. (2000) is limited to a one dimensional setting.

It is remarkable that such properties, ubiquitous in real life social networks, are spontaneously generated within a simple scenario which accounts for a minimal number of ingredients, in the context of a genuine self-consistent formulation.

2.4 Appendix - On the momenta evolution

The aim of this section if to present and sketch the proof of the result used to study the evolution of the momenta of the affinity distribution. We refer the interested reader to Carletti and Fanelli (2008) where a more detailed analysis is presented in a general setting.

For the sake of simplicity, let us label the $N(N-1)$ affinities values α_{ij} by x_l, upon assigning a specific re-ordering of the entries. Hence $\vec{x}$ is a vector with $M = N(N-1)$ elements. As previously recalled (2.5), we assume each x_l to obey a first order differential equation of the logistic type, once time has been rescaled, namely:

$$\frac{dx_l}{dt} = x_l(1 - x_l)\,. \tag{2.9}$$

The initial conditions will be denoted as x_l^0.

Let us observe that each component x_l evolves independently from the others. We can hence imagine to deal with M replicas of a process ruled by (2.9) whose initial conditions are distributed according to some given function. We are interested in computing the momenta of the x variable as functions of time and depending on the initial distribution. The m-th momentum is given by:

$$\langle x^m \rangle(t) = \frac{(x_1(t))^m + \cdots + (x_M(t))^m}{M}\,, \tag{2.10}$$

and its time evolution is straightforwardly obtained deriving (2.10) and making use of Eq. (2.9):

$$\begin{aligned} \frac{d}{dt}\langle x^m \rangle(t) &= \frac{1}{M}\sum_{i=1}^{M}\frac{dx_l^m}{dt} = \frac{m}{M}\sum_{l=1}^{N}x_l^{m-1}\frac{dx_l}{dt} \\ &= \frac{m}{M}\sum_{l=1}^{N}x_l^{m-1}x_l(1-x_l) = m\left(\langle x^m\rangle - \langle x^{m+1}\rangle\right)\,. \end{aligned} \tag{2.11}$$

To solve this equation we introduce the *time dependent moment generating function*, $G(\xi, t)$,

$$G(\xi,t) := \sum_{m=1}^{\infty}\xi^m\langle x^m\rangle(t)\,. \tag{2.12}$$

This is a formal power series whose Taylor coefficients are the momenta of the distribution that we are willing to reconstruct, task that can be accomplished using the following relation:

$$\langle x^m\rangle(t) := \frac{1}{m!}\frac{\partial^m G}{\partial\xi^m}\Big|_{\xi=0}\,. \tag{2.13}$$

By exploiting the evolution's law for each x_l, we shall here obtain a partial differential equation governing the behavior of G. Knowing G will eventually enable us to calculate any sought momentum via multiple differentiation with respect to ξ as stated in (2.13).

On the other hand, by differentiating (2.12) with respect to time, one obtains :

$$\frac{\partial G}{\partial t} = \sum_{m\geq 1}\xi^m\frac{d\langle x^m\rangle}{dt} = \sum_{m\geq 1}m\xi^m\left(\langle x^m\rangle - \langle x^{m+1}\rangle\right)\,, \tag{2.14}$$

where used has been made of Eq. (2.11). We can now re-order the terms so to express the right hand side as a function of G [4] and finally obtain the following non–homogeneous linear partial differential equation:

$$\partial_t G - (\xi - 1)\partial_\xi G = \frac{G}{\xi}\,. \tag{2.15}$$

[4]Here the following algebraic relations are being used:

$$\xi\partial_\xi G(\xi,t) = \xi\sum_{m\geq 1}m\xi^{m-1}\langle x^m\rangle = \sum_{m\geq 1}m\xi^m\langle x^m\rangle\,,$$

Such an equation can be solved for ξ close to zero (as in the end of the procedure we shall be interested in evaluating the derivatives at $\xi = 0$, see Eq. (2.13)) and for all positive t. To this end we shall specify the initial datum:

$$G(\xi, 0) = \sum_{m \geq 1} \xi^m \langle x^m \rangle(0) = \Phi(\xi)\,, \tag{2.16}$$

i.e. the initial momenta or their distribution.

Before turning to solve (2.15), we first simplify it by introducing

$$G = e^g \quad \text{namely} \quad g = \log G\,, \tag{2.17}$$

then for any derivative we have $\partial_* G = G \partial_* g$, where $* = \xi$ or $* = t$, thus (2.15) is equivalent to

$$\partial_t g - (\xi - 1)\partial_\xi g = \frac{1}{\xi}\,, \tag{2.18}$$

with the initial datum

$$g(\xi, 0) = \phi(\xi) \equiv \log \Phi(\xi)\,. \tag{2.19}$$

This latter equation can be solved using the *method of the characteristics*, here represented by:

$$\frac{d\xi}{dt} = -(\xi - 1)\,, \tag{2.20}$$

which are explicitly integrated to give:

$$\xi(t) = 1 + (\xi(0) - 1)e^{-t}\,, \tag{2.21}$$

---

and

$$\begin{aligned} \xi \partial_\xi \frac{G(\xi,t)}{\xi} &= \xi \partial_\xi \sum_{m\geq 1} \xi^{m-1} \langle x^m \rangle = \xi \sum_{m \geq 1} (m-1)\xi^{m-2} \langle x^m \rangle \\ &= \sum_{m\geq 1} (m-1)\xi^{m-1} \langle x^m \rangle \end{aligned}$$

Renaming the summation index, $m-1 \to m$, one finally gets (note the sum still begins with $m = 1$):

$$\xi \partial_\xi \frac{G(\xi,t)}{\xi} = \sum_{m \geq 1} m \xi^m \langle x^{m+1} \rangle\,.$$

where $\xi(0)$ denotes $\xi(t)$ at $t = 0$. Then the function $u(\xi(t), t)$ defined by:

$$u(\xi(t), t) := \phi(\xi(0)) + \int_0^t \frac{1}{1 + (\xi(0) - 1)e^{-s}}\, ds\,, \tag{2.22}$$

is the solution of (2.18), restricted to the characteristics. Observe that $u(\xi(0), 0) = \phi(\xi(0))$, so (2.22) solves also the initial value problem.

Finally the solution of (2.19) is obtained from u by reversing the relation between $\xi(t)$ and $\xi(0)$, i.e. $\xi(0) = (\xi(t) - 1)e^t + 1$:

$$g(\xi, t) = \phi\left((\xi - 1)e^t + 1\right) + \lambda(\xi, t)\,, \tag{2.23}$$

where $\lambda(\xi, t)$ is the value of the integral in the right hand side of (2.22).

This integral can be straightforwardly computed as follows (use the change of variable $z = e^{-s}$):

$$\lambda = \int_0^t \frac{1}{1 + (\xi(0) - 1)e^{-s}}\, ds = \int_1^{e^{-t}} \frac{-dz}{z} \frac{1}{1 + (\xi(0) - 1)z}\,, \tag{2.24}$$

which implies

$$\begin{aligned}\lambda &= -\int_1^{e^{-t}} dz \left(\frac{1}{z} - \frac{\xi(0) - 1}{1 + (\xi(0) - 1)z}\right) = -\log z + \log(1 + (\xi(0) - 1)z)\Big|_1^{e^{-t}} \\ &= t + \log(1 + (\xi(0) - 1)e^{-t}) - \log \xi(0)\,. \end{aligned} \tag{2.25}$$

According to (2.23) the solution g is then

$$g(\xi, t) = \phi\left((\xi - 1)e^t + 1\right) + t + \log \xi - \log((\xi - 1)e^t + 1)\,, \tag{2.26}$$

from which G straightforwardly follows:

$$G(\xi, t) = \Phi\left((\xi - 1)e^t + 1\right) \frac{\xi e^t}{(\xi - 1)e^t + 1}\,. \tag{2.27}$$

As anticipated, the function G makes it possible to estimate any momentum (2.13). As an example, the mean value correspond to setting $m = 1$, reads:

$$\begin{aligned}\langle x\rangle(t) &= \partial_\xi G\Big|_{\xi=0} = \Big[\Phi'\left(1 + (\xi - 1)e^t\right) e^t \frac{\xi e^t}{(\xi - 1)e^t + 1} \\ &+ \Phi\left(1 + (\xi - 1)e^t\right) e^t \frac{(\xi - 1)e^t + 1 - \xi e^t}{(1 + (\xi - 1)e^t)^2}\Big]\Big|_{\xi=0} \\ &= \frac{e^t}{1 - e^t}\Phi(1 - e^t)\,. \end{aligned} \tag{2.28}$$

In the following section we shall turn to considering a specific application in the case of uniformly distributed values of affinities.

Uniform distributed initial conditions

The initial data x_l^0 are assumed to span uniformly the bound interval $[0, 0.5]$, thus the probability distribution $\psi(x)$ clearly reads:

$$\psi(x) = \begin{cases} 2 & \text{if } x \in [0, 1/2] \\ 0 & \text{otherwise} \end{cases}, \tag{2.29}$$

and consequently the initial momenta are [5]:

$$\langle x^m \rangle(0) = \int_0^1 \xi^m \psi(\xi) d\xi = \int_0^{1/2} 2\xi^m \, d\xi = \frac{1}{m+1} \frac{1}{2^m}. \tag{2.30}$$

Hence the function Φ as defined in (2.16) takes the form:

$$\Phi(\xi) = \sum_{m \geq 1} \frac{1}{m+1} \frac{\xi^m}{2^m}. \tag{2.31}$$

A straightforward algebraic manipulation allows us to re-write (2.31) as follows:

$$\sum_{m \geq 1} \frac{y^m}{m+1} = \frac{1}{y} \int_0^y \sum_{m \geq 1} z^m \, dz = \frac{1}{y} \int_0^y \frac{z}{1-z} \, dz = -1 - \frac{1}{y} \log(1-y), \tag{2.32}$$

thus

$$\Phi(\xi) = -1 - \frac{2}{\xi} \log\left(1 - \frac{\xi}{2}\right). \tag{2.33}$$

We can now compute the time dependent moment generating function, $G(\xi, t)$, given by (2.27) as:

$$G(\xi, t) = \frac{\xi e^t}{(\xi-1)e^t + 1} \left[-1 - \frac{2}{(\xi-1)e^t + 1} \log\left(1 - \frac{(\xi-1)e^t + 1}{2}\right) \right], \tag{2.34}$$

[5] We hereby assume to sample over a large collection of independent replica of the system under scrutiny (M is large). Under this hypothesis one can safely adopt a continuous approximation for the distribution of allowed initial data. Conversely, if the number of realizations is small, finite size corrections need to be included(see Carletti and Fanelli, 2008).

and thus recalling (2.13) we get

$$\begin{aligned}
\langle x\rangle(t) &= \frac{e^t}{e^t-1} - \frac{2e^t}{(e^t-1)^2}\log\left(\frac{e^t+1}{2}\right) \\
\langle x^2\rangle(t) &= \frac{e^{2t}}{(e^t-1)^2} + \frac{4e^{2t}}{(e^t-1)^3}\log\left(\frac{e^t+1}{2}\right) + \frac{2e^{2t}}{(e^t-1)^2(e^t+1)}\,.
\end{aligned} \tag{2.35}$$

Let us observe that $\langle x\rangle(t)$ deviates from the logistic growth to which all the single variables $x_i(t)$ do obey.

For large enough times, the distribution of the variable outputs is in fact concentrated around the asymptotic value 1 with an associated variance (calculated from the above momenta) which decreases monotonously with time.

Let us observe that a naive approach would suggest interpolating the averaged numerical profile with a solution of the logistic model whose initial datum $\hat{x}^0$ acts as a free parameter to be adjusted to its best fitted value: as it is proven in Carletti and Fanelli (2008) this procedure yields a significant discrepancy, which could be possibly misinterpreted as a failure of the underlying logistic evolution law. For this reason, and to avoid drawing erroneous conclusions when ensemble averages are computed, attention has to be payed on the role of initial conditions.

Chapter 3

Informational Efficiency under Adaptive Expectations and Confirmatory Bias

3.1 Introduction

> "Investing in speculative assets is a social activity. Investors spend a substantial part of their leisure time discussing investments, reading about investments, or gossiping about others' successes or failures in investing. It is thus plausible that investors' behavior (and hence prices of speculative assets) would be influenced by social movements"
>
> Shiller (1984)

In most economic interactions, individuals possess only partial information about the value of exchanged objects. For instance, when a firm "goes public", i.e. launches an initial public offering of its shares, none of participants in financial market has complete information concerning the future value of the profit stream that the firm would generate. The fundamental question, going back to Hayek (1945), is then: To which extent the market can serve as the aggregator of this dispersed information? In other words, when is the financial market informationally efficient, meaning that the market price converges over time to the value that would obtain if all market participants had full information about the fundamental value of the asset exchanged?

Most studies that address this question build on the assumption that individual market participants are fully rational. Under full rationality, the seminal results on the informational efficiency of centralized markets were established by Grossman (1976), Wilson (1977), Milgrom (1981), whereas for decentralized markets they have been proved by Wolinsky (1990), Blouin and Serrano (2001), and Duffie and Manso (2007).

However, research in experimental economics and behavioral finance indicates that traders do not behave in the way consistent with the full-rationality assumption. For instance, Rabin and Schrag (1999) discuss the evidence that individuals suffer from the so-called confirmatory (or confirmation) bias: they tend to discard the new information that substantially differs from their priors. One model that captures this kind of deviation from full rationality is proposed by Brock and Durlauf (2001). They introduce a setup in which individual utility exhibits social interaction effects: the individuals desire to conform to the behaviour of the social groups to which they belong. One requirement of this approach is that each agent observes the behaviour of a large number of other individuals. As noted by Shiller (1984), individuals actually update their prior information in (mainly bilateral) discussions with others. Therefore conforming to some average 'social' behaviour, or information, is unlikely to occur in an environment consisting of bilateral interactions.

Along a different dimension, Haruvy et al. (2007) find that traders' expectations are adaptive, i.e. they give more importance to the past realized price of the asset than the fully-rational agent would. This constitutes a deviation from full rationality because (under full rationality) past prices cannot serve as predictors of future prices.

Understanding whether (and under which conditions) the financial markets are informationally efficient when agents do not behave fully rationally remains an open question. From the policy perspective, it is important to understand if asset prices bubbles derive from incomplete information (and therefore increasing information flows would solve the problem) or from the irrationality of agents (in which case a different policy approach should be designed).

In this paper, we study the informational efficiency of a market with a single traded asset, in which agents can have both the aforementioned forms of deviation from full rationality. The price, which is a public information, initially differs from the fundamental value, about which the agents have noisy private information (on average, correct). A fraction of traders revise their price expectations in each period giving some weight to the past prices and also exchanging opinions about future prices in a social interaction with another agent. Integrating new information from social interaction is subject to a certain degree of confirmatory bias.

We show that, taken separately, each of the deviations from rationality worsen the information efficiency of the market. However, when the two biases are combined, the degree of informational inefficiency of the market (measured as the deviation of the long-run market price from the fundamental value of the asset) can be non-monotonic both in the weight of the adaptive component and in the degree of the confirmatory bias. In other words for some range of parameters, the two biases tend to mitigate each other's effect, thus increasing the informational efficiency.

The paper is structured as follows. Section 3.2 presents the setup of the model. Section 3.3 derives analytical results for each bias taken separately. In Section 3.4, we present the simulation results when two biases are combined. Finally, Section 3.5 discusses the implications of our results and suggests some future avenues for research.

3.2 The model

Consider a market with N participants, each endowed with an initial level of liquidity $L_0 > 0$. Time is discrete (e.g. to mimic the daily opening and closure of a financial market), denoted with $t = 0, 1, ...$ Market participants trade a single asset, whose price in period t we denote with P_t. This price is public information. Prices are normalized in such a way that they belong to the interval $[0, 1]$.

At the beginning of each period t, every agent i can place an order to buy or short sell 1 unit of the asset, on the basis of her expectation about the price for period t, denoted with $P_t^{e,i}$. Placing an order implies a fixed (small but positive) *transaction cost* c, i.e. $0 < c \ll 1$. At the end of the period, each agent i learns the price P_t at which the trade is settled (as explained below).

The agent i then constructs her price expectation for the next period and decides to participate in the trading in period $t+1$ according to the *expected next-period gain*, i.e. if

$$\left|P_{t+1}^{e,i} - P_t\right| - c > 0\,. \tag{3.1}$$

Moreover, she participates as a *buyer* if her price expectation for the next period exceeds the current price, i.e.

$$P_{t+1}^{e,i} > P_t\,, \tag{3.2}$$

or as a *seller* if, on the contrary,

$$P_{t+1}^{e,i} < P_t \,. \tag{3.3}$$

The way in which agents form their next-period price expectations differs from the standard rational-expectation benchmark in the following way. First deviation is the fact that agents give positive weight to the past public prices, i.e. they have (partially) adaptive expectations. Secondly, they can influence each other's expectations via social interactions with confirmatory bias.

Formally, in every period a fraction, $\gamma \in (0, 1]$, of the agents makes a revision of their price expectations. An agent revises her price expectation by analyzing the past price of the asset and by randomly encountering some other agent (at zero cost), and possibly sharing her own price expectation with this partner. In these encounters, the agents have a *confirmatory bias*, i.e. each agent tends to ignore the information coming from the other agent if it differs too much from her own. If, on the contrary, this difference is not too large , i.e. smaller than some fixed threshold, which we denote with σ, then the agent incorporates this information into her price expectation. The remaining $(1-\gamma)N$ agents do not revise their expectations in the current period.

Summarizing, the expectation formation process of agent i meeting agent j is:

$$P_{t+1}^{e,i} = \alpha P_t + (1-\alpha) \begin{cases} P_t^{e,i} & \text{if } \left| P_t^{e,i} - P_t^{e,j} \right| \geq \sigma \\ \dfrac{P_t^{e,i} + P_t^{e,j}}{2} & \text{otherwise} \end{cases} , \tag{3.4}$$

and it is analogous for $P_{t+1}^{e,j}$. Here, α measures the relative weight of the past price. If $\alpha = 1$, the agents have purely adaptive expectations (and social interactions play no role). If $\alpha = 0$ and $\sigma = 1$ the agents (that revise their expectations) completely disregard the past and fully integrate all the information coming from the social interactions. This latter case corresponds to full rationality.

Our objective is to analyze the price formation under the different values of the parameters α, σ, and γ.

Concerning the market microstructure, we assume that the market is centralized, with a simple price response to excess demand (or supply). In other words, the market mechanism is similar to the Walrasian auctioneer. More precisely, the price formation mechanism functions as follows:

1. There exists a hypothetical price at period $t+1$ that would equate the number of buy-orders and sell-orders. Let us denote it with P^*_{t+1}. From (3.2) and (3.3), P^*_{t+1} is the solution of the equation:
$$n_B(x) = n_S(x)\,,$$
where $n_B(x)$ and $n_S(x)$ are the numbers of buyers and sellers at price x. Whenever there are several solutions to this equation, P^*_{t+1} denotes the average of the values that solve the equation.

2. Out of equilibrium, the price adjustment depends on the size of the excess demand or excess supply relative to the size of the population; in other words, denoting $\beta(x) = |n_B(x) - n_S(x)|/N$, the price adjustment process is:
$$P_{t+1} = \beta(P_t)P^*_{t+1} + (1-\beta(P_t))P_t\,. \tag{3.5}$$
Thus, the deviation from the equilibrium does not disappear instantly. However, the price moves in the direction that eliminates the excess demand or supply, and, moreover, the speed of adjustment depends on the size of disequilibrium (relative to the size of the population) [1].

3. Given that each agent that participates in the market in period t places an order for one unit of the asset, the number of exchanges that occurs is $\min\{n_B(P_t), n_S(P_t)\}$. Then, each seller i updates her liquidity by $L^i_{t+1} = L^i_t + P_t - c$ and, similarly, each buyer by $L^j_{t+1} = L^j_t - P_t - c$.

[1] We avoid the shortcoming of assuming a constant $\beta(P_t)$. As discussed by LeBaron (2001), if $\beta(P_t)$ is assumed to be constant, the behavior of the simulated market is extremely sensitive to the value of β, which makes it difficult to interpret the results.

4. If an agent's i liquidity dries up to zero, then she leaves the market. At her place, at the beginning of the next period enters a new agent with liquidity L_0, and the next-period price expectation equal to P_0^i.

In this setting, consider an initial public offering (IPO) of the asset. At time $t = 0$, the asset gets introduced in the market at some price P_0. Let's also suppose that, *on average*, the agents have the unbiased information about its fundamental value. In particular, suppose that the initial price expectations of the agents is a uniform distribution in the $[0, 1]$ interval, i.e. the fundamental value of the asset is $\frac{1}{2}$. However, the initial price P_0 differs from the fundamental value.

The questions that we pose are:

- Does the market price P_t converge to the fundamental value of the asset?
- If not, how large is the market inefficiency, namely the deviation in the long-run ($t \to \infty$) of P_t from the fundamental value?
- How does this deviation depend on the weight of history α (i.e. the "adaptiveness" of agents' expectations), the confirmatory bias of the traders $(1 - \sigma)$, and the frequency with which agents adjust their expectations, γ?

Figure 4.1 gives a summary of the within-period timing of events of our model.

3.3 Analytical results

While a complete analytical study is beyond reach due to the non-linearity of this model, we can nevertheless characterize analytically the answers to the above questions for some of the values of the parameters. This requires a further assumption that the number of market participants (N) and every agent's initial liquidity (L_0) are sufficiently large.

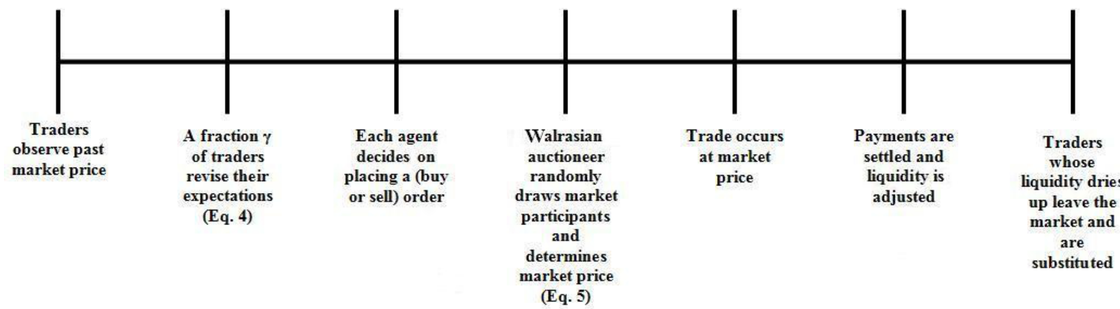

Figure 3.1: Within-period timing of events

Purely adaptive expectations

Consider first the case where agents discard the social interactions and consider only the past price. In other words, $\alpha = 1$ in (3.4). We analyze separately two sub-cases: all agents revise their expectations in every period, i.e. $\gamma = 1$; and only a fraction of agents revise their expectations in every period, i.e. $\gamma < 1$.

All agents revise their expectations ($\gamma = 1$)

In this case Eq. (3.4) simply reduces to:

$$P_{t+1}^{e,i} = P_t \text{ for any } i \text{ and } t.$$

However, this is also the hypothetical price that equates buyers and sellers, i.e. $P_{t+1}^{*} = P_t$, and thus $\beta(P_t) = 0$. Finally, from (3.5) we get $P_{t+1} = P_t$. This means that the market price doesn't evolve: $P_t = P_0$ in every period. Intuitively, if all agents revise their expectations in every period and have purely adaptive expectations, once the initial price P_0 is announced, every agent immediately revises her next-period expectation, substituting it with P_0. Given that every agents does so, no agent is interested in trading, and the price does not evolve.

Only a fraction of the agents revise their expectactions ($\gamma < 1$)

In this second case let's assume $\gamma < 1$, but γN being sufficiently large. Without loss of generality, suppose that $P_0 > 1/2$. We prove that the market reaches the long-run equilibrium, after few periods, with the long-run market price deviating from the initial price by a value smaller than $c(1-\gamma)$. Thus in this case the initial inefficiency remains substantially unchanged.

We need the following preliminary result.

Proposition 1. *Consider a population of agents divided into two groups: agents in the first group, whose size is N_1, have expectations uniformly distributed in $[0,1]$, while agents in the second group, whose size is $N_2 >> N_1$, all have the price expectation equal to some fixed $\hat{P} \in (0,1)$. Then, the price P^* is given by $P^* = \hat{P} - c$ if $\hat{P} > 1/2$ and $P^* = \hat{P} + c$ if $\hat{P} < 1/2$. If $\hat{P} = 1/2$, then $P^* = \hat{P}$.*

Proof. We consider only the case $\hat{P} > 1/2$ (the proof for the case $\hat{P} < 1/2$ is analogous, while the case $\hat{P} = 1/2$ is trivial). Define the functions $\theta_c(x, P)$

$$\theta_c(x,P) = \begin{cases} 1 & \text{if } x > P + c \\ 0 & \text{otherwise} \end{cases},$$

and $\eta_c(x,P) = 1 - \theta_{-c}(x,P)$. For a sufficiently large N_2, the numbers of sellers and buyers at price $x \in [0,1]$ are given by

$$\begin{aligned} n_S(x) &= (x-c)N_1 + N_2\theta_c(x,\hat{P}) \quad \text{and} \\ n_B(x) &= (1-x-c)N_1 + N_2\eta_c(x,\hat{P}). \end{aligned} \tag{3.6}$$

This follows from the trading protocol, given that for a price sufficiently close to $\hat{P}$ (i.e. with a deviation less than c), only the first group of agents participates in the trading, and that the expectations are uniformly distributed in the first group. On the other hand, if $x < \hat{P} - c$ (or $x > \hat{P} + c$), the second group also participates in the trading as buyers (sellers).

Then, the difference in the number of buyers and sellers is

$$\Delta(x) = (1-2x)N_1 + N_2\left(\eta_c(x,\hat{P}) - \theta_c(x,\hat{P})\right), \tag{3.7}$$

and, therefore, P^* is the price at which the sign of $\Delta(x)$ changes (or the average of these values, if more than one exist). We can then easily prove that

$$\Delta_- \equiv \lim_{x\to(\hat{P}-c)^-} \Delta(x) = (1-2\hat{P}+2c)N_1 + N_2 > 0. \tag{3.8}$$

Finally, using the assumption $N_2 >> N_1$, we get $\Delta_- > 0 > \Delta(x)$ for all $x > \hat{P}-c$. This implies that $P^* = \hat{P} - c$. ■

This proposition induces the following corollary:

Corollary 2. *Suppose the assumptions of Proposition 1 hold. If a third group of agents (of arbitrary size) with price expectation $\tilde{P}$, such that $|\tilde{P}-\hat{P}| < c$, joins the market, then P^* does not change.*

We can now analyze the market dynamics under the assumptions $\alpha = 1$ and $\gamma < 1$ with γN large.

During the first period $N_1 = (1-\gamma)N$ agents do not revise their expectations. These expectations are uniformly distributed in $[0,1]$ interval. Contrarily, $N_2 = \gamma N$ agents revise their expectations, which now becomes the IPO price P_0 (i.e. $P_1^{e,i} = P_0$). Proposition 1 ensures that $P_1^* = P_0 - c$. Moreover, the size of market disequilibrium is small: using the definition, we get $\beta(P_0) = (2P_0-1)(1-\gamma)$. Finally the price at the end of period 1 will be

$$P_1 = \beta(P_0)(P_0-c) + (1-\beta(P_0))P_0 = P_0 - \beta(P_0)c. \tag{3.9}$$

Note that this price is c–close to P_0, given that $\beta(P_0)$ is small.

During the next period, γN agents revise their expectations, while $(1-\gamma)N$ agents do not revise. Then, on average, we have that the second group contains $N_2 = \gamma(1-\gamma)N$ agents, for which $P_2^{e,i} = P_1^{e,i} = P_0$, the first group contains $N_1 = (1-\gamma)^2$ agents (who do

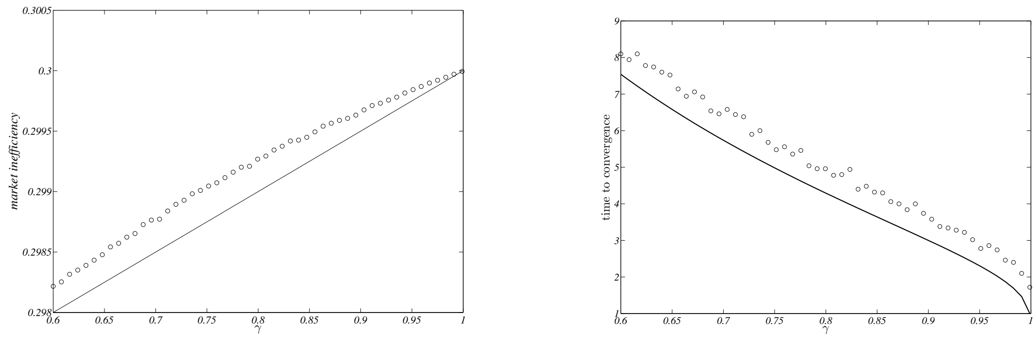

Figure 3.2: Purely adaptive expectations ($\alpha = 1$). Left panel: semilog plot of the market inefficiency as a function of γ. Right panel: time to convergence to the steady state as a function of γ. There are $N = 1000$ agents, the transaction cost $c = 0.005$ and $P_0 = 0.8$. Each point represents the average over 50 simulations. $\bigcirc$: numerical simulations, solid line : analytical results (see Eq. 4.10 and 4.11 respectively)

.

not revise their initial expectations), and, moreover, there exists a third group of agents, of size $N_3 = \gamma N$, for whom $P_2^{e,i} = P_1 = P_0 - \beta(P_0)c$. We can then apply Corollary 2 and conclude that $P_2^* = P_0 - c$. Computing the next-period market disequilibrium $\beta(P_1)$, we can easily observe that $\beta(P_1) \sim (1-\gamma)^2$. Therefore, the next-period price P_2 will be

$$P_2 = \beta(P_1)(P_0 - c) + (1 - \beta(P_1))P_1 \sim P_0 - \beta(P_0)c + \mathcal{O}(1-\gamma)^2 \sim P_0 - c(1-\gamma) + \mathcal{O}(1-\gamma)^2\,. \tag{3.10}$$

Thus, the market price varies as long as there exist agents that have not yet revised their initial expectations. However, the market price does not move too far from P_0. Assuming the extreme-case scenario where in every period the same $(1-\gamma)$ agents happen to be the ones that do not revise their expectations, the number of periods that pass before the market price converges to its steady-state value equals

$$\frac{-\log N}{\log(1-\gamma)} \tag{3.11}$$

Numerical simulations presented in Figure 4.2 confirm our theoretical findings.

Social interactions

We now, consider the setting where agents' expectations have no adaptive component (i.e. $\alpha = 0$), and therefore agents that revise their expectations rely on the social interactions with other agents. Then, the relevant parameter is the extent of the confirmatory bias $(1 - \sigma)$. We derive analytical results for the cases of the extreme form of confirmatory bias $(\sigma << 1)$ and for that of virtually no bias $(\sigma \sim 1)$.

Social interactions: large confirmatory bias

Suppose that whenever two agents meet, neither of them adjusts her price expectation, no matter how close their past-period expectations are. We will prove that the market is fully informationally efficient in the long-run, but convergence to this efficient outcome takes arbitrarily long time.

In every period, γN agents engage in social interactions (without influencing each other's price expectations). This implies that no agent revises her price expectation. Therefore, the mean price expectation (which we denote with $\widetilde{P}^e$) does not change either; namely, from (3.4), $\widetilde{P}^e_{t+1} = \widetilde{P}^e_t$. Under the assumption that the initial price expectations are uniformly distributed, we obtain $P^*_{t+1} = \widetilde{P}^e_{t+1} = 1/2$. The market disequilibrium is thus given by

$$\beta(P_t) = |1 - 2P_t|. \tag{3.12}$$

Therefore, the market price evolves according to the equation

$$P_{t+1} = \frac{|1 - 2P_t|}{2} + (1 - |1 - 2P_t|)\, P_t = P_t + \frac{|1 - 2P_t|(1 - 2P_t)}{2}. \tag{3.13}$$

Let us define the mapping

$$f(P) = \begin{cases} P - \frac{(2P-1)^2}{2} & \text{if } P \geq 1/2 \\ P + \frac{(2P-1)^2}{2} & \text{if } P < 1/2. \end{cases} \tag{3.14}$$

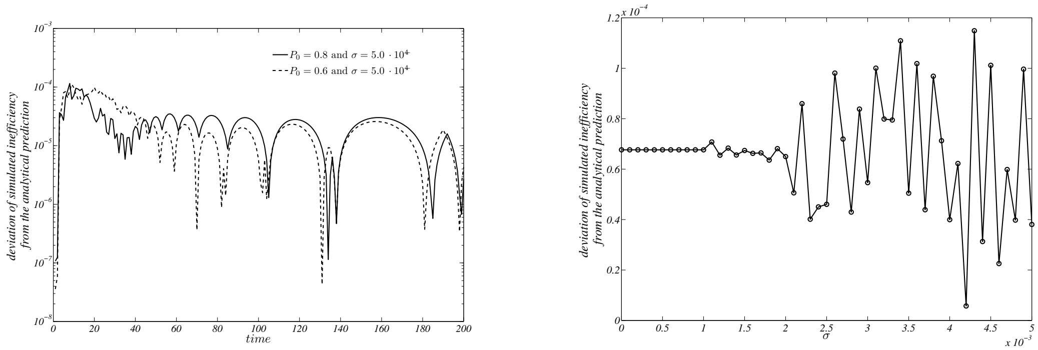

Figure 3.3: Social interaction without adaptive component ($\alpha = 0$ and $\sigma \sim 0$). Left panel: semilog plot of the difference between the simulated and analytical market inefficiency with $\sigma = 5.0 \cdot 10^{-4}$ and $P_0 = 0.6$ or 0.8. Right panel: difference between the simulated and analytical market inefficiency as a function of σ. There are $N = 1000$ agents, the transaction cost $c = 0.005$.

The evolution of the market price is determined by the dynamic system

$$P_{t+1} = f(P_t). \tag{3.15}$$

This mapping has a unique fixed point at $P = 1/2$. Moreover, this is an attractor whose strength decreases the closer we are to the fixed point: $P_t - 1/2 \sim a/t$, the best fit being $a \sim 0.455$. Finally, note that if $P_0 = 1/2$, then Eq. (3.13) implies that $P_t = 1/2$ for all t. Figure 4.3 shows that our theoretical results approximate well the result of numerical simulations.

Social interactions: small confirmatory bias

Let us now consider the opposite extreme, i.e. the agent that integrate into her expectation the information coming from social interactions even when the partner's expectations diverge radically from hers. Assume for the moment that all agents revise their expectations in every period ($\gamma = 1$). We will prove that in this case the market is fully informationally efficient in the long-run, and that convergence to this efficient outcome occurs within in a finite number of periods (that essentially depends on the transaction cost c).

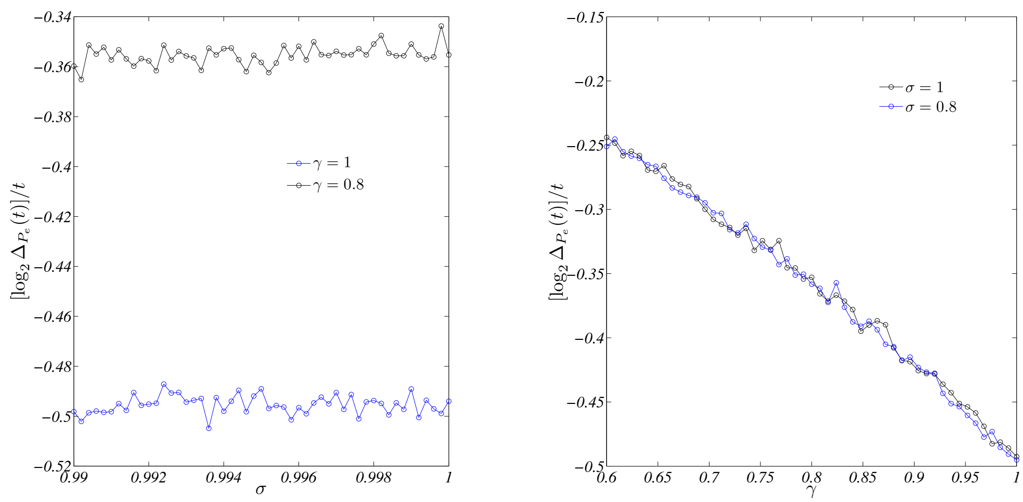

Figure 3.4: Evolution of $[\log_2 \Delta_{P_e}(t)]/t$ as a function of σ (left panel) and as a function of γ (right panel). There are $N = 1000$ agents, the transaction cost $c = 0.005$.

Given that the expectation-revision rule (3.4) with $\alpha = 0$ and $\gamma = 1$ preserves the mean price expectation, we trivially get: $\widetilde{P}^e_{t+1} = \widetilde{P}^e_t = 1/2$.

This follows from the fact that the initial price expectations are uniformly distributed in $[0, 1]$, hence with average value $1/2$, which also equals the hypothetical Walrasian-auctioneer price $P^*_t = 1/2$. Moreover, the equation (3.4) implies that price expectations follow the Deffuant dynamic (Deffuant et al. 2000, Weisbuch et al. 2002). In other words, the dispersion of price expectations, denoted with Δ_{P_e}, shrinks to zero according to the following law (see the left panel of Figure 4.4)

$$\Delta_{P_e}(t) \sim \frac{1}{2^{t/2}}. \tag{3.16}$$

Since the transaction cost is positive, the market activity stops once all the expectations fall inside the interval whose width is smaller than $2c$. This happens after a time $\widehat{T} \sim -2\,(1 + \log_2 c)$.

Let us assume now that the price expectations have a dispersion large enough, so that the market activity still does not stop. Then, we can easily compute the market disequilibrium

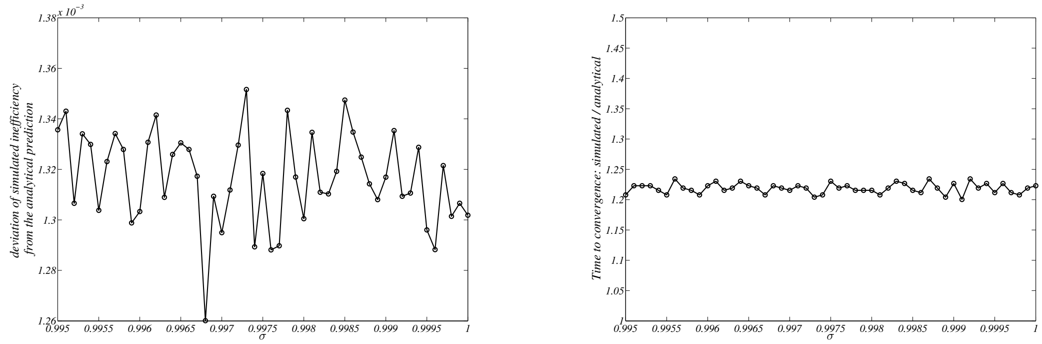

Figure 3.5: Social interaction without adaptive component ($\alpha = 0$ and $\sigma \sim 0$). Left panel: difference between the simulated and analytical market inefficiency as a function of σ. Right panel: the ratio of the simulated time to convergence over the analytical one, σ. There are $N = 1000$ agents, the transaction cost $c = 0.005$.

$\beta(P_t)$:

$$\beta(P_t) = 2^{t/2}|1 - 2P_t|,$$

which implies that the next-period price is given by:

$$P_{t+1} = P_t + 2^{t/2-1}|1 - 2P_t|(1 - 2P_t). \tag{3.17}$$

Let us introduce the auxiliary variable x, defined as $P_t - 1/2 = x_t/2^{t/2}$. This allows us to fully describe the market price evolution with the dynamic system given by the function $g(x)$:

$$g(x) = \begin{cases} \sqrt{2}x - 2\sqrt{2}x^2 & \text{if } x \geq 0 \\ \sqrt{2}x + 2\sqrt{2}x^2 & \text{if } x < 0. \end{cases} \tag{3.18}$$

This mapping has three fixed points: $x = 0$ (unstable), and $x = \pm(2 - \sqrt{2})/4$, that are stable.

We can thus conclude that, as long as the market activity continues, the market price converges to $1/2$, given that in all cases $P_t = \frac{x_t}{2^{t/2}} + \frac{1}{2} \rightarrow \frac{1}{2}$. These findings are supported by numerical simulations, whose results we report in Figure 4.5.

In a similar fashion, we can study the case $\gamma \neq 1$. In this case, in every period γN agents revise their expectations and, because they have a very small confirmatory bias (i.e. $\sigma \sim 1$), they influence substantially each other's expectations. Then, overall there is a tendency for the expectations to converge (because of the process driven by Eq. (3.4) with $\alpha = 0$). We should keep in mind, however, that in every period $(1-\gamma)N$ agents do not revise their expectations.

Zero weight given to the past prices in forming the next-period expectation ($\alpha = 0$) implies that the mean price expectation does not change, $\widetilde{P}^e_{t+1} = \widetilde{P}^e_t$. Hence $\widetilde{P}^e_t = 1/2$ and $P^*_t = 1/2$ for all t. Moreover the expectations are distributed in an interval whose width (denoted with $\Delta_{P_e}(t)$) shrinks to zero, but slower than in the case $\gamma = 1$. The simulations presented in the right panel of Figure 4.4, allow us to see that this narrowing of expectation dispersion follows approximately the law

$$\Delta_{P_e}(t) \sim \frac{1}{2^{q_\gamma t}},$$

where $q_\gamma = a\gamma + b$, $a = 0.61 \pm 0.02$ and $b = -0.13 \pm 0.01$, independent of σ.

Let us observe that this equation for $\gamma = 1$ becomes the Eq. 3.16. Thus, the price disequilibrium can be estimated as

$$\beta(P_t) \sim 2^{q_\gamma t}|1 - 2P_t|,$$

which implies the following price dynamics:

$$P_{t+1} = P_t + 2^{q_\gamma t - 1}|1 - 2P_t|(1 - 2P_t).$$

Introducing a new variable y_t such that $P_t = 1/2 + y_t/2^{q_\gamma t}$, we obtain the following difference equation for the evolution of y_t

$$y_{t+1} = 2^{q_\gamma} y_t - 2^{q_\gamma + 1}|y_t| y_y\,.$$

This mapping has three fixed points, $y = 0$ (unstable), and $y = \pm(1 - 2^{q_\gamma})/2^{q_\gamma + 1}$ (stable). We can finally conclude that, similar to the results above, the market price converges to 1/2 as long as the market activity continues. The market activity stops once all the expectations fall inside the interval whose width is smaller than $2c$. This happens after time $\widehat{T} \sim -(1 + \log_2 c)/q_\gamma$.

3.4 Simulation results

When the price expectations of agents have both the adaptive component and confirmatory bias, obtaining analytical results is beyond reach. We thus proceed by running numerical simulations. In what follows, we vary the values of α and σ, from 0 to 1 (in steps of 0.01). For each pair of values (α, σ) the market is simulated 10 times. The cost of a trading transaction is fixed at $c = 0.005$. Each simulation runs for as many periods as it needs for the market price to converge to steady state. Note that in the simulations we define the steady-state as the situation in which the difference between the market prices in periods t and $t + 1$ differ by a value smaller than 10^{-6}. We then look at the degree of informational inefficiency of the market in the long-run, i.e. how much the market price diverges from the fundamental value of the asset.

The agents have a relatively low level of liquidity. Remember that if the outcomes of the trading strategy of a trader lead to losses that, accumulated over several periods, exhaust her liquidity, she quits the market and is replaced by another trader with the same initial price expectation P_0^i. Given that traders have a relatively low level of liquidity, a certain number of them will quit the market and this implies that the turnover rate of traders is relatively high. This means that some amount of noise gets continuously injected into the market.

Figure 4.6 reports the informational inefficiency of the market (as measured by the divergence of the asymptotic market price from the fundamental value of the asset) for the cases in which the fraction of agents that revise their expectations in every period is $\gamma = 0.2$, 0.5, and 1, respectively (Panels A, B, and C). The market inefficiency is a function of the

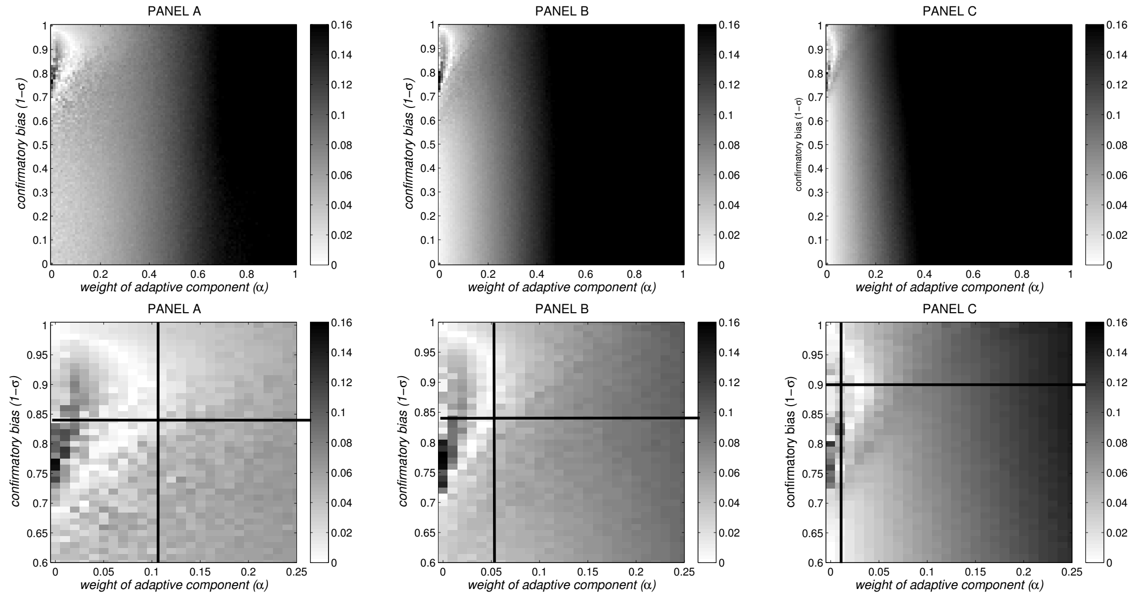

Figure 3.6: The degree of market inefficiency as a function of the adaptive component (α) and the confirmatory bias $(1-\sigma)$. Panel A: $\gamma = 0.2$, Panel B: $\gamma = 0.5$ and Panel C: $\gamma = 1.0$. The initial price is 0.9. In the first row, $(\alpha, 1-\sigma) \in [0, 1]$. In the second row, focus on $\alpha \in [0, 0.25]$, $(1-\sigma) \in [0.6, 1.0]$. The values corresponding to the horizontal and vertical lines are represented in 4.8 and 4.9 respectively.

weight of the adaptive component in the price expectations of traders α and of the degree of confirmatory bias $1-\sigma$. Lighter colors indicate lower level of market inefficiency, while darker ones indicate higher inefficiency. Figure 4.7 shows the average number of traders that exit the market as their liquidity hits the zero bound. In both Figure 4.6 and Figure 4.7 we focus on the interesting parameters set: the ones with large confirmatory bias and low weight of the adaptive component. Let us observe that in the remaining range of parameters α and σ the trend of market inefficiency (data not shown) is trivially monotonic.

Analyzing these figures, we obtain the following findings. Fixing the value of σ, as we move from the extreme-left point ($\alpha = 0$) to the right, the average deviation of the long-run market price from the fundamental value first decreases and then increases - at least for some values of σ (see Figure 4.8 for some characteristic cases). In other words:

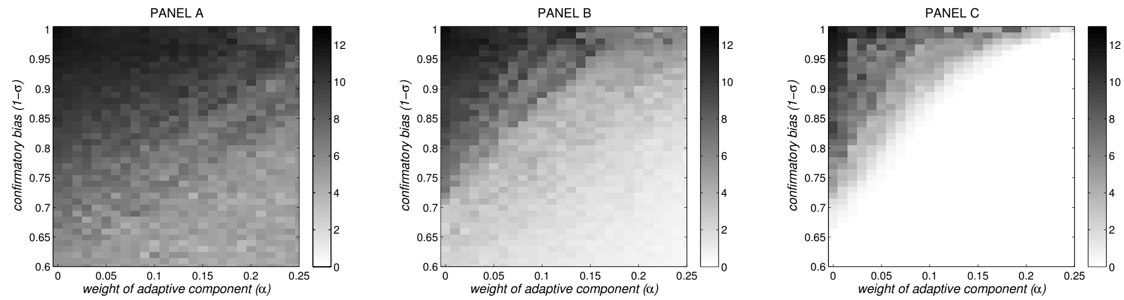

Figure 3.7: Average rate of exiting agents as a function of the adaptive component (α) and the confirmatory bias $(1-\sigma)$. Panel A: $\gamma = 0.2$, Panel B: $\gamma = 0.5$ and Panel C: $\gamma = 1.0$. The initial price is 0.9. Focus on: $\alpha \in [0, 0.25]$, $(1-\sigma) \in [0.6, 1.0]$

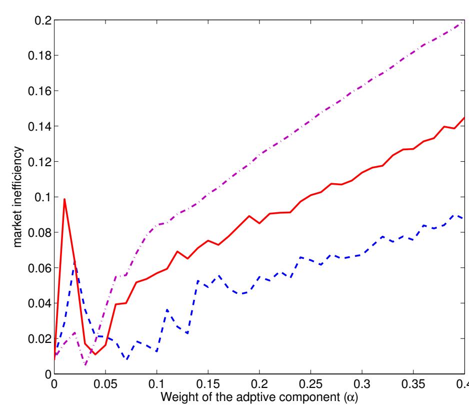

Figure 3.8: The degree of market inefficiency as a function of the the weight of the adaptive component (α). Dashed line: $\gamma = 0.2$, $(1-\sigma) = 0.84$; continuous line: $\gamma = 0.5$, $(1-\sigma) = 0.84$; dash-and-dot line: $\gamma = 1$, $(1-\sigma) = 0.90$. Values correspond to the horizontal black lines in Figure 4.6.

Proposition 3. *Market inefficiency can be non-monotonic in the weight of the adaptive component (α)*

In all three panels of Figure 4.6, we find that for high values of α, the degree of the informational inefficiency of the market is very high. Clearly, when traders put a large weight to the past price in forming expectations, the initial price becomes very important. When receiving information which indicates that the value of the asset is low (even in the absence of confirmatory bias), traders tend to give little weight to it - basically, all that matters is the past price. In this case, the initial price strongly influences the aggregate expectation formation process (the expectations of all agents quickly converge upwards to some point between the initial price and the fundamental value) and, given that in our case the initial price strongly differs from the fundamental value, the long-run market price stays largely above the fundamental value.

Consider now the situation with the most extreme form of confirmatory bias, i.e. all traders completely ignore the information that comes from others. As α declines, the traders give less weight to the past prices and more weight to their own expectations of the previous period. Therefore, the agents whose initial expectations are very low do not move their next-period expectations upwards too much. At the same time, the market price keeps falling, driven by the Walrasian auctioneer (which also implies the downward move in the expectations of the agents whose initial expectations are high). These two inter-related processes - the upward drift of price expectations of initially low-expectation agents and the downward pressure on the market price - converge to some value relatively close to the fundamental one.

As α declines further, we observe that the market inefficiency rises again. This is due to the fact that for the lower values of α, the first process (upward move in expectations of the initially low-expectation agents) becomes slower than the second one (i.e. downward move in the market price). Thus, the low-expectation agents keep making negative profits, eventually hit the zero-liquidity bound, and exit (we can note this by looking at Figure 4.7: the number of agents that exit the market increases for values of the parameters corresponding to the top-left part of the figure). There is a sufficiently high exit rate of these agents from the market so as to soften the downward move in the market price, which means that the long-run price at which the market settles down is higher than in the situation in which the exit of traders is negligible.

As the degree of confirmatory bias of agents becomes smaller (i.e. the value of σ increases), the channel that leads to the exit of low-expectation traders softens, as there is now an additional mechanism that creates an upward pressure on the expectations of those traders: the integration of information that comes from their peers. Notice (Figure 4.7) that the exit rate is lower at the higher values of σ.

Furthermore, comparing across the three panels of Figure 4.6, one notes that as the fraction of agents that revise their expectations in each period (γ) increases, the areas of σ in which the non-monotonicity in α occurs becomes smaller:

Proposition 4. *The tendency of the market inefficiency to be non-monotonic in α is stronger, the lower is the fraction of agents that revise their price expectations in each period.*

This happens because the higher frequency of expectation revision (larger γ) and the likelihood to integrate the information coming from other traders (larger σ) act in a complementary fashion: if the rate of revision of price expectations is relatively low, the "openness of mind" (i.e. low degree of confirmatory bias) has a relatively small effect on the mitigation of the exit channel. It is only when the agents revise their expectations relatively frequently, that the "openness of mind" starts to have a real bite, and the upward-sloping part on the left side of the relation between market inefficiency and α starts to disappear.

Let's now fix the value of $\alpha = 0.1$ on Figure 4.6A, $\alpha = 0.05$ on Figure 4.6B, and $\alpha = 0.01$ on Figure 6C (results for these specific cases are reported in Figure 4.9). As we move from the point at the top ($\sigma = 0$) downwards, the market inefficiency first decreases and then increases. In other words:

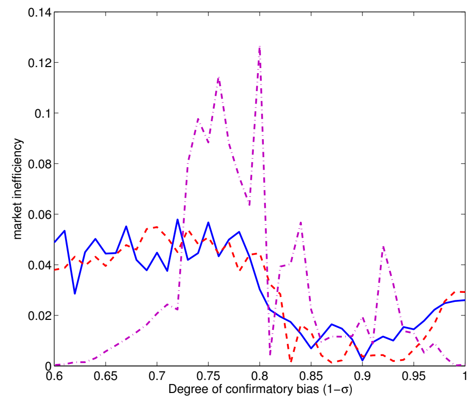

Figure 3.9: The degree of market inefficiency as a function of the confirmatory bias $(1 - \sigma)$. Continuous line: $\alpha = 0.10$, $\gamma = 0.2$; dashed line: $\alpha = 0.05$, $\gamma = 0.5$; dash-and-dot line: $\alpha = 0.01$ $\gamma = 1$. Values correspond to the vertical black lines in Figure 4.6.

Proposition 5. *Market inefficiency can be non-monotonic in the degree of confirmatory bias* $(1 - \sigma)$*.*

The first part of the non-monotonic relationship is easy to explain: as an agent suffers less from confirmatory bias, she starts to integrate at least some of the information about the fundamentals contained in the price expectations of another trader (incidentally, this phenomenon occurs only when the adaptive component in the price expectation is relatively small).

But why the market inefficiency would *rise* as the agents become even more "open-minded"? To understand this, we need to note that this phenomenon occurs only when the adaptive component is not too small. Then, the fact that the initial price differs substantially from the fundamental value plays a key role. The agents have an early-stage upward drift in expectations. At the same time, the market price starts to fall. If the agents are very "open-minded", this implies that they 'excessively' integrate the early upward price drift into their expectations, which, in turn, implies that the price at which the market settles in the long run is relatively high. If, instead, the agents' confirmatory bias is stronger, the decline in the market price is faster than the 'propagation' of the upward-drifting expectations: this is why the market settles at the price relatively close to the fundamental value.

This analysis suggests a very interesting and potentially more general insight: when market participants suffer from more than one deviation from fully rational behavior (in our case, adaptive expectations and confirmatory bias), at least in some range of parameters, the two biases mitigate each other. Given our analysis, it should not be difficult to construct examples of asset markets with traders that have multiple sources of bias, that exhibit the same price behavior as under full rationality, and in which the price behavior would deviate from the full-rationality benchmark as soon as one of the bias sources is eliminated.

Next, looking across the three panels of Figure 4.6, we also note that the values of α at which we have found the non-monotonicity of the market inefficiency decrease at higher values of γ. In other words:

Proposition 6. *The weight of the adaptive component in the expectations (α) at which the non-monotonicity of the market inefficiency in the degree of confirmatory bias gets smaller when the fraction of agents that revise their expectations increases.*

To capture the intuition behind this result, we need to conduct the following thought experiment. Let's fix a point with sufficiently high values of α and σ, for example $(0.4, 0.4)$. Next, let's increase the frequency of revision of expectations by the agents, from $\gamma = 0.2$ to $\gamma = 1$. We then observe that the market inefficiency increases. This indicates the importance of the frequency with which agents revise their price expectations for the propagation of the 'excessive' integration of the upward drift into the expectations, as noted above. In other words, at higher frequency of expectation revision, this 'excessive' integration of the upward drift channel swamps the opposite (i.e. the quantity-of-information) channel more easily. The quantity-of-information channel starts to play a role only when the early-stage upward drift is sufficiently small (i.e., history weighs relatively little in the expectation formation).

If we measure the degree of market inefficiency, while varying σ along a fixed α, in ranges different from those where the non-monotonicity occurs (for example, $\alpha = 0.1$ and $\alpha = 0.3$ on Figure 4.6A), we see that at the higher values of α the relationship between the degree of market inefficiency and σ is negative, while at the lower values of α, this relationship is positive. In other words:

Proposition 7. *The slope of the relationship of market inefficiency in the degree of confirmatory bias (σ) can be of opposite sign at different values of the weight of the adaptive component (α).*

The above discussion has already hinted at the potential explanation why this reversal of the relationship occurs. At sufficiently high values of α, the early-stage upward drift is very important and the smaller confirmatory bias of agents only helps to propagate this drift into the price expectations. At sufficiently low values of α, the early-stage upward drift matters much less and the smaller confirmatory bias becomes beneficial for the informational efficiency of the market, because it helps to integrate more of the relatively unbiased information into the expectations. In other words, in both cases the smaller confirmatory bias (i.e. higher σ) plays the role of the catalyzer; what differs in the two cases is the initial unbiasedness of expectations.

3.5 Conclusion

This paper has studied the informational efficiency of an agent-based financial market with a single traded asset. The price initially differs from the fundamental value, about which the agents have noisy private information (which is, on average, correct). A fraction of traders revise their price expectations in each period. The price at which the asset is traded is public information. The agents' expectations have an adaptive component (i.e. the past price influences their future price expectations to some extent) and a social-interactions component with confirmatory bias (i.e. agents exchange information with their peers and tend to discard the information that differs too much from their priors).

We find that the degree of informational inefficiency of the market (measured as the deviation of the long-run market price from the fundamental value of the asset) can be non-monotonic both in the weight of the adaptive component and in the degree of the confirmatory bias. For some ranges of parameters, two biases tend to mitigate each other's effect, thus increasing the informational efficiency.

Our findings complement the well-known results in the theory of markets showing that the *allocative* efficiency can be obtained even under substantial deviation from individual rationality of agents (Gode and Sunder 1993, 1997). We show that deviations from individual rationality, under certain conditions, can also facilitate the *informational* efficiency of markets. The key condition for this property is that the various behavioral biases that agents possess should mutually dampen their effects on the price dynamics.

Given the potential importance of this insight for financial economics, the natural extension of this work is to test its predictions experimentally. This would require to construct experimental financial markets with human traders, similar to the setting of Haruvy et al. (2007), with the additional feature of allowing agents to share their information (in some restricted form). The outcomes of interest in such an experiment would be both the evolution of market price of the asset and the elicited price expectations of traders.

Chapter 4

Campaign Spending and Rents in a Probabilistic Voting Model

4.1 Introduction

How are the levels of corruption and embezzlement of a political system influenced by electoral campaigns? How rent extraction can be reduced with anti-corruption policies? We answer these questions with a model characterized by the absence of political pressure groups and by the presence of ideological voters whose preferences can be manipulated, by political candidates, through campaign spending. With respect to the present literature, we introduce an important innovation in our model, allowing the voters to directly gain (dis-)utility from the electoral campaign expenditures of the candidates through "participation excitement" (or "disenfranchisement feeling" if the utility is negative).

The problem we study contributes to the analysis of the agency costs of political delegation, addressed mainly by the literature on electoral competition with opportunistic (or rent-seeking) candidates. The debate on this issue has, until now, neglected the study of the effects of campaigns on voters. Indeed, this strand of literature focuses on how the conflicts of interests among heterogeneous voters (with different preferred policies), on one side, and between candidates (that seek to maximize the rent extracted from office) and voters (which would like to see this rents minimized), on the other side, play out and sometimes allow politicians to overtax citizens to finance their own private consumption. On the subject, two schools of thought exists. According to Wittman (1989, 1995), the forces of electoral competition, creating very strong incentives for the candidates to propose policies that please the voters as much as possible [1], are enough to reduce rent extraction to zero. However, as noted by Polo (1998), the efficient outcome strongly depends on the efficiency of the electoral competition, which in turn, is deeply influenced by the type of the information candidates have about the voters' preferences. When this knowledge is not perfect (i.e., in the presence of information asymmetries), rent extraction is possible. Another channel through which political rents can emerge, studied by Persson and Tabellini (1999), is the

[1] This is true, in particular, when the electoral competition is tough, for example when voting happens through a majority rule and the winner does not have to share the power with the loser (winner-take-all).

presence of ideology in voters preferences. Indeed, ideology makes elections more certain, reducing voters elasticity to the quality of party platforms (i.e., increased rent extraction). We build on the setups of Polo and Persson-Tabellini, introducing the possibility for the candidates to do a costly electoral campaign, to analyze how this affects the rent extraction behaviour of the politicians. We use the ideology of voters to introduce an heterogeneity in the effect that the electoral campaign of the candidates has on voters with different political views.

The empirical literature (Kaid, 1987 and Johnson-Cartee, 1989) suggests that campaign spending influences the political bias of the voters thus providing an incentive, for candidates, to use their resources in order to increase the probability of being elected. In this paper, we consider two channels through which expenditures affects voters' choices. First, following the findings of Palda and Palda (1998), the popularity of the candidate that spends more is increased. Second, the campaign affects each single voter utility function. Indeed, following Freie (1997) who observes that an active campaign - in contested and democratic elections - leads to an increased feeling of integration of voters in the political community to which they belong, we consider this second effect a consequence of the presence of "procedural concerns" in the voters. Moreover, in accordance with Matthews (1998) which claims that voters care about the source of campaign expenditures that they observe, our voters will integrate (in their utility functions) more the expenditures of the candidate they favor than those of the one they oppose. While we mainly consider the case in which they enjoy the electoral campaign, our model easily accommodates situations in which voters dislike the campaign spending. In this case, the electoral campaigns bring the voters a disutility which represents the process of disenfranchisement of the electoral body in modern democracies.

Our setup is rather simple. Two rent-seeking candidates decide on binding platforms which are composed of public goods, rent extraction and investments in an electoral campaign. On the base of these platforms the voters, influenced by the electoral campaign, decide rationally which candidate to elect. The latter will then enforce his platform as promised.

Our main finding is that, surprisingly, the electoral campaign does not affect the amount of rent that is ex-ante announced by the candidates (i.e., the gross rent). However, the investment of resources implied by the campaign reduces the amount of rent available, ex-post, for the candidate elected (i.e., the net rent, which is what is left of the gross rent after the electoral campaign has been paid for). Moreover, the political campaign does not influence the ex-post probability of being elected (confirming the stylized fact observed by Levitt, 1994).

The presence of an element of excitement of the electoral body, as a result of the electoral campaign, implies that our model predicts that the expenditures in political advertisement should be higher than in the absence of such element.

Since positive net rents can be observed in our model, we then study which kind of policies can be implemented in order to reduce or eliminate them. Our main policy suggestion is to implement an additional tax aimed at reducing the incentives to rent extraction of the candidate. We find that such tax can, under some conditions, improve the welfare of the voters. Nevertheless, if the tax is too small then it reduces the incentives to campaign, making both candidates better off. Finally, we observe that such a policy is difficult to implement in countries with high income inequalities. On this regard we produce a testable prediction: in the presence of high levels of corruption or high income inequalities, the anti-corruption policies implemented by local politicians will be ineffective or too costly to gain popular support.

Given our assumptions, our analysis applies to democratic countries where, compared to other sources of campaign finance (such as corruption, donations from single citizens or public funding), the money that special interest groups donate (or are allowed to donate) to political parties and candidates is limited. Moreover, concerning corruption it is clear that the empirical predictions and the policy suggestions of this paper are more relevant for countries where this phenomenon is endemic. Finally, for the part relative to the presence of "campaign excitement", the attention should be focused on those countries where the voters have a relatively positive view of politics [2]. In Figure 1 we propose a map of the countries where these conditions apply.

While it is difficult to gather internationally comparable systematic data on the main sources of campaign finance, a study of the electoral laws (we use the *Political Finance Database 2012*, published by I.D.E.A. [132] as proxy for the limitation of the lobbies' role), together with the data on corruption perception (we use the *Corruption Perception Index 2010*, published by Transparency International [133] as proxy for the level of corruption in a country) in the world allows us to select the approximate set of countries to which our model applies. Some of the most fitting examples of this kind of countries are: the new democratic Egypt (a country where surely people are still excited about politics), Russia, Thailand and most of South America (see Figure 4.1).

Given that we assume that the voting takes place with the majority rule, the primary interest of this paper are those counties where the political competition is between two parties. However, the main results hold, with minor modifications, if applied to political systems with more than two parties, provided that two of them dominate.

For the rest of this paper, we proceed as follows: first we discuss the related literature in Section 4.2, then, building on the framework of inefficient electoral competition of Polo (1998), Persson and Tabellini (2000) and detailing the new features discussed, we outline the model in its mathematical details (Section 4.3) and we derive the results (Section 4.4). Finally, we discuss how some public policies can increase the welfare of the players (Section 4.5) and we conclude (Section 4.6). Proofs of the propositions can be found in the Appendix (Section 4.7).

[2] One should not be too restrictive on this issue. Indeed, changing the value of the parameters that measure the effectiveness of campaign spending, we are able to capture results relevant also for countries with a significative level of detachment from politics, or even situations in which political campaign is actively disliked by the voters.

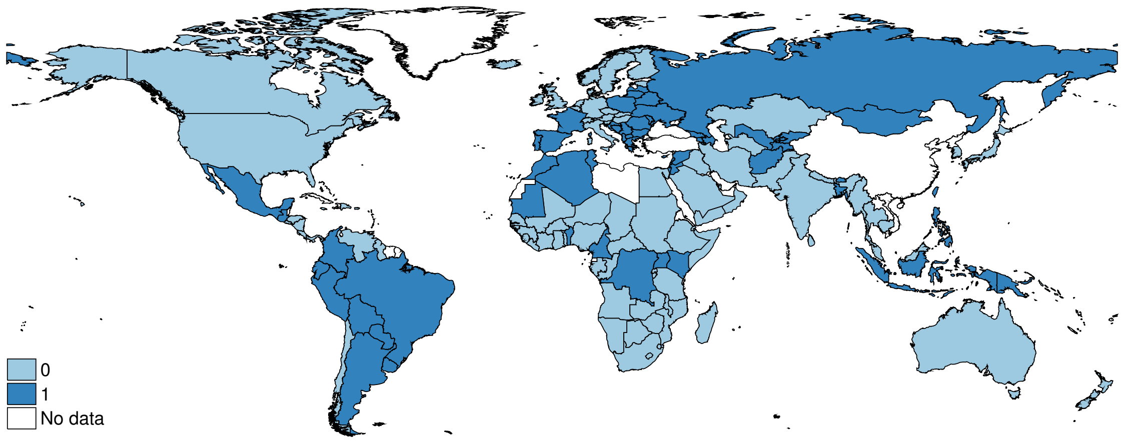

Figure 4.1: This map gives a graphical representation of the countries to which our setup applies best. The countries colored in the darker blue are those where: the public perception of corruption is medium/high (i.e., the Corruption Perception Index 2010 is below 7) and where the electoral law imposes some type of limitation to contribution that political parties can obtain from lobbies. The countries colored in the lighter blue are those that fail to meet this criteria.

4.2 Related Literature

Political competition and more generally representative democracy, where policy choices are delegated to elected officials, is generally considered as an efficient mechanism to aggregate the preferences of large social groups. However, while the possibility of taking policy decisions should be the main reason to run for elective offices, the causality is often reversed: policy announcements are a way to get elected. Several theoretical reasons for this phenomenon have been proposed. In highly competitive elections the satisfaction of the preferences of the median voter (Black, 1948) is required in order to obtain the majority of votes and win the elections (Downs, 1957) so potential candidates with strong policy motivation may opt-out from the political process. Also, candidates could be motivated by the prestigiousness of the office or by the possibility, once elected, of diverting part of the public resources for private use.

This latter approach is the one adopted by the literature on the agency costs of political delegation. Candidates are considered as rent maximizer leviathans with the only objective of being elected and to divert as much as possible of the public budget for private consumption. This literature, assuming a principal-agent approach, focuses on the conflict of interests that arise between the voters (principals) and the candidates (agents). Indeed, while the formers would like the rents to disappear, the latter try to maximize them. Whether the candidates manage to get away with a positive rent or not depends mainly on the information structure of the political game. Wittman (1989, 1995) suggests that the forces of the electoral competition are sufficient to obtain a socially efficient outcome. He shows that the need to get the majority of votes gives a very strong incentive to politicians to trade lower rents with the possibility of election. However, as noted by Polo (1998), this efficient outcome relies heavily on the assumption that the candidates know perfectly the voters' utility (including the median voter's preferences). When this assumption is relaxed, candidates become uncertain about the voters' evaluation of the their characteristics (i.e., about their relative popularity), as observed in Grillo and Polo (1993), Polo (1998) and Myerson (1993) [3] and positive rents can emerge in equilibrium. This is because the uncertainty relaxes the competition for the median voter.

Another channel, through which rents can emerge, is the presence of an ideological component in the preferences of the voters. Indeed, if voters have some idiosyncratic inclination for one of the candidates, as in Persson and Tabellini (1999), their elasticity to changes in the platforms of the candidates becomes lower, so that the marginal cost of less attractive policy choices becomes smaller.

With respect to this literature, our paper introduces several innovations. The first is that candidates do not take their relative popularity as given, but they are able to influence it through a costly electoral campaign. Indeed, Palda and Palda (1998), studying empirically the relationship between electoral results and candidates expenditures, noticed that

[3] This issue is also discussed by Svennson (1997), although in his model the rent seekers are bureaucrats which can afterwards be bailed out by politicians.

outspending the opponent brings significative advantages in terms of election probability, thus creating a strong incentive to invest in campaign. In our setup, the electoral campaign influences the propensity of the electoral body through a mechanism similar to the one introduced by Baron (1994) i.e., moving the distribution of the possible popularities in the direction of the candidate that overspend the opponent. The second innovation we introduce is the fact that the electoral campaign also affects the single voter's utility function directly. Freie (1997) shows that the presence of active campaigns, in free and contested elections, induces the citizens to feel more integrated in the political community, thus increasing their utility. This "excitement of participation" does not correspond to a monetary transfer of resources from candidates to voters, and it is inspired by the concept of "procedural utility", introduced by Frey et Al. (2004) and by Frey and Stutzer (2000, 2005)[4]. These authors argue that economic agents do not only care about outcomes, but also about the way in which those outcomes are achieved. A more democratic system, where plenty of occasions to participate to campaign events are offered and more advertisements are shown, makes the voters feel more involved and, as such, is preferred to an apathetical electoral campaign [5]. However, while the excitement of participation (caused by the expenditures of one candidate) involves all the voters, it is reasonable to assume that the extent to which it is enjoyed depends on their political bias.

A different perspective on the effects of electoral campaign on voters choices is the one of the literature on political advertisements. Most of the contributions deal with the relationship among special interest groups, candidates and voters. They focus on the trade off between the policy cost, imposed by lobbies on citizens as they bias the policy choices of the candidates in their direction, and the informational gains obtained by the voters as a consequence of the interaction between the other two types of agents. Since we neglect the

[4]Notice that the term *procedural fairness* is also used in welfare economics with another meaning, which is not relevant here.

[5]When given the opportunity, as empirically shown by Wielhouwer and Lockerbie (1994), voters do mobilize and participate more to the political process.

role of lobbies, this trade off does not concern us directly. Nevertheless, this literature is interesting for our scopes since it allows us to gain insights into the channels (other than those discussed) through which electoral campaigns affect voting choices. The papers on the issue can be essentially divided in two classes. Some authors (such as Prat, 2001; Wittman, 2007 and Potters, Sloof and Winden, 1997) assume that voters, knowing that candidates just want to be elected, do not believe in the content of the propaganda, thus these advertisements do not directly modify their behaviour. At the opposite, some other authors (such as Austen-Smith, 1987; Coate, 2004 and Grossman and Helpman, 2001) define the electoral campaign of the candidate as informative for the voters.

Belonging to the first class is Prat (2001). According to him, the main effect of political advertisements is to allow voters to gain information about the qualities of the candidates. An expensive electoral campaign signals to the voters that the candidate is efficient in obtaining funds indirectly indicating that he is a good administrator. An alternative, but similar, mechanism through which this information can be transmitted is the direct endorsement of candidates by political pressure groups (Wittman, 2007). A common trait of all these studies is that the only way in which a candidate can obtain funds to run advertisements is through contributions from lobbies which are then shown to be pivotal in solving the signaling problem between candidates and voters. This result is driven by the assumption that the quality of the candidates is, ex-ante, observable by the political insiders. When this is not the case - as assumed in our paper - the informative gains for the population disappear. Moreover, we abstract from the role of lobbies, focusing on the situations in which their contribution to campaigns are relatively limited when compared with those that the candidate can obtain through other sources. This usually happens, as briefly discussed in the introduction, when the role of special interest groups is legally restricted.

The irrelevance of the pressure groups in a context where these are not more informed about candidates' valences than the rest of the voters extends to the situation in which electors consider the advertisements of the politicians to be informative (i.e., to broadcast contents useful for their decision). However, in this case (studied by Austen-Smith, 1987 and Coate, 2004) another mechanism through which electoral campaigns influence voters' choices emerges. In these models, candidates still try to attract funds from lobbies but they do so in order to reduce the "error" with which the voters perceive the candidate's proposals. This reduces the risk, perceived by the risk-adverse voters, that the candidate is poorly qualified. While we do not model voters as explicitly risk-adverse or skeptical about the content of the advertisements, our assumption that voters draw direct utility from the electoral campaign can be interpreted to emulate, in other terms, these two links between campaign expenditures and voters' choices. Indeed, they both result in voters preferring the candidate that invests more in electoral campaign.

This paper draws inspiration from some stylized facts, revealed by the empirical and experimental analysis, and with its results contributes to give a theoretical foundation the some other empirical observation. As noticed by Kaid (1987) and Johnson-Cartee (1989) a *campaign influences the political preferences of voters.* These authors, studying in particular negative campaigning, find that voters recall negative advertisements even if they dislike them. They show that negative advertisements reduce the image evaluation of targeted politician, both for voters with bias toward and against the targeted candidate. As discussed, this stylized fact is integrated here in the way the electoral body is biased in the direction of the candidate that spends more.

Our decision of introducing an heterogeneous reaction of voters to the campaigns of different candidates, is supported by the social psychology literature which presents evidence that *voters do care about the source of the campaign expenditure.* Indeed, Matthews (1998) finds that advertisements have a significative effect on the perception of the member of a group toward the sponsor that paid for it.

Finally, Valentino et al. (2004)'s empirical analysis suggests that, while exposure to political advertisements can be considered informative and may help to reduce informational gaps between informed and uninformed citizens, it *does not produce large shifts in the rank-*

ing of the candidates preferred by a given elector. Similar results are obtained, at a more aggregate level, by Levitt (1994), which confirms that when the candidates spend similar amounts, the absolute level of expenditure has little effect on the probability of election. With the results of our model we reproduce these last two stylized facts thus contributing provide a theoretical justification for them.

4.3 The model

We assume the existence of two candidates, indexed $J = \{A, B\}$, running for an elective office with a majority voting procedure. Each of the candidates must commit to a binding electoral platform (q_J) composed of three elements: the amount of public goods provided (g_J), the amount of rent extracted from office (r_J) and the amount invested in electoral campaign (Ψ_J). The two candidates are assumed to be equal in terms of their ability to provide public goods, to extract rents and in the efficiency of their electoral campaign. The campaigning costs are assumed quadratic (convex). This reflects that, regardless of the amount invested, the candidate has only a limited amount of energy to spend in the campaign (if we think about television advertisements: it takes time to register television appearances and the time on television is limited while, for political rallies, a candidate can participate only to a limited amount of them). Each candidate is motivated uniquely by the perspective of holding the office and extracting rent from his tenure. The extraction of monetary rents is subject to a transaction cost $\gamma \in [0, 1]$, which comes as a consequence of the checks and balances characteristic of the political system, and of the fact that the candidates will need to corrupt other members of the administration in order to transform the overtaxation, that they impose on voters, into private consumption. In our specification, the higher is γ, the higher is the efficiency with which the candidate performs this operation. Finally, from the simple fact of holding the office a politician obtains a non-monetary rent R (ego-rent) which is exogenous and depends on how prestigious is the office. Assuming an electoral competition between two politicians, the objective function of the candidates, which they try to maximize, is:

$$\max_{g_J, r_J, \Psi_J} E[v_J] = p_J(g_J, r_J, \Psi_J)(R + \gamma r_J) - \Psi_J^2 \tag{4.1}$$

where p_J is the probability of election. With respect to Equation 4.1, it should be noted that,while the rent (r_J) is enjoyed only with probability p_J, the candidate has to pay for his electoral campaign even when he is not elected. For simplicity, we assume that candidates are rich enough to advance the amount Ψ_J or that there is some form of public intervention that lends them, without interests, the money required during the campaign (provided that the expenses do not exceed the expected income).

Moreover, the candidates are characterized by a relative popularity (δ) vis-*à*-vis the electorate, about which they are uncertain at the moment they announce their platforms. Indeed, δ is composed of two additively separable parts. The first is a random shock $\tilde{\delta}$, common to all voters, assumed to be drawn from a uniform distribution symmetric around zero and with support $\left[-\frac{1}{2\Omega}, \frac{1}{2\Omega}\right]$. In this distribution the parameter Ω is the density, which measures the degree of uncertainty of the candidates about the electoral body's propensity toward or against them. Higher values for this parameter correspond to lower levels of uncertainty. The second component of δ is deterministic. Indeed, under the assumption that campaign expenditures are effective in modifying the choices of the electorate, this component shifts the support of the previous distribution in the direction of the candidate that spends more in political advertisements. Joining these two components, the popularity of the candidates will be then extracted from the distribution $\delta = \tilde{\delta} + h(\Psi_B - \Psi_A)$, where h is a parameter that expresses the marginal effect of the difference in advertisements spending on the electoral body's propensity to favour one of the candidates. Indeed, if A outspends B the popularity distribution shifts to the left (the direction that favours A's chances of election) and vice versa if B outspends A.

The candidates are elected by the electoral body I composed of a large number of voters

$i \in I$. Each voter is characterized by an income y^i and a political bias σ^i, with the two dimensions assumed to be uncorrelated (or, equivalently, if a correlation exists, candidates are not aware of it). The income determines the preference for the provision of public goods, and it is distributed according to some P.D.F. $F(y_i)$, with full support and of class C^2. The political bias, σ^i, is the preference of the voter for the non-economic ideologies of the parties that express the candidates. We assume that σ^i is distributed in the population according to a uniform distribution centered in zero and with support $\left[-\frac{1}{2\phi}, \frac{1}{2\phi}\right]$. The parameter ϕ is the degree of political polarization in the society. The political bias can be considered as the degree of similarity between the ideological position of the voter and the one of the parties that express the candidates. Indeed, while the candidates are themselves purely opportunistic, they can only decide the economic policy and they are attached to political parties (which are, formally, the same in our model) which express some well characterized position on non-economic issues (e.g., abortion, gay marriages, immigration etc.). In this sense the voters are distributed according to the degree to which they agree or disagree with party ideologies.

However, the decision of a voter to choose one of the candidates is not purely ideological, but also depends on the platform that the latter announce (q). His income is reduced by τ, the taxed levied by the elected candidate and, everything else being equal, his utility becomes smaller when the tax rate increases. The resource gathered with the taxation can be used in two ways: to finance the private consumption of the candidate (r) or to provide public goods (g), thus $\tau = g + r$. All voters agree that r is a bad and all of them prefer platforms which involve less rent extraction. On the contrary, the part of the taxes that goes in the provision of public goods provides a positive utility to voters through a production function $H(g)$. The latter is equal for every voter and assumed to be strictly increasing ($H_g > 0$) and strictly concave ($H_{gg} < 0$).

The campaign expenditures of the candidates (Ψ) influence directly the voters. Specifically, each voter perceives differently the expenditures of the two candidates: he values

values the expenditures more if they come from his favorite candidate than if they come from the opponent. In our specification the weight assigned by a voter to the investment of a candidate grows linearly with the strength of the bias of the latter toward the former [6]. The effectiveness of the campaign in generating excitement in the voter ($\zeta \in [-1, 1]$) is assumed to be equal for all voters and candidates. With this specification of ζ we are able to capture both the case in which campaign generates genuine excitation and participation - thus increasing utilities - and the one in which it creates disutility, progressively detaching voters from politics (this second situation is frequently observed in the more advanced democracies, where is sometimes associated to a drop in the turnout).

Finally, in order to simplify the notation, we transform the political bias of the voter in a value on the segment $[0, 1]$. The relative distance of a voter with bias σ^i with respect to the ideological position of the party that candidate A represents is then given by: $d_i^A = \frac{1}{2} + \sigma^i \phi$. The distance from the ideological position of the party of candidate B is then defined by the complement to one of d_i^A, i.e. $d_i^B = 1 - d_i^A$.

With these elements we can now express the utility function of voter i, assumed to be quasi linear and separable in three parts, vis-à-vis the platforms of the candidate J:

$$W_J^i(g_J, r_J, \Psi_J) = (y - g_J - r_J)\frac{y^i}{y} + H(g) + \zeta(1 - d_i^J)\Psi_J \tag{4.2}$$

Thus, when evaluating the platforms of the two candidates, the voter will compare the following two expressions:

$$W_A^i(g_A, r_A, \Psi_A) = (y - g_A - r_A)\frac{y^i}{y} + H(g) + \zeta(1 - d_i^A)\Psi_A \tag{4.3}$$

$$W_B^i(g_B, r_B, \Psi_B) = (y - g_B - r_B)\frac{y^i}{y} + H(g) + \zeta d_i^A \Psi_B \tag{4.4}$$

and will cast his vote for the candidate whose platform gives him the higher utility, given his income, his political bias and the relative popularity of the two candidates. Thus a voter

[6]Linearity is assumed, but the results would not change using different (sub linear or super linear) specifications.

i will choose A if and only if:

$$W^i(g_A, r_A, \Psi_A) > W^i(g_B, r_B, \Psi_B) + \sigma^i + \delta \tag{4.5}$$

The intuition behind this modeling choice is that voters attribute to each candidate the merit to contribute to the enjoyment of participation (or to the disenchantment) according to the level of electoral campaign made by each of them. However, following the stylized fact discussed, they attribute more weight to the candidate that represent the party with the ideological position nearer to their own than to the opponent.

Since δ is partially stochastic, the result of inequality 4.5 will be internalized by the candidates in probabilistic terms, leading to a probabilistic voting model.

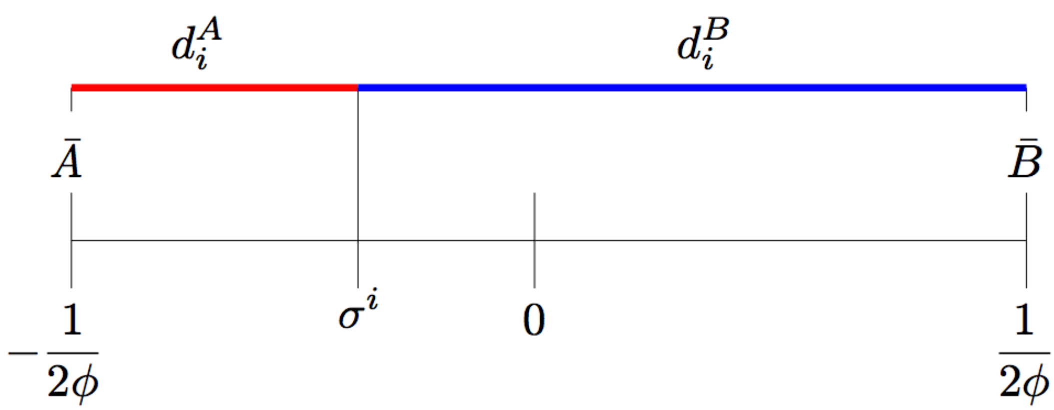

Figure 4.2: A voter i with income y_i and political bias σ^i obtains, as a consequence of the investments of the two candidates an additional (dis-)utility. The amount of utility that he obtains from each of the candidates depends from the investments of the two candidates and from his relative political bias. The latter is transformed in a value on the segment $[0, 1]$ (thick line) with a linear transformation. The proximity to each of the candidates will then determine how much utility is gained (or lost) by i as a consequence of the investments of that candidate. In this graphical example the voter i, being nearer to candidate A, will obtain more utility ($\zeta d_i^B \Psi_A$) from him than from candidate B ($\zeta d_i^A \Psi_B$).

After the elections, one candidate, J, is elected and implements his platform. A general voter i, with income y_i and with bias σ_i will then obtain the following ex-post utility (see Figure 4.2 for a graphical representation of how the third addend is formed):

$$W_J^i = (y - g_J - r_J)\frac{y_i}{y} + H(g) + \zeta(\Psi_A - d_i^A(\Psi_A + \Psi_B)) \tag{4.6}$$

Finally, the timing of the model (summarized in Figure 4.3) is as follows:

1. The two candidates, simultaneously and non cooperatively, announce **binding** electoral platforms $q_A = \{g_A, r_A, \Psi_A\}$ and $q_B = \{g_B, r_B, \Psi_B\}$ respectively. They make their announcements knowing: the distribution of voters' policy preferences (given by $F(y_i)$), the distribution of voters' idiosyncratic bias σ^i and the natural distribution of popularity $\tilde{\delta}$ (as it would be if uninfluenced by the amount invested in political advertisements);

2. The final distribution of δ is calculated using Ψ_A and Ψ_B decided in Step 1;

3. A value is extracted from δ and all uncertainty is resolved;

4. Elections are held, with voters knowing perfectly the platforms of the candidates and their relative popularity;

5. The elected candidate implements his announced policy platform.

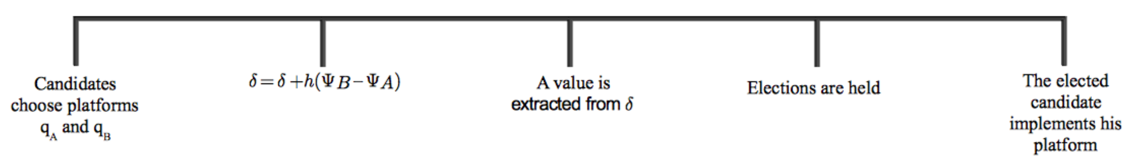

Figure 4.3: Timing of the model

4.4 Results

We can now proceed to solve our model by backward induction. To this end, it is useful to define what constitutes an equilibrium in this model.

Definition 1. *An equilibrium is given by:* $q_J^* = \{g_J^*, r_J^*, \Psi_J^*\}$ *for* $J \in \{A, B\}$ *such that* $\forall J : E[v_J(q_J^*; q_{-J}^*)] \geq E[v_J(q_J; q_{-J}^*)]$ *for all feasible* q_J *and* q_{-J}*, given voters policy preferences, the distribution of voters' idiosyncratic bias* σ^i*, the popularity distribution* δ *and the fact that each voter will vote for the candidate that he likes the most i.e., solving* $argmax_J(W_J|q_A, q_B, \sigma^i, \delta)$ [7]

Without loss of generality, we can study the election probability (p_A) for one of the candidates, say A. The results for his opponent will be then given by $1 - p_A$. Since we assume a competition between two candidates, the election probability of candidate A is equal to the probability that the shares of votes (π_A) he gets is bigger than one half. I.e., under our distributional assumptions:

$$p_A = Prob\left[\pi_A \geq \frac{1}{2}\right] = Prob\left[\sigma \geq 0\right], \tag{4.7}$$

where σ is the idiosyncratic bias of the swing voter, who is exactly indifferent between the two candidates and for whom Equation 4.5 holds with equality. Given the informational structure of our model, the swing voter is not necessarily (and in general he is not) the voter with the median income. His identity is directly (through the utility function) influenced by the platforms (g, r and Ψ) announced by the two candidates and indirectly by the effect of political campaign (Ψ) on the overall distribution of possible popularities of the candidates. Solving Equation 4.5, with equality, for σ we obtain the bias of the swing voter:

$$\sigma = \frac{2(w_A - w_B) + \zeta(\Psi_A - \Psi_B) - 2\delta}{2(1 + \phi\zeta(\Psi_A - \Psi_B))} \tag{4.8}$$

[7] In this definition a variable indexed J indicates a choice variable of the candidate which is taking the strategic decision, while a variable indexed $-J$ refers to the choices of his opponent.

where w_A and w_B (small case) are the welfares of the swing voter (as given by Equation 4.2) net of the direct utility from campaign expenditures. Equation 4.8, stresses the fact that when ζ is positive, and the campaign generates utility, the identity of the swing voter changes in the direction of the candidate that spends more thus generating more excitation of participation. However, everything else being equal, when voters dislike the campaign, the effect of the expenditures on the utility of voters makes the swing voter identity shift in the direction of the candidate that spends less.

Plugging the utility function of the swing voter in Equation 4.7, and solving, we obtain the probability of being elected, i.e.:

$$p_A = \frac{1}{2} + \Omega\left[(r_B - r_A) + (g_B - g_A) + (H(g_A) - H(g_B)) - \left(\frac{\zeta}{2} + h\right)(\Psi_B - \Psi_A)\right] \quad (4.9)$$

Using Equations 4.1 and 4.9, we can derive the response functions of the two candidates and the equilibrium [8]. We will describe in detail the derivations for candidate A, as this gives useful insights in the structure of the model. The same reasoning applies to the other candidate, yielding a symmetric equilibrium.

The first order condition, with respect to g, for the problem of maximization of the candidate A, yields:

$$g_A^* = H_g^{-1}\left(\frac{y}{y}\right) = H_g^{-1}(1) \quad (4.10)$$

This implies that candidate A will always choose the level of public expenditures preferred by the voter with average income. Given the symmetry of the model, we get the standard result that:

Proposition 2. *The candidates reach full policy convergence to the policy g^* preferred by the voter with average income.*

[8] For the problem to be strictly concave (and thus, for the equilibrium to represent a maximum of the candidates' objective function), parameters must satisfy the following condition: for each candidate, it must hold that $(R + \gamma r) > -\frac{\gamma(H_g - 1)^2}{H_{gg}}$. This is always satisfied for a sufficiently prestigious office.

The candidates are purely opportunistic, so they choose their economic policy simply with the aim of maximizing the probability of election. Not having information on the possible correlation between incomes and political bias of the voters they are unable to target their policy to the effective preference of the median voter, so they choose instead its expected position [9]. Moreover, when the income distribution is positively skewed (i.e., the society is composed of more poor than rich), Proposition 2 implies that the public good is overproduced with respect to the preferences of the median voter.

The first order conditions with respect to rent extraction and campaign expenditures yield:

$$p_A = \frac{\Omega}{\gamma}(R + \gamma r_A) \tag{4.11}$$

$$\Psi_A = \frac{\Omega\left(\frac{\zeta}{2} + h\right)(R + \gamma r_A)}{2} \tag{4.12}$$

Therefore, the probability of being elected is increasing in the total desirability of the office $(R + \gamma r_A)$ and in the degree of uncertainty of candidates about their own popularity (Ω). This result match the observation of Polo (1998): uncertainty relaxes the electoral competition so, everything else being equal, it increases the probability of election for any given r_A. Interestingly, p_A is decreasing in the efficiency of transforming rent in utility. Indeed, high γs create stronger incentives to extract rent from the budget, thus reducing the voters' appreciation of the candidate's platform. Equation 4.12 shows that the campaign costs that the politician is willing to pay are equal to the desirability of the office, weighted for the marginal benefits of the electoral campaign (on the swing voter, through ζ, and on the whole population through h), and by the level of uncertainty of the political competition. The higher are those three factors, the higher are the expenses that the candidate is willing to sustain to increase his probability of being elected.

[9] We explicitly exclude the case in which candidates are able to infer, from the income, the likely bias of a voter, as it seems quite unrealistic.

Solving Equation 4.11 for r_A, we obtain $r_A = \frac{1}{\Omega}p_A - \frac{R}{\gamma}$. Thus, the rent extracted depends positively on the probability of being elected, weighted by the degree of uncertainty of the electoral outcome (a high Ω implies the less aggregate uncertainty and thus a more efficient electoral competition, where the rent extraction is more costly in terms of votes lost). Finally, R acts as a substitute for r_A, whose weight is given by γ: candidates are willing to renounce to some of their rents in order to have access to a more prestigious office. Exactly how much they are willing to give up depends on the efficiency of rent extraction.

Since the campaign expenditures are a function of the probability of being elected only through r_A, we first concentrate on the latter. Making explicit the probability of election of A (from Equation 4.9) in Equation 4.11, as a consequence of Proposition 2, we obtain:

$$r_A = \frac{1}{\Omega}\left[\frac{1}{2} + \Omega\left((r_B - r_A) - \left(\frac{\zeta}{2} + h\right)(\Psi_B - \Psi_A)\right)\right] - \frac{R}{\gamma} \tag{4.13}$$

Equation 4.13 shows that outspending the adversary increases the rent extractable. Moreover, given the results of Proposition 2 (the policy convergence of the candidates), Equation 4.13 (together with Equation 4.12) implies that $r_A(r_B, \Psi_A, \Psi_B)$ - the rent of a candidate depends only on the rent of the opponent and on the expenditures of both contenders - and that $\Psi_A(r_A)$ - the expenditure of a candidate depends, directly, only on his rent. So, for Ψ, the strategic responses to the opponent's decisions act directly through the rent and, only indirectly, through both the expenditures in campaign.

Taking the strategies of B as given, r_A becomes:

$$r_A = \frac{1}{4\Omega} - \frac{R}{2\gamma} + \frac{(2h+\zeta)}{4}(\Psi_A - \Psi_B) + \frac{1}{2}r_B \tag{4.14}$$

Therefore, the extractable rent is a function of the sum of two components: one, non-strategic, which depends only on the parameters and one which is a function of rent and expenditures of the adversary and of own expenditures. Equation 4.14 tells us that the candidates' rents are a positive and linear function of each other. Indeed, the more one candidate is planning to divert for private consumption, the more the opponent can extract without reducing his probability of being elected. Moreover, the expected gross rent is a positive function of own campaign expenses and a negative function of those of the opponent. This last relationship uncovers the rationale of rent-seeking candidates for investing in campaign. However, while out of equilibrium the effect of campaign in changing the popularity is twice more important than the effect on the procedural utility of the swing voter, due to the symmetry of the model, under the assumption that $\Psi_A = \Psi_B = \Psi$, there is no equilibrium effect of advertisements on expected rents. This result seems to confirm the empirical observation of Levitt (1994): the absolute level of candidate's expenses has little effect on voter's choices when the two competitors invest similar amounts of resources.

From Equations 4.12 and 4.14 we now calculate the reaction functions of candidate A to the strategies of candidate B, taken as given:

$$r_A = +\frac{4}{\Omega K} - \frac{8R}{\gamma K} + \frac{\Omega R(2h+\zeta)^2}{K} + \frac{8}{K} r_B - \frac{4(2h+\zeta)}{K}\Psi_B \tag{4.15}$$

$$\Psi_A = +\frac{(2h+\zeta)(\gamma + 2\Omega(R+\gamma r_B))}{K} - \frac{\gamma\Omega(2h+\zeta)^2}{K}\Psi_B \tag{4.16}$$

where $K = \left[16 - \gamma\Omega(2h+\zeta)^2\right]$. Assuming enough uncertainty [10], the candidate's rent is decreasing in the expenditures of the opponent and increasing in his rent. As already pointed out, if a candidate increases the rent that he demands, the opponent can do the same without affecting his election probability. However, when a candidate increases the size of his own electoral campaign, he decreases the probability of the opponent to be elected and, in expected terms, his rent.

Moreover, Equation 4.16 shows that the expenditures of one candidate are increasing in the rent that the opponent extracts (since an increase in the rent of B allows A to do the same and extract more money proportionally) and decreasing in his expenditures (more expenditures of B means less expected rent for A, and therefore less expenditures).

[10] I.e., $\Omega < \frac{16}{\gamma(2h+\zeta)^2}$.

Knowing that, by the symmetry of the model, $r_A = r_B$ and $\Psi_A = \Psi_B$, we can obtain the equilibrium values of the decision variables:

$$r_A^* = \frac{1}{2\Omega} - \frac{R}{\gamma} \tag{4.17}$$

$$\Psi_A^* = \frac{\gamma}{8}(2h + \zeta) \tag{4.18}$$

Therefore, we can conclude that in equilibrium:

Proposition 3. *When the candidates are able to influence the voters with their investments in campaign:* $r_A^* = r_B^* = \max\left[0, \frac{1}{2\Omega} - \frac{R}{\gamma}\right]$ *and* $\Psi_A^* = \Psi_B^* = \max\left[0, \frac{\gamma}{8}(2h + \zeta)\right]$. *Thus, on the one side, the extractable rent is increasing and concave in* γ*, decreasing and convex in* Ω *and linearly decreasing in* R *and, on the other side, the equilibrium expenditure is linearly increasing in* h *and* γ *and* ζ*.*

Moreover, r^* *is ortogonal to* Ψ^**.*

Propositions 3 has several implications. The gross rent (r) is not influenced by the possibility of doing an electoral campaign (indeed, it is equal to the one found by Polo, 1998). Thus our model predicts that corruption and campaign expenses, in absence of special interest groups, should be uncorrelated. Empirically, when an increase of the second variable is observed, there should not be a significative effect on the first one.

However, a positive part of the expected rent is, in our model, invested in electoral campaign. The reason for this investment is, on one side, given by the possibility to overspend the adversary, influence the distribution of the voters' bias and therefore increase the chances of being elected and, on the other side, by the attempt of creating enough participation excitation to make some voters change side. In particular, the pressure to do active electoral campaign (and therefore the amount spent) increases when both the channels to influence the outcome are possible (and $\zeta > 0$). Thus, Proposition 3 induces the following corollary.

Corollary 4. *If campaign spending affects positively the utility of the voters ($\zeta > 0$), then the electoral campaign increases in size as a proportion of the amount of resources extracted.*

This is a result of the fact that collective excitation creates an additional incentive to campaign as, to do so, allows to potentiality move more votes. Indeed, under the conditions of Corollary 4, the electoral campaign increases the voters' welfare due to the collective excitation it generates.

Given the orthogonality of rents and campaign expenditures, the investment in the latter effectively transfers some of the extracted rent back to the citizens (at least in a non-monetary sense). Moreover, since both candidates spend an identical amount of money in equilibrium, all the citizens, regardless of their idiosyncratic bias, enjoy the electoral campaign to the same extent (though not all of them will be equally satisfied by the policy g chosen). In this regard, investing in campaign is similar to *pork barrel* spending (a candidate's promise of diverting a - disproportionate - amount of resources in favor of the voters that elect him). However, in this model, since the expenditures in electoral campaign are made before the elections (and not just promised), they are enjoyed by the whole electorate, not just by the voters of the candidate elected.

For candidates, doing an electoral campaign reduces welfare. Both of them would be strictly better off not spending but, in this setting, they are unable to commit themselves to this behavior. Therefore they find themselves in a *Prisoner's dilemma* situation and they both spend a significant amount of their expected rents in order to try to increase their chances of winning.

How much still remains for the candidates, after the electoral campaigns are paid for, is measured by net rents:

$$r_{net} = \frac{1}{2\Omega} - \frac{R}{\gamma} - \frac{\gamma}{8}(2h + \zeta) \tag{4.19}$$

As for the gross rent, the size of the net rent is increasing with the uncertainty about the popularity while, clearly, an increase in the marginal effects of expenditures in campaigns decreases the net rent. Finally, a decrease in the transaction cost γ has different effects on the net rent, depending on the values it assumes. Indeed, r_{net} is an increasing function of γ if $\gamma > \sqrt{\frac{8R}{2h+\zeta}}$ and a decreasing one otherwise. In this respect, notice that this threshold moves toward the unity when the office becomes more prestigious. This implies that, for very prestigious offices, a decrease in transaction cost always decreases the net rents (or leave them equal to zero if $r_{gross} \leq \Psi$). An explanation for this phenomenon follows. If R is big, competition is ceteris-paribus fierce, leaving little space for rent accumulation. However, this also increases the pressure to maximize the probability of election with more spending. In this condition, a decrease in the transaction cost makes the extraction of those little rent more efficient. However the expenditures grow faster than the gross rent thus decreasing the net rent.

Interestingly, Equation 4.19 shows that, under the conditions of corollary 4, the modernization and higher penetration of mass media (which make every dollar spent more effective) and better means of transport and campaigning (which allow voters to participate more to campaigns and feel more involved) increases the efficiency of the electoral competition, reducing the net rent of the candidate.

However, empirically this is not always observed. In some countries it has been shown that a lot of additional information discouraged and disenfranchised voters, who feel less and less involved in politics. Coherently, in this case, our model predicts an inverse relationship between disenfranchisement of voters and campaign expenditures:

Corollary 5. *If campaign spending affect negatively the utility of the voters ($\zeta < 0$), candidates reduce campaign expenses.*

Indeed, in this case the rationale for campaign spending is reduced. While spending is still useful to increase the probability of winning (through its effect on the electoral body as a whole), it also affects negatively the perception of the platform of the candidate.

4.5 Welfare Analysis

We have shown that while the electoral competition, in the presence of uncertainty, leads to positive gross rents, the candidates renounce to some of their rents in order to increase the chances of winning the elections. Moreover, we have shown that the net rents decrease when citizens utility is positively affected by campaign. However, some inefficiency in the electoral competition remains as net rents can be positive. In this section, we study the possibility of introducing public policies aimed at increasing efficiency.

The channel that has been explored by most of the literature on political advertisement to achieve this objective is the application of expenditure limits. Under this policy, the candidates are forbidden from investing in electoral campaign more than a given amount of resources. This kind of policy can be effective when the funding of campaigns comes from special interest groups but, in our setup such limitations are always welfare decreasing for the voters and welfare increasing for the candidates. The proof is intuitive. When voters enjoy the political campaign, reducing the maximum expenditures limit the amount of rent channelled back to the public in the form of utility of participation, without affecting corruption. Therefore the candidates are better off than in the absence of any limitation, but the voters always have their utility reduced by this kind of policy. This result is clearly driven the nature of our model where the candidates self-finance their own campaigns. Analogously, when voters do not enjoy political campaigns ($\zeta < 0$) spending limitations are useless, as the candidates are already willing to reduce their expenses in order to avoid hurting their electorate.

In a situation characterized by rent seeking candidates and by the absence of lobbies, a better policy is one that reduces their ability to extract rents, fighting the corruption of the political system, by abating the incentives to rent accumulation. Let us now assume that a constitutional level authority, or a international institution (that can impose its decisions on the candidates and the politicians), imposes an additional tax on the citizens (called τ_γ in the following) whose size is fixed exogenously. The revenues of this tax are used in order to reinforce the judiciary and police system and to make it harder, for candidates, to extract rents from the public budget, increasing the transaction cost associated to rent extraction (i.e., reducing γ).

The model is therefore modified, introducing an additional proportional tax on voters' income so that total taxation will now be equal to $\tau = g + r + \tau_\gamma$. Moreover, the expected utility of each of the candidates will be $p_J(R + (1-\tau_\gamma)\gamma r_J) - \psi_J^2$ [11] With this kind of policy, while the choice of public services (g), remains fixed at the value preferred by the voter with average income, the equilibrium gross rents and the expenditures become:

$$r^{**}_{A,B} = \frac{1}{2\Omega} - \frac{R}{\gamma(1-\tau_\gamma)} \tag{4.20}$$

$$\Psi^{**}_{A,B} = \frac{1}{8}\gamma(2h+\zeta)(1-\tau_\gamma) \tag{4.21}$$

Given that $0 \leq \tau_\gamma \leq 1$, both the gross rents and the campaign expenditures are reduced by this tax. Whether this policy also increases the efficiency of the political system (reducing net rents), depends on the prestige of the office and on the level of taxation. Indeed:

Proposition 6. *If the office is sufficiently prestigious, i.e. if* $R \geq \frac{\gamma^2}{8}(2h+\zeta)$*, net rents decrease at any level of taxation on corruption. However, if the office is not sufficiently prestigious, then net rents decrease if* $\tau_\gamma \geq 1 - \frac{8R}{\gamma^2(2h+\zeta)}$ *and increase otherwise.*

The level of additional taxation required to reduce net rents increases with the effectiveness of campaign and as the transaction costs decrease (since this creates stronger incentives

[11] This setup is the simplest possible. Indeed, the new tax reduces of $1-\tau_\gamma$ units the utility obtained from every unit of rent extracted. A richer specification could involve a generalization of this formula in which the effectiveness of τ_γ is an arbitrary function $f(\tau_\gamma)$ such that $f(0) = 1$ and $f(1) = 0$. Furthermore, the setup could be further extended taking into account the total amount of resources gathered, which depends on $\int_i y_i dF(y_i)$. This two additional elements would undoubtedly provide a larger wealth of results than those provided here. However, in the interest of clarity and tractability of the results we leave these extensions for future work.

to accumulate rent), and decreases with the prestigiousness of the office. In other words, the prestigiousness of the office acts as a substitute also for taxation on corruption (as discussed in Section 4.4, R makes the electoral process more efficient in reducing the rent extraction). Proposition 6 reflects the qualitative observation that, in countries where political offices are considered more prestigious (or where the civil servant aspect of a politician's job is more important) the amount of resources (monetary but also in terms of the attention of the public opinion) that needs to be dedicated to effectively fight corruption is more limited than in countries where politicians command little respect.

In order to assess the welfare properties of this policy, we now study the threshold levels of taxation that make each of the agents in this model better or worse off. For candidates, it can be proven (see appendix 4.7) that, as a consequence of the previous discussion on the net rents, it holds that:

Proposition 7. *A moderate tax, introduced to fight the rent-seeking behavior, can actually increase candidates' utility. This happens when* $\tau_\gamma < 2 - \frac{16}{\Omega(2h+\zeta)^2\gamma} \equiv \tau_c$

This is a consequence of the form of the utility function of the candidates. When the taxation is small, it reduces the incentives to rent extraction linearly, but it also abates the campaign expenditures quadratically (the campaign costs are assumed to be convex). So, while reducing the absolute size of corruption (gross rents), a moderate taxation also abates expenditures, diminishing the competition between the candidates, which is beneficial for the voters. As long as the candidates decide to make an active electoral campaign the trade-off described remains and ,only if the campaign disappears, the level of taxation for which the candidates are better off tends to zero. This result could also explain why some countries, with high levels of corruption, engage in actives that pretend to fight the phenomenon. A "cosmetic" anti-corruption policy may indeed help the politicians running for the office to overcome part of the prisoner dilemma that makes them overspend in campaign. While creating a façade of honesty these politicians would be actually working to improve their own welfare. Cases of failed corruption policies are widespread in particular in countries where political offices command little or no prestige (see Riley, 1998). Therefore, a policy maker who wants to adopt a tax of this kind, should take into account that if the office is not very attractive, then a limited amount of taxation could actually increase the amount of money that the candidates are able to extract being, de facto, counterproductive.

Let us now turn our attention to the effects of this tax on a given voter. The imposition of an additional tax has three distinct effects: on one side it diminish the welfare of the citizens, abating available income (y_i) and it reduces (see Equation 4.21) the opportunities of participation in political activities, on the other side, the elected candidate extracts a smaller rent, increasing the welfare of the voter. The net balance of these variations of welfare depends crucially on the value of the parameters and, in particular, on the gross income of the citizen (y_i). Indeed, it can be shown (see Appendix 4.7) that:

Proposition 8. *When the tax is sufficiently big, $\tau_\gamma > 1 - \frac{4Ry_i}{\gamma(y\zeta\gamma(2h+\zeta)+4y_i)}$, a voter with income y_i is better off as a result of taxation on corruption. However, if the office is not sufficiently prestigious, i.e. $R < (1-\tau_\gamma)\gamma$, then no voter is better off with the taxation.*

As for every tax, some citizens will be made better off by it, while some other will lose utility. Proposition 8 implies that each voter has a minimal level of taxation, below which the anti-corruption fight worsens his position (see Figure 4.4).

Given this result, it is clear that not every citizen would be willing to support a costly anti-corruption policy. Thus, also in the light of the result of Proposition 7, these results could be interpreted as a partial solution, offered by politicians to the corruption problem. Since only a part of the population is willing to pay to see corruption reduced, the candidates could use this willingness to introduce ineffective reforms.

The result of Proposition 8 can be generalized to the whole electoral body. The overall effect, on the electorate, of a tax-financed anti-corruption campaign is given by the difference between the population welfare with and without the new tax, i.e.:

$$W_{\tau>0} - W_{\tau=0} = \int_0^1 W^i_{\tau>0}\,\mathrm{d}y_i - \int_0^1 W^i_{\tau=0}\,\mathrm{d}y_i \tag{4.22}$$

The solution of Equation 4.22, yields (see Appendix 4.7 for details) the following proposition:

Proposition 9. *The population of voters is overall better off with the introduction of a tax on corruption if* $\tau_\gamma > 1 - \frac{2R}{\gamma(\gamma\zeta y(2h+\zeta)+2)} \equiv \tau_v$.

Everything else being constant, it is clear that an important role is played by y, the average income of the electorate. The higher it is, the higher is the proportional tax rate [12] required for the population to be better off thanks to corruption-fight. This implies that populations with an positively skewed income distribution will be, in principle, better off at lower levels of taxation than a population with relatively more rich.

Will the median voter be better off with or without the taxes that make the whole population better off? This question is relevant, since in most countries, popular support for a policy that acts against the dominant class (here the political class) needs to be strong to allow its implementation. Income distribution turns out to be pivotal in establishing popular support for such measures. Indeed, it is possible to show (see Appendix 4.7) that:

Proposition 10. *For any symmetric distribution the threshold of taxation on corruption that makes the population better off is just equal to the one that makes the median voter better off.*

For any positively (resp. negatively) skewed distribution, at the threshold above which the population is better off with the taxes, the median voter is worse (resp. better) off with the taxes.

Proposition 10 offers an explanation of why, in countries with comparatively many poor (where the median voter's income is below the mean income), it is difficult to effectively

[12] Note that we assume the anti-corruption tax to be a proportional tax. Other types of tax schedule could lead to different results

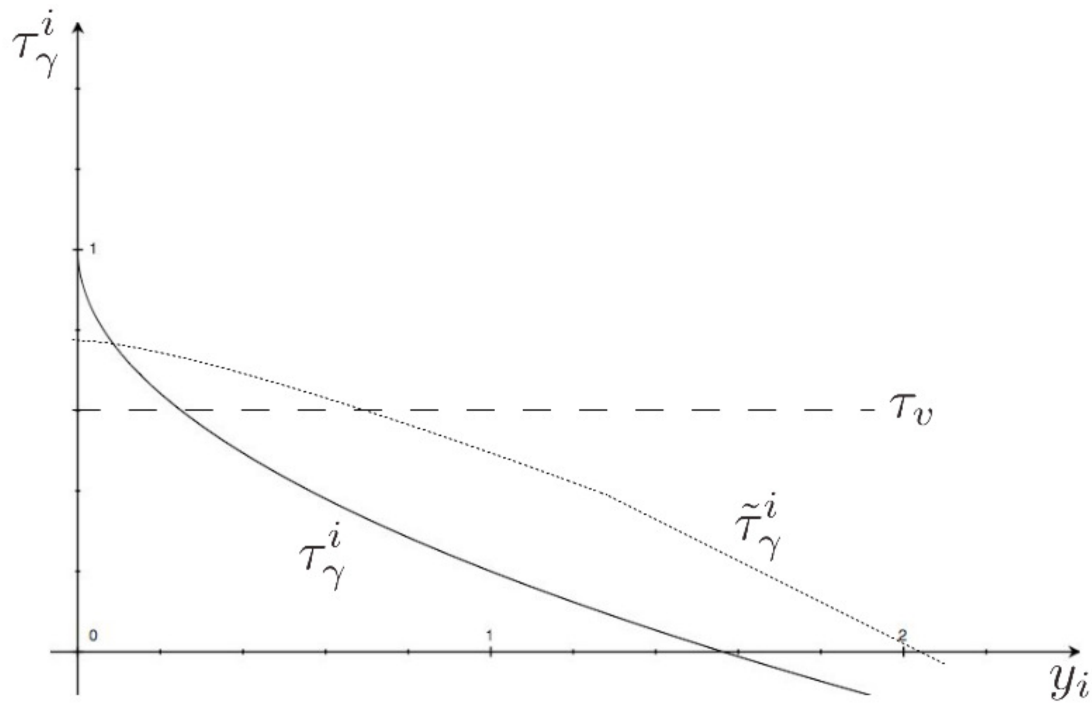

Figure 4.4: Graphical representation of Proposition 8. The continuous line represents the minimal tax rate necessary for a voter with income y_i to increase his welfare. The dotted line instead represents the tax rate above which the agent is better off without any government. Therefore, the area between the two lines is the set of tax rates that makes the citizen better off with the anti-corruption policy. Finally, the dashed line represents the level of taxation on corruption that makes the population as a whole better off (from Proposition 9). At this level, however, some citizens (the poorer ones) will be worse off with the tax.

implement anti-corruption policies of the kind proposed here. Indeed, when the median voter is worse off with the policy, this implies that the level of taxation required to increase the population welfare lacks the popular support needed for the implementation (in a democratic country).

Propositions 7 and 9, taken together, indicate that there is a possible conflict of interest between voters and politicians on the strength of the anti-corrution measures to implement. Indeed, putting together the two thresholds we obtain that $\tau_v > \tau_c$ if $0 < \Omega < \frac{64+16y\zeta(2h+\zeta)\gamma}{(2h+\zeta)^2(4R+\gamma(4+y\zeta\gamma(2h+\zeta)))}$. When this condition is satisfied the choice of an amount of taxation between τ_v and τ_c makes both of the sides better off. However when the condition is not satisfied, the choice is between a level of taxation that makes one side worse off and a level that decreases the welfare of both sides.

Finally, imagine that a benevolent social planner, before the game, has to set the optimal rate of corruption-prevention taxation. He will choose τ_γ in order to maximize the electoral body's welfare, i.e.,

$$\max_{\tau\gamma} \int_0^1 (y - g - r - \tau_\gamma) \frac{y_i}{y} + H(g) + 2\zeta\Psi \quad \mathrm{d}y_i \tag{4.23}$$

Plugging in the equilibrium values from Equations 4.20 and 4.21 and solving the maximization problem we obtain the equilibrium taxation:

$$\tau_\gamma^* = 1 - \frac{\sqrt{2}R}{\sqrt{R\gamma(2 + y\zeta(2h + \zeta)\gamma)}} \tag{4.24}$$

The social optimum in Equation 4.24 is always bigger than the threshold established by Proposition 9, except when the ego-rent associate with the office is particularly low (i.e, if $R < \frac{\gamma}{2}\left[2 + y\zeta\gamma(2h + \zeta)\right]$). In this case, as already discussed, the anti-corruption tax is never able to make the population better off, so the social planner chooses $\tau_\gamma^* = 0$. This condition confirms how, in order to have efficient public policies it is pivotal to establish a minimum level of prestige associated to the office to protect from corruption.

Summarizing, we have shown that the introduction of an anti-corruption tax can make the citizens better off. Surprisingly, it may also make the candidates better off when the gross rent that they extract decrease less than the expenditure in electoral campaign. We finally establish the conditions under which a policy of this kind can achieve the popular support required for an effective implementation.

4.6 Conclusions

Our aim in this paper is to study the relationship between corruption and campaign expenditures in a context characterized by the absence (or the limited relevance) of political lobbies. To this end, we introduce several innovations on the classical framework of the probabilistic voting models (Polo, 1998 and Persson & Tabellini, 2000), with office motivated (rent-seeking) candidates and ideological voters. We explicitly model the campaign spending of the candidates together with the multiple types of influences that it has on the voting decisions of the population. A particular feature of our paper, that distinguishes it from the current literature, is that we establish a direct link between electoral expenditures of candidates and utility of voters. The utility changes (increasing or decreasing depending on which stylized fact we want to capture) if voters observe that candidates are doing an active electoral campaign.

Our main result is that rent extraction and campaign spending are orthogonal variables in equilibrium. This allows us to predict, that in countries where lobbies have a limited role, we should not find a significative relationship between the amount of resources invested in campaign and the scale of corruption of the administration that gets elected. Moreover, we find that in the majority of cases candidates spend a positive part of their expected rent in order to finance their electoral campaign. In itself this result is a qualitative confirmation of the realism of the model. Indeed, very rarely politicians abstain from doing electoral campaign, even when they are forced to self-finance it. On this regard, our model produces a testable prediction: when voters find electoral campaigns exciting, candidates spend more. How to test this prediction? It is difficult to empirically disentangle the difference between h, the effect of advertisement on the general population, and ζ, the effect of campaigning on voter's utility. A possible solution is to consider the level of technology and penetration of mass media as proxies for h and the amount of campaign events in an geographical area as a proxy for ζ. Indeed, mass media advertisements tend to affect a population equally, but do not lead to significative "psychological gains" (Sanders, 2001) in terms of participation opportunities to the election procedure. On the contrary, political rallies and opportunities of employment in the campaign of a candidate, directly stimulate individual citizens to be involved in the political process. Studying the differences between contested (where usually most of the campaign events are concentrated) and uncontested electoral districts, it could be possible (with an empirical strategy similar to the one proposed by Frey and Stutzer 2000, 2005) to estimate the effect of political campaigns on voters contentment with the political system [13]. While a large number of socio-economic controls should be added in order to take into account the local specificities, this seems a promising way of testing our results about "campaign excitement".

In this paper, we also study the effects of the introduction of public policies aimed at reducing the inefficiencies of the electoral competition. In this respect, our model predicts that campaign spending limitations are always welfare decreasing for the voters. This policy should therefore be avoided, at least in contexts where special interest groups have a limited importance. Another way of interpreting this result is to advice policy makers to forbid completely contributions from organized groups as, when campaigns create utility, this policy allows the electoral competition to increase the voters' welfare, without the burden of the policy costs implied by lobbies' contributions.

An alternative policy we study is to fight corruption with investments in police and justice financed by an additional tax paid by the electoral body. A policy of this type may reduce the net rents (and therefore the inefficiency of the electoral competition) when the electoral office to protect is highly prestigious. However, when this condition does not hold, this policy can work in favor of corrupt politicians, increasing their utility. This result provides a theoretical explanation for the behavior of many corrupt politicians who, in countries with

[13]Evidence of the fact that the opportunity of political participation (i.e. working, without wage, in a candidate's campaign) induces positive changes in voters perceived quality of life, or happiness, has been shown by Weitz-Shapiro & Winters (2008) and by Pacheco & Lange (2010). For our purposes it would be important to demonstrate the existence of an effect of political events, such as campaign rallies, on voter's perception of the political system (rather than simply on those that choose to participate to campaigns).

high levels of public embezzlement, often introduce weak anti-corruption policies, that turn out to help them increasing their rent through a reduction of the positive effects of electoral competition. Some of the most problematic democracies, such as Bolivia, Venezuela, Senegal and Russia, present clear examples where politicians claim to fight corruption but use this policy to create a façade of honesty while still allowing themselves to retain a significative share of the public budget for private consumption.

The last important result of this paper is that it uncovers a relationship between the tax-level required, in a given population, to achieve welfare improvements through an anti-corruption policy and the level of popular support that this policy requires in order to be implemented in a democratic setup. We show that it is difficult to implement effective policies in countries with large income inequalities, as in this case the majority of voters will oppose the introduction of a (sufficiently effective) anti-corruption law. In these environments, only an non-elective, external body can hope to reestablish efficiency.

While this paper still has some limitations, our setup is very flexible and, at this stage, we are able to envision several ways to extend it. Regarding the effects of political campaign on voters' utility (the main innovation of this paper), with our parametrization, we are able to capture the excitement that can be generated as well as the feeling of disenfranchisement of voters due to excessive quantities of information. However, in this second case, a possible effect of electoral campaign is neglected: as a consequence of too much information some voters could decide not to participate to elections. Making the turnout endogenous, and directly linking the voters' (dis-)utility from campaign with their participation probability could help us capturing an additional trade-off in the decisions of the candidates: the one between the need to campaign in order to get a maximum of the participants' votes and the need to maximize the share of the population that goes to the ballots.

Finally, a more general specification of the anti-corruption tax, could generate richer results in terms of welfare analysis. In this context, the introduction of re-election concerns could create the disciplinary device required to induce politicians to apply effective anti-corruption policies.

4.7 Appendix

Proof of Proposition 7

In the case of no tax on corruption the expected value of the candidates' utility is given by Equation 4.1, while introducing the tax on corruption, it becomes: $p_J(R+(1-\tau_\gamma)\gamma r_J)-\Psi_J^2$. In equilibrium, given the symmetry of the model, the probability of being elected is always equal to $\frac{1}{2}$. Introducing the relevant equilibrium values for r and Ψ (Equations 4.17 and 4.18 for the case without taxes and Equations 4.20 and 4.21 for the case with anti-corruption taxes), we can now calculate the difference between the expected values in the two cases:

$$E(v_j^{\tau>0}) - E(v_j^{\tau=0}) = -\frac{\gamma\tau_\gamma(16-(2h+\zeta)^2\gamma\Omega(2-\tau_\gamma))}{64\Omega} \tag{4.25}$$

We are interested to know under which conditions on τ_γ the candidates are better off (when Equation 4.25 is positive). Requiring this equation to be bigger than zero and solving for τ_γ, we find the result of Proposition 7.

Proof of Proposition 8

In the absence of anti corruption policies, the welfare of a citizen with income y_i is given by Equation 4.2 while, with the taxation it becomes: $W_i = (y-g-r-\tau_\gamma)\frac{y_i}{y} + H(g) + \zeta(1 - d_i^J\Psi_J$. We can introduce the relevant equilibrium values for g, r and Ψ into the equations of welfare and calculate the difference between the two expressions [14]:

$$W_i^{(\tau>0)} - W_i^{(\tau=0)} = \frac{y_i}{y}\left(\frac{R}{\gamma}\frac{\tau}{1-\tau_\gamma} - \tau_\gamma\right) - \frac{1}{4}\zeta\gamma\tau_\gamma(2h+\zeta) \tag{4.26}$$

[14] By the symmetry of the model $\Psi_A = \Psi_B$. Moreover, since $d_i^A = 1 - d_i^B$, for every voter i, $d_i^A + d_i^B = 1$

In order to compute the conditions for the voter i to be better off with the taxes, we impose that Equation 4.26 must be bigger than zero, i.e.:

$$\frac{y_i}{y}\frac{R-\gamma(1-\tau_\gamma)}{\gamma(1-\tau_\gamma)} > \frac{1}{4}\zeta\gamma(2h+\zeta) \tag{4.27}$$

It is immediately evident that since the RHS of the last equation is always positive, if $R < \gamma(1-\tau_\gamma)$ then the inequality is never satisfied. Finally, solving Equation 4.27 for τ_γ we obtain the other result of Proposition 8.

Proof of Proposition 9

When no anti-corruption taxes are introduced, the ex-post welfare of the population of voters is given by:

$$W_{\tau=0} = \int_0^1 W^i_{\tau=0}\,\mathrm{d}y_i = \int_0^1 \left(y - g^* - \frac{1}{2\Omega} + \frac{R}{\gamma}\right)\frac{y_i}{y} + H(g^*) + \frac{1}{4}\zeta\gamma(2h+\zeta)\,\mathrm{d}y_i \tag{4.28}$$

While when these taxes are present, it becomes:

$$W_{\tau>0} = \int_0^1 W^i_{\tau>0}\,\mathrm{d}y_i = \int_0^1 \left(y - g^* - \frac{1}{2\Omega} + \frac{R}{(1-\tau_\gamma)\gamma} - \tau_\gamma\right)\frac{y_i}{y} + H(g^*) + \frac{1}{4}\zeta\gamma(2h+\zeta)(1-\tau_\gamma)\,\mathrm{d}y_i \tag{4.29}$$

Equations 4.28 and 4.29 can be simplified respectively as:

$$W_{\tau=0} = \frac{1}{4}\left(2 + \frac{2R}{y\gamma} + \gamma\zeta(2h+\zeta) - \frac{1}{y\Omega}\right) \tag{4.30}$$

$$W_{\tau>0} = \frac{2R\Omega + \gamma(\tau_\gamma - 1)(1 + y\Omega(\gamma\zeta(2h+\zeta)(\tau_\gamma - 1) - 2) + 2\tau_\gamma\Omega)}{4y\gamma\Omega(1-\tau_\gamma)} \tag{4.31}$$

To establish our result, we need that $W_{\tau>0} - W_{\tau=0} > 0$. Solving this inequality for τ_γ yields the condition summarized in Proposition 9.

Proof of Proposition 10

Consider a symmetric distribution of income between 0 and 1. By definition $y = y^M = 0.5$. Therefore, the threshold taxation that makes the median voter better of (the one of

Proposition 8 with $y_i = y^M = 0.5$) becomes:

$$\tau_\gamma > 1 - \frac{2R}{\gamma(0.5\gamma\zeta(2h+\zeta)+2)} \tag{4.32}$$

which, given that $y^M = y$, is equal to the level of taxation that makes the whole population better off (see Proposition 9).

Consider now a positively skewed distribution bounded between zero and one [15]. For any such distribution $y^M < y < 0.5$. Proceeding as above we can find the threshold that makes the median voter better off:

$$\tau_\gamma^M > 1 - \frac{4Ry^M}{\gamma(y\gamma\zeta(2h+\zeta)+4y^M)} \tag{4.33}$$

From Proposition 9 we know the threshold of the population (let us call it τ_γ^P). We want to show that, at τ_γ^P, the median voter is worse off with the taxes. For this to be true it must hold that his threshold is higher than the one of the population: $\tau_\gamma^M > \tau_\gamma^P$. This inequality is satisfied for $y^M < \frac{1}{2}$. Since we know that for every positively skewed distribution between 0 and 1, $y^M < y < 0.5$, the inequality is always satisfied. In the same way, since for any negatively skewed distribution $y^M > y > 0.5$, we can prove that the median voter is always better off with taxes at the level at which the population is indifferent.

[15] The Kumaraswamy Distribution is an example of this kind of distribution.

Bibliography

Introduction

[1] Aldashev G., Carletti T., Righi, S. (2011), **Follies subdued: Informational Efficiency under Adaptive Expectations and Confirmatory Bias**, in *Journal of Economic Behavior and Organization*, volume 80, Issue 1, pp. 110-121

[2] Bagnoli, F. and Carletti, T. and Fanelli, D. and Guarino, A. and Guazzini, A. (2007), **Dynamical affinity in opinion dynamics modeling**, *Physical Review E*, Vol 76, pp. 066105.

[3] Biondi Y., Giannoccolo P., Galam S. (2011) **The formation of share market prices under heterogeneous beliefs and common knowledge**, arXiv:1105.3228v1 [physics.soc-ph], `http://arxiv.org/abs/1105.3228`

[4] Carletti, T. (2009), **Stochastic Weighted Fractal Networks**, *preprint.*

[5] Carletti, T., Fanelli D., Righi S., (2011) **Emerging structures in social networks guided by opinions exchanges**, in *Advances in Complex Systems*, volume 14, Issue 1, pp. 13-30

[6] Carletti T., Righi S. (2008) **Shaping Opinions in a Social Network**, in *Proceedings of Wivace 2008 Italian Workshop of Artificial Life and Evolutive Computing*, `http://perso.fundp.ac.be/~srighi/papers/cproceed/shaping.pdf`

[7] Carletti, T., Righi S. (2010) **Weighted Fractal Networks**, *Physica A*, volume 389, pp. 2134-2142

[8] Cattuto C. et al. (2010), **Dynamics of person-to-person interactions from distributed RFID sensor networks**, *PLoS ONE*, Vol. 5 (7): e11596

[9] Granovetter, M. (1983), **The strength of weak ties: A Network theory revised**, *Social Theory*, Vol. 1, pp. 201.

[10] Hayek, F. (1944), **The Road to Serfdom**, *Routledge*, London.

[11] Hayek, F. (1945), **The uses of knowledge in society**, *American Economic Review*, Vol. 35, pp. 519-530.

[12] Mandelbrot, B.B. (1975) **Les Object Fractals: formes, hasards et dimension**, Flammarion, Paris.

[13] Righi S., Carletti T. (2009a), **The influence of social network topology in a opinion dynamics model**, in *Proceeding of the European Conference on Complex systems 2009*, `http://perso.fundp.ac.be/~srighi/papers/cproceed/topology.pdf`

[14] Righi S., Carletti T., (2009b), **How Opinion Dynamics Shapes Social Networks Topology**, in *Proceedings of AI*IA Conference (Italian Association for artificial Intelligence)*, `http://perso.fundp.ac.be/~srighi/papers/cproceed/opinion.pdf`

[15] Righi S., Carletti T., Aldashev G. (2011), **Behavioral Biases and Informational Inefficiency in an Agent-Based Financial Market**, *preprint*, `http://perso.fundp.ac.be/~srighi/papers/wprogress/behavioral.pdf`

[16] Watts, D.J. and Strogatz, S.H. (1998), **Collective dynamics of 'small-world' networks**, *Nature*, Vol. 393, pp. 440.

Weighted Fractal Network

[17] Albert, R. and Barabási, A.-L. (2002), **Statistical mechanics of complex networks**, *Reviews of Modern Physics*, Vol. 74, pp. 47.

[18] Almeida, M.P., Andrade Jr., J.S., Buldyrev, S.V., Cavalcante, F.S.A., Stanley, H.E. and Suki, B. (1999), **Fluid flow through ramified structures**, *Physical Review E*, Vol. 60 (5), pp. 5486.

[19] Andrade, J.S. et al. (1998), **Asymmetric flow in symmetric branched structures**, *Physical Review Letters*, Vol. 81(4), pp. 926.

[20] Banavar, J.R., Maritan, A. and Rinaldo, A. (1999), **Size and form in efficient transportation networks**, *Nature*, Vol. 399, pp. 130.

[21] Barabási, A.-L. and Albert, R. (1999) **Emergence of Scaling in Random Networks**, *Science*, Vol. 286, pp. 509.

[22] Barabási, A.-L., Buldyrev, S.V., Stanley, H.E. and Suki, B. (1996), **Avalanches in lung: a statistical mechanical model**, *Physical Review Letters*, Vol. 76(12), pp. 2192.

[23] Barabási, A.-L., Ravasz, E. and Vicsek, T. (2001), **Deterministic scale–free networks**, *Physica A*, Vol. 299, pp. 559.

[24] Barnsley, M. (1988), **Fractals everywhere**, *Academic Press London.*

[25] Barrat, A., Barthélemy, M., Pastor-Satorras, R. and Vespignani, A. (2004), **The architecture of complex weighted networks**, *Proceedings of the National Academy of Sciences*, Vol. 101, pp. 3747.

[26] Barrat, A., Barthélemy, M. and Vespignani, A. (2004), **Modeling the evolution of weighted networks**, *Physical Review E*, Vol. 70, pp. 066149.

[27] Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. and Hwang, D.-H. (2006), **Complex networks: Structure and dynamics**, *Physics Reports*, Vol. 424, pp. 175.

[28] Carletti, T. (2009), **Stochastic Weighted Fractal Networks**, *preprint.*

[29] Erdős, P. and Rényi, A. (1959), **On random graphs**, *Publicationes Mathematicae Debrecen*, Vol. 6, pp. 290.

[30] Dorogovtsev, S.N., Goltsev, A.V. and Mendes, J.F.F. (2002), **Pseudofractal scale–free web**, *Physical Review E*, , Vol. 65, pp. 066122.

[31] Dorogovtsev, S.N. and Mendes, J.F.F. (2004), **Minimal model of weighted scale–free networks**, preprint *cond-mat/0408343v2.*

[32] Edgar, G.A. (1990), **Measure, Topology and fractal geometry**, *UTM*, Springer–Verlag, New York .

[33] Fortunato, S. (2010), **Community detection in graphs**, *Phys. Rep.*, Vol. 486, pp. 75.

[34] Guan, J., Wu, Y., Zhang, Z., Zhou, S. and Wu, Y. (2009), **A unified model for Sierpinski networks with scale-free scaling and small-world effect**, *Physica A*, Vol. 388, pp. 2571-2578.

[35] Jung, S., Kim, S. and Kahng, B. (2002), **Geometric fractal growth model for scale–free networks**, *Physical Review E*, Vol. 65, pp. 056101.

[36] Kamiya, A. and Takahashi, T. (2007), **Quantitative assessments of morphological and functional properties of biological trees based on their fractal nature**, *Journal of Applied Physiology*, Vol. 102, pp. 2315.

[37] Kitaoka, H. and Suki, B. (1997), **Branching design of the bronchial tree based on diameter–flow relationship**, *Journal of Applied Physiology*, Vol. 82, pp. 968.

[38] Mäkinen, V. **Himmeli, a free software package for visualizing complex networks**, available at *http://www.artemis.kll.helsinki.fi/himmeli*.

[39] Mandelbrot, B.B. (1982), **The Fractal Geometry of Nature**, *W.H. Freeman and Company*, New York.

[40] Moreira, A.A., Andrade Jr., J.S., Herrmann, H.J. and Indekeu, J.O. (2009), **How to make a fragile network robust and vice versa**, *Physical Review Letters*, Vol. 102, pp. 018701.

[41] Radicchi, F., Ramasco, J.J., Barrat, A. Fortunato, S. (2008),**Complex Networks Renormalization: Flows and Fixed Points**, *Physical Review Letters*, Vol. 101, pp. 148701.

[42] Radicchi, F., Barrat, A., Fortunato, S. and Ramasco, J.J. (2009), **Renormalization flows in complex networks**, *Physical Review E*, Vol. 79, pp. 026104.

[43] Ravasz, E. and Barabási, A.-L. (2003), **Hierarchical organization in complex networks**, *Physical Review E*, Vol. 67, pp. 026112.

[44] Saramäki, J., Kivelä, M., Onnela, J.-P., Kaski, K. and Kertész, J. (2007), **Generalizations of the clustering coefficient to weighted complex networks**, *Physical Review E*, Vol. 75, pp. 027105.

[45] Song, C., Havlin, S. and Makse, H.A. (2005), **Self-similarity of complex networks**, *Nature*, Vol. 433, pp. 392.

[46] Song, C., Havlin, S. and Makse, H.A. (2006) **Origin of fractality in the growth of complex networks**, *Nature Physics*, Vol. 2, pp. 275.

[47] Suki, B., Barabási, A.-L., Hantos, Z., Peták, F. and Stanley, H.E. (1994), **Avalanches and power–law behaviour in lung inflation**, *Nature*, Vol. 368, pp. 615.

[48] Watts, D.J. and Strogatz, S.H. (1998), **Collective dynamics of 'small-world' networks**, *Nature*, Vol. 393, pp. 440.

[49] West, G.B., Brown, J.H. and Enquist, B.J. (1997), **A General Model for the Origin of Allometric Scaling Laws in Biology**, *Science*, Vol. 276, pp. 122.

[50] Yook, S.H., Jeong, H. and Barabási, A.-L. (2001), **Weighted evolving networks**, *Physical Review Letters*, Vol. 86(25), pp. 5835.

[51] Zhang, Z., Guan, J., Chen, L., Yin, M. and Zhou, S. (2010), **Mapping Koch curves into scale-free small-world networks**, Journal of Physics A, Vol.43(39), pp 1-16.

[52] Zhang, Z., Zhou, S., Su, Z., Zou, T. and Guan, J. (2008) **Random Sierpinski network with scale-free small-world and modular structure**, *The European Physical Journal B*, Vol. 65, pp. 141-147.

Endogenous Social Network under Opinion Exchange

[53] Amblard, F. and Deffuant, G. (2004), **The role of network topology on extremism propagation with the relative agreement opinion dynamics**, *Physica A*, Vol. 343, pp. 725

[54] Albert, R. and Barabási, A.-L. (2002), **Statistical mechanics of complex networks**, *Reviews of modern physics*, Vol. 72, pp.47.

[55] Bagnoli, F. and Carletti, T. and Fanelli, D. and Guarino, A. and Guazzini, A. (2007), **Dynamical affinity in opinion dynamics modeling**, *Physical Review E*, Vol 76, pp. 066105.

[56] Barrat, A. and Barthélemy, M. and Pastor-Satorras, R. and Vespignani, A. (2004), **The architecture of complex weighted networks**, *Proceedings of the National Academy of Sciences*, Vol. 101, pp. 3747.

[57] Berlow, E.L. (1999), **Strong effects of weak interactions in ecological communities**, *Nature*, Vol. 398, pp. 330.

[58] Boccaletti, S. and Latora, V. and Moreno, Y. and Chavez, M. and Hwang, D.-U. (2006), **Complex networks: Structure and dynamics**, *Physics Reports*, Vol. 424, pp. 175.

[59] Bordogna, C.M. and Albano, E.V. (2007), **Dynamic behavior of a social model for opinion formation**, *Physical Review E*, Vol. 76, pp. 061125.

[60] Carletti, T. and Fanelli, D. and Grolli, S. and Guarino, A. (2006), **How to make an efficient propaganda**, *Europhysics Letters*, Vol. 74(2), pp. 222.

[61] Carletti, T. and Fanelli, D. and Guarino, A. and Bagnoli, F. and Guazzini, A. (2008a), **Birth and death in a continuous opinion dynamics model - The consensus case**, *European Physical Journal B*, Vol. 64(2), pp. 285.

[62] Carletti, T. and Fanelli, D. and Guarino, G. and Guazzini, A. (2008b), **Meet, Discuss and Trust each other: large versus small groups**, in *Proc. WIVACE2008 Workshop Italiano Vita Artificiale e Computazione Evolutiva*, Sept. 2008, Venice, Italy.

[63] Carletti, T. and Fanelli, D. (2008), **Collective observables in repeated experiments of population dynamics**, *arxiv:0810.5502v1 [cond-mat.stat-mech].*

[64] Castellano, C. and Fortunato, S. and Loreto, V. (2009), **Statistical physics of social dynamics**, *Reviews of Modern Physics*, Vol. 81, pp. 591.

[65] Csermely, P. (2009), **Weak links**, Springer Verlag, Heidelberg, Germany.

[66] Deffuant, G. and Neau, D. and Amblard, F. and Weisbuch, G.(2000), **Mixing beliefs among interacting agents**, *Advances in Complex Systems*, Vol. 3(4), pp. 87.

[67] Evans, L.C. (1998), **Partial Differential Equations**, *American Mathematical Society*, Providence.

[68] Festinger, L. and Carlsmith, J.M. (1959), **Cognitive consequences of forced compliance**, *Journal of Abnormal and Social Psychology*, Vol. 58, pp. 203.

[69] Fortunato, S. (2010), **Community detection in graphs**, *Physics Report*, Vol. 486 , pp. 75.

[70] Galam, S. (2008), **A review of Galam models**, *International Journal of Modern Physics*, Vol. 19(3), pp. 409.

[71] Granovetter, M. (1983), **The strength of weak ties: A Network theory revised**, *Social Theory*, Vol. 1, pp. 201.

[72] Gross, T. and Blasius, B. (2008), **Adaptive coevolutionary networks: a review**, *Journal of the Royal Society Interface*, Vol. 5(20) , pp. 259.

[73] Holme, P. and Newman, M.E.J. (2006), **Nonequilibrium phase transition in the coevolution of networks and opinions**, *Physical Review E*, Vol. 74, pp. 056108.

[74] Klimek, P. and Lambiotte, R. and Thurner, S. (2008), **Opinion formation in laggard societies**, *Europhysics Letters*, Vol. 82, pp. 28008.

[75] Kozma, B. and Barrat, A. (2008), **Consensus formation on adaptive networks**, *Physical Review E*, Vol. 77, pp. 016102.

[76] Milgram, S. and Traver, J. (1969), **An Experimental Study of the Small World Problem**, *Sociometry*, Vol. 32(4), pp. 425 .

[77] Newman, M.E.J. and Watts, D.J. (1999), **Scaling and percolation in the small-world network model**, *Physical Review E*, Vol. 60(6), pp. 7332.

[78] Onnela, J.-P. and Saramaki, J. and Hyvonen, J. and Szabó, G. and Lazer, D. and Kaski, K. and Kertész, J. and Barabási, A.-L. (2007), **Structure and tie strengths in**

mobile communication networks, *Proceedings of the National Academy of Sciences*, Vol. 104, pp.7332.

[79] Schweitzer, F. and Hołys, J.A. (2000), **Modelling Collective Opinion Formation by Means of Active Brownian Particles**, *European Physical Journal B*, Vol. 15, pp. 723 .

[80] Watts, D.J. and Strogatz, S.H. (1998), **Collective dynamics of small-world networks**, *Nature*, Vol. 393, pp. 440 .

[81] Zimmermann, M.G. and Eguíluz, V.M. and San Miguel, M. (2004), **Co-evolution of dynamical states and interactions in dynamic networks**, *Physical Review E*, Vol. 69, pp. 065102.

Informational Efficiency under Adaptive Expectations and Confirmatory Bias

[82] Blouin, M., and Serrano, R. (2001), **A decentralized market with common values uncertainty: Non-steady states** *Review of Economic Studies*, Vol. 68, pp. 323-346.

[83] Brock, W.A., Durlauf, S.N. (2001), **Discrete choice with social interactions**, *Review of Economic Studies*, Vol. 68, pp. 235-260.

[84] Deffuant, G., Neau, D., Amblard, F., and Weisbuch, G. (2000), **Mixing beliefs among interacting agents**, *Advances in Complex Systems*, Vol. 3, pp. 87-98.

[85] Duffie, D., and Manso, G. (2007), I**Information percolation in large markets** *American Economic Review*, Vol. 97, pp. 203-209.

[86] Gode, D., and Sunder, S. (1993), **Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality** *Journal of Political Economy*, Vol. 101, pp. 119-137.

[87] Gode, D., and Sunder, S. (1997), **What makes markets allocationally efficient?**, *Quarterly Journal of Economics*, Vol. 112, pp. 603-630.

[88] Grossman, S. (1976), **On the efficiency of competitive stock markets when traders have diverse information**, *Journal of Finance*, Vol. 31, pp. 573-585.

[89] Haruvy, E., Lahav, Y., and Noussair, C. (2007), **Traders' expectations in asset markets: Experimental evidence**, *American Economic Review* , Vol. 97, pp. 1901-1920.

[90] Hayek, F. (1945), **The uses of knowledge in society**, *American Economic Review*, Vol. 35, pp. 519-530.

[91] LeBaron, B. (2001), **A builder's guide to agent-based financial markets**. *Quantitative Finance*, Vol. 1, pp. 254-261.

[92] Milgrom, P. (1981), **Rational expectations, information acquisition, and competitive bidding** *Econometrica*, Vol. 49, pp. 921-943.

[93] Rabin, M., and Schrag, J. (1999), **First impressions matter: A model of confirmatory bias** *Quarterly Journal of Economics*, Vol. 114, pp. 37-82.

[94] Shiller, R. (1984), **Stock Prices and Social Dynamics**, *Brookings Papers on Economic Activity*, Vol. 2, pp. 457-510.

[95] Weisbuch, G., Deffuant, G., Amblard, F., and Nadal, J. (2002) **Meet, discuss and segregate!**, *Complexity*, Vol. 7, pp. 55-63.

[96] Wilson, R. (1977), **Incentive efficiency of double auctions**, *Review of Economic Studies*, Vol. 44, pp. 511-518.

[97] Wolinsky, A. (1990), **Information revelation in a market with pairwise meetings**. *Econometrica*, Vol.58, pp. 1-23.

Campaign Spending and Rents in a Probabilistic Voting Model

[98] Austen-Smith D. (1987), **Interest groups, campaign contributions and probabilistic voting**, *Public Choice*, 54, pp. 123-139.

[99] Abrams B.A., Settle R .E. (1976), **A modest proposal for election reform**, *Public Choice*, Vol. 26, No. 1 (Winter), pp 37-54.

[100] Black D. (1948), **On the Rationale of Group Decision-making**, *Journal of Political Economy*, Vol. 56, No. 1, pp. 23-34.

[101] Coate S. (2004), **Pareto-Improving Campaign Finance Policy**, *American Economic Review*, Vol. 94, No. 3, pp 628-655.

[102] Downs A. (1957), **An Economic Theory Democracy**, Harper, New York.

[103] Freie J.F. (1997), **The effects of campaign participation on political attitude**, *Political Behaviour*, Vol 19, No. 2.

[104] Grillo M., Polo M. (1993), **Political Exchange and the Allocation of Surplus a Model of Two-Party Competition**, in A. Breton and G. Galeotti (eds.), *Preferences and Democracy*, Dordrecht, Kluwer.

[105] Frey S.B., Benz M., Stutzer A. (2004), **Introducing Procedural Utility: Not Only What, but Also How Matters**, *Journal of Institutional and Theoretical Economics*, Mohr Siebeck, Tbingen, Vol. 160, No. 3, pp. 377-401.

[106] Frey S.B., Stutzer A. (2000), **Happiness Prosper in democracy**, *Journal of Happiness Studies*, Vol. 1, pp. 79-102.

[107] Frey S.B., Stutzer A. (2005), **Beyond outcomes: measuring procedural utility**, *Oxford Economic Papers*, Vol. 57, No. 1, pp. 90-111.

[108] Grossman G.M., Helpman, E. (2001), **Special Interest Politics**, Cambridge and London: MIT Press.

[109] Johnson-Cartee K.S. (1989), **Southern voters' reaction to negative political ads in the 1986 election**, *Journalism Quarterly.*

[110] Kaid, LL. (1987), **An experimental study of the effectiveness of negative political advertisements**, *Communication Quarterly.*

[111] Levitt, D. (1994), **Using Repeat challengers to estimate the effect of campaign spending on election outcomes in the US House**, *Journal of Political Economics*, Vol. 102, pp. 777-798.

[112] Matthews D. (1998), **The Black-Sheep Effect: How Positive and Negative Advertisements Affect Voters' Perceptions of the Sponsor of the Advertisement**, *Journal of Applied Social Psychology*, Vol. 28, No. 20, pp. 1903-1915.

[113] Myerson R.B. (1993), **Effectiveness of Electoral Systems for Reducing Government Corruption: A Game-Theoretic Analysis**, *Games and Economic Behavior*, Elsevier, Vol. 5, No. 1, pp 118-132.

[114] Pacheco G., Lange T. (2010), **Political participation and life satisfaction: a cross-European analysis**, *International Journal of Social Economics*, Vol. 37, No. 9, pp.686 - 702.

[115] Palda F., Palda K. (1998), **The impact of campaign expenditure on political competition in the French legislative elections of 1993**, *Public Choice*, Vol. 94, pp. 157-174.

[116] Prat A. (2002), **Rational Voters and Political Advertising**, in *The Oxford Handbook of Political Economy (Oxford Handbooks of Political Science)*, of Weingast B.R., Wittman D. A. (eds.), Oxford University Press, USA (2006).

[117] Prat A. (2002), **Campaign Advertising and Voter Welfare**, Review of Economic Studies, 69, 999-1017.

[118] Persson T., Tabellini G. (2000), **Political Economics: Explaining Economic Policy**. MIT Press.

[119] Polo M., (1998), **Electoral Competition and political Rents**, *IGIER Working paper*, nr. 144, pp 28.

[120] Potters J.J.M., Sloof R., Winden F.A.A.M. van (1997), **Campaign Expenditures, Contributions and Direct Endorsements: The Strategic Use of Information and Money to Influence Voter Behavior**, *Discussion Paper 1997-27*, Tilburg University, Center for Economic Research.

[121] Putnam, R.D. (1995), **Tuning In, Tuning Out: the strange disappearance of social capital in America**, Policial Science and Politics, vol. 28, No. 4, pp. 664-683.

[122] Riley S.P. (1998), **The Political economy of anti-corruption strategies in Africa**, *The European Journal of Development Research*, Vol. 10, No. 1, pp. 129-159.

[123] Sanders L. (2001). **The psychological benefits of political participation**. Paper presented at the annual meeting of the American Political Science Association, San Francisco, CA.

[124] Skaperdas S., Grofman B. (1995), **Modeling Negative Campaigning**, *American Political Science Review*, Vol. 89, No. 1

[125] Svensson J. (1998), **Controlling Spending: Electoral Competition, Polarization and Endogenous Platforms**, Stockholm University, *mimeo.*

[126] Valentino N.A., Hutchings V.L., Williams D. (2004), **The Impact of Political Advertising on Knowledge, Internet Information Seeking, and Candidate Preference**, *Journal of Communication*, Vol. 54, No. 2, pp. 337-354.

[127] Wielhouwer P.W., Lockerbie B. (1994), **Party Contacting and Political participation**, *American Journal of Political Science*, Vol.38, No. 1, pp. 211-229.

[128] Weitz-Shapiro R., Winters, M. S. (2008). **Political participation and quality of life**, *Research Department Working Papers*; No. 638. Inter-American Development Bank. 1300 New York Avenue, N.W. Washington, DC 20577.

[129] Wittman D.A. (1989), **Why democracies produce efficient results**, *The Journal of Political Economy*, Vol 97, No. 6.

[130] Wittman D.A. (1995), **The Myth of Democratic Failure: Why Political Institutions Are Efficient**, *University of Chicago Press.*

[131] Wittman D.A. (2007), **Candidate quality, pressure group endorsements and the nature of political advertising**, *European Journal of Political Economy*, Vol. 23, No. 2, pp. 360378.

[132] `http://www.idea.int/political-finance/index.cfm`

[133] `www.transparency.org/policy_research/surveys_indices/cpi/2010/results`