CICLO DE SEMINARIOS 1995

DEPARTAMENTO DE ECONOMIA

Experimentation in an Unknown Environment

Marcelo Clerici-Arias

QUAERERE VERUM

Universidad de San Andres

Universidad de
San Andres
DEPARTAMENTO DE ECONOMIA

Experimentation in an Unknown Environment

Marcelo Clerici-Arias

CICLO DE SEMINARIOS 1995
Día: Martes 21 de noviembre
9:00 hs.

Experimentation in an Unknown Environment

Marcelo Clerici-Arias

Department of Economics
Stanford University
Stanford, CA 94305
(415) 497-4221
email: marcelo@leland.stanford.edu

First version: October 1995

This version: November 6, 1995

Abstract: This paper presents a model where boundedly rational agents explore an unknown environment using only random encounters with other individuals and the observation of the most successful strategies in the past period. Agents do not calculate best responses, but rather imitate, with a certain probability, the characteristics of the most successful players. Also, once in a while they experiment with some new or relatively unsuccessful strategy. In the long run we find a unique stationary distribution of the states played by the agents. The results are applied to several games and show that perfect rationality or complete knowledge of the payoff matrix are not necessary to reach "rational" equilibria.

Journal of Economic Literature Classification: C72, C73, D83, C62, C63, C61

Keywords: Long-run equilibria, experimentation, mutation, adaptation, imitation, evolutionary game theory.

1. Introduction

For many years game theory has mostly concentrated on the study of agents with an amazing capacity for calculating their optimal strategies. In the last few years we have witnessed an increase in the number of models with boundedly rational agents. We will focus our attention on evolutionary game theory, which is characterized by (1) players with limited rationality, (2) an explicit dynamic process that describes how players adjust their strategies through time, and (3) the addition of new strategies through mutation.

The first characterization of evolutionary equilibria is due to Maynard Smith and Price (1973) who introduced the concept of evolutionary stable strategy. Maynard Smith and Price described a societal game where individuals are randomly matched to play a symmetric two-person game, and they can only play a certain pure or mixed strategy. If we now allow a relatively small population of mutants to invade the original population, then if the incumbent strategy survives and satisfies certain mathematical conditions, it is an evolutionary stable strategy.

Foster and Young (1990) replaced the one-time mutation of the evolutionary stable strategy with a repeated mutation process which results in a stricter equilibrium concept. Kandori, Mailath and Rob (1993), Kandori and Rob (1995) and Young (1993) among others applied this framework of repeated mutations resulting in the concepts of long-run equilibrium and stochastically stable equilibrium, respectively. This paper follows an approach similar to Kandori and Rob (1995).

The main contributions of this paper are:

(1) the replacement of the mutation rate by experimentation rates, since the concept of mutation and its biological interpretation are understandably rejected by many economists as not suitable for economic agents.

(2) the experimentation rate is endogenous. Agents have different experimentation rates, a relaxation of the assumption of independence of the mutation rate across players adopted in the literature. In the long run, the diverse experimentation rates converge toward a stationary distribution.

(3) agents need significantly less information. Particularly, they discover better strategies through imitation and experimentation without knowing the payoff matrix.

(4) agents need less computational capabilities than in previous models. For example, during the adaptation stage each player simply compares its payoff with the highest average payoff in the population, instead of calculating the best response to the frequency of strategies in the population.

(5) the discrete and finite state space used in the literature can be generalized to countable discrete state spaces or even continuous state spaces.

(6) the imitation/adaptation rate is endogenous. Agents have different imitation rates and, in the long run, evolution results in them converging toward a stationary distribution.

(7) a future version of this paper will (i) explicitly extend the discrete state space to general state spaces, (ii) allow for a changing (and unknown) environment, (iii) study the effect of local interactions, and (iv) include more applications of the general model to specific games. I am also working in using this model to explain some phenomena that appear in insurance markets.

2. Model

First let me summarize the chain of events in the model so we have a map to guide us when we analyze the complete, detailed version.

(1) **initialization**: the M agents in the population are randomly assigned one strategy of a two-person game and an experimentation rate.

(2) **confrontation**: all individuals are randomly matched, and each pair plays the two-person game with their pre-assigned strategies.

(3) **adaptation**: agents imitate with probability η_i the strategies played by the most successful individuals (on average) during the confrontation stage.

(4) **experimentation**: agents experiment playing random strategies with probability e_i, their individual experimentation rate. Individuals can also experiment using new imitation and experimentation rates. After experimentation, the players go back to the confrontation stage.

The first part of the model was built on Kandori, Mailath and Rob's (KMR from now on) model, but there are several crucial differences that I will point out along the way. For simplicity I start considering a symmetric two-person game with n strategies, but we can extend the model to an asymmetric N-player game with a continuum of strategies (or more general spaces) without affecting the main results. The fact that we can consider games with a continuum of strategies is the first departure from KMR's model, since they only consider discrete strategies. This will

3. Equilibrium

If the Markov chain were characterized only by the strategies and the adaptation part of the dynamic process, then the initial configuration would determine where the system would settle or cycle. However, experimentation introduces a positive probability of reaching any final state from any initial state. The system is irreducible and aperiodic since we have limited the imitation and experimentation rates to be strictly positive, and the theory of Markov chains allows us to conclude that there is a unique, globally stable and ergodic stationary distribution (see definition 1 and proposition 1).

Definition 1: $\mu \in \Delta^{|Z|}$ is a stationary distribution if $\mu P = \mu$, where $\Delta^{|Z|}$ is the $|Z|$ - 1 dimensional simplex and $|Z| = nl^2$ is the number of elements in the state space.

Proposition 1: [2]

(1) **uniqueness**: there exists a unique stationary distribution μ.

(2) **global stability**: for any initial distribution q, the distribution converges to μ.

(3) **ergodicity**: μ represents the long-run proportion of time spent in each state.

---

[2]Kandori and Rob (1995) use the assumption of finiteness of the state space and cite Hoel, Port and Stone (1972) to obtain these results. However, we can obtain the same results using more general state spaces. For example, discrete spaces do not need to be finite (see Romanovsky, 1970). If we use continuous strategies and/or experimentation rates (and even if we use a combination of discrete and continuous strategies), we can model the Markov chain as a positive Harris recurrent chain and get uniqueness, global stability and ergodicity (see Meyn and Tweedie, 1993).

KMR use a small mutation rate $\varepsilon \rightarrow 0$ to obtain a unique limit distribution. Note that the lower the mutation rate, the longer the time the system will take to reach the long-run equilibria. That is why KMR argue that their analysis is most relevant for a small population, though Ellison (1993) has shown that local interaction, in which each agent is allowed to interact only with its neighbors, reduces the amount of time necessary to achieve equilibrium.

In our model the experimentation rates can be significantly larger than KMR's mutation rate and the imitation process depends only on the highest average payoff achieved by the players during the confrontation stage—and not on the frequency of individuals in the population playing each strategy—, so these two factors accelerate the convergence of the strategies to the long-run Nash equilibrium. Thus, even a large population can reach the strategy portion of the long-run equilibrium reasonably fast, though the imitation and experimentation rates will take a longer time before they converge to their final endogenous value.

In the following section I present results for some specific games.

4. Applications

I have analyzed a number of two-person symmetric games with different characteristics to test the results of the societal game described in this paper.

4.1 Pure coordination games

Kandori, Mailath and Rob (1993) apply their model to pure coordination games—Kandori and Rob (1995) also study supermodular games—so it seems natural to start by analyzing numerical results for the following two-strategy pure coordination game.[3]

	S1	S2
S1	1	0
S2	0	2

Table 1

Two-strategy pure coordination games have two pure-strategy Nash equilibria. Both KMR's agents and our imitating-experimenting individuals under the complete-imitation rule mostly play the strategy that results in the Pareto optimal equilibrium (S2 in table 1). Figure 1 shows the stationary distribution of strategies under the complete-imitation rule, while figure 2 shows the

[3]The matrix shows the payoffs to player 1 (the row player). Since the game is symmetric we do not need to show the payoffs to player 2.

same distribution under the strategy-imitation rule. As you can see, the strategy-imitation rule does not result in the selection of a pure-strategy Nash equilibrium due to the lack of a link between payoff and imitation-experimentation, as explained in section 2.

The same observation applies to the experimentation and imitation rates. Under the complete-imitation rule the experimentation rate of those agents playing the highest-average-payoff strategy converges to its lowest strictly positive limit (figure 3) while higher imitation rates are rewarded (figure 4) due to the nature of the payoffs in a pure coordination game, but different results are observed under the strategy-imitation rule (figures 5 and 6). In this case, the stationary distributions of the imitation and experimentation rates simply reflect what happens when you randomly alter the imitation and experimentation rates over a long period of time and, in fact, under the strategy-imitation rule the same stationary distributions are obtained for all the games studied in this section. These observations show that the strategy-imitation rule, though simpler than the complete-imitation rule, yields stationary distributions not only far from theoretical predictions for perfectly rational players, but also far from experimental results. Thus, in the rest of this section we will restrict our attention to individuals who copy the whole organizational structure—the triple (strategy, imitation rate, experimentation rate)—of the agents playing the strategies with the highest average payoffs.

Figure 7 shows an typical example of how fast the population converges to playing mostly strategy 2 under the complete-imitation rule. It is interesting to notice that we do not need hundreds or thousands of periods to get close to the stationary distribution of the strategies, even though the population size in this particular simulation is quite large (M = 500 individuals).[4]

4.2 Coordination games

Two-strategy non-pure coordination games (for example, see Table 2) have two pure-strategy Nash equilibria, and the agents choose the risk dominant one—in this case, S2-S2—under the complete-imitation rule (figure 8). Again the population quickly converges to the equilibrium distribution (figure 9).

	S1	S2
S1	50	49
S2	0	60

Table 2

Even though this is a coordination game, high imitation rates are punished—not rewarded—after the risk dominant equilibrium is found. Once in a while experimentation and random matching will result in two agents playing S2 against each other for a payoff of 60, higher than the risk dominant payoff. Individuals with high imitation rates will quickly switch to S2, but they will soon find out that whenever they meet someone playing S1 they get a payoff of zero and they have to imitate one of the agents that continued playing S1. Since most of the agents who did not switch to S2 had low imitation rates, everyone will eventually end up with a low imitation rate

[4]The software used to calculate the stationary distributions and all runs are available upon request.

(figure 10). The same rationale applies to the experimentation rate: if you have a high experimentation rate and switch to S2, very soon you will have to copy the organizational structure of those who remained playing S1 and you will also get a low experimentation rate (figure 11).

4.3 Prisoner's dilemma

As in the prisoner's dilemma with perfectly rational agents and known payoff matrix, most of the individuals choose to defect (strategy 2) given the present societal game (figure 12). This behavior is expected to change when we add the possibility of local interaction to the model—some cooperation should follow.

	S1	S2
S1	50	20
S2	60	27

Table 3

4.4 Hawk-dove games

The population of agents reaches a proportion close to that predicted by the evolutionary stable strategy with known payoff matrix and replicator dynamics. For the game shown in table 4, 83% of the agents should be playing strategy 2 (dove). Compare that to 80.3% playing dove in our model (figure 13) without knowledge of the payoff matrix and without using the replicator dynamics, a tool inherited from evolutionary biology.

	S1	S2
S1	-5	2
S2	0	1

Table 4

5. Conclusions

This paper eliminates the system-wide mutation prevalent in the literature by reinterpreting mutation as experimentation—an explanation much more palatable for economic agents who are learning while they interact with other players in an unknown environment than the explanation used in biology. This model also allows agents to experiment and imitate with different probabilities, and lets the experimentation and imitation rates evolve to endogenous values instead of imposing them exogenously.

Note that under the complete-imitation rule the experimentation rate converges to its lowest strictly positive value allowed in all games analyzed. This fact suggests that KMR's assumption of a small system-wide mutation rate going to zero is reasonable for two-person symmetric games, but I believe we will find a quite different stationary distribution of experimentation rates in asymmetric games with only mixed strategy equilibria, such as the matching pennies game.

This paper also provides some insight on what is necessary and what is not enough to obtain equilibria similar to those obtained with perfectly rational agents who know the payoff matrix. Imitating the observed strategies of the most successful players is not enough; we need to copy their whole organizational structure in order to get reasonable results.

Future improvements in this paper will include:

(1) the explicit extension of the theoretical framework to more general state spaces,

(2) the analysis of the effect of a changing environment on the equilibria and the out-of-equilibrium path for both strategies and experimentation rates,

(3) local interaction;

(4) costs of imitation and experimentation, and

(5) more applications, including asymmetric games. I am currently working on models of the insurance industry (earthquake, life, and auto insurance).

6. References

Ellison, G. (1993): "Learning, Local Interaction, and Coordination," *Econometrica*, 61, 1047-1071.

Freidlin, M. I., and A. D. Wentzell (1984): *Random Perturbations of Dynamical Systems*. New York: Springer-Verlag.

Hoel, P. G., S. Port, and C. J. Stone (1972): *Introduction to Stochastic Processes*. Boston: Houghton Mifflin Company.

Kandori, M., G. J. Mailath, and R. Rob (1993): "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica*, 61, 29-56.

Kandori, M., and R. Rob (1995): "Evolution of Equilibria in the Long Run: A General Theory and Applications," *Journal of Economic Theory*, 65, 383-414.

Maynard Smith, J., and G. R. Price (1973): "The Logic of Animal Conflict," *Nature*, 246, 15-18.

Meyn, S.P., and R. L. Tweedie (1993): *Markov Chains and Stochastic Stability*. New York : Springer-Verlag.

Romanovsky, V. (1970): *Discrete Markov Chains*. Groningen: Wolters-Noordhoff.

Young, H. P. (1993): "The Evolution of Conventions," *Econometrica*, 61, 57-84.

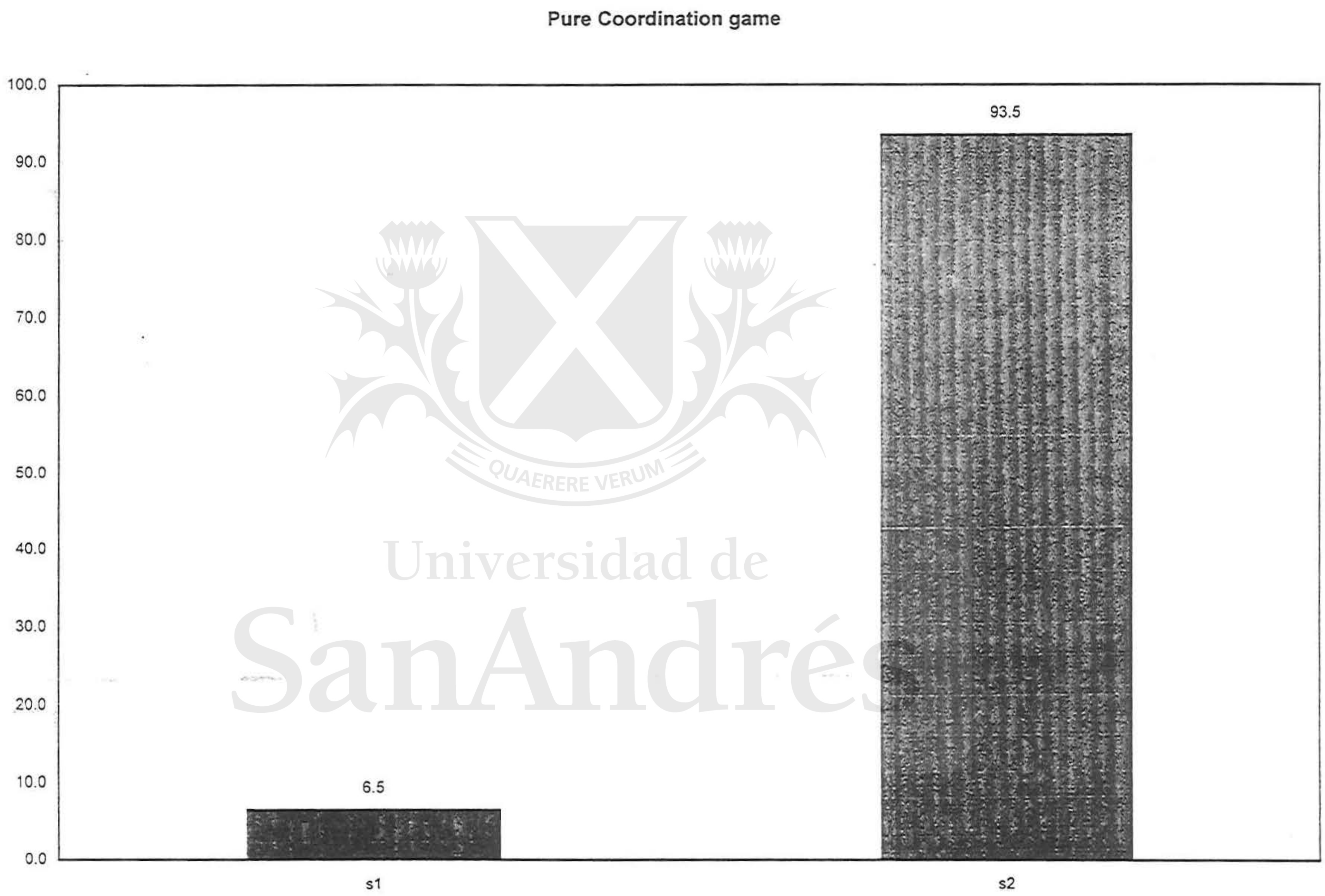

Pure Coordination game
100.0
90.0
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
93.5
6.5
s1
s2

Figure 1

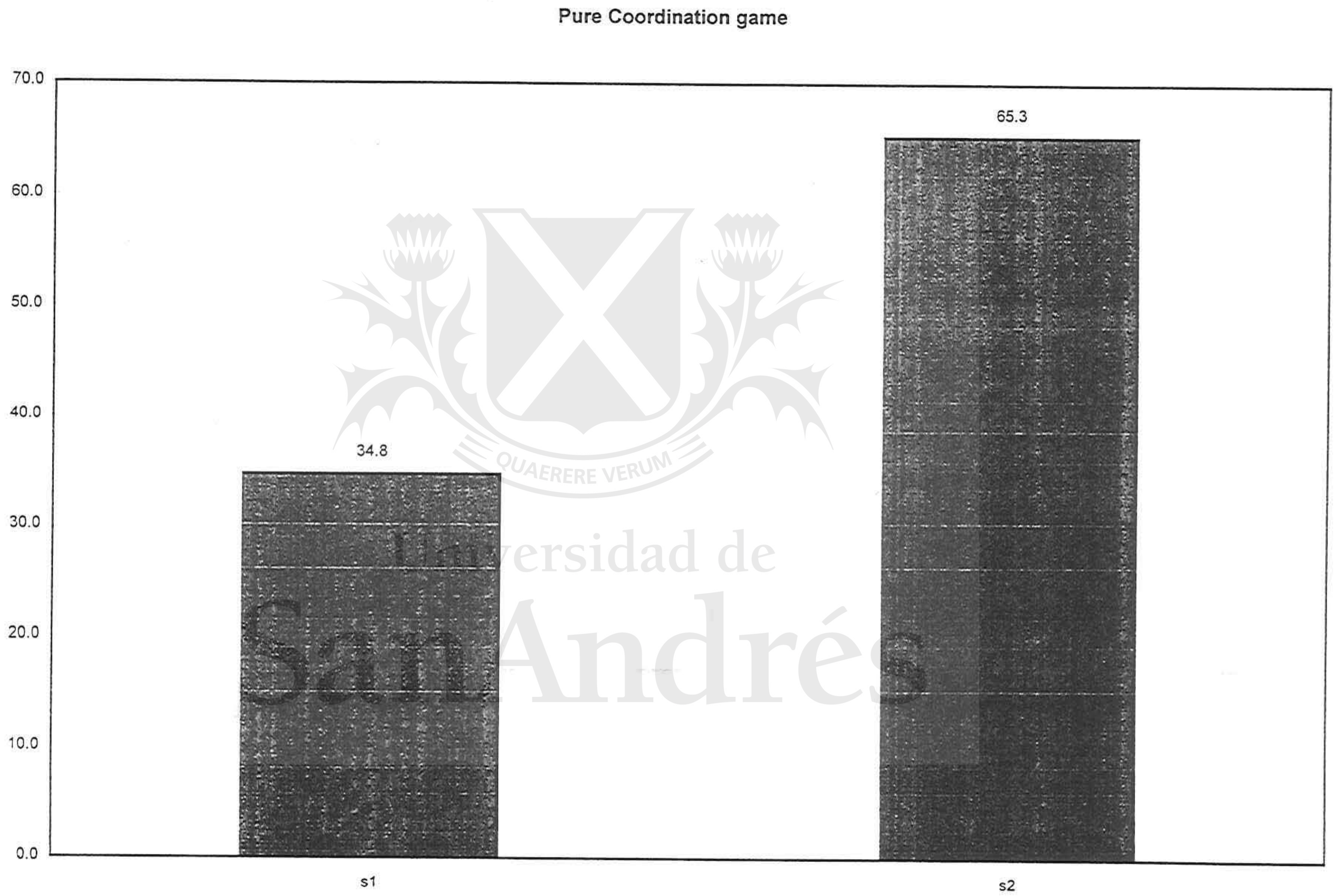

Pure Coordination game
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
34.8
65.3
s1
s2

Figure 2

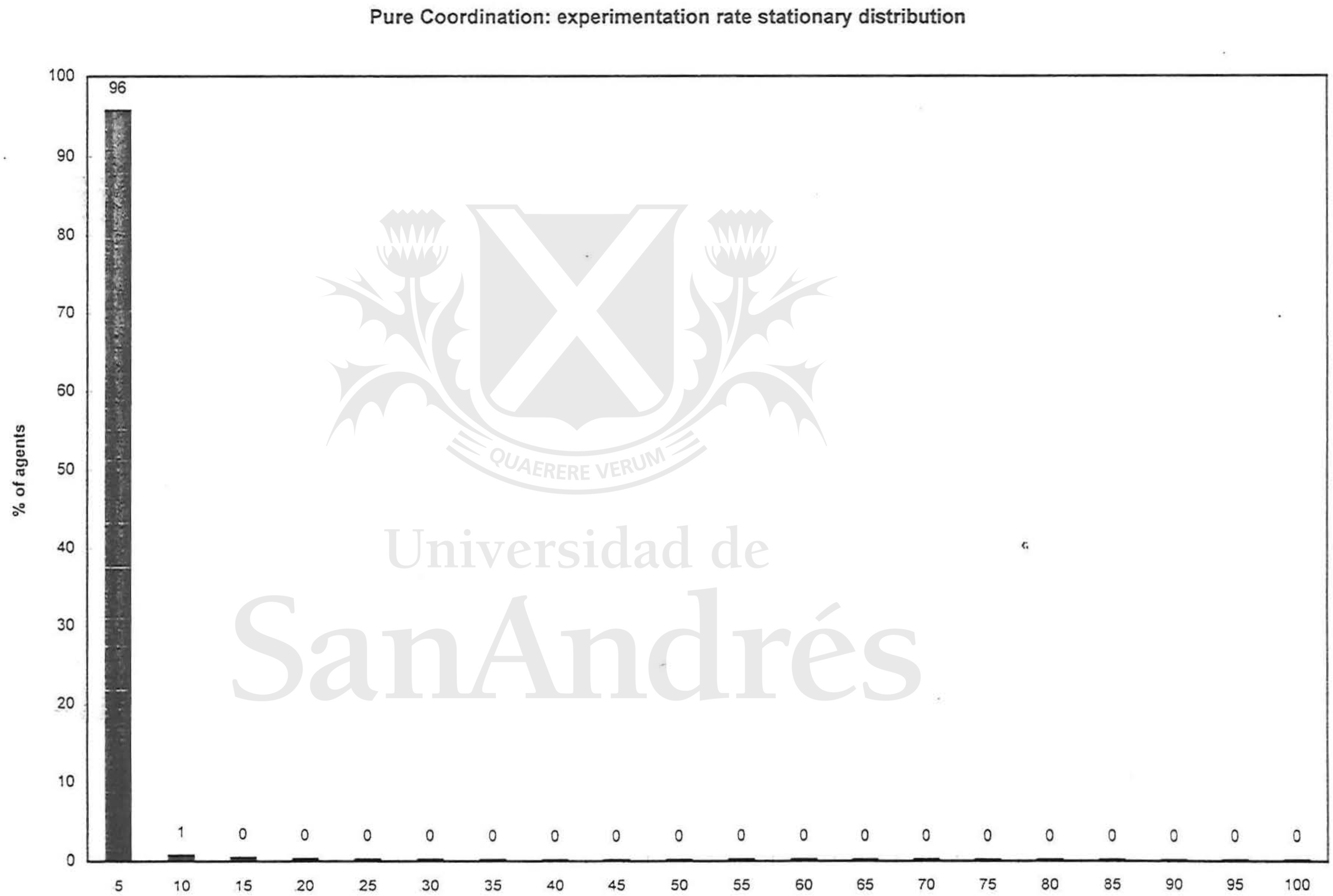

Pure Coordination: experimentation rate stationary distribution
% of agents
100
90
80
70
60
50
40
30
20
10
0
96
1
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100

Figure 3

Pure Coordination: imitation rate stationary distribution

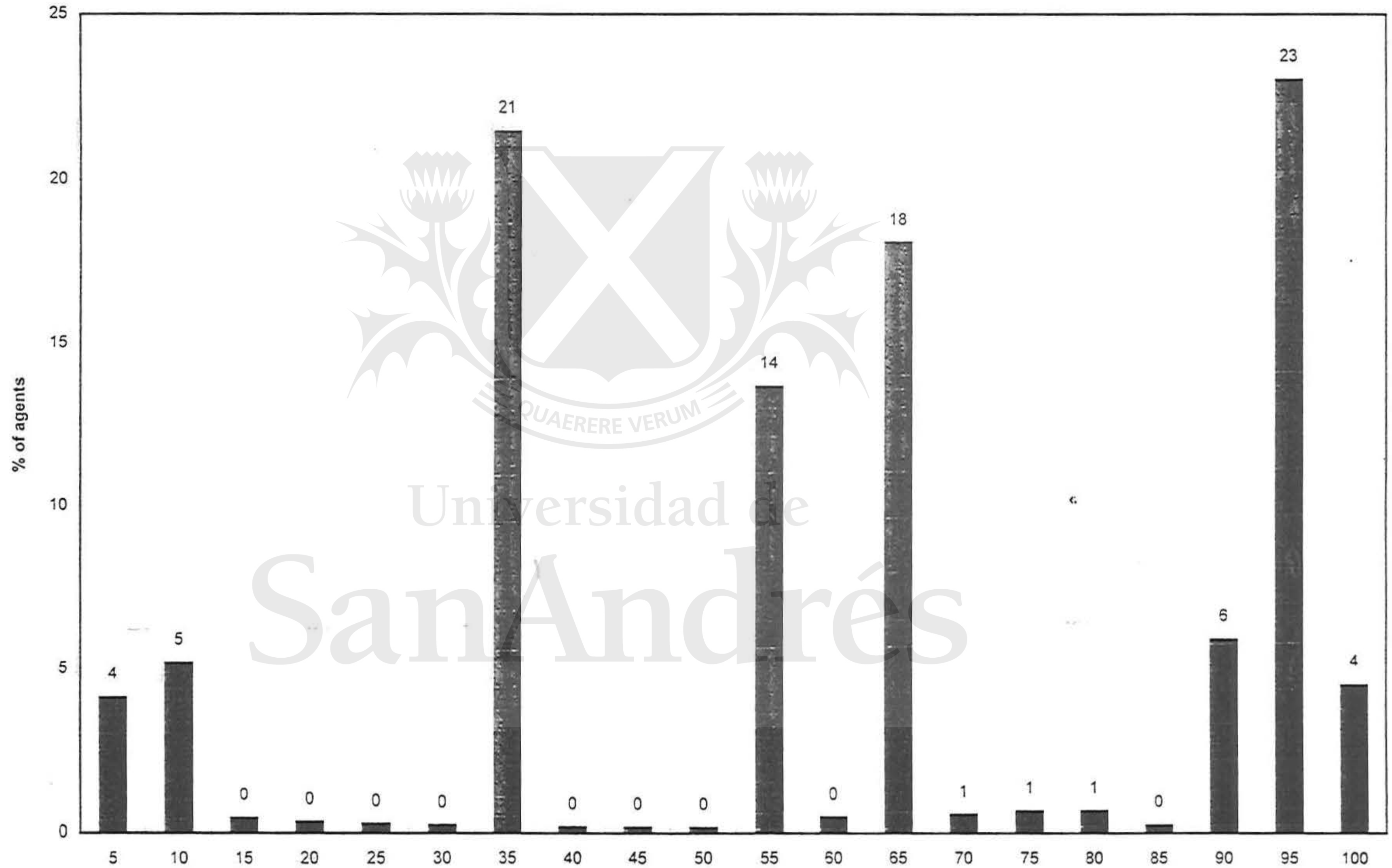


Figure 4

Pure Coordination: experimentation rate stationary distribution

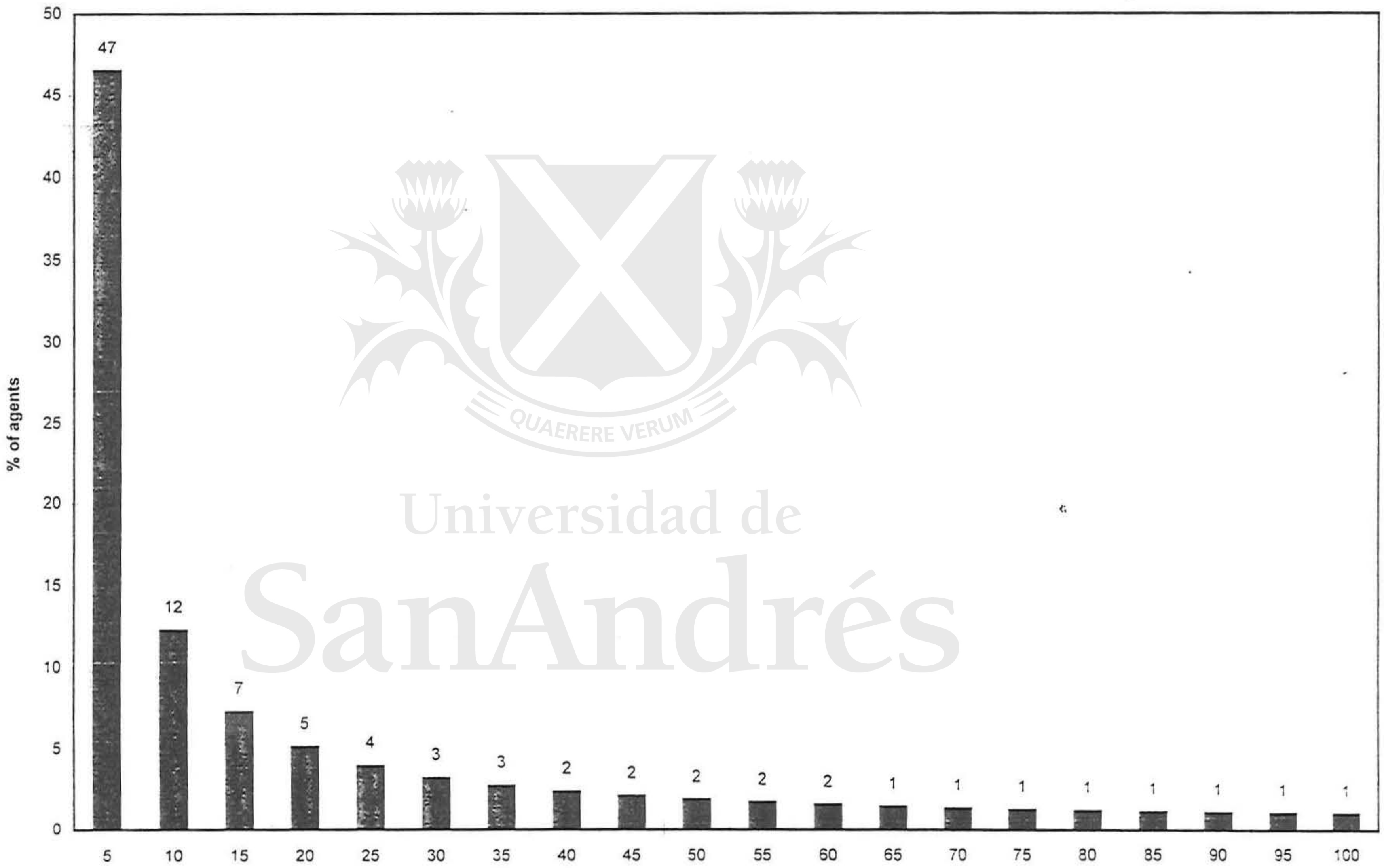
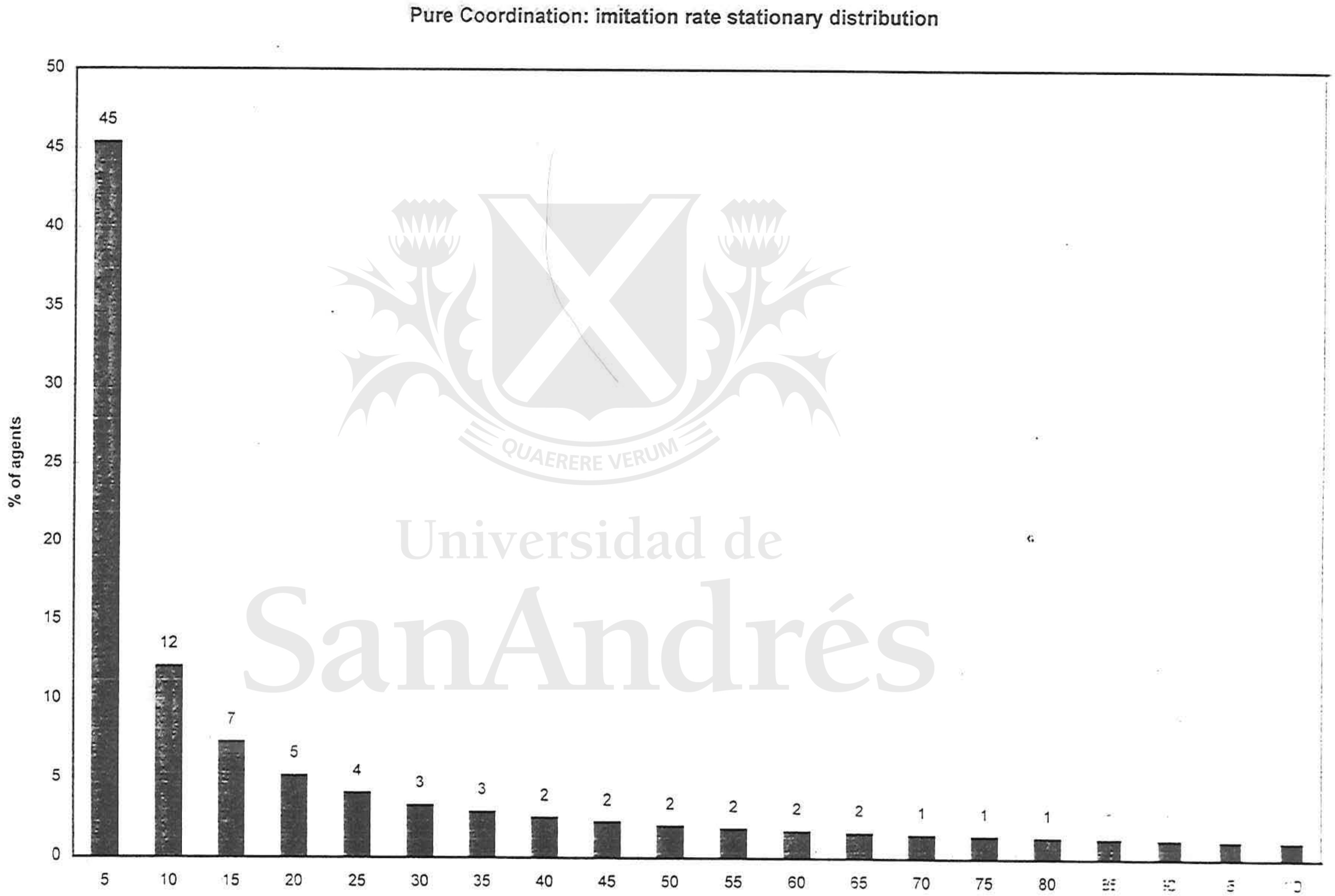


Figure 5



Pure Coordination: imitation rate stationary distribution
% of agents
50
45
40
35
30
25
20
15
10
5
0
45
12
7
5
4
3
3
2
2
2
2
2
2
1
1
1
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80

Figure 6

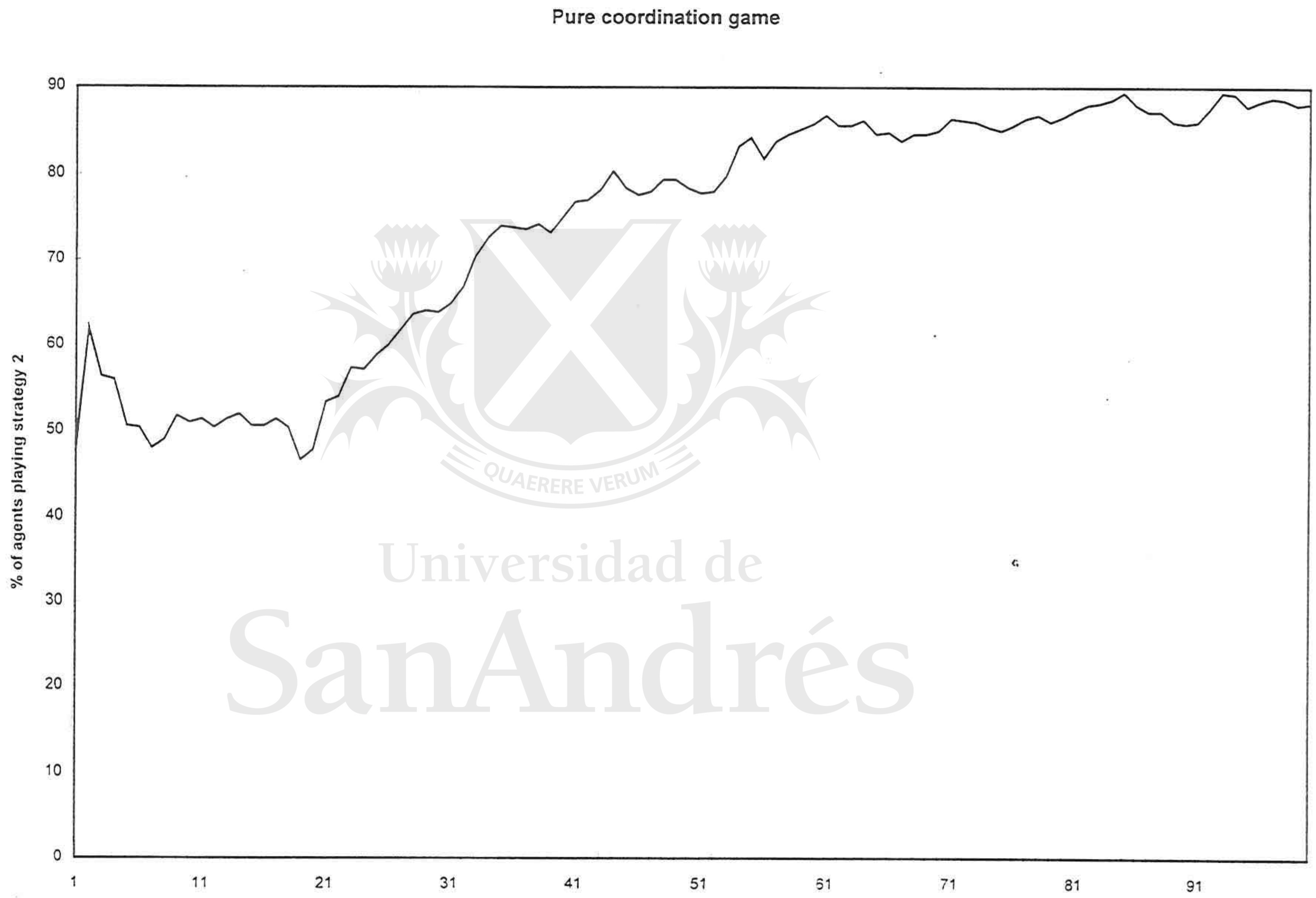

Pure coordination game
% of agents playing strategy 2
90
80
70
60
50
40
30
20
10
0
1
11
21
31
41
51
61
71
81
91

Figure 7

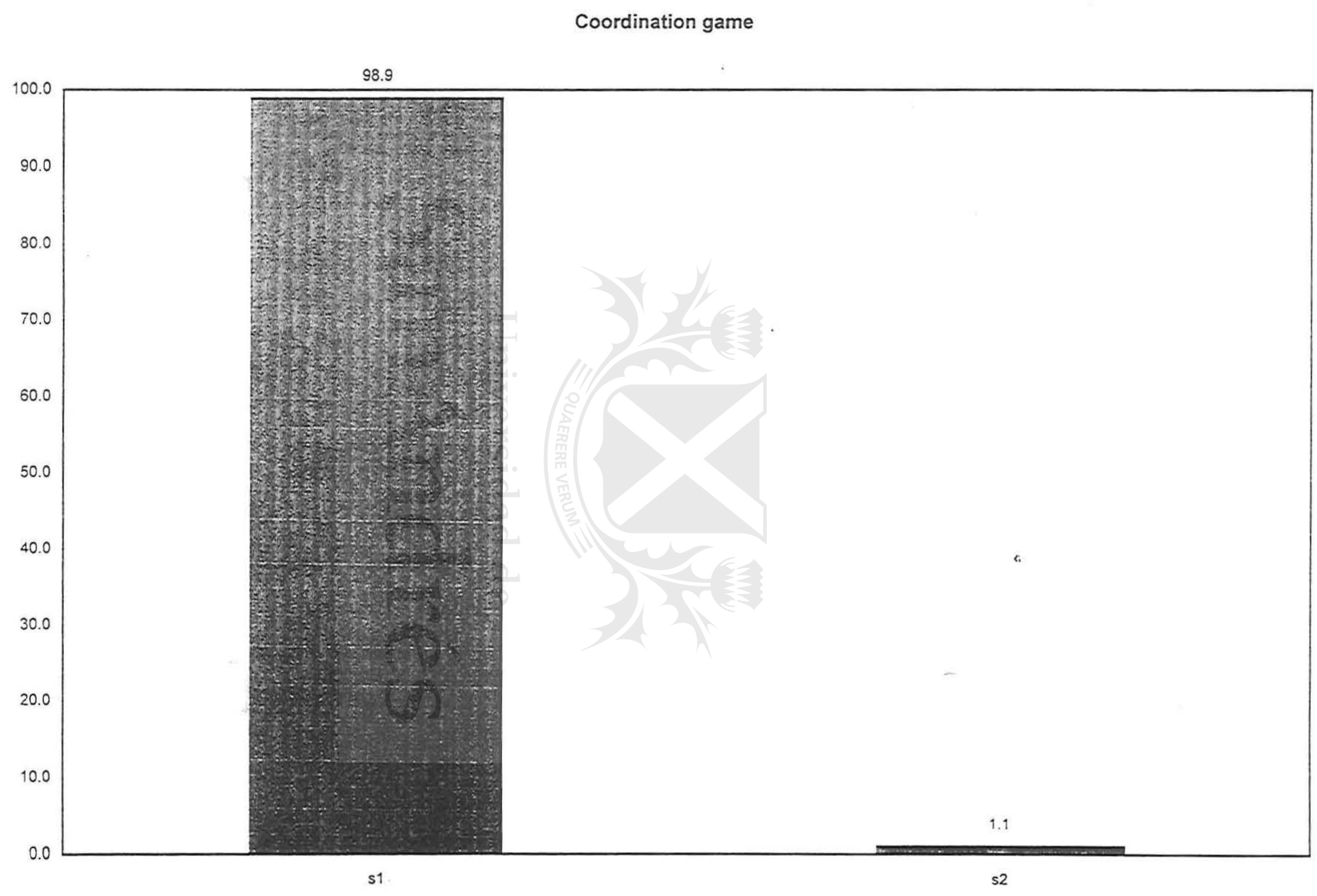

Coordination game
98.9
1.1
100.0
90.0
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
s1
s2

Figure 8

Coordination game

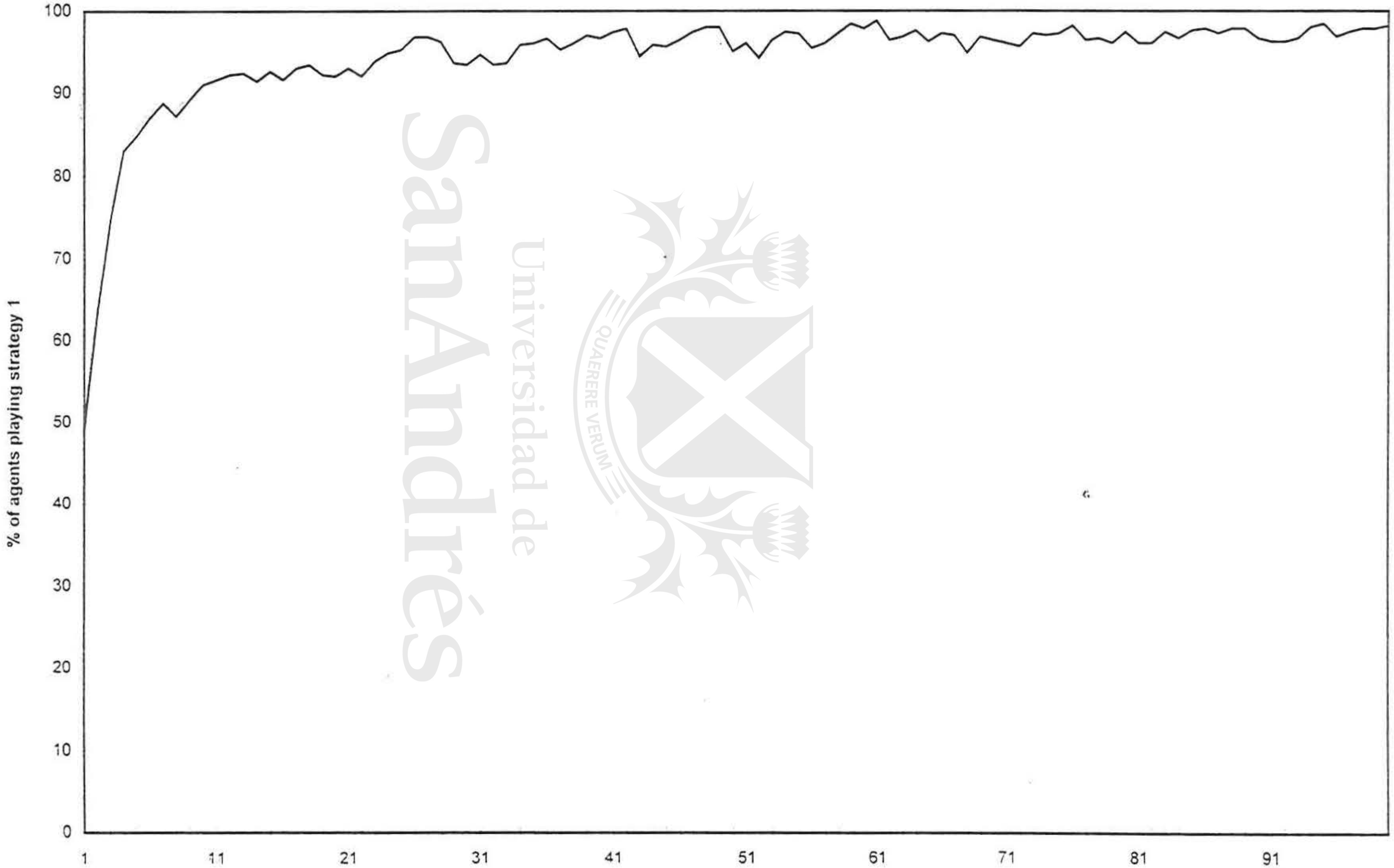


Figure 9

Coordination: imitation rate stationary distribution

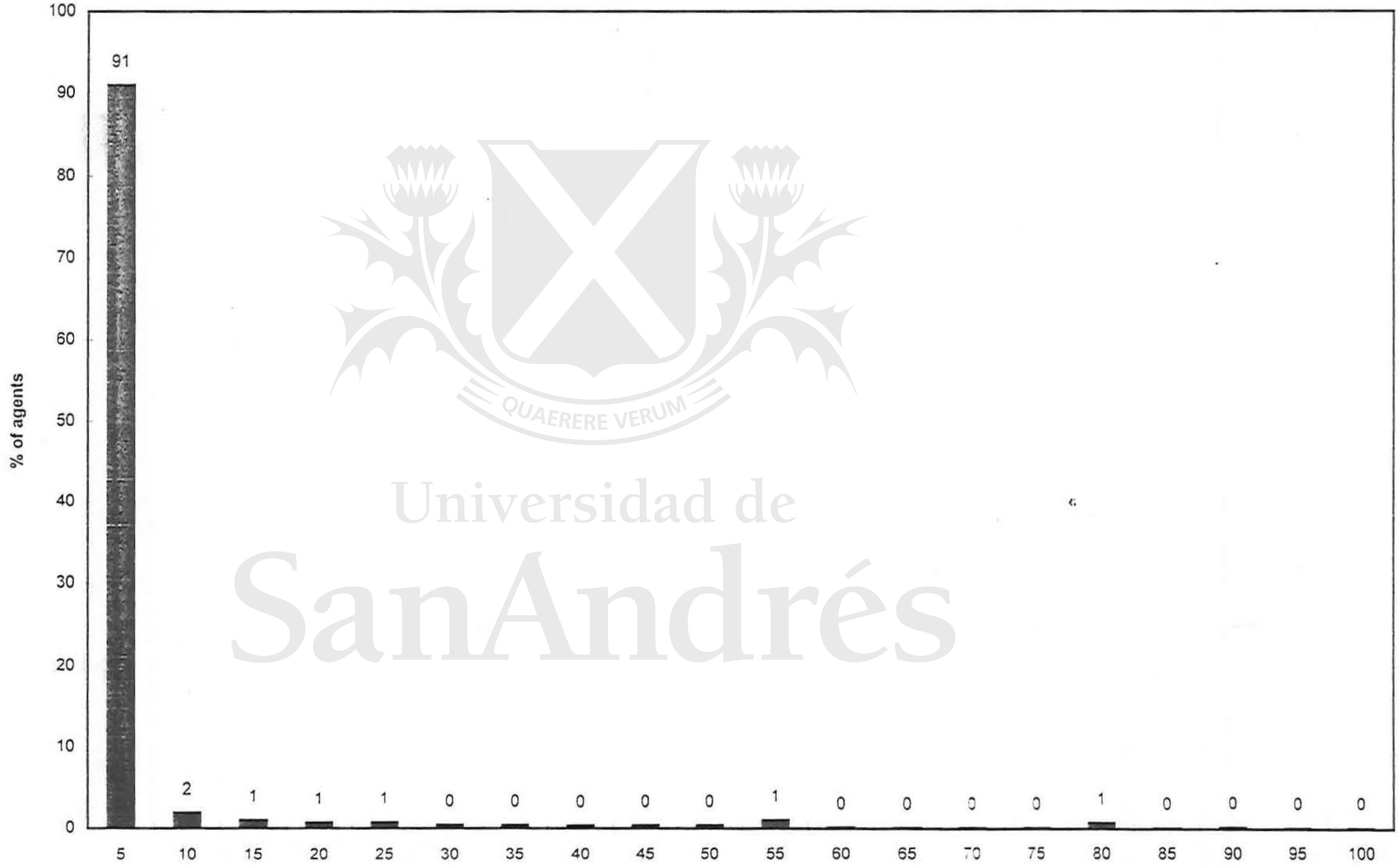


Figure 10

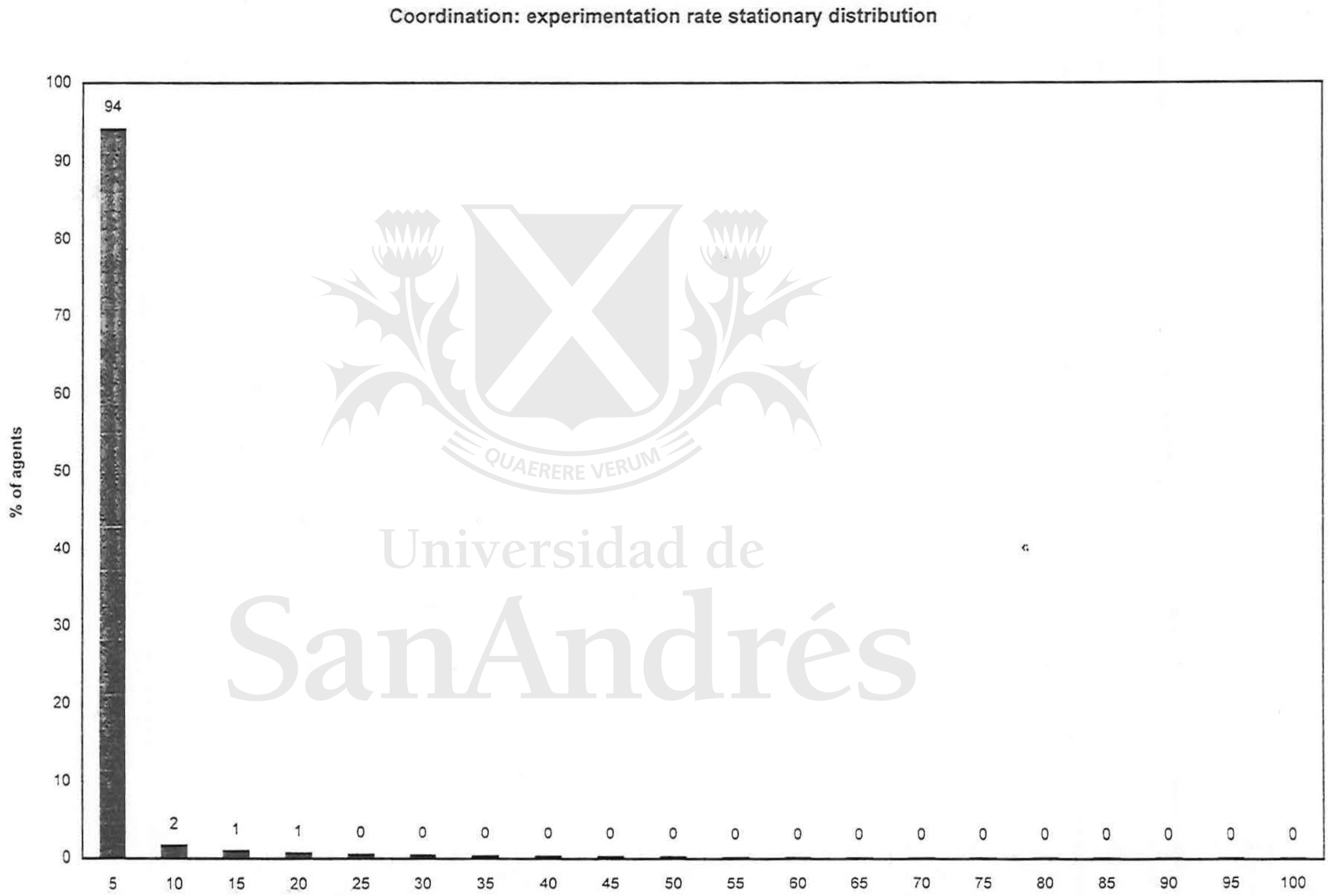

Coordination: experimentation rate stationary distribution
% of agents
100
90
80
70
60
50
40
30
20
10
0
94
2
1
1
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100

Figure 11

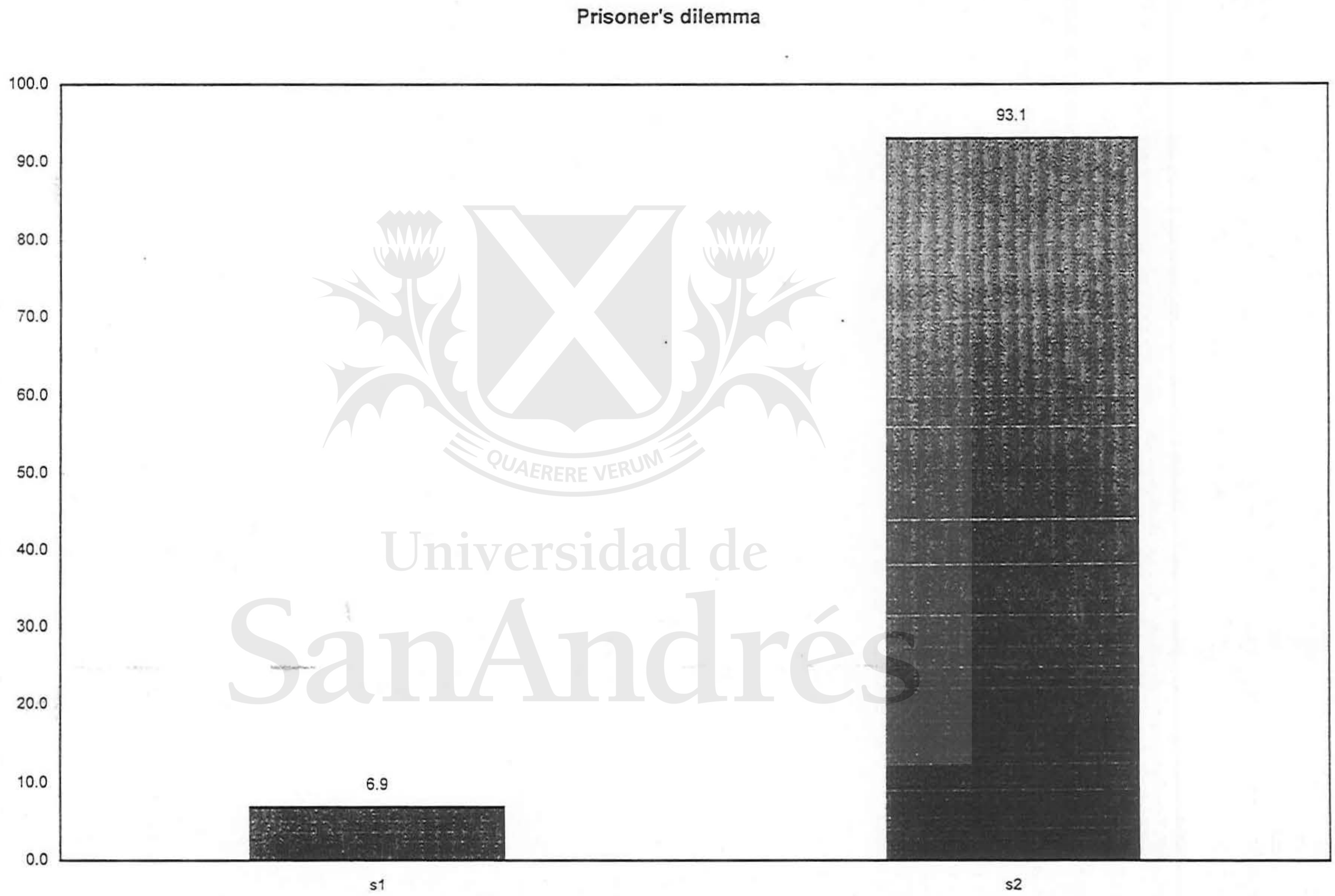

Prisoner's dilemma
100.0
90.0
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
93.1
6.9
s1
s2

Figure 12

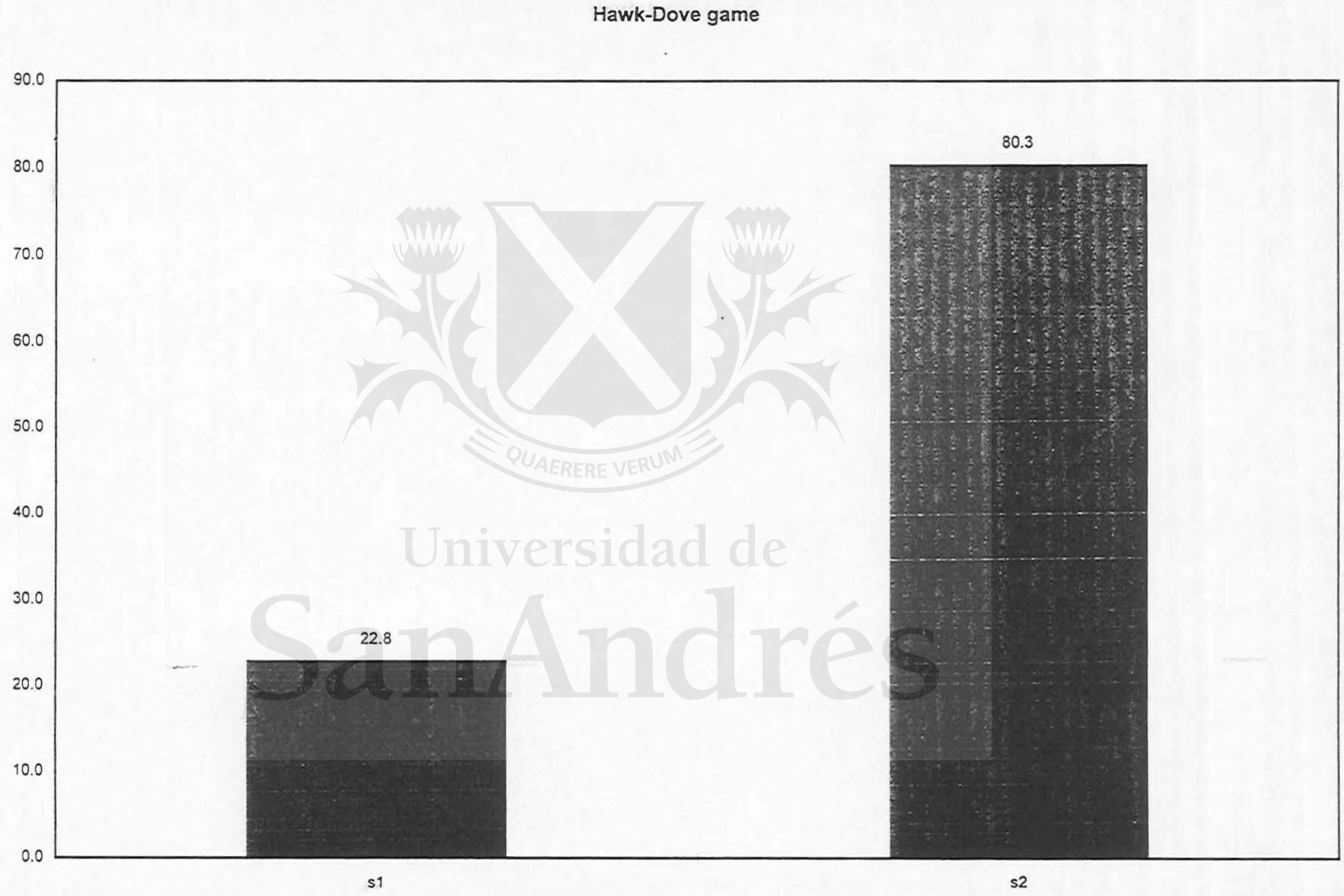

Hawk-Dove game
90.0
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
80.3
22.8
s1
s2

Figure 13