

Dangerous tangents: an application of Γ -convergence to the control of dynamical systems

Rosario Maggistro¹ · Paolo Pellizzari² · Elena Sartori³ · Marco Tolotti⁴

Abstract

Inspired by the classical riot model proposed by Granovetter in 1978, we consider a parametric stochastic dynamical system that describes the collective behavior of a large population of interacting agents. By controlling a parameter, a policy maker seeks to minimize her own disutility, which in turn depends on the steady state of the system. We show that this economically sensible optimization is ill-posed and illustrate a novel way to tackle this practical and formal issue. Our approach is based on the Γ -convergence of a sequence of mean-regularized instances of the original problem. The corresponding minimum points converge toward a unique value that intuitively is the solution of the original ill-posed problem. Notably, to the best of our knowledge, this is one of the first applications of Γ -convergence in economics.

Keywords Dynamical systems · Finite population dynamics · Γ -convergence · Saddle-node bifurcations · Social interaction

✉ Rosario Maggistro
rosario.maggistro@deams.units.it

Paolo Pellizzari
paolop@unive.it

Elena Sartori
esartori@math.unipd.it

Marco Tolotti
tolotti@unive.it

¹ Department of Economics, Business, Mathematics and Statistics 'Bruno de Finetti', University of Trieste, Via Valerio 4/1, 34127 Trieste, Italy

² Department of Economics, Ca' Foscari University of Venice, Cannaregio 873, 31121 Venice, Italy

³ Department of Mathematics "Tullio Levi-Civita", University of Padova, 63, Via Trieste, 35121 Padova, Italy

⁴ Department of Management, Ca' Foscari University of Venice, Cannaregio 873, 31121 Venice, Italy

1 Introduction

Understanding and controlling collective behavior is a challenging, subtle, and potentially very useful endeavor [see Granovetter (1978)]. The mere attempt to know agents' preferences, motives, and norms is fraught with practical and conceptual difficulties: Data are sometimes scarce or difficult to obtain, and people may have blurred incentives and may - whether consciously or not - hide or misrepresent the drivers of their behavior. Moreover, even when a reasonable model for individual actions is agreed, the results can be puzzling. One example of this is shown in Thomas Schelling's segregation model, where *slight* homophilic preferences for neighbors of the same race lead in equilibrium to *massive* residential segregation, Schelling (1971). These collective models are known to produce very diverse and intriguing outcomes (simple *vs.* complex attractors, bifurcations, stable polarized *vs.* non-polarized equilibria,...). For this reason, a policy maker (or any external mastermind) could be interested in *controlling* the behavior of the population by setting some parameters at a certain *optimal level*.

We consider a problem that is characterized by two components: a dynamical system describing the evolution of some state variable that represents, in aggregate terms, agents' collective behavior; and a social planner who *ex-ante* sets the parameters of such dynamics to minimize her disutility, which in turn depends on the steady state of the system. Interestingly, in some cases, the problem turns out to be *ill-defined* in that the objective function is bounded from below, but does not admit a minimum. However, we show that by defining a proper sequence of auxiliary stochastic problems, it is possible to formally prove a convergence result whose limit is well-defined and helps shed light on the original problem.

To fix ideas and keep the formalism to a minimum for the moment, let us consider the dynamical system

$$r(t + 1) = F(r(t); \sigma),$$

where the future state $r(t + 1)$ depends on the current state $r(t)$ due to the (iterated) application of F , as well as on a parameter σ . Let us now assume that for a given initial condition $r(0) = r_0$, a limit state $\rho(\sigma)$ is reached.¹ The social planner attempts to solve an optimization problem as follows:

$$\min_{\sigma} f(\sigma), \text{ where } f(\sigma) = c(\sigma) - \rho(\sigma). \quad (1)$$

The fact that the social planner's payoff depends on the long-run outcome of the system (i.e., on $\rho(\sigma)$) seems reasonable for many applications. Let us consider, for

¹ Formally, this means that $\rho(\sigma) = \lim_{t \rightarrow \infty} F^t(r_0; \sigma)$, where F^t is the composition of F with itself for t times.

example, the case where $r(t)$ represents the market share of a durable good and σ is a parameter related to the quality of that good. In this respect, $\rho(\sigma)$ can be interpreted as the equilibrium market share related to a specific quality strategy implemented by the company. In the payoff, the component, $c(\sigma)$, plays the role of the (potential) cost to implement this strategy and, mathematically speaking, completes the definition of a specific parametric dynamical system indexed by σ .

Examples of modeling approaches of individual decisions giving rise, at the aggregate level, to dynamics resembling (1) can be found in Blume and Durlauf (2003) and Barucci and Tolotti (2012a, b). Models of supply and demand emerging in financial markets, where a similar structure manifests itself, are presented in Gordon et al. (2013). The diffusion of innovations and adoption models of durable goods as pioneered in Bass (1969) and, more recently, in Peres et al. (2010) share a microfoundation that is similar in spirit to the one presented here. In addition, epidemiological models of SI (Susceptible Infected) type exhibit dynamics that resemble the formalism proposed in (1). Finally, as already mentioned, a celebrated model in this context goes back to Granovetter (1978); in that study, a mastermind *manipulates* the mood of crowds to trigger riots that eventually involve a large proportion of the population. For the sake of clarity and simplicity, the latter riot model will hereafter be used to motivate our treatment and exemplify our results.

The previous situations all share a similar conceptual structure: contingent on the value of the parameter, some equilibrium is likely to endogenously appear as the final outcome of the dynamics, and an external agent is interested in shaping or controlling this outcome. While the intuition appears to be quite natural, the explicit setup of this “optimization problem over a dynamical system” is not very frequent, and we explore some of the intrinsic difficulties that may arise when defining such problems.

In fact, we show that problems formalized as in (1) can be *ill-posed* in that the objective function is discontinuous, it has no minimum, and it only admits an infimum. Essentially, this is due to a *saddle-node bifurcation* characterizing the dynamical system: One of the fixed points of the equation $r = F(r; \sigma)$ disappears when σ increases beyond some critical value σ_c .² Note that $\rho(\sigma)$ is one such fixed point (the smallest one in our setup, in fact) and that the presence of a bifurcation makes its value jump abruptly at the bifurcation value. Moreover, as already noted in Granovetter (1978), the bifurcation value σ_c is intuitively the value that the mastermind is looking for. Roughly speaking, the system is purposely steered to reach a bifurcation point (because this is advantageous).

However, as a matter of fact, this value is not a minimum point for f , but rather a minimizer. Note that from now on we refer to the “minimum point” if a minimum is attained, and “minimizer” if it is not attained. As a consequence, this optimization problem is intrinsically ill-posed. Summarizing, in contrast to their natural appeal, problems such as (1) can be fraught with technical difficulties that can impede formal and numerical treatments.

Technically speaking, one of the main goals of this paper is to overcome such a technical issue by examining the problem from a different angle. Instead of looking at

² A brief description of saddle-node bifurcations is given in Appendix A. See Strogatz (2015) for an exhaustive analysis of bifurcations in dynamical systems.

the steady states of the one-dimensional (deterministic) dynamical system as expressed in (1), we introduce a regularized stochastic version of it, where the number of agents in the population is large but finite. Despite being stochastic, this problem is now well-defined and admits a unique minimum point σ_N^* for all N , where N is the size of the population. Finally, we are able to provide a formal limit for N going to infinity and show that the sequence of minimum points, $(\sigma_N^*)_{N \geq 1}$, converges exactly to σ_c , the value that is “expected” to be a solution of the ill-posed original problem. Interestingly, it turns out that the aforementioned convergence holds in a Γ -sense: We will formally prove that the sequence of objective functions f_N of the auxiliary N -dimensional problems Γ -converges to a well-defined f_∞ (see (13)). This Γ -limit turns out to be equal to the original objective function f *almost everywhere*, except at the optimal level σ_c .³ This finding sheds light on one rigorous way to deal with the ill-posedness of the original problem suggested in Granovetter’s work. More recently, singularities similar to the one described above are also detected in Nadal et al. (2005) and Gordon et al. (2013). In these works, the authors model a monopolist in charge to set the optimal price to foster demand and maximize profits. With language and notation borrowed often from physics, these papers contain a model that has similarities with ours. They also contain discussions of the “*epistemic uncertainty*” inherent in selecting a price to maximize the monopolist’s profits. Uncertainty due to the presence/disappearance of multiple equilibria is acknowledged by sentences such as “[*the optimal solution*] lies very near the critical price value at which such high demand no more exists” (Gordon et al. (2013)). The singularity recognized (albeit not “solved”) in their model may be formally tackled through the use of a Γ -convergent sequence of problems whose limit turns out to be identical to the original singular model almost everywhere.

In terms of methodology, to the best of our knowledge, papers in the field of mathematical economics and social sciences that employ Γ -convergence are quite rare. Ghisi and Gobbino (2005) describes a variational problem arising from a generalization of the well-known monopolist’s problem introduced in Rochet and Choné (1998). In this model, the monopolist proposes a set of products and looks for the optimal price list, which minimizes costs and therefore maximizes profit. This leads to a minimum problem for functionals H (the “pessimistic cost expectation”) and G (the “optimistic cost expectation”), which are in turn defined through two nested variational problems. The authors prove that the minimum of G exists using an approximating sequence that Γ -converges to G and where such a minimum coincides with the infimum of H . An economic model of monopoly has also been studied in a general setting by Monteiro and Page (1998) and under convexity assumptions by Carlier (2001). In the latter, the author studies a principal- agent model with adverse selection and characterizes incentive-compatible contracts in terms of an envelope property called h -convexity. Using this characterization, the principal’s problem is written as a non-standard variational problem (with h -convexity constraint) for which the existence of a solution is proved. Monteiro and Page (1998) provides similar general existence results for the principal-agent problem with adverse selection. However, their approach to the problem differs to the one in Carlier (2001). After all, the authors consider budget

³ We refer the reader to Appendix B for a brief recap on Γ -convergence and to Braides (2002) for an extensive description.

constraints that force prices to remain in a given compact set, and their results rely on a *nonessentiality assumption* (i.e., nonessentiality of some goods relative to others), which is not required in Carlier (2001).

In this paper, we show how Γ -convergence can be used to deal with a natural but ill-posed problem faced by the decision maker. Indeed, the economic interest of our approach stems from the observation that it is the decision maker herself who rationally pushes the system to configurations where some equilibria disappear. This fact also poses severe difficulties when seeking numerical solutions, especially if higher-dimensional systems are under investigation.

This paper is organized as follows. Section 2 describes in detail the optimization problem as stated in (1). In Sect. 3, we provide a stochastic version of the same problem, where the number of actors in the economy is now finite. This stochastic approach, while seeming more cumbersome on the surface, provides a double benefit: Besides allowing us to take advantage of probabilistic tools, it naturally leads the modeler to simulation and numerical methods. We will see that this finite-population approach is crucial to analytically setting the proper convergence scheme. Section 4 is devoted to the analysis of numerical simulations, and conclusions are drawn in Sect. 5. Appendix A contains the technical proofs, and Appendices B and C summarize the main concepts of Γ -convergence and continuity employed in the paper, respectively.

2 The deterministic dynamic model

Inspired by Granovetter (1978), we first consider a (infinite) population of actors facing the decision of whether to take part in a riot. The agents are heterogeneous, they are aware of the current proportion of people involved in the riot, and they decide whether or not to join in based on a personal (random) activation threshold. If the proportion of active agents is above that threshold, they will join the crowd; otherwise they will not. Technically speaking, the random thresholds are independent copies of a random variable with distribution F , which, from now on, is assumed to be Gaussian with fixed mean $\mu = 0.25$ and standard deviation $\sigma \in [\sigma_{\min}, \sigma_{\max}]$ with $\sigma_{\min} > 0$ and σ_{\max} large enough but finite.⁴ Recall that σ will be set by the social planner. In this respect, by choosing σ , she is fixing a specific Gaussian distribution F and, therefore, a specific *behavioral trait* of the underlying population of potential rioters. Practically, this can be interpreted as a deliberate modification of the volatility of the crowd to ignite a riot. We define $r(t)$ as the proportion of agents taking part in the riot at time t , where $t \in \mathbb{N}$ and fix $r(0) = 0$.⁵ People with a threshold below $r(t)$ decide to join the riot at time $t + 1$; moreover, once they have joined the riot, they cannot withdraw from it. Therefore, $r(t + 1)$ accounts for the people who were already involved in the riot at the previous time plus the proportion of *newcomers*. Put in other terms, this is exactly $F(r(t); \sigma)$, i.e., the probability that the threshold is below $r(t)$. Hence, we obtain the

⁴ The Gaussian distribution with $\mu = 0.25$ was suggested by Granovetter (1978). Note that both the lower and the upper bound σ_{\min} and σ_{\max} are imposed for technical reasons related to the proof of the Γ -convergence.

⁵ We are interested in studying the case where the initial condition is zero (or small) as a benchmark for applications where the social phenomenon is analyzed from its outset.

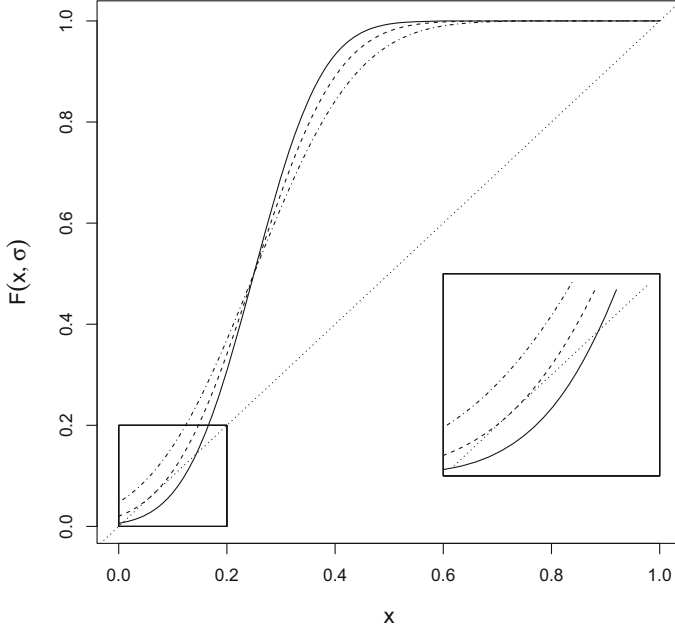


Fig. 1 The graph of $x \mapsto F(x; \sigma)$ for three different values of σ . Specifically, $\sigma = 0.1$ (solid line), $\sigma = 0.122$ (dashed line), and $\sigma = 0.15$ (dot-dashed line).

recursive equation of Granovetter (1978):

$$r(t + 1) = F(r(t); \sigma). \quad (2)$$

We denote by $\rho(\sigma)$ the smallest solution⁶ of the fixed point equation

$$r = F(r; \sigma). \quad (3)$$

In the case that F admits a unimodal density function (as in our example based on Gaussian thresholds), it is possible to characterize such an equilibrium. In fact, it is well known that under this minimal assumption, there are at most three solutions to (3), and the number of such solutions depends on the value of σ . Moreover, when there are three equilibria, the intermediate one is always unstable, whereas the two extreme ones are (locally) stable. As already stated, we are interested in the case where the initial condition is zero and, therefore, the dynamical system always converges to the smallest solution of (3). This family of dynamical systems exhibits what is called a *saddle-node bifurcation*.⁷ Keeping $\mu = 0.25$, the black line in Fig. 1 shows three fixed points of $r = F(r; \sigma)$, occurring when $\sigma = 0.1$. The intermediate fixed point, located

⁶ In this respect, $\rho(\sigma)$ can be interpreted as the *lowest fixed point* function in the sense of Milgrom and Roberts (1994).

⁷ For a formal analysis of saddle-node bifurcations and related facts, we refer the reader to Lemma A.1 in Appendix A.

at about 0.15, is unstable and, hence, the limit state for any initial condition below this fixed point would converge to the lower fixed point. For all such initial conditions, the limiting fraction of the population involved in the riot would be about 0. Suppose now that the social planner has the power to increase σ (at a cost): This has the effect of sliding the graph of F upward, up to the point where the two smaller fixed points merge when F is tangent to the bisector line. (The case is depicted by the dashed line in Fig. 1, obtained when $\sigma = 0.122 \approx \sigma_c$, where σ_c denotes the exact tangency value.) Now, the limit state when the initial proportion of citizens involved in the riot is an infinitesimal fraction would increase to about 6.2%. But even more importantly, a further tilt to F beyond σ_c triggers the occurrence of a saddle-node bifurcation, and the limit state abruptly jumps to 1; see the dot-dashed curve in the picture, relative to $\sigma = 0.15$.

As strongly pointed out in Granovetter (1978), finite populations drawn from the same F , when $\sigma \approx \sigma_c$, may lead to very different limit states under iteration: “*There is no obvious sociological way to explain why a slight perturbation of the normal distribution around the critical standard deviation should have a wholly discontinuous, striking qualitative effect. This example shows again how two crowds whose average preferences are nearly identical could generate entirely different results.*”⁸

The revelation that minor perturbations in individual features could produce large aggregate effects was probably a part of the *zeitgeist* of the seventies, as similar ideas are present in Schelling (1971) concerning racial segregation. Likewise, Allen and Sanglier, examining dynamic models of urban growth in 1979, distilled their results, observing that “*small perturbations of density, perhaps of random origin, are amplified by the interactions between the elements of the system and lead to a qualitative change in the macroscopic structure of the spatial distribution.*”⁹ We consider these insights to be of continued relevance, despite the passage of four decades.

To make this argument formal, we deviate from the original problem and introduce an *external agent* acting as a social planner. Suppose that a *rabble-rouser* is keen on minimizing the cost needed to fuel a massive participation in the uprising. Of course, due to the cost of stirring up people, the policy maker faces the following decision problem: “*How can I optimally settle the mood of the population in order to obtain a large-scale riot?*”. Considering a unitary cost $k > 0$ for a unit of standard deviation, her optimization problem can be written as

$$\min_{\sigma} f(\sigma), \text{ where } f(\sigma) = k\sigma - \rho(\sigma). \quad (\text{P1})$$

Note that a linear cost $k\sigma$ has been chosen for tractability reasons, and that if k is too large, the interest in the problem will be lost since the objective function reaches a proper minimum at some point $\sigma \leq \sigma_c$, possibly at $\sigma = \sigma_{\min}$. A similar argument applies also to the expected value μ related to the distribution function F .¹⁰

Figure 2 depicts the effect of the saddle-node bifurcation on the target function $f(\sigma) = k\sigma - \rho(\sigma)$, for $k = 0.5$. Clearly, as σ increases beyond σ_c , the disappearance

⁸ See Granovetter (1978, pp. 1427–1428).

⁹ See Allen and Sanglier (1979, p. 257).

¹⁰ More details on such threshold levels for k and μ are provided in Appendix A.

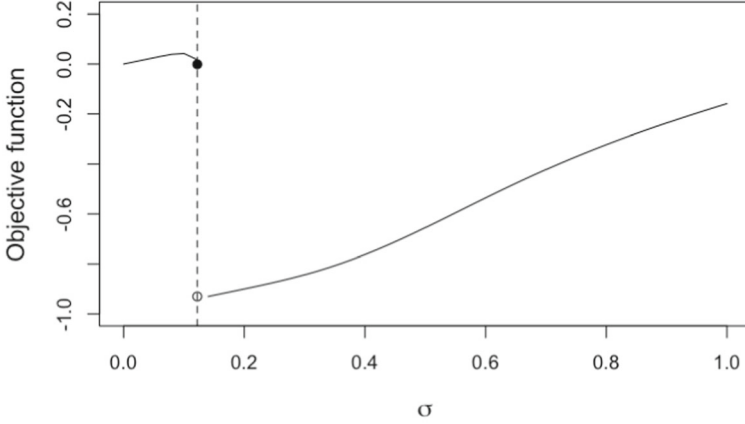


Fig. 2 The graph of $k\sigma - \rho(\sigma)$ as a function of σ , for $k = 0.5$. The discontinuity occurs at $\sigma_c \approx 0.122$

of two fixed points generates a jump in the limit state that will be reached through iteration of F . As pointed out in Strogatz (2015), the term “saddle-node” is not entirely consolidated in the dynamical systems literature; for instance, such a bifurcation is quite imaginatively called a “blue sky bifurcation” in some cases to stress that, reversing the direction of the change of σ , a new equilibrium can be created “out of the blue.”

Although no closed-form solution for this problem exists, the objective function satisfies the following lemma.¹¹

Lemma 2.1 *Define σ_c as the value of σ that makes the graph of the map $x \mapsto F(x; \sigma)$ tangent to the bisector line. Define also*

$$\tilde{f} = \lim_{\sigma \rightarrow \sigma_c^+} f(\sigma) = k\sigma_c - \tilde{r}_c, \quad \tilde{r}_c := \lim_{\sigma \rightarrow \sigma_c^+} \rho(\sigma). \quad (4)$$

Suppose finally that $k < k^{th} := \frac{\tilde{r}_c - \rho(\sigma_c)}{\sigma_c}$ and $\mu < \frac{1}{2}$. Then the function

$$f(\sigma) = k\sigma - \rho(\sigma),$$

where $\rho(\sigma)$ is the smallest solution to (3), has the following properties:

- i) *it is left-continuous with a discontinuity point at σ_c ;*
- ii) *it admits a finite infimum \tilde{f} such that $\tilde{f} < f(\sigma)$ for every $\sigma \in [\sigma_{\min}, \sigma_{\max}]$.*

In other words, f has no minimum and the inf is reached, for $\mu = 0.25$, at $\sigma_c \approx 0.122$. On the other hand, σ_c is exactly the value at which the curve $F(x; \sigma_c)$ is tangent to the bisector line (see Fig. 1). Algebraically, σ_c is the unique value of the parameter

¹¹ All proofs are given in Appendix A.

such that the following system is solvable for some $r \in [0, 1]$:

$$\begin{cases} F(r, \sigma) = r, \\ \frac{dF}{dr}(r; \sigma) = 1. \end{cases}$$

As a matter of fact, the problem is clearly ill-defined: No solution exists, even though the “optimal point” has a very clear geometrical interpretation in terms of the distribution function of the random thresholds.

To deal with this issue, we propose an auxiliary optimization problem in the next section. In particular, we let the population of actors be finite, of size N . We will see that this makes the problem stochastic but, on the other hand, well-posed for any finite N . Finally, we will show that the sequence of objective functions of such finite population problems converges in a Γ -sense to a function f_∞ , which, almost everywhere will be equal to the original objective function f . However, in contrast to f , f_∞ admits a unique minimum; therefore, the new auxiliary problem will also be well-defined for $N \rightarrow \infty$.

3 The (stochastic) problem with N agents

Let us consider N actors faced with the following decision problem: “to participate or not in a *riot*.” The state variable is thus binary. We define $y_i(t) \in \{0, 1\}$ for $i = 1, \dots, N$, where $y_i(t) = 1$ means that agent i is participating in the riot at time t . Once part of the riot, the rioter cannot withdraw. Moreover, we assume that $y_i(0) = 0$ for all i . This choice is intended to represent the fact that in the original infinite-population problem, we analyze the social phenomenon from its outset. It is assumed that no agent is initially rioting, setting $r(0) = 0$.

Being a decision under *collective behavior*, the reward for participating in the riot depends on the present number (or better, the proportion) of people involved. This quantity is given by¹²

$$r_N(t) = \frac{\sum_i y_i(t)}{N}. \quad (5)$$

Any actor decides to join the riot as soon as the quantity r_N is large enough. We model *random thresholds* X_i , $i = 1, \dots, N$ as N independent copies of a random variable X with absolutely continuous distribution function F . The rule is straightforward:

$$y_i(t + 1) = 1 \quad \Leftrightarrow \quad X_i \leq r_N(t). \quad (6)$$

¹² We could also consider the quantity $r_N^{(i)} = \frac{\sum_{j \neq i} y_j}{N-1}$. When N becomes large (infinite), the contribution of y_i is negligible, thus the two problems have exactly the same limiting behavior.

Summing up and normalizing over N ,

$$r_N(t+1) = \frac{1}{N} \sum_{i=1}^N y_i(t+1) = \frac{1}{N} \sum_{i=1}^N \mathbb{I}_{\{(-\infty, r_N(t)]\}}(X_i). \quad (7)$$

The right-hand side of Eq. (7) is nothing but $F_N(r_N(t); \sigma)$, the N -dimensional empirical distribution of the random thresholds. We thus obtain the following recursive equation characterizing the N -dimensional system

$$r_N(t+1) = F_N(r_N(t); \sigma). \quad (8)$$

We now have a population of N agents that evolves according to (8). In addition, the social planner aims to control the behavior of the population by setting the *optimal* σ so as to minimize a measurable function of the stochastic process r_N . More specifically, the policy maker considers the *steady states* of such dynamics as a natural outcome to consider.

Should the social planner look at the *finite* (stochastic) population or rather implement some theoretical asymptotic results to rely on some simpler deterministic limit dynamics? Technically speaking, this translates in the order in which we calculate the limits of $N \rightarrow \infty$ and $t \rightarrow \infty$. To deal with a deterministic system, we first take the N -limit; the computation of steady states is then straightforward, but, as just seen above, the optimization problem is ill-posed. Conversely, if we first perform the time limit, the observable steady states are stochastic, and the modeler relies on statistical tools to remove the randomness inherent in the finite-population system. In the remainder of the section, we formalize the two approaches, highlighting how different they are in their outcomes.

3.1 Weak convergence of stochastic processes

The first approach relies on the classical asymptotic theory developed in Ethier and Kurtz (2009). The convergence of the stochastic process r_N to a well-posed limiting process r , when the number of actors N tends to infinity, is ensured by the following result.¹³

Proposition 3.1 *Assume the recursion given in (8), with $r_N(0) = 0$ for all N . When $N \rightarrow \infty$, the process $(r_N(t))_{t \geq 0}$ weakly converges toward $(r(t))_{t \geq 0}$, characterized by the following asymptotic recursion*

$$r(t+1) = F(r(t); \sigma), \quad r(0) = 0, \quad (9)$$

where F is the distribution of X_i , $i = 1, \dots, N$ as defined in (6).

¹³ The proof of this rather classical result is omitted. We refer the reader to Ethier and Kurtz (2009) for more details.

Since, by assumption, $r(0) = 0$, the dynamical system necessarily converges to $\rho(\sigma)$, i.e., the smallest among the (possibly multiple) solutions of

$$r = F(r; \sigma), \quad (10)$$

obtained considering the limit for $t \rightarrow \infty$ of (9). In particular, as discussed above, there is no solution to the social planner minimization problem.

Example 3.2 (The Granovetter setting) Assume that $X \sim \mathcal{N}(\mu, \sigma)$, where μ is fixed at the level 0.25. Regarding the standard deviation, there exists a critical level σ_c such that there is only one solution of (3) for $\sigma > \sigma_c$. For $\sigma < \sigma_c$, the solutions are three, and for $\sigma = \sigma_c$, the function $r \rightarrow F(r; \sigma)$ is tangent to the bisector and the solutions are exactly two.

3.2 Γ -convergence of regularized operators

The second approach is based on the stochastic version of the dynamical system given in (8). In contrast to what was done previously, we first take the time limit of the dynamics. As a matter of fact, the steady state is now a random variable, measurable w.r.t. the N -dimensional sample X_1, \dots, X_N . Let the random variable R_N represent the (random) steady state of the dynamics in (8). Formally,

$$R_N = \lim_{t \rightarrow \infty} r_N(t). \quad (11)$$

To properly define an optimization problem for any finite N , we rely on the expected value of R_N as the observable to be investigated by the social planner:

$$\rho_N(\sigma) = \mathbb{E}^\sigma [R_N]. \quad (12)$$

It turns out that the auxiliary problem based on such a new observable is now *well-posed*. Furthermore, it is possible to characterize the limit of the sequence $(\rho_N(\sigma))_N$ in terms of $\rho(\sigma)$, the steady state related to the original and ill-posed optimization problem. These claims are stated in the following

Theorem 3.3 *Fix N and consider ρ_N as defined in (12). Then the optimization problem*

$$\min_{\sigma} f_N(\sigma), \quad f_N(\sigma) := k\sigma - \rho_N(\sigma), \quad (P2)$$

is well-posed and admits a minimum point σ_N^ . Moreover, for $N \rightarrow \infty$,*

$$f_N \xrightarrow{\Gamma} f_\infty,$$

where

$$f_\infty(\sigma) = \begin{cases} k\sigma - \rho(\sigma) & \text{if } \sigma \neq \sigma_c \\ k\sigma_c - \tilde{r}_c & \text{if } \sigma = \sigma_c \end{cases} \quad (13)$$

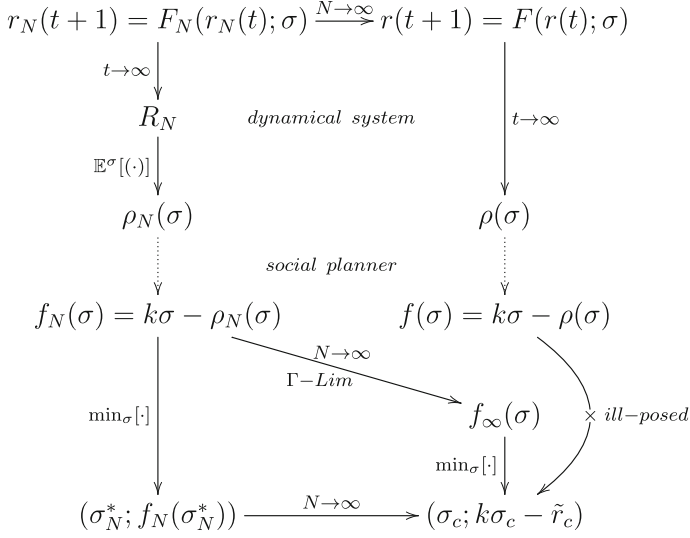


Fig. 3 The scheme of convergences in the model

and \tilde{r}_c is defined in (4). As a consequence, the sequence of minimum points $(\sigma_N^*)_N$ related to (P2) converges to σ_c and $(f_N(\sigma_N^*))_N$ converges to $k\sigma_c - \tilde{r}_c$.

The limit of f_N is therefore taken in a Γ -sense and f_∞ is exactly the Γ -limit of the sequence of real functions $(f_N)_N$. Instead, it can be proven that the convergence does not work properly in a classical *almost sure* or *uniform* sense. Note, finally, that f_∞ is nothing but the lower-semicontinuous envelope of f (the original objective function of problem (P1)); the two functions only differ to each other for the value at σ_c , which turns out to be exactly the minimum point for the problem where f_∞ is now used as an (auxiliary) objective function.

A summary of the different convergence schemes is graphically visualized in Fig. 3. The starting point is the top-left stochastic dynamical system expressed by (8). In the first approach, described in Sect. 3.1, we first move to the right by taking an N -limit, in the sense of the “convergence of stochastic processes”; then we move down (time limit), obtaining $\rho(\sigma)$. Moving in the area of the social planner, we come across the *ill-posed* problem (P1) as depicted in the bottom part of the right column. Conversely, by starting at the top-left corner and implementing the second approach described in Sect. 3.2, we first move down along the left column (time limit), and later introducing the expectation operator, we arrive at the bottom-left point where σ_N^* is defined as the minimum point of f_N for any fixed N . Finally, as f_N converges in a Γ -sense to f_∞ , we have that σ_N^* converges to σ_c , the minimum point of f_∞ .

In the next section, we collect some numerical findings and give a sense of the previous asymptotic result by solving the problem for increasing values of N . Moreover, we show that the numerical results are notably affected by the accuracy in computing an estimator for the expectation of the random variable R_N .

4 Numerical findings

In this section, we numerically analyze the riot problem where, for the sake of concreteness, we set $k = 0.5$. We simulate M stochastic instances of the optimization problem (P2) in Theorem 3.3 to obtain M couples minimum point-minimum $({}_m\hat{\sigma}_N^*; {}_m\hat{f}_N^*)$, $m = 1, \dots, M$, for a given N . Then we increase N to visually and numerically show the Γ -convergence given in Theorem 3.3.

The key insight of the previous section was that convergence can be achieved by taking the expectation, $\rho_N(\sigma)$, of the random steady state R_N defined in (11). Since a closed-form computation of $\rho_N(\sigma)$ appears to be impossible, we approximate it with a sample mean of S independent generations of steady states. The numerical results are obtained as follows:

Given N ;

for $m = 1, \dots, M$;

Provide an estimate of the solution to Problem (P2), that is, evaluate $f_N(\sigma)$ as many times as it is needed to get the optimal solution.¹⁴

1. To estimate $f_N(\sigma)$ for fixed σ :
 - 1.1 for $s = 1, \dots, S$:
 - 1.1.1 sample individual activation thresholds ${}_s\tilde{X}_N$ from $\mathcal{N}(\mu, \sigma)$;
 - 1.1.2 compute the random equilibrium ${}_sR_N$;¹⁵
 - 1.2 compute the sample mean $\hat{\rho}_N(\sigma) = \frac{1}{S} \sum_{s=1}^S {}_sR_N$;
 - 1.3 return the approximated value $\hat{f}_N(\sigma) = k\sigma - \hat{\rho}_N(\sigma)$;
2. Varying σ , identify the minimum ${}_m\hat{f}_N^*$ and the minimum point ${}_m\hat{\sigma}_N^*$.

Output: an M -dimensional sequence of pairs $({}_m\hat{\sigma}_N^*; {}_m\hat{f}_N^*)$, $m = 1, \dots, M$.

Since the sample mean converges to the correct average for $S \rightarrow \infty$, the higher the number S of simulated steady states, the closer our approximation will be to $\rho_N(\sigma) = \mathbb{E}^\sigma[R_N]$. Hence, it is important to realize that the numerical method to solve the social planner's optimization problem relies on an additional parameter, S , due to the need to replace the mathematical expectation with a sample mean. This is equivalent to saying that whenever an evaluation of the objective function in (P2) is needed, a set of N thresholds will be independently resampled (over and over) for S times, as seen in the above pseudocode to compute $f_N(\sigma)$.

Figure 4 shows some representative graphs of the objective functions of the optimization problem (P2). The top panel shows, for fixed $S = 100$, the functions $\hat{f}_N(\sigma)$ relative to $N = 100, 1000$, and $10,000$. It can be seen that larger values of N lead to smoother behavior and, more importantly, produce more accurate approximations of the function f_∞ defined in Theorem 3.3.

The bottom panel illustrates the role of S , for given $N = 1000$: when $S = 10$, admittedly a case in which the true expectation is poorly estimated, the function $\hat{f}_N(\sigma)$ is noisy and shows jumps that would make any minimization effort difficult or impossible. In contrast, when $S = 100$ or $S = 1000$, the graphs are smoother and almost overlapped; this seems to suggest that such numbers of sample observations are

¹⁴ We use the R function `optimize`.

¹⁵ The random equilibrium ${}_sR_N$ is obtained through a simple repeated iteration of (8) till convergence. Since iteration begins at zero, the smallest fixed point of F_N is always reached, usually in very few steps (with the obvious exception of $\sigma \approx \sigma_c$).

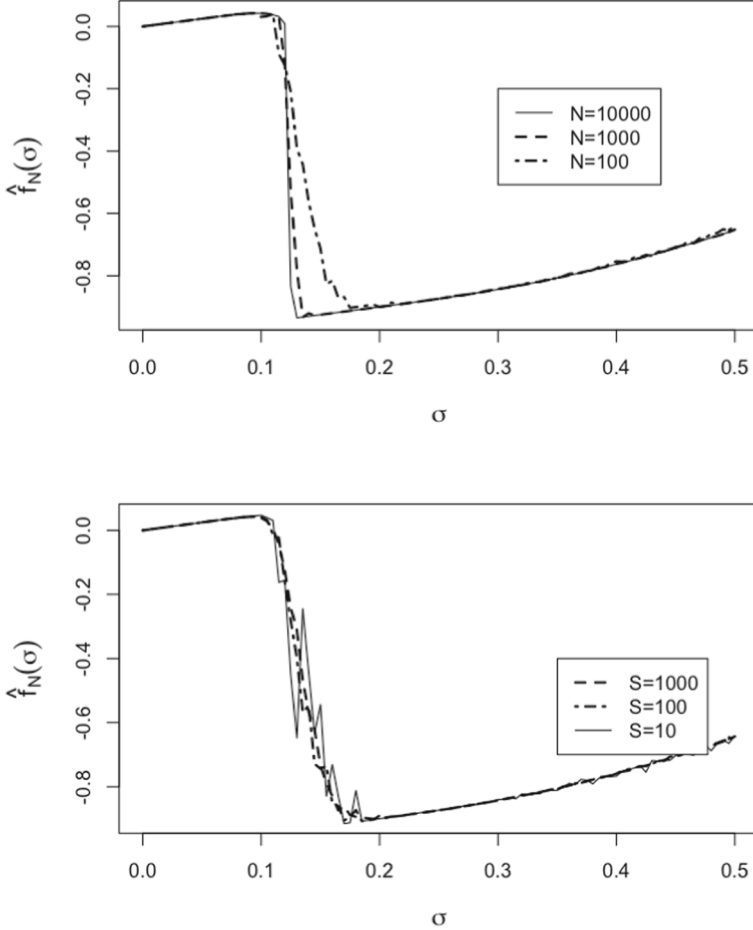


Fig. 4 The objective function $f_N(\sigma)$. Top panel: for fixed $S = 100$, the cases $N = 100, 1000, 10,000$ are shown. Bottom panel: for fixed $N = 1000$, the cases $S = 10, 100, 1000$ are displayed

appropriate to reasonably approximate σ_N^* . Practically speaking, some balance must be struck and, if N is to be increased to explore asymptotics, S should be limited to avoid an excessive computational burden.

Based on the above argument, we set $S = 100$ and solve $M = 100$ independent instances of the optimization problems, for increasing values of $N = 100, 500, 1000, 5000, 10,000, 50,000, 100,000$. Figure 5 shows the boxplots of $m\sigma_N^*$, $m = 1, \dots, M$, in various cases. The boxes depict interquartile ranges (IQR); the whiskers extend up to 1.5 IQR to provide evidence of outliers (circles), if any; the horizontal line is the median value. While for small values of N , the true minimum point $\sigma_c \approx 0.122$, referred to in Theorem 3.3, is overestimated and there are some outlying results, as N increases, most of the minimum points of the (stochastic) problems are quite close to the correct result and, clearly, boxplots support the convergence rigorously proven in the previous section. In other words, if N is too small, the min-

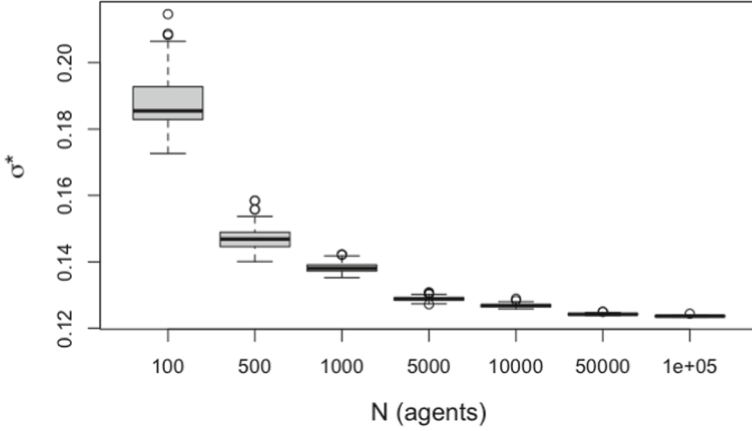


Fig. 5 Boxplots of M minimum points $m\hat{\sigma}_N^*$, $m = 1, \dots, M$, for $N = 100, 500, 1000, 5000, 10,000, 50,000, 100,000$ and fixed $S = 100$

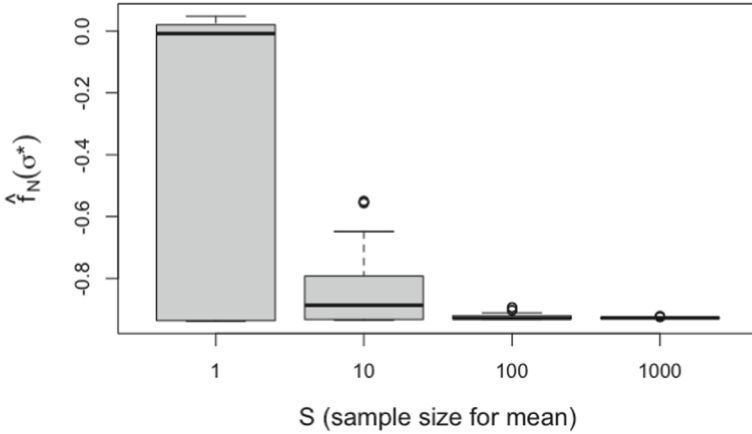


Fig. 6 Boxplots of M minima $m\hat{f}_N^*$, $m = 1, \dots, M$, for $S = 1, 10, 100, 1000$ and fixed $N = 1000$

imum points obtained by numerical optimization inaccurately span a wide interval, which can be roughly located at $[0.17, 0.21]$. As N increases, the numerical results are increasingly close to σ_c and are no longer dispersed. The plotted data were obtained using the routine `optimize` in R, R Core Team (2018); the computations took about 18 min on a 51-core 2.2 GHz Linux cluster setup with the library `parallel`. It may be worth noting that, say, the boxplot for $N = 100,000$ required to sample an order of $M \cdot S \cdot N = 100 \cdot 100 \cdot 100,000 = 10^9$ individual thresholds.¹⁶

Not surprisingly, the value of S also has a conspicuous impact on the accuracy of the M solutions of the problems found numerically in our simulations. Figure 6 depicts, for fixed $N = 1000$, the boxplots of the minima $m\hat{f}_N^*$, $m = 1, \dots, M$, for

¹⁶ As `optimize` typically evaluates the objective function about 20 times, the sampled thresholds exceed 10^{10} .

different values of S : for $S = 1$ the computed results are essentially unreliable and spread over a large interval. Note that the boxplot relative to $S = 1$ represents well the ill-posedness depicted in Fig. 3: Unless the mean operator is used, moving down along the right side of the diagram by increasing N does not remove the discontinuity of $f(\sigma)$. As a consequence, numerically computed minima are most often very far from $k\sigma_c - \tilde{r}_c$. In contrast, by increasing S , i.e., by taking the mean on the left side of the diagram, and then minimizing the objective function f_N such that $f_N \xrightarrow{\Gamma} f_\infty$, produces a sequence of pairs of real numbers $(\sigma_N^*, k\sigma_N^* - \rho_N(\sigma_N^*))$ that converges to $(\sigma_c, k\sigma_c - \tilde{r}_c)$.

Improving the accuracy of the sample mean by using a larger S reduces the likelihood of error and visually reconfirms that S in the range 100–1000 appears to be numerically satisfactory. Going into greater depth, the figure again portrays the importance of the smoothing effect provided by the average operator: problems can be solved more accurately, the ill-posedness of the original formulation is removed, and tractability is obtained in the form of Γ -convergence.

5 Conclusions

We have studied a class of “optimization problems over a dynamical system” and shown that care is needed to deal with ill-posedness. This was motivated by the classical riot example discussed in Granovetter (1978): intuitively, at some critical value for one parameter, the resulting endogenous equilibrium can abruptly jump due to small and natural sampling variability in the activation thresholds of agents. More formally, in the presence of a saddle-node bifurcation of the large-scale deterministic dynamics, some equilibria disappear, and the objective of the policy maker turns out to be discontinuous, breaking down conventional optimization approaches. The problem cannot be eliminated by merely increasing, or taking the limit of, the number N of agents and solving the deterministic version of the problem due to the tangency condition.

For this reason, we considered a sequence of finite-population problems, where the steady states of the dynamical system are random variables, and we proved that the objective functions of such problems are now continuous thanks to the *regularizing effect* of the expectation operator. Moreover, the sequence of the minimum points converges to the critical value of the parameter, i.e., the value where the bifurcation happens, and the sequence of minima converges to the infimum of the (ill-posed) original problem. Technically speaking, this convergence applies in a Γ -sense, and we believe that our work may be one of the very few practical applications of this methodology to economically relevant problems. Specifically, the sequence of the objective functions f_N Γ -converges toward f_∞ , which is proven to be the lower-semicontinuous envelope of the original objective function f . Interestingly, f_∞ and f only differ at the value σ_c , which is the limit of the sequence of minimum points, and in turn the point where the social planner (and we, as rational and conscious external observers) expected a minimum to exist in the ill-posed problem. In this respect, our approach endows the social planner with a mathematically precise argument to identify the unique optimal policy.

From a theoretical point of view, the use of the expectation operator appears to be crucial in the definition of a sequence of problems that converges to a regularized right-continuous objective function, where the minimum is now well-defined and coincides with the infimum of the original problem.

Our numerical results also make clear that the use of an expectation is not simply an astute device to obtain some usable approximated result but, more fundamentally, poses the basis for the Γ -convergence (in the number of agents N) to the theoretical model. This, together with a proper *Law of Large Numbers* in S , the sample size of the estimator of the expectation, ensures also convergence of the numerical simulations.

We have demonstrated, in this specific case, that sample sizes of the order of one hundred to one thousand provide enough smoothing to obtain reasonably accurate numerical results (with the help of a multi-core processor and parallel computations). We believe that our exemplification shows the need to incorporate proper sampling schemes to similar problems or whenever, say, a numerical treatment is the only viable option due to a lack of closed-form expressions for the steady state of the dynamical systems.

Acknowledgements We thank Marco LiCalzi for his insightful discussions and comments. Paolo Dai Pra brought our attention to Γ -convergence. The work was funded in part by the ITN ExSIDE European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 721846.

A Proofs

All proofs are given in this appendix.

Proof of Lemma 2.1

We first state and prove a technical result on bifurcation theory of a dynamical system.

Lemma A.1 (Saddle-node bifurcation) *Consider the dynamical system*

$$x(t+1) = F(x(t); \sigma); \quad x(0) = x_0 \in [0, 1], \quad (14)$$

where $x \in [0, 1]$ and $F(\cdot; \sigma)$ is a continuous distribution function with standard deviation σ , admitting a unimodal density function. The set of steady states of (14) is given by the solutions to

$$x = F(x; \sigma). \quad (15)$$

There exists a threshold level σ_c , such that:

- i) if $\sigma > \sigma_c$, (15) admits a unique solution x ;
- ii) if $\sigma = \sigma_c$, (15) admits two solutions $x^l < x^h$;
- iii) if $\sigma < \sigma_c$, (15) admits three solutions $x^l < x^m < x^h$.

If the solution x is unique, it is a globally stable attractor for the dynamical system in (14). In case of three equilibria, if $x_0 < x^m$, then

$$\lim_{t \rightarrow \infty} x(t) = x^l.$$

In the opposite case, if $x_0 > x^m$, then

$$\lim_{t \rightarrow \infty} x(t) = x^h.$$

Proof Note that $F(0; \sigma) > 0$ and $F(1; \sigma) < 1$ for any σ . Therefore, at least one solution x to (15) exists. Since F is S-shaped, it is convex for small x and concave for large x . A maximum of three solutions to (15) can therefore appear. The number of solutions to (15) depends on σ . This is an example of a saddle-node bifurcation.¹⁷ By definition, σ_c identifies the unique situation in which F is tangent to the bisector line at some point. In this case, exactly two different solutions to (15) exist. As soon as we take a value $\sigma > \sigma_c$, such a tangency point disappears. In the opposite case, for $\sigma < \sigma_c$, there are three intersections. We check the stability of the steady states by looking at the linearized version of the system (14). In this way, it is not difficult to see that if the equilibrium \bar{x} is unique, then $F'(\bar{x}; \sigma) < 1$, so it is linearly stable. In case of three equilibria, $F'(x^m; \sigma) > 1$, whereas $F'(x^l; \sigma) < 1$ and $F'(x^h; \sigma) < 1$. \square

Returning to the proof of Lemma (2.1), we are exactly in this situation, since F is Gaussian. The graph of the distribution function F intersects the bisector either three times, twice, or once. A visual representation of the three different cases is reported in Fig. 1. When $\sigma = \sigma_c$, we are in the tangency situation: the graph of F intersects the bisector line at one point x^l , when the curves are tangent, and at one second point x^h , when they are not. If the expected value μ of the distribution is such that $\mu < 1/2$, then it is easy to see that $x^l < 1/2 < x^h$. Consider finally that we take $r_0 = 0$ so that the dynamical system always converges to x^l , that is, the smallest among the possible solutions to (15). Let us call this equilibrium $\rho(\sigma)$.

The continuity of $f(\sigma)$ on $[\sigma_{\min}, \sigma_c]$ and on $(\sigma_c, \sigma_{\max}]$ immediately follows from the continuity of the map $\sigma \mapsto \rho(\sigma)$ on the same intervals. However, the proof of this latter property is not trivial at all, and is postponed to dedicated Appendix C.

It remains to show that $f(\sigma) = k\sigma - \rho(\sigma)$ is bounded from below and that the infimum of f is exactly \bar{f} defined as $\lim_{\sigma \rightarrow \sigma_c^+} f(\sigma)$. It is convenient to separate the study on the two intervals $[\sigma_{\min}, \sigma_c]$ and $(\sigma_c, \sigma_{\max}]$. Concerning the latter, we show that on this interval, $\sigma \mapsto \rho(\sigma)$ is decreasing, so that f is increasing on the same interval. As noted above, when $\sigma > \sigma_c$ there exists a unique x solving (15), with $x > \mu$. Moreover, still for $\sigma > \sigma_c$ and $x > \mu$, $F(x; \sigma)$ is concave and increasing. Let $\sigma_1, \sigma_2 \in (\sigma_c, \sigma_{\max}]$ be such that $\sigma_1 < \sigma_2$. Then $F(\cdot; \sigma_1)$ and $F(\cdot; \sigma_2)$ satisfy

$$F(x; \sigma_2) < F(x; \sigma_1) \quad \forall x > \mu. \quad (16)$$

¹⁷ See Sect. 3.1 in Strogatz (2015) for a general discussion on saddle-node bifurcations. Figure 3.1.7. provides an example that is similar in spirit to our situation.

Let us call ξ_1 the unique solution to $F(x; \sigma_1) = x$. Then, by (16),

$$F(\xi_1; \sigma_2) < F(\xi_1; \sigma_1) = \xi_1.$$

Since $F(\cdot; \sigma_2)$ is increasing, there exists $\xi_2 < \xi_1 : F(\xi_2; \sigma_2) = \xi_2$. Note that $\xi_1 = \rho(\sigma_1)$ and $\xi_2 = \rho(\sigma_2)$. Therefore, $\rho(\sigma)$ is decreasing on $(\sigma_c, \sigma_{\max}]$. Concerning the interval $[\sigma_{\min}, \sigma_c]$, a similar argument shows that, in this case, $\sigma \mapsto \rho(\sigma)$ is increasing up to a level $\rho(\sigma_c)$ such that $\lim_{\sigma \rightarrow \sigma_c^+} \rho(\sigma) := \tilde{r}_c > \rho(\sigma_c)$. Since, by assumption,

$$k < k^{th} := \frac{\tilde{r}_c - \rho(\sigma_c)}{\sigma_c},$$

we have that:

1. on $(\sigma_{\min}, \sigma_c)$,

$$k < \frac{\tilde{r}_c - \rho(\sigma_c)}{\sigma_c} < \frac{\tilde{r}_c - \rho(\sigma_c)}{\sigma_c - \sigma} < \frac{\tilde{r}_c - \rho(\sigma)}{\sigma_c - \sigma},$$

where the latter is due to the monotonicity of ρ . Thus,

$$k < \frac{\tilde{r}_c - \rho(\sigma)}{\sigma_c - \sigma} \iff k(\sigma_c - \sigma) < \tilde{r}_c - \rho(\sigma) \iff k\sigma_c - \tilde{r}_c < k\sigma - \rho(\sigma);$$

2. for $\sigma = \sigma_{\min}$, $k\sigma_c - \tilde{r}_c < k\sigma_{\min} - \rho(\sigma_{\min}) := f(\sigma_{\min})$;

3. for $\sigma = \sigma_c$, $-\tilde{r}_c < -\rho(\sigma_c)$ and hence $\tilde{f} < f(\sigma_c)$.

Summarizing, we have proved that: (i) f is left-continuous with a discontinuity at σ_c ; (ii) the function f is bounded from below and admits a finite infimum \tilde{f} , which is not a minimum. \square

Proof of Theorem 3.3

We first state and prove five technical lemmas related to R_N as defined in (11) and to $\rho_N(\sigma)$ as defined in (12).

Lemma A.2 *For each N , R_N is a measurable and bounded function of the finite sample $\tilde{X} \equiv (X_1, \dots, X_N)$. Moreover, the function $\sigma \mapsto \rho_N(\sigma)$ is continuous.*

Proof

$$\rho_N(\sigma) = \mathbb{E}^\sigma[R_N] = \int_{\text{supp}(\tilde{X})} R_N(\tilde{x}) d\tilde{F}(\tilde{x}; \sigma), \quad (17)$$

where $\tilde{F}(\tilde{x}; \sigma) = \prod_{i=1}^N F(x_i; \sigma)$ and where $\tilde{x} = (x_1, \dots, x_N)$ is a realization of the sample \tilde{X} . Note that the integrand function $R_N(\tilde{x})$ is a measurable and bounded function of the sample. As a consequence, the integral is well-defined; moreover, it is continuous in σ as soon as F is continuous in σ . \square

Lemma A.3 For any $\sigma \neq \sigma_c$ and $\varepsilon > 0$,

$$\lim_{N \rightarrow \infty} \mathbb{P}(|R_N - \rho(\sigma)| > \varepsilon) = 0,$$

where $\rho(\sigma) = \min\{x : F(x; \sigma) = x\}$. Moreover,

$$\mathbb{E}^\sigma[|R_N - \rho(\sigma)|] \rightarrow 0.$$

As a consequence, $\rho_N(\sigma) := \mathbb{E}^\sigma[R_N] \rightarrow \rho(\sigma)$.

Proof Fix $\sigma \neq \sigma_c$. We show separately that, for N large enough and with probability one, $R_N < \rho(\sigma) + \varepsilon$ and $R_N > \rho(\sigma) - \varepsilon$ for any $\varepsilon > 0$. We start with the former inequality.¹⁸ To this end, we consider an alternative and equivalent definition for R_N :

$$R_N = \min\{x : F_N(x; \sigma) = x\}.$$

We show that there exists $\varepsilon_0 > 0$ such that for all $\varepsilon > \varepsilon_0$, there exists N such that $F_N(x; \sigma) < x$ for $x = \rho(\sigma) + \varepsilon$.¹⁹ This latter inequality states exactly that $R_N < \rho(\sigma) + \varepsilon$. By way of contradiction, suppose that there exists $\varepsilon > 0$ such that for all N , $F_N(\rho(\sigma) + \varepsilon; \sigma) \geq \rho(\sigma) + \varepsilon$; then

$$F_N(\rho(\sigma) + \varepsilon; \sigma) \geq \rho(\sigma) + \varepsilon > F(\rho(\sigma); \sigma),$$

where the latter inequality comes from the fact that $F(\rho(\sigma); \sigma) = \rho(\sigma)$. Now this is a contradiction, since F is continuous in its first argument and $\sup_x |F_N(x; \sigma) - F(x; \sigma)| \rightarrow 0$ by virtue of the classical Glivenko–Cantelli Theorem.

To prove that $R_N > \rho(\sigma) - \varepsilon$, we use a similar argument. We show that there exist $\varepsilon_0 > 0$ such that for all $\varepsilon > \varepsilon_0$, there exists N such that $F_N(x; \sigma) > x$ for $x = \rho(\sigma) - \varepsilon$. By way of contradiction, suppose that there exists $\varepsilon > 0$ such that for all N , $F_N(x; \sigma) \leq x$; then

$$F_N(\rho(\sigma) - \varepsilon; \sigma) \leq \rho(\sigma) - \varepsilon < F(\rho(\sigma); \sigma).$$

Finally, note that for all $x \leq \rho(\sigma) - \varepsilon$, there exists \bar{N} such that $F_{\bar{N}}(x; \sigma) > x$ for sure. Suppose there exists $\tilde{x} < \rho(\sigma) - \varepsilon$ such that $F_N(\tilde{x}; \sigma) \leq \tilde{x}$ for all N . Since $F(\tilde{x}; \sigma) > \tilde{x}$, again we find a contradiction for the Glivenko–Cantelli Theorem. Therefore, $R_N > \rho(\sigma) - \varepsilon$ for sure. Note now that the sequence of random variables R_N is uniformly bounded by Lebesgue’s dominated convergence theorem, $\mathbb{E}^\sigma[|R_N - \rho(\sigma)|] \rightarrow 0$ and we obtain convergence in mean. \square

Lemma A.4 The sequence of derivatives $(\rho'_N(\sigma))_N$ exists and is uniformly bounded on $[\sigma_{\min}, \sigma_{\max}]$.

¹⁸ To ease readability, in the remainder of the proof we omit to say that the inequalities dealing with random variables hold almost surely (i.e., with probability one).

¹⁹ If $\sigma = \sigma_c$, no matter of the value of ε , it is not possible to identify a strict inequality due to the fact that F is tangent to the bisector line at the value $\rho(\sigma)$ when $\sigma = \sigma_c$.

Proof We use the expression for ρ_N as in (17) and derive it w.r.t. σ . To this end, note that the random variable R_N does not explicitly depend on σ ; we recall, moreover, that

$$d\tilde{F}(\tilde{x}; \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(\tilde{x}-\mu)^2}{\sigma^2}} d\tilde{x}.$$

A simple calculation gives:

$$\frac{d\rho_N(\sigma)}{d\sigma} = -\frac{1}{\sigma}\rho_N(\sigma) + \frac{1}{\sigma^3} \int_{\text{supp}(\tilde{X})} R_N(\tilde{x})(\tilde{x} - \mu)^2 d\tilde{F}(\tilde{x}; \sigma).$$

Therefore,

$$\left| \frac{d\rho_N(\sigma)}{d\sigma} \right| \leq \left| \frac{1}{\sigma}\rho_N(\sigma) \right| + \left| \frac{1}{\sigma^3} \int_{\text{supp}(\tilde{X})} R_N(\tilde{x})(\tilde{x} - \mu)^2 d\tilde{F}(\tilde{x}; \sigma) \right|. \quad (18)$$

Concerning the integral, since $R_N(\tilde{x})(\tilde{x} - \mu)^2 \leq R_N^2(\tilde{x}) + (\tilde{x} - \mu)^4$, we have

$$\begin{aligned} 0 &\leq \int_{\text{supp}(\tilde{X})} R_N(\tilde{x})(\tilde{x} - \mu)^2 d\tilde{F}(\tilde{x}; \sigma) \\ &\leq \int_{\text{supp}(\tilde{X})} R_N^2(\tilde{x}) d\tilde{F}(\tilde{x}; \sigma) + \int_{\text{supp}(\tilde{X})} (\tilde{x} - \mu)^4 d\tilde{F}(\tilde{x}; \sigma). \end{aligned}$$

Returning to (18), and noting that all the expressions on the r.h.s. are positive, we have

$$\left| \frac{d\rho_N(\sigma)}{d\sigma} \right| \leq \frac{\rho_N(\sigma)}{\sigma} + \frac{1}{\sigma^3} \int_{\text{supp}(\tilde{X})} R_N^2(\tilde{x}) d\tilde{F}(\tilde{x}; \sigma) + \frac{1}{\sigma^3} \int_{\text{supp}(\tilde{X})} (\tilde{x} - \mu)^4 d\tilde{F}(\tilde{x}; \sigma).$$

Since $\rho_N(\sigma) \leq 1$, $R_N^2 \leq 1$ for all N , and

$$\int_{\text{supp}(\tilde{X})} (\tilde{x} - \mu)^4 d\tilde{F}(\tilde{x}; \sigma) = \sigma^4 \text{Kurt}(\tilde{X}) = 3\sigma^4,$$

we have that

$$\left| \frac{d\rho_N(\sigma)}{d\sigma} \right| \leq \frac{1}{\sigma} + \frac{1}{\sigma^3} + 3\sigma \leq K,$$

where K is a suitable constant, possibly depending on σ_{\min} and σ_{\max} but independent of σ . □

Lemma A.5 *The sequence $(\rho_N)_N$ converges uniformly to ρ on the two disjoint intervals $[\sigma_{\min}, \sigma_c - \delta]$ and $[\sigma_c + \delta, \sigma_{\max}]$ for every $\delta > 0$ ²⁰.*

²⁰ Obviously $\delta < \min\{\sigma_c - \sigma_{\min}, \sigma_{\max} - \sigma_c\}$, otherwise the statement is meaningless.

Proof We use the fact that if a sequence of continuous real-valued functions $(F_n)_n$ converges in L^p for some $p \in [1, +\infty)$ on a closed and finite interval to a limit function F and, in addition, the sequence of derivatives F_n' exists and is uniformly bounded, then the sequence converges to F also uniformly.²¹ We apply this result to the sequence $(\rho_N)_N$. Convergence in L^1 on the disjoint intervals follows from Lemma A.3. The fact that the sequence of derivatives ρ_N' is uniformly bounded on the entire domain $[\sigma_{\min}, \sigma_{\max}]$ has been proved in Lemma A.4. \square

Lemma A.6 For all $\varepsilon > 0$ and for N large enough,

$$\rho_N(\sigma_c) \leq \tilde{r}_c + \varepsilon.$$

Proof Fix $\varepsilon > 0$, there exist $\delta > 0$, such that

$$|\rho_N(\sigma_c) - \tilde{r}_c| \leq \varepsilon_\delta < \varepsilon.$$

Take $\sigma = \sigma_c + \delta$, $\delta > 0$. Then

$$\begin{aligned} |\rho_N(\sigma_c) - \tilde{r}_c| &\leq |\rho_N(\sigma_c) - \rho_N(\sigma_c + \delta)| + |\rho_N(\sigma_c + \delta) - \tilde{r}_c| \\ &\leq |\rho_N(\sigma_c) - \rho_N(\sigma_c + \delta)| + |\rho_N(\sigma_c + \delta) - \rho(\sigma_c + \delta)| + |\rho(\sigma_c + \delta) - \tilde{r}_c| \\ &\leq \varepsilon_\delta := \varepsilon_\delta^{(1)} + \varepsilon_\delta^{(2)} + \varepsilon_\delta^{(3)}, \end{aligned}$$

where

$$|\rho_N(\sigma_c) - \rho_N(\sigma_c + \delta)| \leq \varepsilon_\delta^{(1)}$$

is due to the continuity of ρ_N ,

$$|\rho_N(\sigma_c + \delta) - \rho(\sigma_c + \delta)| \leq \varepsilon_\delta^{(2)}$$

follows from the fact that $\rho_N \rightarrow \rho$ for any $\sigma \neq \sigma_c$, and, finally,

$$|\rho(\sigma_c + \delta) - \tilde{r}_c| \leq \varepsilon_\delta^{(3)}$$

follows from the fact that $\lim_{\sigma \rightarrow \sigma_c^+} \rho(\sigma) = \tilde{r}_c$. \square

Returning to the statement of Theorem 3.3, to prove the well-posedness of the problem (P2), simply note that, according to Lemma A.2, $f_N(\sigma) = k\sigma - \rho_N(\sigma)$ is continuous, $\rho_N(\sigma)$ is bounded, and, finally, $\lim_{\sigma \rightarrow \sigma_{\max}} f_N(\sigma) = M$, with $M > 0$ large enough. The objective function is therefore continuous and bounded from below; hence, it admits a minimum and a minimum point σ_N^* .

For the second part of the statement, we use the tool of the Γ -convergence which, under suitable conditions, implies convergence of minimum values and minimizers, which in our case are also minimum points.

²¹ See, for example, Corollary 3 in Ford and Penline (2007).

It is easy to see that the sequence of objective functions $(f_N)_N$ defined on \mathbb{R}^+ is *equi-mildly coercive*.²² Moreover, given the function $f(\sigma) = k\sigma - \rho(\sigma)$, we consider its lower-semicontinuous envelope scf (see Definition B.4), that is, for every $\sigma \in [\sigma_{\min}, \sigma_{\max}]$,

$$scf(\sigma) = \begin{cases} k\sigma - \rho(\sigma) & \text{if } \sigma \neq \sigma_c \\ k\sigma_c - \tilde{r}_c & \text{if } \sigma = \sigma_c \end{cases} \quad (19)$$

with \tilde{r}_c as in (4).

As seen in Lemma A.5, $\rho_N(\sigma)$ converges uniformly to $\rho(\sigma)$ for any $\sigma \in [\sigma_{\min}, \sigma_c - \delta] \cup [\sigma_c + \delta, \sigma_{\max}]$ for every $\delta > 0$ (see footnote 20 (Page 24)). It is now immediate to derive that $f_N(\sigma) \rightarrow f(\sigma)$ uniformly on the open set $U_\delta := (\sigma_{\min}, \sigma_c - \delta) \cup (\sigma_c + \delta, \sigma_{\max}) \forall \delta > 0$. As a consequence, it converges uniformly on the set $U := \bigcup_{\delta > 0} U_\delta = (\sigma_{\min}, \sigma_c) \cup (\sigma_c, \sigma_{\max})$. Then, by applying Proposition B.5, we obtain that

$$\Gamma - \lim_N f_N(\sigma) = scf(\sigma) = f(\sigma) \text{ on } U.$$

It remains to study the value of Γ - $\lim_N f_N(\sigma)$ for the values of σ at the frontier of U , namely, $\sigma \in \{\sigma_{\min}, \sigma_c, \sigma_{\max}\}$. We now show that at $\sigma = \sigma_c$,

$$\Gamma - \lim_N f_N(\sigma_c) = scf(\sigma_c).$$

Given that the Γ -limit of a sequence of functions, if it exists, is necessarily lower-semicontinuous [see, e.g., Proposition 1.28 in Braides (2002)] and it is unique, $scf(\sigma_c)$, as in (19), is a good candidate to be the Γ -limit value we seek. In the following, we prove that $scf(\sigma_c)$ satisfies both conditions (26) and (28) of the definition of Γ -limit (see Definition B.2), hence the Γ - $\lim_N f_N(\sigma_c)$ exists and is unique.

- 1) (liminf inequality). By way of contradiction, suppose that there exists a sequence $(\tilde{\sigma}_N)_N, \tilde{\sigma}_N \rightarrow \sigma_c$ such that $scf(\sigma_c) > \liminf_N f_N(\tilde{\sigma}_N)$. On the other hand, since for every ℓ , f_ℓ is continuous, then $f_\ell(\sigma_c) \leq \liminf_j f_\ell(\sigma_j)$ for every sequence $(\sigma_j)_j$ converging to σ_c . Such a condition holds, in particular, for the sequence $(\tilde{\sigma}_N)$ identified above, and letting ℓ be large; therefore, we can write

$$\lim_\ell f_\ell(\sigma_c) \leq \liminf_N \lim_\ell f_\ell(\tilde{\sigma}_N). \quad (20)$$

Applying Definition B.1 to the right-hand side of (20), we get

$$\liminf_N \lim_\ell f_\ell(\tilde{\sigma}_N) = \inf \left\{ \lim_N \lim_\ell f_\ell(\tilde{\sigma}_N) : \tilde{\sigma}_N \in \mathbb{R}, \tilde{\sigma}_N \rightarrow \sigma_c, \exists \lim_N \lim_\ell f_\ell(\tilde{\sigma}_N) \right\}. \quad (21)$$

²² See Definition B.7. In particular, our sequence $(f_N)_N$ is equi-mildly coercive since all functions f_N are bounded from below. All infima are reached in $[\sigma_{\min}, \sigma_{\max}]$, which is compact.

We can now take ℓ growing as N , and by (21) it follows

$$\begin{aligned} \liminf_N \lim_N f_N(\tilde{\sigma}_N) &= \inf \left\{ \lim_N f_N(\tilde{\sigma}_N) : \tilde{\sigma}_N \in \mathbb{R}, \tilde{\sigma}_N \rightarrow \sigma_c, \exists \lim_N f_N(\tilde{\sigma}_N) \right\} \\ &= \liminf_N f_N(\tilde{\sigma}_N), \end{aligned}$$

where the last equality is precisely Definition B.1. Then, for ℓ growing as N , the inequality (20) becomes

$$\lim_N f_N(\sigma_c) \leq \liminf_N f_N(\tilde{\sigma}_N).$$

Therefore,

$$scf(\sigma_c) > \liminf_N f_N(\tilde{\sigma}_N) \geq \lim_N f_N(\sigma_c).$$

As a consequence, for N large enough,

$$scf(\sigma_c) > f_N(\sigma_c).$$

Substituting the definition of $scf(\sigma_c)$ and $f_N(\sigma_c)$ in the previous inequality, it follows that

$$k\sigma_c - \tilde{r}_c > k\sigma_c - \rho_N(\sigma_c),$$

or, equivalently, $\tilde{r}_c < \rho_N(\sigma_c)$. This latter inequality is a contradiction due to Lemma A.6, hence the liminf inequality is satisfied.

- 2) (existence of a recovery sequence). We choose as a converging sequence $\sigma_N = \sigma_c + 1/N$ and show that

$$scf(\sigma_c) = \lim_N f_N(\sigma_N). \quad (22)$$

For (22) to be fulfilled, it is sufficient to prove that $\tilde{r}_c = \lim_N \rho_N(\sigma_N)$. To this end, we note that

$$|\rho_N(\sigma_N) - \tilde{r}_c| \leq |\rho_N(\sigma_N) - \rho(\sigma_N)| + |\rho(\sigma_N) - \tilde{r}_c| < \varepsilon.$$

This result follows from the uniform convergence of ρ_N to ρ in $(\sigma_c, \sigma_{\max}]$ (inferred by the arguments following Eq. (19)) and due to the fact that $\lim_N \rho(\sigma_N) = \tilde{r}_c$. Hence, the second requirement is also satisfied.

Concerning the Γ - $\lim_N f_N(\sigma)$ for $\sigma \in \{\sigma_{\min}, \sigma_{\max}\}$, we can proceed as in the case of $\sigma = \sigma_c$. Indeed, by using both the regularity of ρ_N and ρ and the convergence of ρ_N to ρ , we get that Γ - $\lim_N f_N(\sigma) = scf(\sigma)$ for $\sigma \in \{\sigma_{\min}, \sigma_{\max}\}$.

In conclusion, we have ensured that $\Gamma\text{-}\lim_N f_N(\sigma) = scf(\sigma)$ for every $\sigma \in [\sigma_{\min}, \sigma_{\max}]$. Renaming $scf(\sigma)$ in (19) as $f_\infty(\sigma)$, we have that

$$f_N(\sigma) \xrightarrow{\Gamma} f_\infty(\sigma).$$

Then, relying on Theorem B.8, we get in our case

$$\exists \min_{\mathbb{R}^+} f_\infty(\sigma) = \lim_{N \rightarrow +\infty} \inf_{\mathbb{R}^+} f_N(\sigma) = \lim_{N \rightarrow +\infty} \min_{\mathbb{R}^+} f_N(\sigma). \quad (23)$$

Moreover, since all functions f_N admit a minimum point σ_N^* (which exists by virtue of Lemma A.2), then, up to subsequences, σ_N^* converges to a minimum point of f_∞ . According to (19), the only minimum point of f_∞ is σ_c . Hence, by (23), it follows that

$$\lim_{N \rightarrow +\infty} k\sigma_N^* - \rho_N(\sigma_N^*) = k\sigma_c - \tilde{r}_c. \quad (24)$$

Accordingly,

$$\sigma_N^* \rightarrow \sigma_c.$$

□

B Some basics of Γ -convergence

In this section, we introduce some abstract notions and results on Γ -convergence. We start by recalling the concepts of lower and upper limits and of lower-semicontinuous functions to introduce the definition of the Γ -limit. We also define the lower-semicontinuous envelope of a function and provide an example of computation of the Γ -limit by noting how this can be different from the pointwise limit. Finally, we show that, under suitable conditions, Γ -convergence implies convergence of minimum values and minimizers.

From now on, unless otherwise specified, X will be a metric space equipped with the metric d .

Definition B.1 Let $f : X \rightarrow \overline{\mathbb{R}}$. We define the lower limit (liminf for short) of f at x as

$$\begin{aligned} \liminf_{y \rightarrow x} f(y) &= \inf\{\liminf_j f(x_j) : x_j \in X, x_j \rightarrow x\} \\ &= \inf\{\lim_j f(x_j) : x_j \in X, x_j \rightarrow x, \exists \lim_j f(x_j)\}, \end{aligned}$$

and the upper limit (limsup for short) of f at x as

$$\begin{aligned}\limsup_{y \rightarrow x} f(y) &= \sup\{\limsup_j f(x_j) : x_j \in X, x_j \rightarrow x\} \\ &= \sup\{\lim_j f(x_j) : x_j \in X, x_j \rightarrow x, \exists \lim_j f(x_j)\}.\end{aligned}$$

Definition B.2 A function $f : X \rightarrow \overline{\mathbb{R}}$ is lower-semicontinuous at $x \in X$ if, for every sequence (x_j) converging to x , we have

$$f(x) \leq \liminf_j f(x_j), \quad (25)$$

or, in other words, $f(x) = \min\{\liminf_j f(x_j) : x_j \rightarrow x\}$. We will say that f is lower-semicontinuous on X if it is l.s.c. at all $x \in X$.

Definition B.3 (Γ -convergence) A sequence $f_j : X \rightarrow \overline{\mathbb{R}}$ Γ -converges in X to $f_\infty : X \rightarrow \overline{\mathbb{R}}$ if for all $x \in X$ we have

(i) (liminf inequality) for every sequence (x_j) converging to x

$$f_\infty(x) \leq \liminf_j f_j(x_j); \quad (26)$$

(ii) (limsup inequality) there exists a sequence (x_j) converging to x such that

$$f_\infty(x) \geq \limsup_j f_j(x_j). \quad (27)$$

The function f_∞ is called the Γ -limit of (f_j) , and we write $f_\infty = \Gamma\text{-}\lim_j f_j$. Condition (ii) can be substituted by the following:

(ii') (existence of a recovery sequence) there exists a sequence (x_j) converging to x such that

$$f_\infty(x) = \limsup_j f_j(x_j). \quad (28)$$

Definition B.4 Let $f : X \rightarrow \overline{\mathbb{R}}$ be a function. Its lower-semicontinuous envelope scf is the greatest lower-semicontinuous function not greater than f , that is, for every $x \in X$

$$scf(x) = \sup\{g(x) : g \text{ l.s.c.}, g \leq f\}.$$

Proposition B.5 If $f_j \rightarrow f$ uniformly on an open set U , then

$$\Gamma\text{-}\lim_j f_j = scf \quad \text{on } U.$$

Proof See Remark 1.38 of Braides (2002).

Below, we report an example that highlights the different roles of the *limsup* and *liminf* inequalities. It is also useful to visualize in a simple case of a sequence of real functions the difference between the classical pointwise (or uniform) limit and the Γ -limit.

Example B.6 Let $f_j(t)$ be a sequence of functions, where

$$f_j(t) = \begin{cases} \pm 1 & \text{if } t = \pm 1/j, \\ 0 & \text{otherwise.} \end{cases}$$

Note that $f_j \rightarrow 0$ pointwise, but $\Gamma\text{-}\lim_j f_j = f_\infty$, where

$$f_\infty(t) = \begin{cases} -1 & \text{if } t = 0, \\ 0 & \text{if } t \neq 0. \end{cases}$$

Indeed, the sequence f_j converges pointwise (and hence also Γ -converges) to 0 in $\mathbb{R} \setminus \{0\}$, while the optimal sequence for $t = 0$ is $t_j = -1/j$, for which $f_j(t_j) = -1$.

Definition B.7 (*Coerciveness conditions*) A function $f : X \rightarrow \overline{\mathbb{R}}$ is mildly coercive if there exists a nonempty compact set $K \subset X$ such that $\inf_X f = \inf_K f$. A sequence (f_j) is equi-mildly coercive if there exists a nonempty compact set $K \subset X$ such that $\inf_X f_j = \inf_K f_j$ for all j .

Theorem B.8 Let (X, d) be a metric space, let (f_j) be a sequence of equi-mildly coercive functions on X , and let $f_\infty = \Gamma\text{-}\lim_j f_j$; then

$$\exists \min_X f_\infty(x) = \lim_{j \rightarrow +\infty} \inf_X f_j(x). \quad (29)$$

Moreover, if all functions f_j admit a minimizer x_j^* , then, up to subsequences, x_j^* converges to a minimum point of f_∞ .

Proof See Theorem 1.21 and Remark 1.22 in Braides (2002).

C Continuity of $\sigma \mapsto \rho(\sigma)$

Given $X \equiv [0, 1]$ and $\Sigma \equiv [\sigma_{\min}, \sigma_{\max}]$, define $G : X \times \Sigma \rightarrow \mathbb{R}$ as $G(x, \sigma) = F(x; \sigma) - x$, where F has been introduced in Sect. 2. Then, for any $\sigma \in \Sigma$, we can define $D_\sigma = \{x \in X : G(x, \sigma) \leq 0\} \subseteq X$. With these new notations, $\rho(\sigma)$ as defined in Lemma 2.1 can be rephrased equivalently as

$$\rho(\sigma) = \min_{x \in D_\sigma} x,$$

because $F(0; \sigma) > 0$ for every $\sigma \in [\sigma_{\min}, \sigma_{\max}]$. Note that the relation $\sigma \mapsto D_\sigma$ is formally a correspondence, mapping each $\sigma \in \Sigma$ into a subset of X . We now introduce the definition of upper(lower)-semicontinuity for a correspondence as defined in Sundaram (1996) and prove a technical lemma.

Definition C.1 A correspondence $\Phi : \Sigma \rightarrow P(X)$, where $P(X)$ denotes the power set of X , is said

i) upper-semicontinuous (usc) at σ if, for all open sets V such that $\Phi(\sigma) \subset V$, there exists an open set U containing σ , such that $\sigma' \in U$ implies $\Phi(\sigma') \subset V$.

We say that Φ is usc on $S \subseteq \Sigma$ if it is usc for each $\sigma \in S$;

ii) lower-semicontinuous (lsc) at σ if, for all open sets V such that $V \cap \Phi(\sigma) \neq \emptyset$, there exists an open set U containing σ , such that $\sigma' \in U$ implies $V \cap \Phi(\sigma') \neq \emptyset$.

We say that Φ is lsc on $S \subseteq \Sigma$ if it is lsc for each $\sigma \in S$.

Lemma C.2 *The correspondence $\sigma \mapsto D_\sigma$ is compact valued; moreover, it is both upper and lower-semicontinuous on the intervals $[\sigma_{\min}, \sigma_c]$ and $(\sigma_c, \sigma_{\max}]$. Therefore, on the same intervals, it is continuous.*

Proof Compactness is easy to see, since D_σ is a closed subset of X which is bounded. The closeness is due to the fact that D_σ is the preimage of a closed set through a continuous function. We now prove upper-semicontinuity on $(\sigma_{\min}, \sigma_c)$. To this end, fix $\sigma \in (\sigma_{\min}, \sigma_c)$ and take any open $V \in \mathbb{R}$ containing D_σ . Now define $U = (\sigma - \delta, \sigma + \delta) \subset (\sigma_{\min}, \sigma_c)$, for $\delta > 0$, and consider σ' such that $\sigma' \in U$. By way of contradiction, suppose that $D_{\sigma'} \not\subset V$; put differently, $D_{\sigma'} \cap V^c \neq \emptyset$. Then there exists $x \in X$ such that $x \in V^c$ and $G(x, \sigma') \leq 0$. Since G is not constant and $D_{\sigma'}$ is not a singleton, we can assume $G(x, \sigma') < 0$. On the other hand, $x \in V^c$ implies $x \notin V$, hence $x \notin D_\sigma$. Therefore, $G(x, \sigma) > 0$. As a consequence, we can find ε small enough such that $|G(x, \sigma) - G(x, \sigma')| > \varepsilon$; this latter inequality contradicts the continuity of G in σ , since, by assumption, $\sigma' \in U = (\sigma - \delta, \sigma + \delta)$. To prove usc for $\sigma = \sigma_{\min}$, we use the same argument, where now $U = (\sigma_{\min}, \sigma_{\min} + \delta)$, $\delta > 0$. Similarly, for $\sigma = \sigma_c$, we can take $U = (\sigma_c - \delta, \sigma_c)$, $\delta > 0$. The usc on the open interval $(\sigma_c, \sigma_{\max})$ is proved using the same argument, as well as the usc for $\sigma = \sigma_{\max}$, considering $U = (\sigma_{\max} - \delta, \sigma_{\max})$, $\delta > 0$.

To prove the lower-semicontinuity on the open interval $\sigma < \sigma_c$, we consider V such that $V \cap D_\sigma \neq \emptyset$. Consider, as before, $U = (\sigma - \delta, \sigma + \delta)$, for $\delta > 0$. Take any $\sigma' \in U$. By way of contradiction, suppose that $V \cap D_{\sigma'} = \emptyset$. This means that there exists at least one $x \in V \cap D_\sigma$ that does not belong to $D_{\sigma'}$. Since $\sigma < \sigma_c$, D_σ is not a singleton and therefore we can assume $G(x, \sigma) < 0$; moreover, $G(x, \sigma') > 0$. As before, we contradict the continuity of G . Consider now $\sigma = \sigma_{\min}$, the argument holds for $U = (\sigma_{\min}, \sigma_{\min} + \delta)$, $\delta > 0$.

Concerning $\sigma = \sigma_c$, by Lemma A.2, we know that there exist exactly two solutions to the equation $G(x, \sigma_c) = 0$; the smallest, x^l , is such that $x^l < \mu$, whereas the largest one is $x^h > \mu$. Moreover, for all $\sigma' \in (\sigma_c - \delta, \sigma_c)$, $G(x^l, \sigma') < 0$ and therefore $x^l \in V \cap \Phi(\sigma')$ for any V such that $x^l \in V$. Therefore, for such σ' , the definition of lsc is guaranteed. Finally, lsc on $(\sigma_c, \sigma_{\max})$ is proved using the same argument as for the open set $(\sigma_{\min}, \sigma_c)$, while the lsc at $\sigma = \sigma_{\max}$ is obtained by considering $U = (\sigma_{\max} - \delta, \sigma_{\max})$.

To provide evidence that the correspondence $\sigma \mapsto D_\sigma$ is not continuous at σ_c (from the right), we show that for $\sigma = \sigma_c$ the lsc fails when considering the open interval $U = (\sigma_c, \sigma_c + \delta)$. As stated, in case $\sigma = \sigma_c$, there exist two solutions to the equation $G(x, \sigma_c) = 0$ such that $x^l < \mu < x^h$. Consider now V such that $x^l \in V$ but

$V \cap [\mu, 1] = \emptyset$. In this way, $V \cap D_{\sigma_c} \neq \emptyset$. Now define $U = (\sigma_c, \sigma_c + \delta)$, for $\delta > 0$ and take any $\sigma' \in U$. In this case, $V \cap D_{\sigma'} = \emptyset$.²³ This contradicts lower-semicontinuity. \square

Proposition C.2 *The map $\sigma \mapsto \rho(\sigma)$ is continuous on $[\sigma_{\min}, \sigma_c]$ and on $(\sigma_c, \sigma_{\max}]$.*

Proposition C.2 is now a straightforward corollary of Theorem 9.14 of Sundaram (1996). The assumptions of that theorem are that the target function is continuous and that the constraint, defined through the correspondence $\sigma \mapsto D_\sigma$, is compact valued and continuous. Our target function is clearly continuous and the assumptions on $\sigma \mapsto D_\sigma$ are ensured by Corollary C.2. For convenience, we report below the statement of Theorem 9.14 taken from Sundaram (1996), adapting it to our notation. Note that $f(x, \sigma)$ and $f^*(\sigma)$, as in the statement of such a theorem, correspond, respectively, to the identity function $x \mapsto x$ related to the first component and to our map $\rho(\sigma)$.

Theorem C.3 (Theorem 9.14, Sundaram) *Let $f : X \times \Sigma \rightarrow \mathbb{R}$ be a continuous function and let $D_\sigma : \Sigma \rightarrow P(X)$ be a compact-valued, continuous correspondence. Let $f^* : \Sigma \rightarrow \mathbb{R}$ be defined by*

$$f^*(\sigma) = \min\{f(x, \sigma) | x \in D_\sigma\}.$$

Then f^ is a continuous function on Σ .*

References

- Allen, P.M., Sanglier, M.: A dynamic model of growth in a central place system. *Geogr. Anal.* **11**(3), 256–272 (1979)
- Barucci, E., Tolotti, M.: Social interaction and conformism in a random utility model. *J. Econ. Dyn. Control* **36**(12), 1855–1866 (2012a)
- Barucci, E., Tolotti, M.: Identity, reputation and social interaction with an application to sequential voting. *J. Econ. Interac. Coord.* **7**(1), 79–98 (2012b)
- Bass, F.M.: A new product growth for model consumer durables. *Manage. Sci.* **15**(5), 215–227 (1969)
- Blume, L., Durlauf, S.: Equilibrium concepts for social interaction models. *Int. Game Theory Rev.* **5**(03), 193–209 (2003)
- Braides, A.: Γ -Convergence for Beginners, Oxford Lecture Series in Mathematics and its Applications, 22. Oxford University Press, Oxford (2002)
- Carlier, G.: A general existence result for the principal-agent problem with adverse selection. *J. Math. Econom.* **35**(1), 129–150 (2001)
- Ethier, S.N., Kurtz, T.G.: *Markov Processes: Characterization and Convergence* (Vol. 282). Wiley (2009)
- Ford, W.F., Pennline, J.A.: When does convergence in the mean imply uniform convergence? *Am. Math. Mon.* **114**(1), 58–60 (2007)
- Ghisi, M., Gobbino, M.: The monopolist's problem: existence, relaxation, and approximation. *Calc. Var. Partial Differ. Equ.* **24**(1), 111–129 (2005)
- Gordon, M., Nadal, J.-P., Phan, D., Semeshenko, V.: Entanglement between demand and supply in markets with bandwagon goods. *J. Stat. Phys.* **151**(3–4), 494–522 (2013)
- Granovetter, M.: Threshold models of collective behavior. *Am. J. Sociol.* **83**(6), 1420–1443 (1978)
- Milgrom, P., Roberts, J.: Comparing equilibria. *Am. Econ. Rev.*, 441–459 (1994)

²³ This latter relation is due to the fact that, when the solution x_0 to $G(x, \sigma) = 0$ is unique, i.e., for $\sigma > \sigma_c$, then $x_0 > \mu$.

- Monteiro, P., Page, F.H., Jr.: Optimal selling mechanisms for multiproduct monopolists: incentive compatibility in the presence of budget constraints. *J. Math. Econom.* **30**(4), 473–502 (1998)
- Nadal, J.-P., Phan, D., Gordon, M.B., Vannimenus, J.: Multiple equilibria in a monopoly market with heterogeneous agents and externalities. *Quant. Finance* **5**(6), 557–568 (2005)
- Peres, R., Muller, E., Mahajan, V.: Innovation diffusion and new product growth models: a critical review and research directions. *Int. J. Res. Mark.* **27**(2), 91–106 (2010)
- R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/> (2018)
- Rochet, J.C., Choné, P.: Ironing, sweeping and multidimensional screening. *Econometrica* **66**, 783–826 (1998)
- Schelling, T.C.: Dynamic models of segregation. *J. Math. Sociol.* **1**(2), 143–186 (1971)
- Strogatz, S.H.: *Nonlinear Dynamics and Chaos. With Applications to Physics, Biology, Chemistry, and Engineering.* Second edition. Westview Press, Boulder, CO (2015)
- Sundaram, R.K.: *A First Course in Optimization Theory.* Cambridge University Press, Cambridge (1996)