Social Learning about Consumption*

Isabelle Salle[†] & Pascal Seppecher[‡]

January 12, 2015

*We are grateful to two anonymous referees and an associate editor, the participants to the internal seminar at CeNDEF, University of Amsterdam, on February 15, 2013, the participants to the internal seminar at GREDEG, University of Nice-Sophia Antipolis, on October 16, 2013, and the participants to the 2014 CEF conference in Oslo, Norway for helpful discussions and feedbacks. We also thanks Christopher Carroll, Cars Hommes, Hendrick Hagedorn, Murat Yildizoglu and Martin Zumpe for interesting comments on an earlier draft of this paper. Isabelle Salle acknowledges financial support from GREThA, University of Bordeaux and from the EU FP7 RAstaNEWS project, grant agreement No. 320278.

†CeNDEF, Amsterdam School of Economics - University of Amsterdam - Roetersstraat 11, NL-1018 WB Amsterdam & Tinbergen Institute - The Netherlands & GREThA, UMR CNRS 5113 - University of Bordeaux - Avenue Léon Duguit - 33608 Pessac - France; Corresponding author: I.L.Salle@uva.nl

‡Centre d'Economie de Paris Nord (CEPN) - UMR CNRS 7234 - University of Paris Nord & Groupe de Recherche en Droit, Economie, Gestion (GREDEG), UMR CNRS 7321 - University of Nice Sophia Antipolis - France.

Social Learning about Consumption

Proof to be sent to : Isabelle Salle, CeNDEF, Amsterdam School of Economics - University of Amsterdam - Roetersstraat 11, NL-1018 WB Amsterdam & Tinbergen Institute - The Netherlands, +31 20 52 54 214, I.L.Salle@uva.nl

Abstract: This paper applies a social learning model to the optimal consumption rule of Allen & Carroll (2001), and delivers convincing convergence dynamics towards the optimal rule. These findings constitute a significant improvement regarding previous results in the literature, both in terms of speed of convergence and parsimony of the learning model. The learning model exhibits several appealing features: it is frugal, easy to apply to a various range of learning objectives, and requires few procedures and little information. Particular care is given to behavioural interpretation of the modelling assumptions in light of evidence from the fields of psychology and social science. Our results highlight the need to depart from the genetic metaphor, and account for intentional decision-making, based on agents' relative performances. By contrast, we show that convergence is strongly hindered by exact imitation processes, or random exploration mechanisms, which are usually assumed when modelling social learning behaviour. Our results suggest a method for modelling bounded rationality, which could be interestingly tested in a wide range of economic models with adaptive dynamics.

Keywords: learning, bounded rationality, evolutionary algorithms, consumption rule.

JEL-classification: D83; D91; C63; E21.

We are continually living a solution of problems that reflection cannot hope to solve.

Van den Berg (1955)

1 Introduction

The standard way to model how individuals deal with intertemporal decisions is to assume that they use a dynamic stochastic optimization procedure, based on a complete set of elements of information for the problem. Yet, empirical and experimental evidence questions the plausibility of the framework of substantive – or unbounded – rationality, and the corresponding optimization under constraints and rational expectation assumptions, especially concerning complex dynamic decisions under uncertainty (see, notably, the extensive literature initiated by Kahneman and Tversky and their colleagues, and Simon 1996, Chap. 3 and 4). A leading example is the theory of lifetime utility maximization under labour income uncertainty and liquidity constraints. Carroll (1997, 2001) demonstrates that the solution to the optimal consumption problem implies that agents follow a "buffer stock" rule: they target a level of cash-on-hand, and use it to smooth their consumption path in face of unanticipated variations in their labour income. Even if the computation of this optimal rule requires an astonishing amount of information and mathematical ability, empirical findings suggest that consumers do exhibit similar buffer-stock saving behaviour (see, for example, Deaton 1991, Carroll 1997). The question that immediately arises is the following: how can individuals, in real life, come to this optimal rule if we, quite realistically, hypothesise that they are not endowed with extremely sophisticated mathematical skills and powerful computers?[1]

This question is often addressed by means of the Friedman (1953) "as if" postulate, according to which agents are, of course, not involved in very demanding optimization programs but, instead, can roughly learn the resulting optimal behaviour by a process of trial and error. At first glance, this argument sounds plausible, all the more so that Allen & Carroll (2001) prove that the exact optimal non-linear buffer-stock rule can be very efficiently approximated by a simple linear rule involving two parameters, whose interpretation is straight-forward : the intercept stands for the target level of cash-on-hand, and the slope for the speed at which consumers try to return to that target when they have moved away from it. This simplified form of the rule is thus a natural candidate to serve as a benchmark situation to test learning models of intertemporal choice under uncertainty.

However, attempts to test M. Friedman's assumption within this particular framework yield rather disappointing results, and the literature fails to construct convincing learning processes to explain how agents come to make use of this rule. This suggests that the theoretical question, whether boundedly rational agents can actually learn to behave in a way predicted by models of unbounded rationality, is far from being completely resolved.

Allen & Carroll (2001) find that individual agents would need an absurd amount of iterations (roughly four million) to approximate this rule by a simple trial/error process. The difficulty arises because today's consumption choices have consequences on consumption for later periods, so that the performance of a given rule is not immediately observable. Palmer (2012) demonstrates that the number of iterations can be drastically reduced if the problem is parallelized among agents. This mechanism still requires the exhaustive exploration of a discrete space of strategies, and the issue of coordination is not raised, as it is assumed that all agents adopt the optimal rule as soon as it has been discovered by one of them. Two other contributions consider reinforcement learning, i.e. a process which selects an action rule within a set of rules, with a probability that increases with the relative past performance of this rule. Lettau & Uhlig (1999) use a classifier system to discriminate between different consumption rules, including the optimal one, and find a selection bias towards rules that yield the highest performances in periods with high incomes. This bias arises because the selection mechanism is only based on past performances of the strategies. Moreover, they consider a very simplified intertemporal choice, by allowing agents to make only discrete and binary consumption choices. Howitt & Özak (2013) provide more encouraging results, and show that consumers can quickly discover the optimal rule, but only by adding a very sophisticated updating derivative-based mechanism, which involves, among other elements, the computation of marginal utility. They further show that imitation strongly enhances the speed of convergence towards the optimal rule, but again by using a very sophisticated local imitation process. The most encouraging results can be found in Yıldızoğlu et al. (2014), in which agents develop a mental model of their environment thanks to an artificial neural network, and use it to form adaptive expectations. With this behaviour, agents are able to approximate and even, under certain circumstances, converge towards the optimal consumption rule. By contrast, they show that individual learning modelled through a genetic algorithm, even augmented with an imitation process between consumers, yields unconvincing results. The main progress initiated by their framework is to offer explicit modelling of adaptive expectations and forward-looking behaviour in a bounded rationality context. However, their network involves many parameters (almost ten), and convergence is quite long (around 1200 iterations).

This paper overcomes these two weak points by means of a parsimonious social learning model, and delivers convincing convergence dynamics towards the optimal consumption rule of Allen & Carroll (2001) within a limited amount of time. Modelling learning under bounded rationality proves to be a very challenging task. While models of substantive rationality appear to form a self-contained and unified framework, there is no consensus about the way bounded rationality should be represented, and several attempts have been developed. This paper is related to two strands of this literature.

The first strand applies evolutionary algorithms, initially developed to optimize non-linear and sophisticated functions, to model adaptive behaviour, be it collective or individual (see, for example, Sargent 1993, Arifovic 1995, Dawid & Hornik 1996, Arifovic 2000, Vallée & Yıldızoğlu 2009). Usually, this form of behaviour is inspired by genetic algorithms (see, notably, Holland 1975, Goldberg 1989, Holland et al. 1989, Holland & Miller 1991)[2]. One criticism that can be addressed to those algorithms is that they treat agents as automata, while individuals make deliberate choices (Rubinstein 1998, p. 2). As noticed by Waltman et al. (2011), little attention is usually paid to the economic interpretation of these algorithms.

Closer to the frontier with psychology, the second line of research also explores the

Simon (1955) concept of bounded rationality. In this strand of literature, agents tend to simplify complex decision problems, and much human decision-making can be described by simple algorithmic process models, called heuristics. These heuristics specify the cognitive process that leads agents to a satisficing solution. They are depicted as *fast*, because they require a low amount of information, and hence allow for quick decision making, and *frugal*, because they involve but few parameters in the design of the process and, hence, avoid overfitting issues in the case of small learning samples (Hoffrage & Reimer 2004). Furthermore, their predictive power has been proved to be at least as good as that of more sophisticated algorithms, such as standard statistical procedures (see Hutchinson & Gigerenzer 2005 for a variety of examples). Accordingly, Gigerenzer & Selten (2001) develop the concept of the "adaptive toolbox", which is a repertoire of specific-purpose heuristics designed to make decisions under uncertainty, by dispensing with optimization and calculations of probabilities and utilities. Heuristics and bounded rationality are not envisioned in the sense of Kahneman & Tversky (1996), that is as a rationale to the observed systematic deviations from standard probability laws, nor as optimization under constraints of time, knowledge or resources. Rather they are conceived as simple procedures that "make us smart", and allow us to make decisions with realistic mental resources.

In line with the evolutionary learning literature, this paper starts by implementing a basic tournament evolutionary algorithm (see e.g. Vriend 2000), and then proceeds through a series of successive improvements. Simple heuristics are progressively introduced into the learning model in order to depart from the genetic metaphor, and this paper culminates in a simple social learning model, which exhibits several appealing features. First, its functioning appears to be *intuitive*, and its behavioural interpretation is made easy. While our model is derived from an evolutionary algorithm, it models deliberate decision-making in intelligent agents, and economic and behavioural interpretation is discussed regarding evidence from experiments and observations in psychology and social science[3]. Second, it is particularly *frugal*, and only involves two free parameters. Third, procedures and information requirements are extremely limited. This model can thus be applied to a various range of problems, either static or dynamic, either in a discrete or in a continuous search space, either in a one-dimensional or in a multi-dimensional set of strategies. Furthermore, it is well in tune with Simon's (1955) concept of bounded rationality.

Our main results can be summarized as follows. First, we obtain sound and particularly stable consumption behaviour, without any unrealistic erratic fluctuations, and we even obtain good convergence to the optimal consumption rule. Importantly, we overcome one of the weaknesses of previous results within this framework, by drastically reducing the time taken to converge to the optimal solution. This evidence suggests that the proposed learning model does a very good job in describing the collective learning of agents with realistic cognitive capacities, acting under bounded rationality, without being able to see the whole picture of their environment.

Second, we highlight the key features of convergence towards optimal behaviour under learning. The tension between *exploration* of the search space and *exploitation* of collected information is a major feature of choices under uncertainty. In that respect, we emphasize the social dimension of learning, by allowing for a collective approach of that tension[4]. Our learning algorithm combines *experience sharing* between agents and *oriented search* in the space of potential strategies. This combination allows agents to learn how to smooth their consumption path over time, even through this intertemporal problem is known to be particularly hard to learn.

We highlight the need to depart from the genetic metaphor when designing learning models, and to account for intentional and oriented decision-making, based on agents' relative performances. By contrast, learning performances are strongly hindered by exact imitation processes ("copycat operator"), or global exploration mechanisms, which are usually assumed when using EAs to model agents' behaviour under bounded rationality. Furthermore, we find that increasing the selectivity of agents' relationships sharply improves their ability to learn what the optimal solution is. In this light, we offer an answer to the question asked by Allen & Carroll (2001, p. 13), and provide a proof of what they sense as being the "most plausible answer": "*the most interesting question to be addressed in a future literature on social learning about intertemporal choice is under what circumstances the population does and does not settle on a reasonably good set of rules*".

The rest of the paper is organized as follows: Section 2 exposes the setting of Allen & Carroll (2001) and the learning objective. In Section 3, we derive the social learning model that we intend to test within this setting. Section 4 describes the simulation protocol, Section 5 reports the results, and Section 6 presents our conclusions.

2 The general set-up : the learning objective

We use the original intertemporal consumption problem of Allen & Carroll (2001) as a framework.

The representative consumer solves the following maximization problem under liquidity constraint:

$$\max_{\{C_s\}_t^\infty} E_t \left[\sum_{s=t}^{\infty} \beta^{s-t} u(C_s) \right]$$
$$s.t. \quad X_{s+1} = X_s - C_s + Y_{s+1} \tag{1}$$
$$C_s \leq X_s \quad \forall s$$

where the utility function is derived only from consumption C, and is CRRA, $u(C) = \frac{C^{1-\rho}}{1-\rho}$ with $\rho = 3$. X_s is the total sum of resources available for consumption for period s. The labour income Y follows a stochastic process, and takes the values 0.7, 1 and 1.3, respectively with probability 0.2, 0.6 and 0.2, so that $E(Y) = 1$. Consumers cannot borrow but can save at a zero interest rate, and discount future utility at a rate $\beta = 0.95$.

In this framework, under mild conditions, Carroll (1997) shows that the optimal consumption rule can be rewritten as :

$$C^*(X_t) = 1 + f\left(X_t - \bar{X}^*\right) \tag{2}$$

with $f(.)$ a function with specific properties but no analytical expression and, more importantly, $\bar{X}^* \geq 1$ the target level of cash-on-hand. A first-order Taylor expansion of (2) around the point $X = \bar{X}^*$ yields :

$$C^*(X_t) \approx 1 + \gamma^*(X_t - \bar{X}^*) \tag{3}$$

where $\gamma^* \equiv f'(0)$. The approximation (3) delivers a complete plan of consumption that is only characterized by two parameter values, $\bar{X}^*$ and γ^*.

If we define a strategy $\theta = (\gamma, \bar{X})$ over a two-dimensional set Ω, taking into account the liquidity constraint, a consumption rule θ is given by:

$$C^\theta(X) = \begin{cases} 1 + \gamma(X - \bar{X}) & \text{if } 1 + \gamma(X - \bar{X}) \leq X, \\ X & \text{otherwise.} \end{cases} \tag{4}$$

As underlined in the introduction, the two parameters in θ find a natural interpretation: $\bar{X}$ is a cash-on-hand target, and γ indicates the speed at which consumers return to that target when away from it. Consequently, the rule (4) provides a good heuristic: consumers have a target level for a buffer-stock of liquid assets $\bar{X}$, that they use to smooth consumption in face of an uncertain income stream. They consume less than the expected income ($E(Y) = 1$) if the buffer-stock falls below the target, and vice-versa, the degree of adjustment depending on the coefficient γ.

In this setting, Allen & Carroll (2001) consider the complete set of strategies $\Omega = [0.05, 1] \times [1, 2.9]$, and demonstrate that the best approximation of the exact optimal non-linear rule over Ω is[5]:

$$\theta^* = (\gamma^*, \bar{X}^*) = (0.233, 1.243) \tag{5}$$

Figure 1 provides a 3-dimensional representation of the fitness landscape of this optimization problem. Starting with a given cash-on hand $X_0 \in \{0, 1, 2\}$, Allen & Carroll (2001) investigate whether consumers can discover this optimal strategy θ^* through systematic and individual exploration of the strategy space Ω, but arrive at very disappointing conclusions: they identify the fact that a "reasonably good" consumption rule requires a search time of roughly one million iterations. This therefore leads them to a discussion of the potential role of social learning and of *collective* exploration of the space Ω:

> *If there were a mechanism by which all of that information could be efficiently combined, the number of model periods required for finding the optimal rule could surely be radically reduced.* [...] *A potential mechanism to accomplish this purpose is "social learning".* [...] *Even if the social learning process is less than perfectly efficient it still seems plausible that it might lead a population of consumers to converge on the optimum relatively quickly.* (Allen & Carroll 2001, p. 13).

The aim of the next sections is to propose a simple social learning process, and to demonstrate that their conjecture is right.

[Figure 1 about here.]

3 Modelling social learning

The trade-off between the *exploitation* of high pay-off strategies (that already have been already discovered) and the *exploration* of the search space (looking for new actions that may potentially improve utility) has been identified as a major feature of iterated choice problems (Arthur 1991). Beginning with a standard Evolutionary Algorithm (hereafter, EA), we proceed by successive improvements in how that trade-off is dealt with, and end up with a social learning model, that provides a formalization of Simon's (1955) concept of procedural rationality.

At the beginning of each simulation, consumers' initial wealth is drawn out of the set $\{0, 1, 2\}$ with equal probability. For each period t, the labour income $Y_{i,t}$ is drawn independently for each consumer i, $i \in \{1, ...n\}$, and takes the values 0.7, 1 and 1.3, respectively with probability 0.2, 0.6 and 0.2, so that consumers receive heterogeneous labour income flows and a common initial wealth[6]. For each period, each consumer is endowed with a single strategy $\theta_{i,t}$, so that the population of strategies always contains n elements. The fitness of any strategy θ for any period t is given by the current utility $u(C_t^\theta)$ of the consumers, who are using the strategy θ in t.[7] Moreover, strategies are not binary encoded; rather we use real-valued numbers for γ and $\bar{X}$ values. This method allows for a direct behavioural interpretation of the learning model, while avoiding pitfalls associated with the use of elements for which the economic interpretation is unclear (see Waltman et al. 2011 for a critical discussion of this point).

3.1 EA1: a basic tournament EA with global exploration

EAs have been used as a collective approach to model the learning of interacting agents under bounded rationality. In this case, the population of strategies evolves through two operators: social learning between individuals (crossover operator) and random experimenting by some agents (mutation operator) (Vallée & Yıldızoğlu 2009). EA1 corresponds to a basic Tournament Evolutionary Algorithm (hereafter, TEA) with those two operators.

Crossover is implemented using a tournament selection of size $m < n$.[8] Each consumer i randomly draws a pool (the so-called tournament) of m other consumers among the whole population of n consumers, observes their current strategies and consumption levels, selects the two fittest consumers (say consumers k and l, $k, l \neq i$) to be the two mates, and uses crossover to combine their strategies θ_k and θ_l to update his own strategy θ_i. We assume an average crossover, i.e. consumers simply adopt the barycenter of the two selected strategies (see, for instance, Yıldızoğlu et al. 2014):

$$\theta_{k,t} = (\gamma_{k,t}, \bar{X}_{k,t}) \text{ and } \theta_{l,t} = (\gamma_{l,t}, \bar{X}_{l,t}) \longrightarrow \theta_{i,t+1} = (\gamma_{i,t+1}, \bar{X}_{i,t+1}) \equiv \left(\frac{\gamma_{k,t} + \gamma_{l,t}}{2}, \frac{\bar{X}_{k,t} + \bar{X}_{l,t}}{2} \right) \quad (6)$$

We further assume that crossover occurs with a fixed probability P_{co} for each period t and each consumer i. The crossover operator captures the idea of exchanges of information between agents (Arifovic 2000). This operator allows consumer i to *exploit* information contained in strategies θ_k and θ_l: assuming that the strategies of the two mates perform well, a combination of these two strategies is also likely to yield a high utility, or even a better one. Interestingly, this crossover operator has also the ability to introduce some heterogeneity into the population of rules. The new rule has elements that did not belong to either of the two mates used to create it[9]. However, crossover does not allow other strategies to be explored beyond the convex set of existing strategies. For this reason, a global exploration operator is also introduced, in order to ensure that all strategies may potentially be reached (the so-called ergodicity property), see e.g. Waltman et al. (2011) for a comparable mutation operator. With a fixed probability P_{mut} for each period, each consumer randomly draws a new strategy $\theta \in \Omega$[10]. Box 3.1 summarizes the implementation of EA1.

The choice of a tournament selection is justified in the light of evidence from social science and psychology[11]. Bounded rationality involves cognitive limitations in processing information, as well as in social interactions and organisational capabilities. As stressed by Simon (1962), limits exist to the number of people simultaneously involved in most forms of social interaction. For this reason, in our set-up, agents are supposed to be endowed with both bounded rationality, and what may be called "bounded sociability". This concept is translated into the choice of a fixed and small number m of agents in the tournament, that each agent is assumed to be able to observe. Moreover, endowing agents with a social "network" of this type provides an explanation for the information requirements of the crossover, whose interpretation may appear problematic in population-based models (see Fudenberg & Levine 1998).

This choice may also be justified from the perspective of experimental evidence in the field of psychology. For example, Tversky & Shaar (1982) show that when individuals have an arresting signal by which to discriminate between two options, they do not try to extend the decisions set. In our framework, consumers objectively discriminate on the basis of the relative observed utility of the members of their tournament, so that we can assume that they do not feel the need to enlarge the pool of candidates. More generally, social psychologists report that people imitate the actions of those who appear to have expertise (see, e.g., Bikhchandani et al. 1998). This evidence illustrates why agents select the fittest individuals in the tournament to update their strategies. This is also fully in line with the concept of *upward comparison* developed in the field of social psychology, according to which individuals tend to choose a comparison-target who slightly outperforms them as a means of self-improvement[12]. It is interesting to note that the tournament selection procedure is in line with findings in the field of behavioural biology too. Janetos (1980) argues that female animals follow simple rules of thumb to achieve good, but not optimal, matching, and select the best candidate male in a pre-set number of N males, after a process of sequential assessment (the so-called "best-of-N" rule).

TEA is inspired by the analogy with genetic heredity and crossover, and with genetic mutations. Exploration of the search space is purely stochastic, and can reach all points of the set Ω with equal probability (uniform draw). This algorithm exhibits efficient exploratory properties, but also constantly introduces disturbances in the population of strategies and, hence, may hinder coordination onto (or close to) the optimal solution. Furthermore, a major difference between a learning process and a natural selection process is that learning agents make intentional decisions and deliberate experimentation, while natural mutations are purely random (see, for example, Penrose 1952 for such a distinction). TEA is not able to fully account for such deliberate decision-making processes. Consequently, an oriented search operator is now introduced into the TEA with a twofold objective : allowing for local exploration, and modelling conscious behaviour in intelligent consumers.

Box 3.1 : learning algorithm (under EA1, EA2 and EA3)

Initialization

1. Endow each consumer i with a strategy θ_i.
2. Endow each consumer i with a pool of $m \in \{2, ..., n-1\}$ other consumers, indexed by $j \in \{1, ..., m\}$, with $j \neq i$.

Execution

3. For each period $t \leq T$ (T being the length of the simulation) and for each consumer $i \in \{1, ..., n\}$:
   (a) *Cross-over*: with an exogenous probability P_{co} (under EA1 and EA2) or whenever consumer i is the less fit of the tournament, i.e. whenever $u(C_{i,t}) \leq u(C_{j,t})$, $j \in \{1, ..., m\}$, with $j \neq i$ (under EA3) :
      i. Sort the m consumers of the pool by decreasing utility.
      ii. Take the first two agents of the pool to become the two mates (indexed by k and l),
      iii. Compute the new strategy $\theta_{i,t+1}$ (through Equation (6) under EA1, or through Equation (7) under EA2 and EA3).
      iv. Renew the tournament: randomly draw m new consumers among the whole population of $n-1$ consumers.
   (b) *Mutation* (EA1 only): with an exogenous probability P_{mut}, draw a new strategy $\theta_{i,t+1} \in \Omega$.

3.2 EA2: searching in promising regions

The selection procedure of the two mates k and l remains unchanged, but EA2 brings exploitation and exploration together into a single oriented search operator, which implements a distance-proportionate exploration (we follow here Eshelman & Schaffer 1993, Lux & Schornstein 2002)[13]. We assume that consumers explore the neighbourhood of their two mates around the barycenter $\{\frac{\gamma_k+\gamma_l}{2}, \frac{\bar{X}_k+\bar{X}_l}{2}\}$. The scale of the search area is proportional to the distance between the two mates (up to a scale factor $d \in]0,1]$)[14]. The intuition behind that modelling assumption can be stated as follows. Assuming that the two mates perform well, consumers explore the mates' region because they assume that this region is promising. The more distant the two mates, the greater the level of uncertainty regarding the position of the optimal strategy in the search area. Conversely, the closer the mates, the more promising the region, and the less incentive to move away from it.

Formally, for each period t, with a fixed probability P_{co}, each consumer i randomly draws a new strategy $\theta_{i,t+1}$ in a neighbourhood of his two mates k and l. This neighbourhood is defined as follows :

$$\theta_{k,t} = (\gamma_{k,t}, \bar{X}_{k,t}) \text{ and } \theta_{l,t} = (\gamma_{l,t}, \bar{X}_{l,t}) \longrightarrow \theta_{i,t+1} = (\gamma_{i,t+1}, \bar{X}_{i,t+1}) \text{ with}$$
$$\gamma_{i,t+1} \hookrightarrow \mathcal{U}\left(\frac{\gamma_k + \gamma_l}{2} - d \mid \gamma_{k,t} - \gamma_{l,t} \mid, \frac{\gamma_k + \gamma_l}{2} + d \mid \gamma_{k,t} - \gamma_{l,t} \mid\right)$$
$$\text{and } \bar{X}_{i,t+1} \hookrightarrow \mathcal{U}\left(\frac{\bar{X}_k + \bar{X}_l}{2} - d \mid \bar{X}_{k,t} - \bar{X}_{l,t} \mid, \frac{\bar{X}_k + \bar{X}_l}{2} + d \mid \bar{X}_{k,t} - \bar{X}_{l,t} \mid\right) \quad (7)$$

where the nested case $d = 0$ corresponds to the averaging crossover (6). The random draw introduces noise in the learning operator, which plays the role of randomness during the exploration process[15].

Figure 2 illustrates that mechanism : consumers k and l, whose strategies are $\theta_k = (\gamma_k, \bar{X}_k) = (0.7, 1.3)$ and $\theta_l = (\gamma_l, \bar{X}_l) = (0.5, 1.7)$, have been selected to be the two mates. By implementing EA1, the new strategy would be the barycenter $\theta = (\gamma, \bar{X}) =$

$(0.6, 1.5)$, see Equation (6). Under EA2, the consumer explores a uniform area *around* the barycenter, the precise size of this area depending on the absolute distance between the two mates k and l, and on the value of parameter d. The distance between k and l is 0.2 for strategy γ and 0.4 for strategy $\bar{X}$, so that the new strategy θ is randomly drawn in $(0.6 \pm 0.2d, 1.5 \pm 0.4d)$. The higher d, the wider the exploration area. For instance, this area is $[0.55, 0.65] \times [1.4, 1.6]$ if we set $d = 0.25$, and $[0.4, 0.8] \times [1.1, 1.9]$ if we set $d = 1$.

This modelling device clearly corresponds with the concept of *selective trial and error*, which lies at the heart of Simon's explanation of human problem solving: " *The trial and error is not completely random or blind; it is, in fact, rather highly selective.*" (Simon 1962, p. 472). "*Selectivity, based on rules of thumb or "heuristics", tends to guide the search into promising regions, so that solutions will generally be found after search of only a tiny part of the total space.*" (Simon (1978, p. 362)). This gives a strong rationale to the design of learning processes based on EAs, but deliberately departing from the biological analogy.

[Figure 2 about here.]

3.3 EA3: enhancing upward comparison

Under EA1 and EA2, the rate of occurrence of the learning operators is assumed to be exogenously fixed (i.e. P_{co} and P_{mut} are set by the modeller). The rationale behind this feature may be seen to be weak as soon as we intend to explicitly model deliberate decision-making in intelligent agents. As already mentioned in Sub-section 3.1, psychology and social sciences provide evidence of upward comparison among people concerned with self-improvement. In our set-up, consumers aim at finding the optimal consumption rule, in order to maximize their utility. We reinforce the upward comparison component of EA2, by assuming that strategy changes are an endogenous decision, following a simple routine. For each period, consumers draw a new strategy as described in Equation (7) whenever they are the less fit person in their tournament. Otherwise, they leave their strategy unchanged. In this case, they act according to the *satisficing principle*, in the sense that they consider that they have met an acceptable utility threshold and, consequently, retain their current strategy (see Simon 1976)[16]. The rest of the algorithm is unchanged, most particularly the tournament procedure, as does the selection of mates. We call this learning model EA3.

It should be underlined that this modification of the EA provides parsimony benefits, by ruling out the exogenous parameter P_{co}. EA3 is very frugal: it involves only two free parameters, i.e. the scale of the exploration area d and the tournament size m. This algorithm is therefore able to address a recurrent criticism levelled against agent-based models, which rightly points out that such models involve a high number of free parameters, making calibration and sensitivity analyses a challenging task (see Judd 2006 for a discussion).

3.4 EA4: endowing agents with a social memory

The selection of the two mates appears as a crucial part of our learning schemes, because both information exploitation and exploration of the search space depend on the two preselected candidates k and l. EA1, EA2 and EA3 all imply that the tournament be randomly and entirely renewed for each period. The concept of *bounded sociability* introduced in Subsection 3.1 would suggest that individuals could engage in some kind of networking, and keep track of the members of their past tournaments. The learning process could thus be augmented by memory[17], as stressed by Simon (1962, p. 473): "*various paths are tried out, the consequences of following them are noted, and this information is used to guide further research*". In our set-up, consumers are information carriers, in the sense that they convey records on pairs of implemented consumption rules/resulting utility. What is essential in the learning process is how this information flows among individuals.

Von Hippel et al. (2009, p. 3) develops the idea of a *pyramiding search* in order to find people with a rare attribute among a large population : "*Pyramiding is a search process based upon the idea that people with a strong interest in a topic tend to know people more expert than themselves*". This search model provides an appealing organisational model for the population of consumers, according to which consumers keep track of agents who are fitter than they are, and rule out those who are less fit, in order to reinforce the selectivity of the learning process.

For that reason, we now introduce a simple routine of tournament selection, mainly based on the assumption that agents have a (limited) knowledge and memory of the other consumers. The tournament is only partially renewed for each period, so that consumers tend to memorize agents with a relatively high level of utility: only the less fit consumer is removed from the tournament for each period, and randomly replaced by another, so that the tournament size remains constant (see Box 3.4). This pyramidal organisation is fully decentralized: consumers autonomously constitute their own tournament in a dynamic fashion, as it is updated for each period according to the relative fitness of each consumer and the tournament members. We stress that consumers keep in their tournament other consumers associated with successful strategies, and not these strategies themselves, so that a consumer stays as a member of the tournament even if his strategy has changed, unless he becomes the less fit of the tournament. This feature is a crucial part of EA4, and we further discuss this point in Subsection 5.1. The rest of the learning model EA3 remains unchanged. We then obtain EA4, which we refer to as the *pyramiding TEA*.

This model exhibits three interesting features. i) It is clearly consistent with the concept of bounded rationality and sociability, as information requirements are very limited: consumers only need to know the search space Ω, their current strategies and their resulting utility, and they are assumed to be able to observe m other strategies and corresponding utilities. ii) It selects the best strategies in terms of utility among a subset of the population, which in turn tends to select the best ones among the whole population, through a constant adaptation of that subset according to recorded fitnesses. iii) It is particularly frugal: as underlined above, besides the number of consumers n, there are only two free parameters, i.e. the size of exploration d and the tournament size m.

Box 3.4 : learning algorithm under EA4

Initialization

1. Endow each consumer $i \in \{1, ..., n\}$ with a strategy θ_i.
2. Endow each consumer i with a pool of $m \in \{2, ..., n-1\}$ other consumers, indexed by $j \in \{1, ..., m\}$, with $j \neq i$.

Execution

3. For each period $t \leq T$, and for each consumer $i \in \{1, ..., n\}$:
   (a) Sort the m consumers by decreasing utility.
   (b) Consider one of the alternatives:
      i. *Either* consumer i is the less fit of the tournament, *i.e.* $u(C_{i,t}) \leq u(C_{j,t})$, $\forall j \in \{1, ..., m\}$, with $j \neq i$, then implement the learning operator:
         A. Sort the m consumers of the tournament by decreasing utility.
         B. Take the first two consumers of the tournament to become the two mates (indexed by k and l),
         C. Compute the new strategy $\theta_{i,t+1}$ (through Equation (7)).
      ii. *Or* there exists agents in the tournament with a lower utility than consumer i, *i.e.* $\exists j \in \{1, ..., m\}$, so that $u(C_{j,t}) \leq u(C_{i,t})$, then:
         A. Remove from the tournament the less fit agent, *i.e.* consumer j, for who $u(C_{j,t}) \leq u(C_{k,t})$, $k \neq j$.
         B. Randomly draw one new consumer among the $n - m - 2$ other consumers to obtain m different consumers in the tournament.

We now assess whether these four EAs are able to deliver interesting learning dynamics within the optimal consumption rule framework.

4 Simulation protocol

We use the agent-based model of Yıldızoğlu et al. (2014), whose pseudo-code is given in Appendix A[18], as the underlying economic model of our simulations, and assess whether these four EAs may allow consumers to discover the optimal consumption rule of Allen & Carroll (2001). Table 1 in Appendix B provides a summary of the features of the four EAs and the calibration for the `baseline` scenario, as well as for the sensitivity analyses that we consider in Sub-section 5.3. In the baseline scenario, we set $n = 200$ consumers, the tournament size is set to $m = 10$, and the exogenous probabilities of mutation P_{mut} (under EA1) and crossover P_{co} (under EA1 and EA2) are randomly drawn from uniform distributions for each simulation: $P_{mut} \in [0.01, 0.1]$ and $P_{co} \in [0.05, 0.4]$. These values correspond to standard values in the literature[19]. Under EA2, EA3 and EA4, the scale of the exploration area d is drawn out of a uniform distribution $]0, 1]$ for each simulation. We run the simulations for $T = 200$ periods. We launch 100 simulations for each EA, and collect data every 10 periods.

5 Performances of social learning

5.1 Overview of the results

Figure 3 compares the average distances of $\bar{X}$ and γ values to optimal values at the end of the simulations (i.e. in $t = 200$) for the four EAs. Figure 4 provides the dynamics of those average distances over the whole period $[0, 200]$, as well as the evolution of the variance of the strategies among consumers, in order to assess their coordination.

[Figure 3 about here.]

[Figure 4 about here.]

We clearly see that neither EA1 nor EA2 exhibits convincing learning dynamics, strategy values are fairly stable during the simulations. Under EA2, consumers coordinate their strategies (their variance sharply decreases over time), but distances to optimal values remain high. This observation indicates that the exploration operator leads to premature convergence under EA2. The global exploration mechanism avoids that pitfall under EA1, and preserves the diversity of strategies, but this is obtained at the expense of consumer coordination. Accordingly, Figure 5 (left panel) shows the influence of the exogenous probabilities of mutation and crossover on the average distance to optimal consumption at the end of the simulations in the baseline scenario under EA1. The role of the probability of crossover is not striking[20]. On the contrary, a higher probability of mutation strongly deteriorates convergence performances under EA1, due to more heterogeneity between strategies. Figure 6 (left panel) reports similar conclusions in EA2.

[Figure 5 about here.]

[Figure 6 about here.]

By contrast, EA3 and EA4 allow strategies and the resulting consumption behaviour to move closer to their optimal values throughout time, which is an obvious sign of learning. The major difference between EA1 and EA2 on the one hand, and EA3 and EA4 on the other hand, lies in the occurrence of the exploration operator: it is implemented with an exogenous probability P_{co} in EA1 and EA2, while consumers activate it whenever they are the less fit among their tournament in EA3 and EA4. This feature directly connects the choice of strategy with the relative performances of the consumers, and appears as the major improvement of the EA in order to obtain interesting learning dynamics.

EA4 displays the best convergence pattern: not only *average* consumption behaviour moves closer to that obtained under the optimal rule[21] but the variance of strategies among consumers reaches the lowest level among the four algorithms after 200 periods, which indicates that *individual coordination* between consumers is particularly salient under EA4. It should be underlined how quick convergence and coordination are, typically within the first 100 periods. This is an improvement considering previous results in the literature (see Yıldızoğlu et al. 2014)[22].

The key difference between EA3 and EA4 lies in the procedure for tournament renewal. Consequently, our findings show that endowing agents with a dynamic social memory, based on the model of pyramiding search, provides the means to coordinate consumer behaviour with (or, at least closely with) the optimal consumption rule. This device efficiently balances selection pressure (which is increased under EA4 compared with EA3, as less fit agents are ruled out from the pool of potential mates), and strategy diversity (which is broadly the same under EA3 and EA4, as the learning operator is implemented according to the same procedure).

EA4 performances deserve further explanations. It should be recalled that the main challenge behind the consumption rule problem is that the pay-off of a given rule is not immediately observable, because consequences of consumption decisions today spread out over subsequent periods. Moreover, current utility at any period also depends on income, which is a stochastic and individual process. This is the reason why Allen & Carroll (2001) assume that each consumer has to evaluate each rule over a fixed amount of periods to obtain a rough idea of the associated utility level. Our learning mechanism is different, as consumers evaluate rules using only *current* utility. This assumption is appealing because it requires few knowledge and cognitive abilities in consumers, and fits well into the concept of bounded rationality. Furthermore, it allows for interesting learning dynamics, and we now provide intuitions behind this result.

Under EA4, the renewal procedure implies that consumers using good rules are less likely to be ruled out from the tournament, but consumers who do not smooth out their consumption are gradually excluded : even if they occasionally obtain high utilities in high-income periods by using bad rules, they perform poorly as soon as they face a low income level, and are removed from the tournament. Our implementation of tournament hence discriminates between good rules and "luck". Consequently, each consumer gradually constitutes a *library of experiences* through his tournament, in which members are those who have durably managed to obtain high utilities by smoothing out their consumption path, and therefore are using promising consumption rules. Furthermore, this library is dynamic, as it is updated with better and better performing rules every time consumers realise that the rule they have been using so far performs poorly (i.e. every time consumers are the less fit of their tournament). By combining *experience sharing* and *oriented search* for new rules, EA4 allows consumers to coordinate on (or at least close to) the optimal rule. Interestingly, this is in line with the mechanism suggested by Allen & Carroll (2001, p. 13) in their discussion about the potential role of social learning: "*rather than relying solely on their own (insufficient) experience, people observe the experiences of others and can learn from such observation and direct social communication.*"

We illustrate our findings in Figures 7-10. A picture of one run under each EA is reported, with specific learning parameter values (we further document the role of parameters in Sub-section 5.2). The 200 yellow points represent the distribution of the 200 consumers over the strategy space Ω, and the red cross indicates the optimal strategy θ^*.

From the left to the right, pictures display that distribution respectively for the initial period $t = 0$, and for periods $t = 50$, $t = 100$, and at the final stage of the simulation $t = 200$. We clearly see the poor performances under EA1 and EA2, the lack of coordination under EA1 and the premature convergence under EA2. EA3 and EA4 exhibit much more satisfying learning dynamics, and EA4 obviously displays the best convergence between consumers over the four EAs.

[Figure 7 about here.]

[Figure 8 about here.]

[Figure 9 about here.]

[Figure 10 about here.]

5.2 The role of the learning parameters

We now analyse which features of the EAs are key to ensuring convergence towards the optimal consumption rule, and we focus on EA3 and EA4, as these models exhibit the most convincing learning dynamics. The two learning parameters d and m may play an ambiguous role. Figures 11 and 12 display how these two parameters affect consumers' ability to move closer to the optimal rule in EA3 and EA4. Whatever the values of d, distances to optimal strategies are small compared with previous yet encouraging results in the literature[23].

[Figure 11 about here.]

[Figure 12 about here.]

Parameter d measures the exploratory strength of EA3 and EA4: the higher d, the wider the search space around the two mates (see Figure 2 and Equation (7)). However, wide search spaces may hinder coordination between individuals[24]. It appears that moderate values of d, i.e. between 0.4 and 0.8, deliver the most successful convergence.

Interestingly, our results are in line with some studies in the computer science literature. In particular, Herrera et al. (1998) perform a series of experiments with different crossover mechanisms, and find that the type of crossover that we implement under EA2, EA3 and EA4 outperforms a wide range of alternatives, including the barycenter crossover (implemented in EA1). They also find that the most efficient exploration and exploitation relationship is induced with moderate values of d, typically close to 0.5. This provides an additional rationale behind the use of the oriented-search operator, beside its behavioural interpretation as Simon's description of human problem solving.

The tournament size m measures the selection pressure of the EAs: the larger m, the stronger the selection pressure, so that agents with weak fitness are less likely to be selected as mates, and a deeper exploitation of the information embedded in the consumer population is made possible. On the other side, we should remember that in EA3 and EA4 the learning operator is only implemented when consumers perform less than all other m consumers in their tournament (see Box 3.4). Consequently, the higher m, the less frequent the exploration, and this effect may weaken the learning dynamics, and hinder the learning process.

In EA3, small-sized tournaments allow for a better convergence towards the optimal rule. This finding suggests that frequent exploration (corresponding to a low value of m) is preferable to a stronger selection pressure (corresponding to a high value of m). Figures 5 and 6 show a similar effect of m in EA1 and EA2. By contrast, learning dynamics under EA4 are less sensitive to tournament size, because the renewal procedure of the tournament is more selective than it is in EA3. As explained above, under EA4, by ruling out the less fit consumer for each period, the tournament gradually excludes consumers with low utilities. These consumers do not manage to smooth out their consumption: even if they may occasionally obtain a high utility in high-income periods, they face a significant drop in their consumption and, hence, in their fitness as soon as they have to cope with a decrease in their income. Even if exploration is less frequent (due to a high value of m), it is based on mates with a better fitness than under EA3, because the tournament itself is more selective. The pyramiding EA4 thus provides the additional appealing feature of being less sensitive to parameter values.

5.3 Robustness checks

Finally, we perform robustness checks in the `baseline` scenario regarding changes in the number of consumers n, the introduction of imitation through a probabilistic copying process, and heterogeneous initial endowments among consumers. Results are reported in Figure 13.

[Figure 13 about here.]

Changing the size of consumer population This paper emphasizes the importance of social learning. We thus naturally ask whether decreasing or increasing the size of the population may change our results. We find that learning performances are strongly robust to changes in the number of agents: dynamics are fairly comparable with a population of $n = 100$ or 400 consumers (see the first two panels of Figure 13). This result contrasts with the social learning mechanism implemented in Palmer (2012), in which increasing the size of the population allows more rules to be evaluated in parallel and mechanically decreases the time necessary for discovering the optimal rule.

Introduction of imitation Social learning is often modelled by implementing a "copy-cat operator", i.e. by allowing agents to copy the best strategy in the population with some probability for each period (the so-called *selection* operator, see Waltman et al. (2011) for a discussion in the social learning context). We therefore incorporate some imitation in the four EAs, with a 0.15 probability for each period and each consumer. On one side, exact imitation may enhance learning dynamics, by allowing the best strategies to spread among the population, and therefore increasing the selectivity of the search process[25]. On the other side, it can potentially lead to a premature loss of diversity in the existing strategies, by purely reproducing the best ones at a given period, to the expense of others and, hence, may impede convergence towards the optimal rule.

Our results show that this second negative effect clearly dominates (see the third panel of Figure 13). Imitation strongly hinders learning dynamics in all four EAs, probably because it leads to a premature convergence towards existing good-performing strategies, instead of using those strategies to further explore the search space. This negative effect is especially salient under EA3 and EA4.

Following a reviewer's suggestion, we perform an additional robustness check under EA1 in order to test whether an increase in imitation alongside an increase in mutation may improve the diffusion of strategies, and, hence, may address the issue of the lack of coordination between consumers highlighted in Sub-section 5.1. In order to do so, we run 100 additional simulations by drawing the probability of imitation from the uniform distribution $[0.2, 1[$ and the probability of mutation from $[0.01, 0.1]$ for each of these 100 simulations. We further set $P_{co} = 0.25$, $m = 10$, $n = 200$. Figure 14 reports the effects of mutation and imitation on the average distance to the optimal consumption level after $t = 200$ periods. We clearly see that more mutation strongly deteriorates the performances of EA1, whatever the level of imitation, and more imitation does not improve the performances of EA1, whatever the level of mutation. We conclude that no interaction effect between diffusion (imitation) and diversity (mutation) are likely to improve the poor performances of EA1.

[Figure 14 about here.]

This negative result also echoes the disappointing conclusion of Yıldızoğlu et al. (2014) as to the role of exact imitation within the same set-up : agents are heterogeneous regarding their income and wealth, so that sharing current rules corresponding to different current situations does not make learning more efficient. This finding emphasises the importance of heterogeneity for social learning efficiency, and should caution us against the use of an imitation process of this type when designing learning algorithms[26].

Heterogeneous initial wealth Up until this point, we have considered that all consumers start the simulation with the same initial endowment $X_{0,i} = 0$, 1, or 2, $\forall i$. Palmer (2012) suggests that considering heterogeneous initial wealth may hinder the social learning process, thereby making the process for learning the optimal consumption rule more challenging. We then allow for such a set-up : each consumer draws his own initial wealth $X_{0,i}$ in $\{0, 1, 2\}$, so that initial endowments are heterogeneous. The rest of the `baseline` scenario remains unchanged. Our results are once again strongly robust to that change (see the last panel of Figure 13).

6 Conclusions

Starting from a basic evolutionary algorithm, this paper progressively departs from the genetic metaphor, and derives a simple social learning model which might be referred to as a *pyramiding tournament evolutionary algorithm* after the concept of pyramiding search initiated by Von Hippel et al. (2009). This model exhibits several appealing features. i) It is especially frugal, as it involves only two free parameters, to which learning dynamics are little sensitive. ii) Its interpretation is relatively intuitive in terms of the intentional decision-making of intelligent agents, who try to adapt their behaviour in a complex dynamic environment of which they cannot see the whole picture. In that respect, specific attention is paid to the behavioural interpretation of the modelling assumptions, regarding evidence in psychology and social science. iii) It is easy to apply to a various range of learning objectives, as only the definition of the search space of strategies, whether discrete or continuous, is required. iv) It is parsimonious in terms of procedures and information requirements, thereby complying with the cognitive limitations implied by bounded rationality.

We apply this model to the optimal consumption rule of Allen & Carroll (2001), and we come up with two main results. First, we obtain convergence within a limited number of periods (less than 200), which is a significant improvement regarding previous attempts in the literature. We demonstrate that the intuition set forward by Allen & Carroll (2001) on the potential role of social learning in the coordination of consumers on this optimal rule was right. We explain consumers' ability to learn how to smooth their consumption path by the use of simple and plausible mechanisms of social comparison and selection, without the need to solve sophisticated optimization computations. Second, we highlight the key features of an efficient way of dealing with the trade-off between *exploration* of the search space and *exploitation* of information, in order to obtain convergence towards optimal behaviour under learning. In that respect, we emphasize the social dimension of learning, and the need to depart from the genetic metaphor when designing learning models, by accounting for intentional and oriented decision-making. Our learning algorithm is able to combine *experience sharing* between agents and oriented search in the space of potential strategies. This ability turns out to be very useful for dealing with consumption smoothing, even through this problem is known to be particularly challenging to solve. By contrast, learning performances are strongly hindered by exact imitation processes (copying), or global exploration mechanisms, which are usually assumed when using evolutionary algorithms to model agents' behaviour under bounded rationality.

However, even if the social learning model is able to coordinate agents on or, at least, close to the optimal consumption rule, agents are not endowed with an optimal behaviour rule beforehand, coordination takes time, so that agents are not capable of behaving optimally in the short- or medium-run. That would have been of minor importance if the economic environment were static and the optimum remained permanently unchanged, as is the case in the simple learning framework considered in this paper. In reality, new elements are constantly introduced into the environment, and optimal behaviour rules are modified as a consequence of these new elements, for example after a policy shock. Whether social learning, or any learning model, could allow agents to optimally react in the face of such changes within a limited amount of time remains an open question. Our results provide no reason to think that this could be the case.

Nevertheless, our results suggest a promising way of modelling social learning and bounded rationality, which could be tested most interestingly in more complex environments. General equilibrium learning problems, in which the average choices and beliefs of agents in turn affect individual pay-off require reactive learning schemes. Designing learning models suited for such situations constitutes a challenging task for future research.

Another extension of this work concerns the implementation of the social learning model in laboratory experiments with real subjects, in order for it to be confronted and calibrated with observed human behaviour. This exercise may provide further justification for promoting this model as a convincing candidate for the formalization of H. Simon's concept of procedural rationality in order to represent agents' behaviour when departing from the rational and optimizing agent benchmark.

Notes

[1]A similar question arises in the case of optimal portfolio selection during life cycles, see Binswanger (2011).

[2]Other algorithms have also been contemplated, notably classifier systems (see Arthur 1991) and artificial neural networks (see, for example, Salmon 1995, Yıldızoğlu 2001).

[3]See also Waltman et al. (2011) for a sensible discussion of the economic interpretation of evolutionary algorithm operators.

[4]It is interesting to note that, in a closely related experimental setting, Brown & Camerer (2009) demonstrate that social learning enhances human subject ability to converge towards optimal savings behaviour. The beneficial role of social learning in coordination has also been studied in evolutionary game theory, notably in the context of the choice between which technologies to use, see Ellison & Fudenberg (1993, 1995).

[5]Precisely, the implied sacrifice ratio is very low, around 0.003.

[6]This assumption is made to keep our set-up comparable to the one in Allen & Carroll (2001) and Yıldızoğlu et al. (2014). Robustness checks performed in Subsection 5.3 clearly show that our results remain unchanged if consumers receive heterogeneous initial wealth.

[7]Evaluating strategies over a longer period of time, by means of their average performance over the last *GARate* periods, has also been considered in the literature (see e.g. Vriend (2000)). In this case, the EA is only applied every *GARate* periods, and this setting mechanically increases the amount of simulation periods. This also introduces an artificial synchronisation of agents' learning steps which is difficult to justify regarding human problem solving in reality. For these reasons, we choose the fitness criterion to be the current utility. Subsection 5.1 clearly shows that this fitness criterion is enough to obtain convincing learning dynamics, and further elaborates on this point.

[8]See notably Bullard & Duffy (1998) or Vriend (2000) for a use of tournament selection in social learning models.

[9]An alternative to the average crossover could be an "exchange" crossover, which takes one element

from the first mate ($\bar{X}$ or γ) and the other one from the other mate to create a new rule. Additional simulations (not displayed in the paper, but available on request) have shown that results under EA1 deteriorate with this alternative crossover, because the only introduction of diversity stems from large and random mutations, which are disconnected from the relative performances of the mutated strategy. The trade-off between exploitation and exploration is therefore less efficient under an exchange crossover than under an average crossover.

[10]An alternative to this global exploration mechanism could be to perform mutations as random (possibly normal) draws around the strategy to be mutated, with an (exogenously fixed) standard deviation. This alternative does not improve our results under EA1.

[11]Another popular selection procedure is the roulette-wheel procedure, see *inter alia* Goldberg (1989), Arifovic (1995), Waltman et al. (2011). For the reasons we discuss here, we believe that tournament selection results in an easier behavioural interpretation.

[12]See Suls & Wheeler (2000) for a general statement, see Huguet et al. (2001) for evidence of upward comparison and performance-enhancing effects among children in a classroom.

[13]As we assume a single learning operator, we rule out global exploration, and set $P_{mut} = 0$.

[14]Exploration is bounded, so that strategies θ always remain within Ω during the whole simulation.

[15]We assume a uniform draw but results are robust to different specifications, e.g. a draw in a normal distribution around the barycenter. The point is that we assume a *random* draw, and thus allow for exploration, beyond the sole exploitation of the mates' strategies.

[16]This routine is similar in two ways to the one introduced in Dawid (1997) within an evolutionary game theoretical framework: it prescribes a switching rule that is based on the satisficing principle and only on current pay-offs. However, we assume a less rudimentary response. Interestingly, despite the very simple nature of this rule and limited information about the game, the author shows that players are able to find a Nash equilibrium. As will become clear in Section 5, we also find convergence dynamics thanks to the use of a routine of this nature.

[17]For a general statement of memory into EAs for dynamic optimization problems, see e.g. Franke (1999) or Yang (2008).

[18]The *Netlogo* program is available on request, the code is also available in *Java*.

[19]See, for example, Arifovic (1995), Bullard & Duffy (1998), Arifovic et al. (2013) or Yıldızoğlu et al. (2014).

[20]This result is in line with previous ones in the GA literature, see e.g. Lux & Schornstein (2005). Additional sensitivity analyses (available on request) show that the learning dynamics under EA1 are rather insensitive to a variation of P_{co} from 0.05 to 0.9.

[21]A Student test at 5% leads us to strongly reject the null hypothesis that average distance of consumption to optimal value in period 200 is equal under EA3 and EA4, against the alternative that it is smaller under EA4 (p-value $1.285e-05$).

[22]Howitt & Özak (2013) attain negligible welfare losses after less than 100 periods, but the learning scheme they implement is much more sophisticated that those we set forward.

[23]For instance, in Yıldızoğlu et al. (2014), average distances to optimal consumption remain mainly above 0.06.

[24]Figures 5 and 6 in also suggest the negative role of exploration in EA1, as discussed in Sub-section 5.1.

[25]It can be interpreted as a situation where the tournament size equals the population size, i.e. $m = n$, and only the information of the fittest individual in the tournament is exploited through copy.

[26]See also Salle et al. (2012) for an another discussion of harmful effects of imitation in a general equilibrium set-up.

References

Allen, T. W. & Carroll, C. (2001), 'Individual learning about consumption', *Macroeconomic Dynamics* **5**, pp. 255 – 271.

Arifovic, J. (2000), 'Evolutionary Algorithms in Macroeconomic Models', *Macroeconomic Dynamics* **4**, 373–414.

Arifovic, Y. (1995), 'Genetic algorithms and inflationary economies', *Journal of Monetary Economics* **36**(1), pp. 219–243.

Arifovic, Y., Bullard, J. & Kostyshyna, O. (2013), 'Social learning and monetary policy rules', *The Economic Journal* **123**(567), 38–76.

Arthur, W. B. (1991), 'Designing Economic Agents that Act Like Human Agents: A Behavioral Approach to Bounded Rationality', *American Economic Review* **81**(2), pp. 353–59.

Bikhchandani, S., Hirshleifer, D. & Welch, I. (1998), 'Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades', *Journal of Economic Perspectives* **12**(3), 151–170.

Binswanger, J. (2011), 'Dynamic decision making with feasibility goals: A procedural-rationality approach', *Journal of Economic Behavior & Organization* **78**(3), pp. 219–228.

Brown, A., C. Z. & Camerer, C. (2009), 'Learning and Visceral Temptation in Dynamic Saving Experiments', *Quarterly Journal of Economics* **124**(1), 197–223.

Bullard, J. & Duffy, J. (1998), 'A model of learning and emulation with artificial adaptive agents', *Journal of Economic Dynamics and Control* **22**(2), 179–207.

Carroll, C. (1997), 'Buffer stock saving and the life cycle permanent income hypothesis', *Quarterly Journal of Economics* **CXII**(1), pp. 1–56.

Carroll, C. (2001), 'A theory of the consumption function, with and without liquidity constraints', *Journal of Economic Perspectives* **15**(3), pp. 23–46.

Dawid, H. (1997), 'Learning of equilibria by a population with minimal information', *Journal of Economic Behavior & Organization* **32**, 1–18.

Dawid, H. & Hornik, K. (1996), 'The dynamics of genetic algorithms in interactive environments', *Journal of Network and Computer Applications* **19**, 5–19.

Deaton, A. (1991), 'Saving and liquidity constraints', *Econometrica* **59**(5), pp. 1221 – 1248.

Ellison, G. & Fudenberg, D. (1993), 'Rules of Thumb for Social Learning', *Journal of Political Economy* **101**(4), 612–643.

Ellison, G. & Fudenberg, D. (1995), 'Word-of-Mouth Communication and Social Learning ', *The Quarterly Journal of Economics* **110**(1), 93–125.

Eshelman, L. & Schaffer, J. (1993), Real-Coded Genetic Algorithms and Interval-Schemata, *in* 'Foundations of Genetic Algorithms 2', San Mateo: Morgan Kaufman.

Franke, J. (1999), Memory enhanced evolutionary algorithms for changing optimization problems, *in* 'Evolutionary Computation. proceedings of the 1999 Congress on', Vol. 3.

Friedman, M. (1953), *Essays in Positive Economics*, University of Chicago Press.

Fudenberg, D. & Levine, D. (1998), *Theory of Learning in Games*, Cambridge, MA : MIT Press.

Gigerenzer, G. & Selten, R. (2001), *Bounded Rationality: The Adaptive Toolbox*, the MIT Press.

Goldberg, D. E. (1989), *Genetic algorithms in search, optimization, and machine learning*, Addison-Wesley.

Herrera, F., Lozano, M. & Verdegay, J. (1998), 'Tackling real-coded genetic algorithms: operators and tools for behavioural analysis', *Artificial Intelligence Review* **12**, 265–319.

Hoffrage, U. & Reimer, T. (2004), 'Models of Bounded Rationality: The Approach of Fast and Frugal Heuristics', *The International Review of Management Studies* **15**(4), pp. 437–459.

Holland, J. (1975), *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*, U Michigan Press, Oxford, England.

Holland, J., Goldberg, D. & Booker, L. (1989), 'Classifier Systems and Genetic Algorithms ', *Artificial Intelligence* **40**, pp. 235–289.

Holland, J. & Miller, J. (1991), 'Artificial Adaptive Agents in Economic Theory', *AER Papers and Proceedings* **91**(2), pp. 365–370.

Howitt, P. & Özak, O. (2013), 'Adaptive Consumption Behavior', *Journal of Economics Dynamics & Control* **39**(C), 37–61.

Huguet, P., Dumas, F., Monteil, J. M. & Genestoux, N. (2001), 'Social comparison choices in the classroom further evidence for students' upward comparison tendency and its beneficial impact on performance', *European Journal of Social Psychology* **31**, 557–578.

Hutchinson, J. M. & Gigerenzer, G. (2005), 'Simple heuristics and rules of thumb: Where psychologists and behavioural biologists might meet', *Behavioural Processes* **69**, pp. 97–124.

Janetos, A. C. (1980), 'Strategies of female mate choice : a theoretical analysis', *Behavioral Ecological Sociobiology* **7**, pp. 107–112.

Judd, K. (2006), Computationally Intensive Analyses in Economics, *in* L. Tesfatsion & K. Judd, eds, 'Handbook of Computational Economic, vol. 2', North-Holland, chapter 17, pp. 881–894.

Kahneman, D. & Tversky, A. (1996), 'On the reality of cognitive illusions', *Psychological Review* **103**, pp. 582–591.

Lettau, M. & Uhlig, H. (1999), 'Rules of thumb versus dynamic programming', *The American Economic Review* **89**(1), 148–174.

Lux, T. & Schornstein, S. (2002), Genetic learning as an explanation of stylized facts of foreign exchange markets, Discussion Paper Series 1: Economic Studies 2002,29, Deutsche Bundesbank, Research Centre.

Lux, T. & Schornstein, S. (2005), 'Genetic learning as an explanation of stylized facts of foreign exchange markets', *Journal of Mathematical Economics* **41**(1-2), 169–196.

Palmer, N. (2012), Learning to Consume: Individual versus Social Learning. mimeo, George Mason University.

Penrose, E. T. (1952), 'Biological Analogies in the Theory of the Firm', *American Economic Review* **42**(5), pp. 804–809.

Rubinstein, A. (1998), *Modeling Bounded Rationality*, MIT Press.

Salle, I., Zumpe, M., Yıldızoğlu, M. & Sénégas, M.-A. (2012), Modelling Social Learning in an Agent-Based New Keynesian Macroeconomic Model, Technical Report 2012-20, Cahiers du GREThA, University of Bordeaux.

Salmon, M. (1995), Bounded Rationality and Learning: Procedural Learning, *in* 'Learning and rationality in economics', Oxford : Basil Blackwell edn, Kirman, A. P. and Salmon, M., chapter 8, pp. 236–275.

Sargent, T. (1993), *Bounded Rationality in Macroeconomics*, Oxford University Press.

Simon, H. (1955), 'A behavioural model of rational choice', *Quarterly Journal of Economics* **69**, pp. 99–118.

Simon, H. (1962), The architecture of complexity, *in* 'Proceedings of the American Philosophical Society', Vol. 106, pp. 467–481.

Simon, H. (1976), From substantial to procedural rationality, *in* 'Method and Appraisal in Economics', cambridge university press edn, S. J. Latsis, pp. 129–148.

Simon, H. (1978), Rational decision-making in business organizations. Nobel Memorial Lecture.

Simon, H. A. (1996), *The Sciences of the Artificial*, third ed. edn, Cambridge, Mass. : MIT Press.

Suls, J. & Wheeler, L., eds (2000), *Handbook of social comparison: Theory and research*, the plenum series in social/clinical psychology edn, Kluwer Academic Publishers. Dordrecht, Netherlands.

Tversky, A. & Shaar, E. (1982), 'Choice Under Conflict: The Dynamics of the Deferred Decision', *Psychological Science* **3**, 358–361.

Vallée, T. & Yıldızoğlu, M. (2009), 'Convergence in the finite Cournot oligopoly with social and individual learning', *Journal of Economic Behavior & Organization* **72**(2), 670–690.

Van den Berg, J. (1955), *The phenomenological approach to psychiatry: an introduction to recent phenomenological psychopathology*, Springfield, Ill.,Thomas.

Von Hippel, E., Franke, N. & Prügl, R. W. (2009), 'Pyramiding: Efficient search for rare subjects', *Research Policy* **38**(9), 1397–1406.

Vriend, N. (2000), 'An illustration of the essential difference between individual and social learning, and its consequences for computational analyses', *Journal of Economic Dynamics and Control* **24**, 1–19.

Waltman, L., van Eck, N. J., Dekker, R. & Kaymak, U. (2011), 'Economic modeling using evolutionary algorithms: the effect of a binary encoding of strategies', *Journal of evolutionary economics* **21**, 737–756.

Yang, S. (2008), 'Genetic Algorithms with Memory- and Elitism- Based Immigrants in Dynamic Environments', *Evolutionary Computation* **16**(3), 385–416.

Yıldızoğlu, M. (2001), 'Connecting adaptive behaviour and expectations in models of innovation: The Potential Role of Artificial Neural Networks ', *European Journal of Economic and Social Systems* **15**(3), 51–65.

Yıldızoğlu, M., Sénégas, M.-A., Salle, I. & Zumpe, M. (2014), 'Learning the optimal buffer - stock consumption rule of Carroll', *Macroeconomic Dynamics* **18**(4), 727–752.

A The pseudo-code of the computational model, based on Yıldızoğlu et al. (2014)

Initialization

1. Choose a learning model: EA1, EA2, EA3 or EA4.
2. Set parameter values.
3. Create n consumers, each consumer $i \in \{1, ..., n\}$ is endowed with:
   - a strategy $\theta_{i,0} = \{\gamma_{i,0}, \bar{X}_{i,0}\} \in \Omega$,
   - the common initial cash-on-hand $X_{i,0} \in \{0, 1, 2\}$, $\forall i$.
   - a tournament of $m \in \{2, ..., n-1\}$ other consumers, indexed by $j \in \{1, ..., m\}$, with $j \neq i$.
4. For each consumer i :
   - compute the corresponding consumption $C_{i,0} = \min\{X_{i,0}, 1 + \gamma_{i,0}(X_{i,0} - \bar{X}_{i,0})\}$,
   - compute the corresponding utility $u(C_{i,0})$,
   - compute the remaining cash-on-hand $X_{i,1} = X_{i,0} - C_{i,0} \geq 0$.
   - save the corresponding statistics for the analysis (*observer only*):
     - the optimal consumption level that would have been obtained if the optimal strategies θ^* had been used: $C^*_{i,0} = \min\{X_{0,i}, 1 + \gamma^*(X_{i,0} - \bar{X}^*)\}$,
     - the corresponding optimal utility $u(C^*_{i,0})$,
     - the corresponding remaining cash-on-hand $X^*_{i,1} = X_{i,0} - C^*_{i,0} \geq 0$
     - the distance between the observed levels and the optimal values $Z - Z^*$ for variables $Z \equiv \gamma$, $\bar{X}$, C, X and u.
5. Compute all other aggregate statistics from individual ones (*observer only*).

Execution

6. For each period $t \leq T$ (T being the length of the simulation):
   (a) For each consumer i:
   - draw a new income $Y_{i,t} \in \{0.7, 1, 1.3\}$ with probability $\{0.2, 0.6, 0.2\}$,
   - update the cash-on-hand $X_{i,t} = X_{i,t-1} + Y_{i,t} \geq 0$,
   - compute the corresponding optimal flow $X^*_{i,t} = X^*_{i,t-1} + Y_{i,t} \geq 0$ (*observer only*).
   - update the tournament.
   - update the strategy $\theta_{i,t-1}$:
     i. with a probability P_{imit}, take as the new strategy $\theta_{i,t}$ the strategy of the fittest consumer in the population in period $t-1$,
     ii. with a probability P_{co} under EA1 and EA2, or if consumer i is the less fit of the tournament (i.e. $u(C_{i,t-1}) < u(C_{j,t-1})$, $j \in \{1, ..., m\}$, with $j \neq i$) under EA3 and EA4, perform cross-over according to the chosen EA in Step 1 (Equation (6) under EA1, Equation (7) under EA2; EA3 and EA4).
     iii. with a probability P_{mut}, randomly draw a new strategy $\theta_{i,t} \in \Omega$,
     iv. otherwise, keep the strategy unchanged, i.e. $\theta_{i,t} = \theta_{i,t-1}$.
   - Execute Steps 4 and 5 using $\theta_{i,t}$ and $X_{i,t}$.

   (b) Update all other aggregate statistics from individual ones (*observer only*).

N.B. : *observer only* indicates the computation of indicators (either aggregate or individual ones) that we require to perform the results analysis, but that consumers do not observe.

B Calibration of the model

[Table 1 about here.]

	EA1	EA2	EA3	EA4
	basic TEA	TEA with distance-proportionate exploration	TEA with distance-proportionate and endogenous exploration	*pyramiding* TEA
tournament updating	renewed for each period			only the less fit is renewed for each period
P_{mut}: probability of global exploration	$\in [0.01, 0.1]$	0		
P_{co}: probability of cross-over	$\in [0.05, 0.4]$		endogenous (only when being the less fit of the tournament)	
d: size of exploration for cross-over	0	$\in [0, 1]$		
P_{imit}: probability of pure imitation	`baseline`: 0, sensitivity analysis: 0.15			
n: total number of consumers	`baseline`: 200, sensitivity analysis: $\{100, 400\}$			
m: tournament size	`baseline`: 10, sensitivity analysis: $\{5, 20\}$			

Table 1: Summary of the EAs and calibration

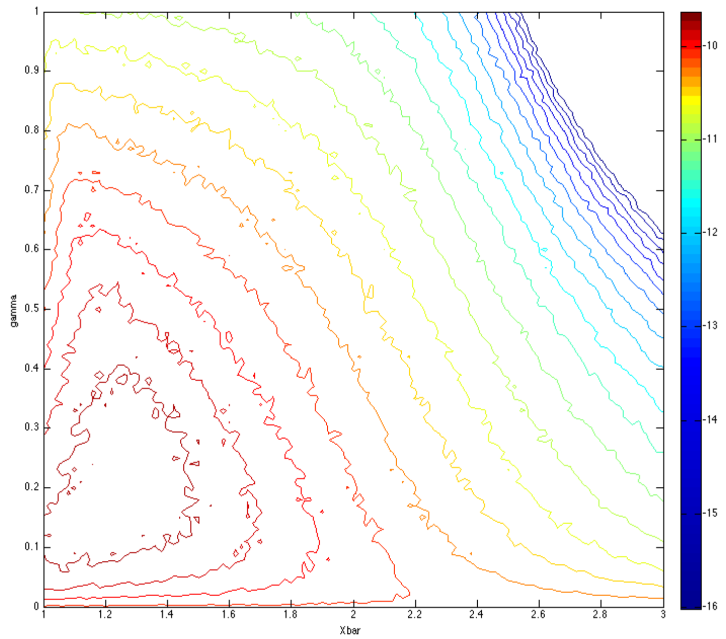


Figure 1: Discounted utility flow over 200 periods

Utility flows are evaluated at $101 \times 101 = 10,201$ points, which are uniformly distributed over $(\bar{X}, \gamma) \in \Omega = [1, 3] \times [0, 1]$, and averaged among 1000 agents at each point ($X_0 = 1$).

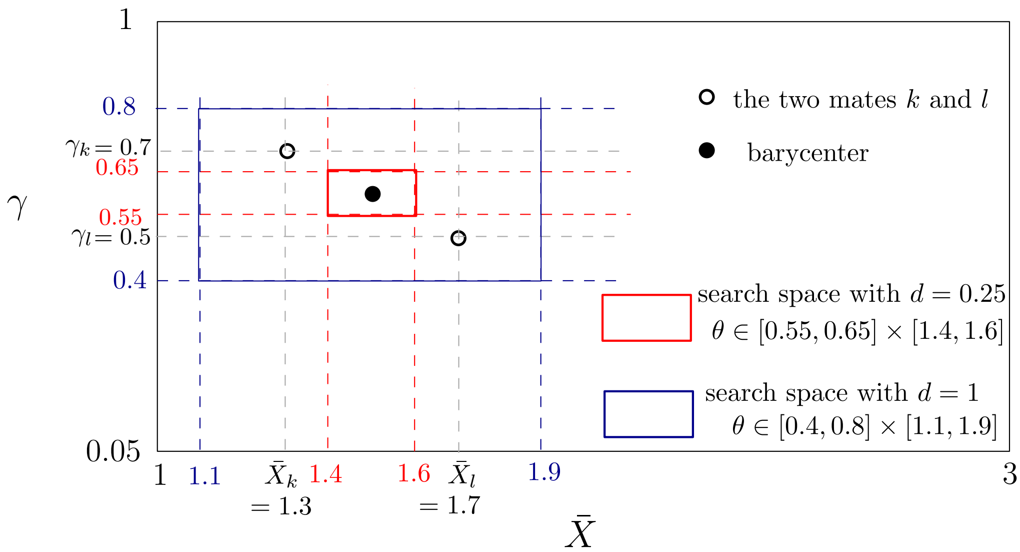


Figure 2: Examples of local exploration in EA2, EA3 and EA4.

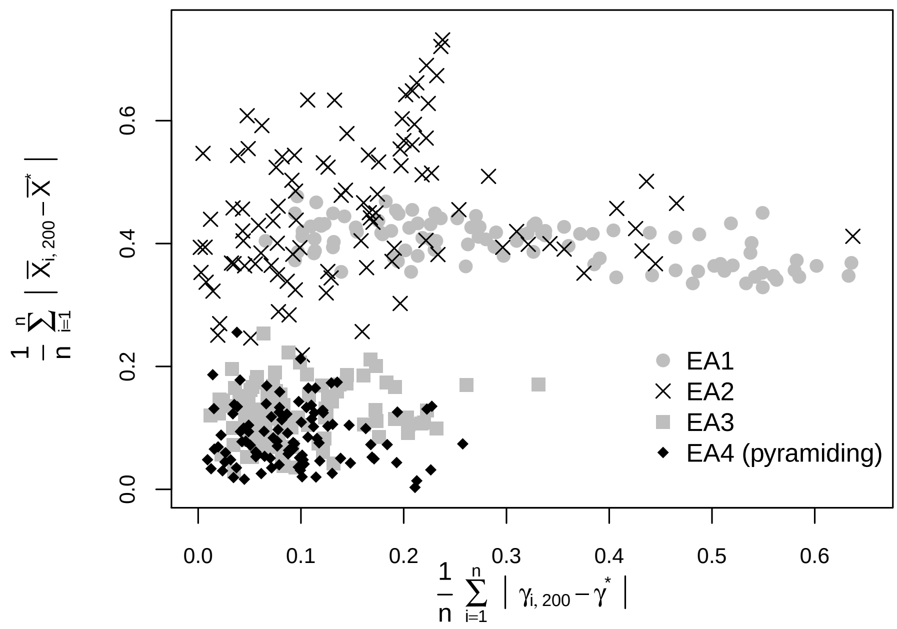


Figure 3: Average distances of $\bar{X}$ and γ values to optimal ones at the end of the simulations in the four EAs

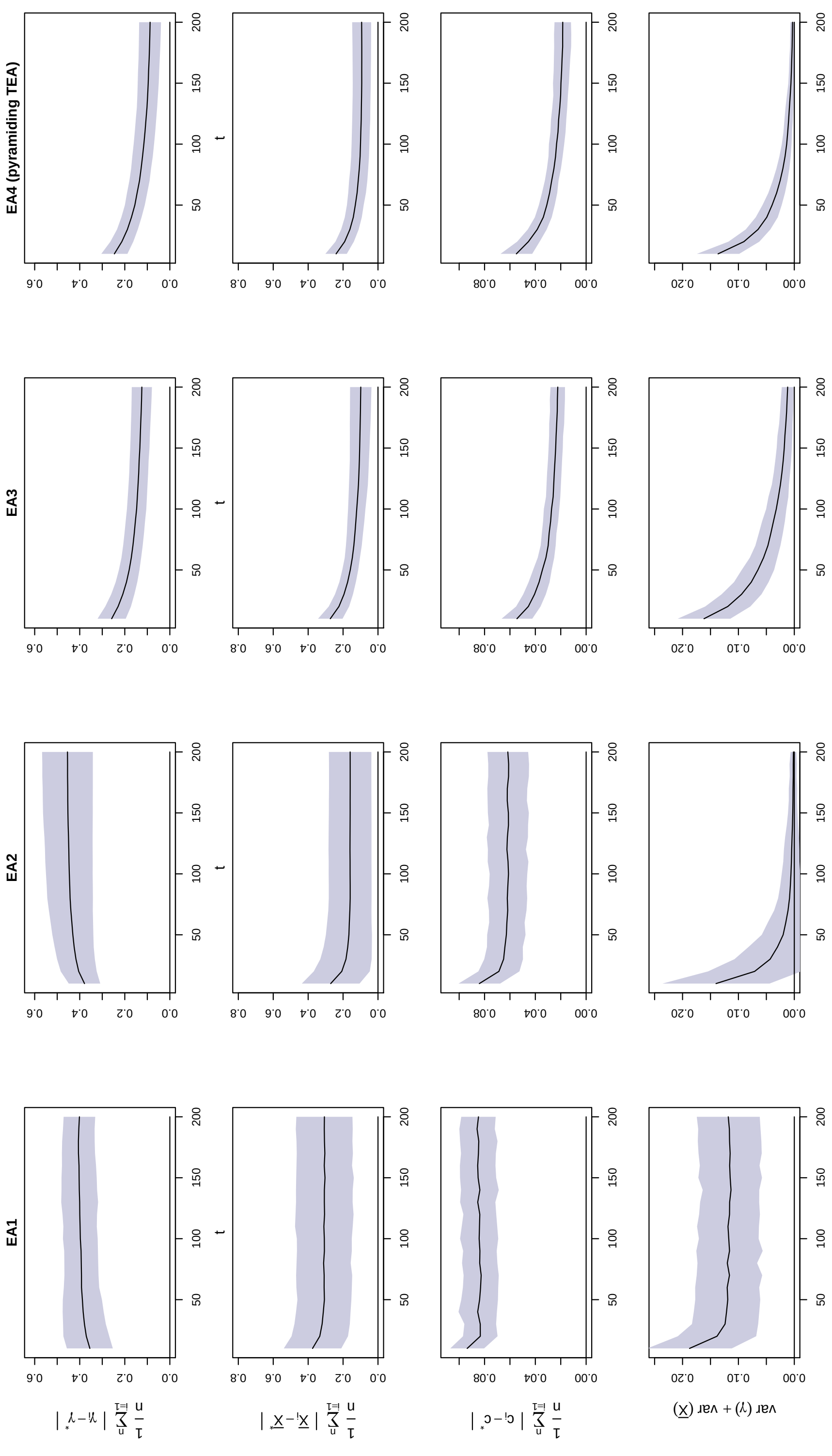
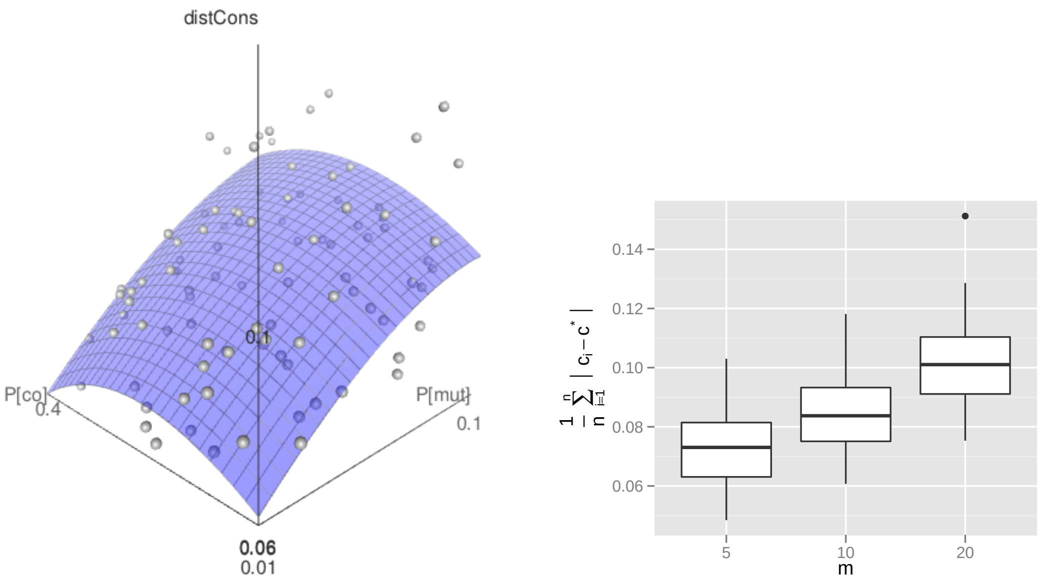


Figure 4: Baseline scenario: from the top panel to the bottom panel, for each of the four EAs: distances to optimal values for strategies γ, $\bar{X}$ and the resulting consumption c and variance of the strategies γ and $\bar{X}$ among the n consumers: average over the 100 runs, $+/-$ 1 standard deviation.



(a) Role of probabilities

(b) Role of the tournament size

Figure 5: Role of the learning parameters in EA1 on the distance to optimal consumption at $t = 200$ (z-axis).

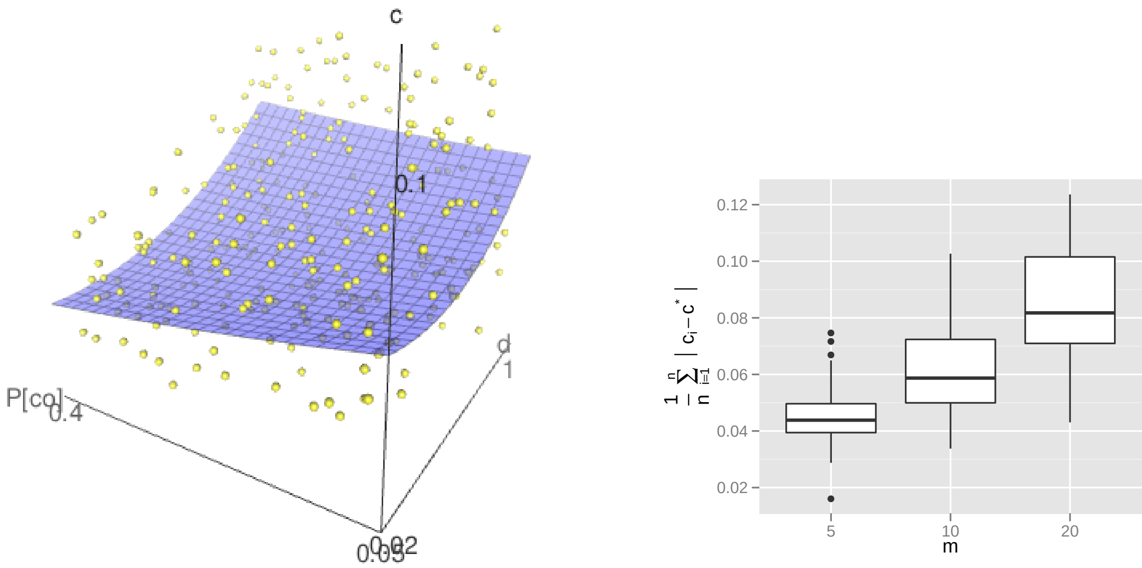


Figure 6: Role of the learning parameters in EA2 on the distance to optimal consumption at $t = 200$ (z-axis).

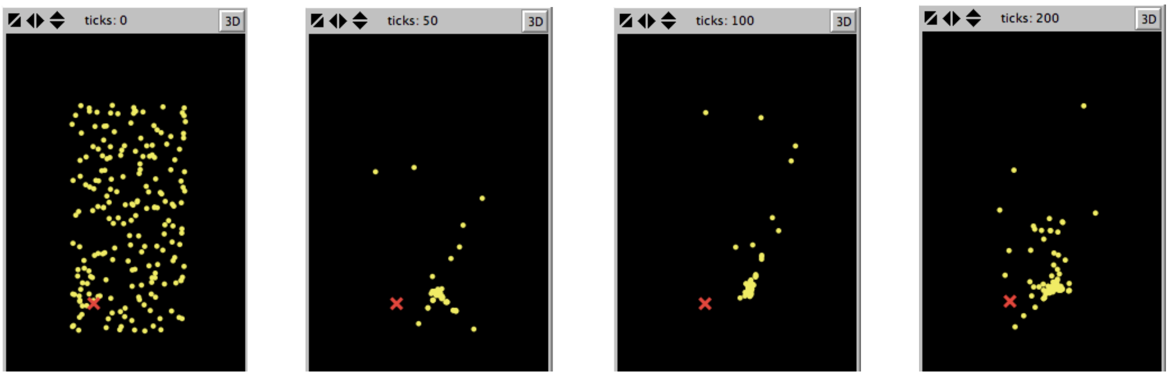


Figure 7: Illustrative run of EA1 ($P_{mut} = 0.01$, $P_{co} = 0.25$, $m = 10$).

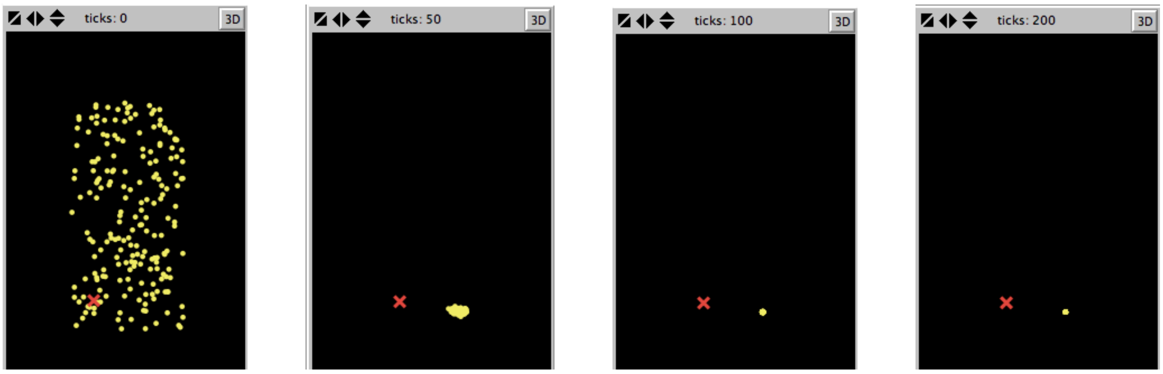


Figure 8: Illustrative run of EA2 ($d = 0.6$, $P_{co} = 0.25$, $m = 10$).

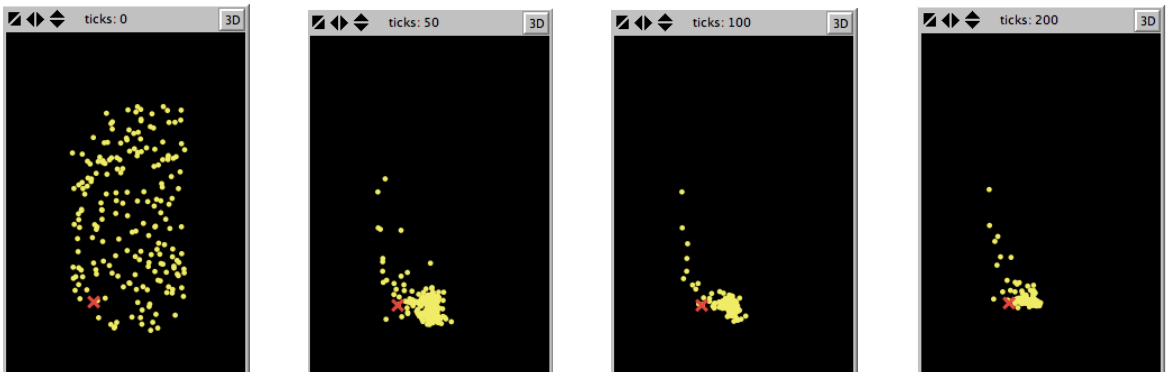


Figure 9: Illustrative run of EA3 ($d = 0.6$, $m = 10$).

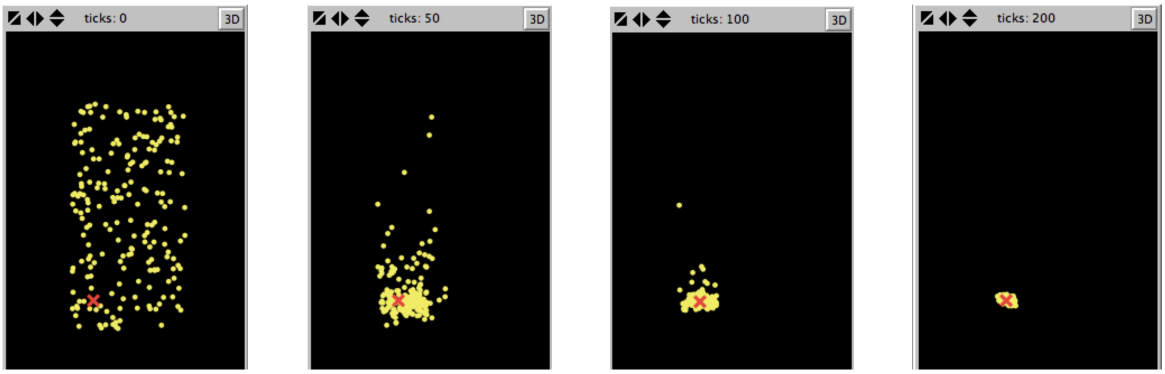


Figure 10: Illustrative run of EA4 ($d = 0.6$, $m = 10$).

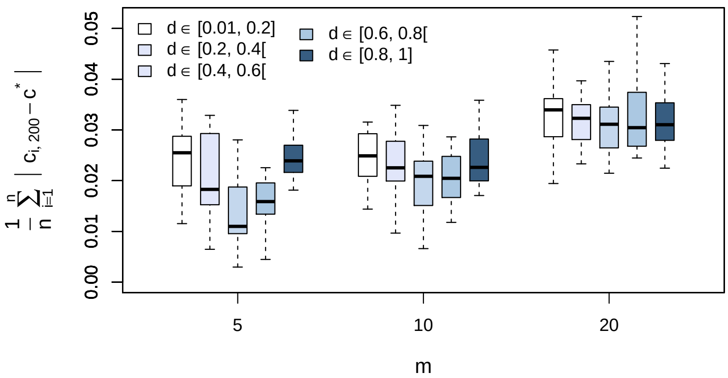


Figure 11: Average distances to optimal consumption at the end of simulations according to m and d values in EA3, 100 simulations per parameter configurations – `baseline` scenario

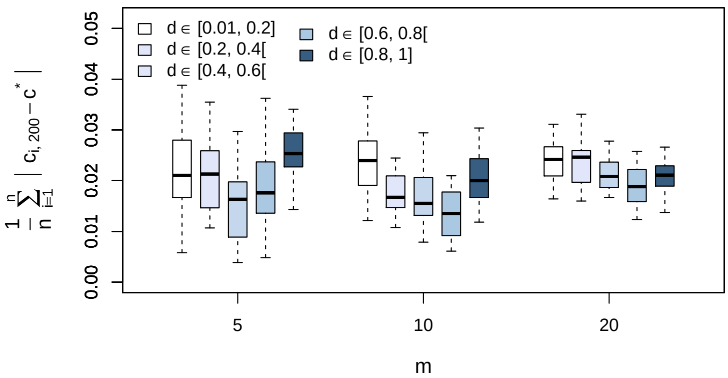


Figure 12: Average distances to optimal consumption at the end of simulations according to m and d values in EA4, 100 simulations per parameter configurations – `baseline` scenario

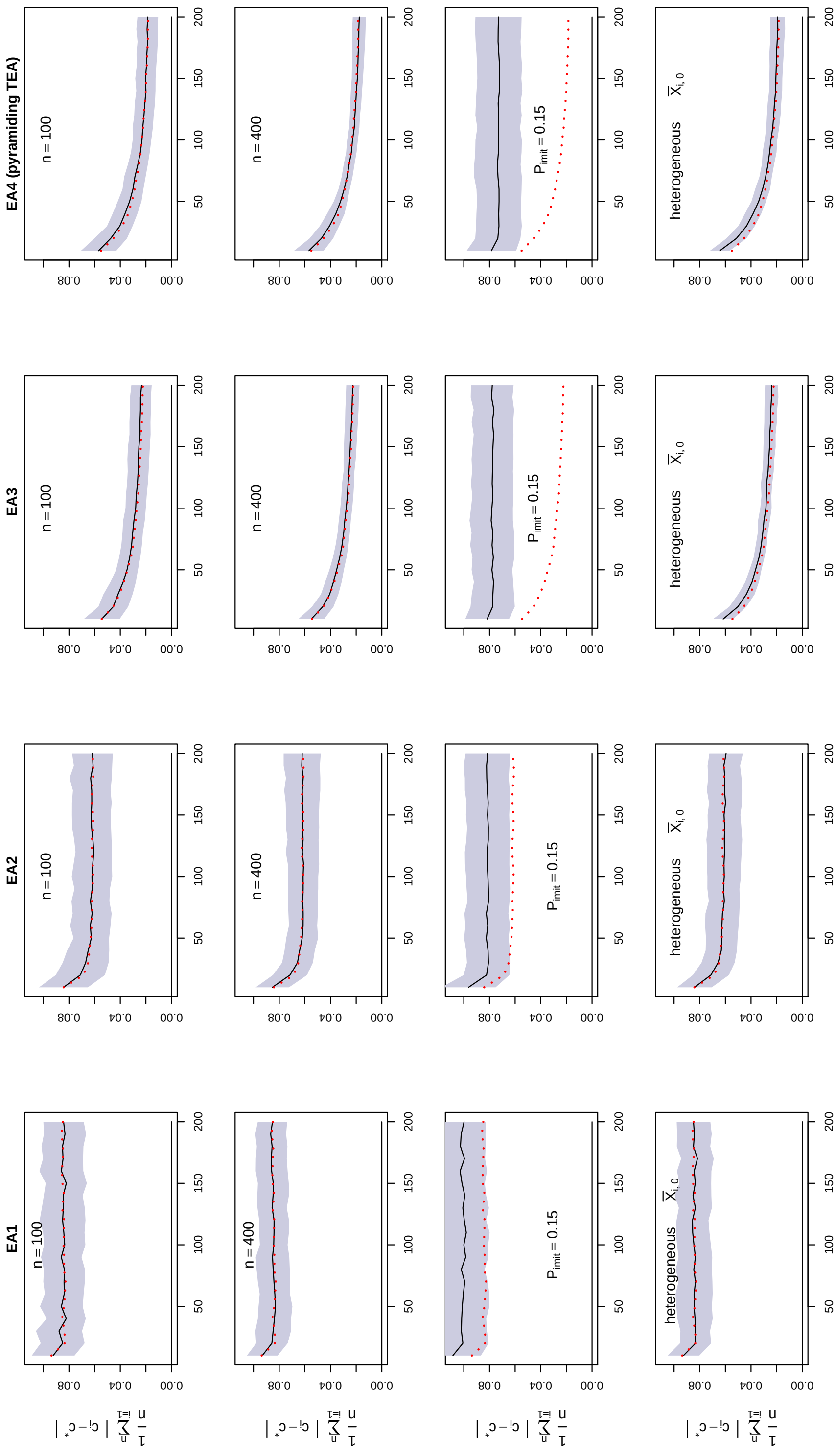


Figure 13: Robustness checks: distances to optimal values of the consumption level c among the n consumers : average over the 100 runs, +/- 1 standard deviation, the dotted line represents the baseline scenario.

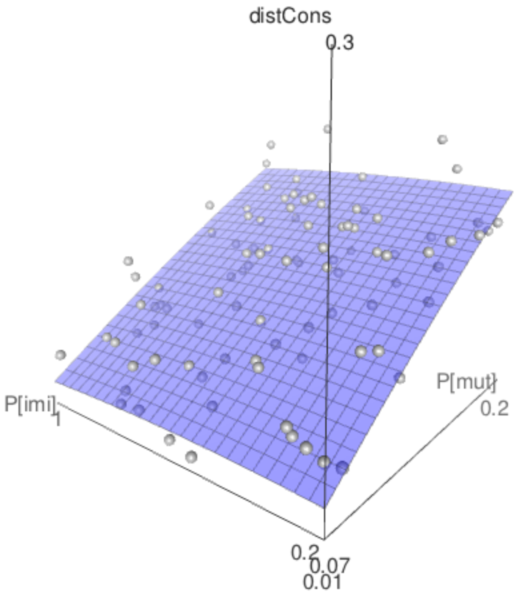


Figure 14: Robustness check in EA1: $P_{co} = 0.25$, $m = 10$, $n = 200$, $P_{mut} \in [0.01, 0.1]$, $P_{imit} \in [0.2, 1[$, 100 simulations, average distance to optimal consumption after $t = 200$ periods.