Laboratory of Economics and Management
Sant'Anna School of Advanced Studies
Piazza Martiri della Libertà, 33 - 56127 PISA (Italy)
Tel. +39-050-883-343 Fax +39-050-883-344
Email: lem@sssup.it Web Page: http://www.lem.sssup.it/

LEM

Working Paper Series

Repeated Choices under Dynamic Externalities

Giulio Bottazzi*
Angelo Secchi *

* Scuola Superiore Sant'Anna, Pisa, Italy

2007/08 December 2011

Repeated Choices under Dynamic Externalities

Giulio Bottazzi[a] and Angelo Secchi[b]

[a]Institute of Economics and LEM, Scuola Superiore Sant'Anna
[b]Paris School of Economics - Université of Paris 1 Panthéon Sorbonne and LEM, Scuola Superiore Sant'Anna

December 19, 2011

Abstract

We consider an economy in which a heterogeneous population of agents have to choose among a common set of alternatives. The utilities associated to the different alternatives posses a common component and an individual component, which reflect differences in the underlying structure of agents preferences. The common components are characterized by a fixed term which describe the intrinsic utility of each choice, and by a social component which depends on the actual distribution of agents across the different alternatives. In particular, we analyze the case of linear positive externalities. Assuming a simple Markovian process for the revision of the selection process, we derive the equilibrium distribution of the population of agents. We analyze in details the extremal cases of few choices and large population of agents. The proposed models can be applied to different domains of economics, like technological adoption, location of production activities, co-evolution of business models or financial decision rules. The resulting self-reinforcing dynamics can be considered an alternative formulation of the Polya urn scheme developed by Brian Arthur et al. (1986) when the possibility of choice revision is taken into account. We analyze the differences and similarity of the two approaches.

JEL codes: C1, L6, R1

Keywords: Industrial Location, Agglomeration, Dynamic Increasing Returns, Markov Chains, Polya Urns.

1 Introduction

Over the last decades economists have increasingly recognized that individuals, even in choosing among fixed alternatives, very often experience uncertainties and inconsistencies. Hence, the ideal situation in which individuals have perfect discriminatory power, unlimited information and are able to rank all the alternatives in a well-defined and consistent way is not an adequate description of human behavior (cfr. Anderson et al. (1992) and the references therein).

These natural constraints to a full and complete exertion of the rationality of agents suggest to interpret the outcome of their choice procedure as a random variable. Indeed, the idea that the choice behavior of agents is better described in term of probabilistic processes has a very old tradition in the psychological literature. At the beginning of the past century, Thurstone, in a seminal work (Thurstone, 1927), suggests to describe the perceived values associated to different alternatives as "discriminal processes", that are stochastic variables agents compare in order to produce their choice.

In economics, the probabilistic nature of individual choice behaviors has been acknowledged through two diverse classes of models, resting on two different interpretations of the underling cognitive mechanism. The first tradition considers models in which the agents decision rule is stochastic while their utility function is deterministic (Luce, 1959; Tversky, 1972). The second family of models takes the opposite approach: here the decision rule is deterministic while the utility associated with a given alternative is stochastic (McFadden, 1984).

The main objective of the present paper is to formulate a discrete choice model with social interactions in which agents repeatedly choose among several alternatives whose perceived utilities are influenced by the choices of other agents. We model a simple economy in which a population of agents has to choose among a finite set of alternatives. In the spirit of Thurstone (1927) and McFadden (1984) the perceived utility of each alternative is stochastic and it is composed by two terms: a term which captures the common, to all agents, benefits associated to the observable characteristics of the various alternatives and an idiosyncratic term which accounts for all the unobservable and agent-specific characteristics of the different alternatives. The effect of other agents' decision on that made by each single agent is modeled assuming that the common component in the utility function contains a social term according to which the attractiveness of a given alternative increases linearly with the number of times it has been chosen in the past. Moreover, using a simple random selection mechanism, we allow the possibility for agents to revise their previous choices.

The introduction of social terms in individual utilities has recently proven fruitful in a variety of contexts in economics. Different types of nonmarket interactions are incorporated in models illustrating the functioning of labor markets (cfr. (Montgomery, 1991; Topa, 2001)), in models describing the diffusion of innovations (Brian Arthur, 1989), in endogenous growth models with human capital accumulation (Benabou, 1996; Durlauf, 1996) and, also, in the vast literature on the localization choices of firms (among many others see (Fujita et al., 1999; Bottazzi et al., 2007)).

Our approach extends the existing literature in two directions: we provide an explicit discrete choice model that is valid for an arbitrary large number of possible alternatives and, together, we introduce a random procedure of choice revision. Moreover, in departing from the original framework proposed in Brian Arthur et al. (1986), we do not take any large economy limit and we solve the model for a finite population of active agents. In this way, the aggregate state of the economy is uniquely and completely specified, at each point in time, by a vector containing the number of agents who have chosen any of the available alternatives.

Our model proves to have a number of interesting analytical properties. First, it generates a stationary distribution of agents across alternatives that can be compared to empirical distributions in order to estimate the parameters of the model. Second, it provides an explicit expression for the transition probabilities between different states of the economy at equilibrium which can be used to assess the degree of short term stability of the observed distribution. Third, the ergodic nature of the model allows to run comparative static exercises to investigate the effects, on the equilibrium probability of the different possible states of the economy, of changes both in the number of active agents and in the long term structural parameters of the model.

One may interpret our model as a multi-choice dynamic extension of the framework developed in Brock and Durlauf (2001). In their analytic approach they describe a binary choice problem that is "genuinely" static, since it amounts to find the equilibrium distribution of agents across two alternatives given a set of interdependent utility functions. No explicit reference is made to any choice procedure. In our case, instead, we model a dynamic choice procedure in which a population of agents may choose among an arbitrary large number of alternatives.

In a slightly different perspective, the model discussed in this paper may be considered an ergodic reformulation of the non-Markovian urn processes proposed in the literature originated by Brian Arthur et al. (1986). In particular, since then, the formal tool of *Generalized Urn Schemes* has been applied to a variety of situations characterized by the interactions of individual behaviors of agents who have incomplete information about their environment and its mechanisms of evolution (among others cfr. Brian Arthur (1994); Dosi et al. (1994); Dosi and Kaniovski (1994)). Formally the generalized Urn Schemes represent non stationary Markov Chains with a growing number of states enabling one to handle positive and/or negative feedbacks possibly coexisting in the same process. In this case an explicit choice structure is in general presented, but the derived results only refer to the asymptotic distribution generated by an infinite stream of choices with progressively decreasing marginal relevance. Conversely, in our model we consider a Markov process with a finite number of alternatives and with reversability of choices. In this framework we are able to derive the ergodic equilibrium distribution generated by the cumulative effect of repeated choices of the agents. This provide a simple approach to the estimation of the magnitude of the externality effects induced by social interactions.

The remainder of the paper is organized as follows. Section 2 sets the stage presenting the basic assumptions underlying the model and shows that the outcome of the decision process does not depend, in probabilistic terms, from the idiosyncratic component of agents preferences. In Section 3 we describe how we specify the common term in the utility function to introduce social interactions effects and we derive the main analytical properties of the model. In order to allow comparisons with the existing literature Section 3 studies the effects of switching off the choice revision process in the model. Section 4 concludes and suggests lines for further investigations.

2 The model

In this section we present a model of individual choice which incorporates social effects. We study an economy in which a population of heterogeneous and boundedly rational agents has to choose a single alternative among a set of predetermined possibilities. At each time step new agents enter the economy while incumbents may leave it. Each agent, when entering the economy, chooses the alternative which is expected to provide him the highest utility. The description of the dynamics of such an economy requires the preliminary specification of two fundamental aspects of the decision process: the mechanism selecting at each time step the agent called to choose and the utility function on which his choice is based.

Regarding the first aspect, there are, at least in principle, many ways to design the procedure to single out the agent allowed to make his choice and each one may describe a different economic situation. One can imagine that the different agents are called to choose according to a pre-determined and fixed list, organized for example in alphabetical order, or based on age, weight or other peculiar characteristic identifying each agent. Diversely one may relate the probability of picking a given agent to the outcome, in terms of utility, of the previous choices. Again one has many possibilities ranging from situations in which who got more utility from the past choices has also higher probabilities to be called for a new choice to other situations in which the probability of picking a given agent depends (positively or negatively) on the number of times the same agent has been selected in the past. Finally one may assume more complicated settings in which a topology is defined over the set of available agents and the probability of choosing a particular individual is defined as a function of his distance from the agent selected in the last time step. In what follows, in order to isolate the properties of our model due to individual choice behaviors from the ones induced by the selection mechanism, we decide to describe the latter in the simplest way: we assume it as a pure stochastic selection mechanism.

Formally we consider a framework in which N different agents choose among a set of L distinct alternatives, labeled by integers between 1 and L and we assume

Assumption 1. *At each time step one agent is randomly selected to exit the economy. All incumbent agents have the same probability to be selected.*

Regarding the second aspect, as mentioned in the Introduction, we choose to follow the approach inspired by the work of Luis L. Thurstone. Agents are considered heterogeneous with respect to their preferences due to problems of asymmetric information or cognitive biases. Since we are interested in the aggregate dynamics of the economy, this heterogeneity is, at this stage, modeled as a probabilistic effect. We assume that the preference structure of different agents over the available alternatives builds on two terms: a common factor and an idiosyncratic component. The common factor affects the decision of any possible agent and is meant to represent the common "perceived" advantage of picking a certain alternative. The idiosyncratic component captures the individual preferences of that particular agent.

Formally, we assume the following

Assumption 2. *Let $\mathcal{F}$ be the population of potential entrants and let $c_l \geq 0, l \in \{1, \ldots, L\}$ stand for the common (to all agents) benefits from selecting the alternative l.*

When a new agent enters the economy is selected at random from $\mathcal{F}$ and chooses the alternative l which satisfies

$$l = \arg\max_j \{c_j + e_j | j \in \{1, \ldots, L\}\} \quad ,$$

where $(e_1, \ldots, e_L)$ represents the individual preferences of the agents.

To sum up, at each time step an agent leaves the economy according to Assumption 1 and, after such an exit, a new single agent is allowed to enter according to Assumption 2. Notice that the new entrant may well choose an alternative different from the one chosen by the agent who left. Thus, the model is designed to capture both the genuine entry of new agents as well the reversability of decisions of incumbent agents.

Essentially, Assumption 2 postulates that the choice dynamics is defined by the probability distribution $F(\boldsymbol{e})$ of individual preferences $\boldsymbol{e} = (e_1, \ldots, e_L)$ on the population of agents $\mathcal{F}$. The probability p_l that the next agent called to the choice, chooses location l is indeed

$$p_l = \text{Prob}\{c_l + e_l \geq c_j + e_j \forall j \neq l | \boldsymbol{c}, F(\boldsymbol{e})\} \ .$$

The dynamical process implied by this assumption[1] is essentially undetermined until one provides a precise definition of the distribution F, a difficult task as it requires to model the (private and unexpressed) preferences of the whole population of agents.

However, it is possible to substantially simplify this problem without restricting too much the generality of our approach. Indeed, either by introducing a minimal degree of structure in the decision process or, alternatively, by assuming a simple but plausible structure of the economy it suffices to show that the decision is, in probability, only driven by the common component of the utility function.

The first approach is recovered by interpreting the value of the available alternatives as "discriminal processes" which agents compare to determine their preferred alternative (Thurstone, 1927). In this

[1]Notice that this is exactly the same entry process assumed in Brian Arthur (1990).

case, it seems plausible to assume that the resulting choice is invariant under a uniform expansion of the choice set itself: if we increase the number of alternatives in the economy, by adding, for each alternative, an identical number of new possibilities of the same type, then the probability of choosing a given alternative of each type should be invariant. This simple assumption is enough to guarantee a notable (and desirable) simplification of the problem.

Formally one has

Proposition 2.1. *Consider an economy E with L alternatives and a population of agents $\mathcal{F}$. Now consider an expanded economy, obtained by adding to E, for each alternative l, $k-1$ new identical possibilities. The obtained economy, denoted with E_k, has exactly kL alternatives. Let p_{kl} the probability that a "type l" alternative (that is, an alternative identical to l) is selected by an agent according to the rule in Assumption 1 and Assumption 2. Then if*

$$p_{kl} = p_l \quad \forall k \in \mathbb{N}$$

it is

$$p_l = \frac{c_l}{\sum_j c_j} \ . \tag{2.1}$$

Proof. The proof essentially amounts to show that the only discriminal process compatible with the uniform expansion of the choice set is the one that assumes double-exponentially distributed random utilities. An elegant proof of this property is provided in Yellot (1977), Theorem 6, Section 4. It is then known that a double-exponential distribution of relative utilities assures that the choice probability follows Luce's Choice Axiom (2.1) (Luce, 1959). □

An analogous result can be obtained by looking at the problem from a different perspective. Assume that each available alternative l is actually composed of a number of sub-alternatives. All the sub-alternatives of l possess the same common expected utility c_l, but different agents have, in general, different preferences for the different sub-alternatives. When the number of sub-alternatives becomes large, irrespectively of the particular distribution of individual preferences F, the probability that each alternative is chosen is given by (2.1). Formally we have

Proposition 2.2. *Consider an economy with L alternatives and a population of agents $\mathcal{F}$. Let M_l be the number of sub-alternatives of l and $e_{l,j}$, with $j \in \{1, \ldots, M_l\}$, the individual preferences associated to sub-alternative j. Moreover, agents choose the sub-alternative j of l if*

$$c_l + e_{l,j} = \max \{c_i + e_{i,h} | i \in \{1, \ldots, L\}, h \in \{1, \ldots, M_i\}\} \quad .$$

Then, if the individual preferences $(e_{1,1}, \ldots, e_{L,M_L})$ are i.i.d. *random variables which follow a common distribution F and this distribution has an upper tail which decays sufficiently fast when* $\min \{M_1, \ldots, M_L\} \to \infty$*, then the probability that alternative l is chosen follows* (2.1)*.*

Proof. The general proof, and a complete discussions of the assumptions, is in Jaïbi and ten Raa (1998). The "sufficiently fast" of the proposition means faster than exponential (for instance, a Gaussian distribution). □

The two previous results are sufficient to guarantee that, from a probabilistic point of view, the result of the decision process of heterogeneous agents is completely characterized by the vector of common attractiveness $\boldsymbol{c}$. Notice that the two approaches just outlined get to the same conclusion even if they start from highly different premises. In either case the information processing abilities of the agents and, together, their abilities to specify their "fine-grained" preferences are different. Still it is reassuring to notice that both approaches simplify our dynamical process in exactly the same way, thus adding plausibility to the assumptions underling equation (2.1).

3 Linear externalities

To recall, the model we have introduced in the previous section attempts to describe the distribution of choices of a population of heterogeneous agents among a set of alternatives allowing for different regimes of (positive or negative) social externalities. We have shown that, under plausible simplifying assumptions, the outcome of the decision process does not depend, in probabilistic terms, from the idiosyncratic component of agents' preferences but, on the contrary, is completely characterized by the vector of common attractiveness $\boldsymbol{c}$ of the various alternatives. Hence in order to complete the specification of the model one has to provide an analytic expression for c_l, the common attractiveness component of the alternative l.

We assume that the choice of agents is affected by two factors: by the "intrinsic benefit" associated with each alternative and by a "social benefit" representing the effect of the actual distribution of the entire population among all the possible alternatives on the individual choice.

For sake of tractability, we begin by describing the social effect with a simple linear relationship and we assume the following

Assumption 3. *The common expected utility c_l from choosing the alternative l at time t is given by*

$$c_l = a_l + b_l n_l \quad ,$$

where n_l represents the number of agents that have already selected l at the time of choice and $a_l \geq 0$, $b_l \geq 0$.

Each alternative $l \in \{1, \ldots, L\}$ is then characterized by an "intrinsic attractiveness" parameter a_l and by a "social externality" parameter b_l. The coefficient a_l captures the intrinsic gains that an agent obtains by choosing alternative l, net of any social externality effects. The parameter b_l captures the strength of the externality effect, induced by social interactions, of the alternative l: it is the amount by which the advantages obtained by choosing l increases as a function of the number of agents already chose the same alternative l. The larger is the value of b_l the higher is the incentive for agents to select l as the number of agents that have already chosen the same alternative increases.

Let us summarize assumptions and results discussed above in the following

Proposition 3.1. *At the beginning of each time period t an agent is chosen among the N incumbents to leave the economy according to Assumption 1. Let $m \in \{1, \ldots, L\}$ be the alternative previously chosen by the exiting agent. After the exit takes place a new agent enters the economy. The probability p_l to pick alternative l conditional to the exit occurred in m, according to Assumption 3 and (2.1), is defined as*

$$p_l = \frac{a_l + b_l \left(n_{l,t-1} - \delta_{l,m}\right)}{A + \boldsymbol{b} \cdot \boldsymbol{n} - b_m} \quad , \tag{3.1}$$

where $A = \sum_{l=1}^{L} a_l$, $\boldsymbol{b} \cdot \boldsymbol{n} = \sum_{l=1}^{L} b_l n_l$ and the Kronecker delta $\delta_{x,y}$ is 1 if $x = y$ and 0 otherwise.

In (3.1) $n_{l,t-1}$ is the number of agents who selected l at the previous time step $t-1$ while Kronecker delta $\delta_{l,m}$ in (3.1) implies that it is the number of agents choosing l after the revision that affects the probability that the new choice of the agent will be l. Notice that the assumption of non negative b coefficients introduces in our model a "tendency for conformity" similar, in the spirit, to the one discussed in Brock and Durlauf (2001). The assumption of non-negative b coefficients implies non-decreasing dynamic returns and, whenever $b > 0$, linearly positive externalities.

If $n_{l,t}$ is the number of agents choosing l at time t (with $\sum_{l=1}^{L} n_{l,t} = N$, $\forall t$) the occupancy vector $\boldsymbol{n}_t = (n_{1,t}, ..., n_{L,t})$ completely defines the state of the economy at this time. Due to the stochastic nature of the dynamics (as implied by Proposition 3.1), the only possible description of the evolution of the economy is in terms of probability of observing, at a given point in time, one particular occupancy vector among the many possible ones. Let $\boldsymbol{a} = (a_1, \ldots, a_L)$ and $\boldsymbol{b} = (b_1, \ldots, b_L)$ be the L-tuples containing the parameters for intrinsic attractiveness and for the externality strength of alternative $\{1, \ldots, L\}$. The characterization of the stochastic dynamics of the model is formally provided by the following

Proposition 3.2. *The dynamics of the system described in Assumption 3.1 is equivalent to a finite Markov chain with state space*

$$S_{N,L} = \{\boldsymbol{n} = (n_1, \ldots, n_L) | n_l \geq 0, \sum_{l=1}^{L} n_l = N\} \quad .$$

If $p_t(\boldsymbol{n}; \boldsymbol{a}, \boldsymbol{b})$ is the probability that the economy is in the state $\boldsymbol{n}$ at time t, the probability that the economy is in state $\boldsymbol{n}'$ at time $t+1$ is given by

$$p_{t+1}(\boldsymbol{n}'; \boldsymbol{a}, \boldsymbol{b}) = \sum_{\boldsymbol{n} \in S_{N,L}} P(\boldsymbol{n}'|\boldsymbol{n}; \boldsymbol{a}, \boldsymbol{b}) P_t(\boldsymbol{n}; \boldsymbol{a}, \boldsymbol{b}) \quad ,$$

where $P(\boldsymbol{n}'|\boldsymbol{n}; \boldsymbol{a}, \boldsymbol{b})$ represents the generic element of the Markov chain transition matrix.

Let $\boldsymbol{\delta}_h = (0, ..., 0, 1, 0, ...0)$ be the unitary L-tuple with h-th component equal to 1. Then

$$P(\boldsymbol{n}'|\boldsymbol{n}; \boldsymbol{a}, \boldsymbol{b}) = \begin{cases} \frac{n_m}{N} \frac{a_l + b_l\,(n_l - \delta_{l,m})}{C(\boldsymbol{n}, \boldsymbol{a}, \boldsymbol{b})} & \text{if } \exists l, m \in (1, \ldots, L) \text{ s.t. } \boldsymbol{n}' = \boldsymbol{n} - \boldsymbol{\delta}_m + \boldsymbol{\delta}_l \\ 0 & \text{otherwise} \quad , \end{cases} \tag{3.2}$$

where

$$C(\boldsymbol{n}, \boldsymbol{a}, \boldsymbol{b}) = A + (1 - \frac{1}{N}) \boldsymbol{b} \cdot \boldsymbol{n} \quad . \tag{3.3}$$

Proof. See Appendix A.1. □

The state space of the Markov chain that describes the evolution of the model is the set of all the L-tuples of non-negative integers whose sum of elements is equal to N. The number of elements of the state space, i.e. the dimension of the Markov chain, is

$$\dim S_{N,L} = \binom{N + L - 1}{N} \quad .$$

Note that when the number of alternatives L and/or of agents N increase, the dimension of the Markov chain becomes soon very large. For instance, for $N = 50$ and $L = 10$ the state space contains more than a billion states. On the other hand, according to Assumption 3.1, at most one agent is allowed to choose at each time steps. This implies that the transition matrix of the chain contains many zeros and all transitions happen between very similar states, i.e. states that differ by the location of a single agent. The number of non-zero possible transitions from a given state are at most[2] $L(L-1)+1$. The fraction of non-zero entries in the transition matrix goes to zero when $L, N \to +\infty$.

Assumption 3 allows for an alternative l to have zero intrinsic attractiveness ($a_l = 0$). This kind of alternative is peculiar because, if at some point in time nobody is choosing it, it will never be chosen again. Indeed, according to (3.1), if $a_l = 0$ and $n_l = 0$ the probability of alternative l to be selected by an agent is $p_l = 0$. One can think of this alternative as if it had disappeared from the economy. Since the probability that any chosen alternative looses an agent is always positive, one should expect that, asymptotically, all alternatives with zero intrinsic attractiveness become empty. This is actually the case. More formally, the following applies

Proposition 3.3. *Consider the set of states $\mathcal{S}'_{N,L} \subset \mathcal{S}_{N,L}$ obtained considering only occupancy vectors with no agents choosing alternatives with attractiveness equal to* 0

$$\mathcal{S}'_{N,L} = \{(n_1, \dots, n_L) | n_l \geq 0, n_l = 0 \; if \, a_l = 0, \; \sum_{l=1}^{L} n_l = N\}$$

and let $\mathcal{T}_{N,L} = \mathcal{S}_{N,L}/\mathcal{S}'_{N,L}$ be its complement. Then all states in $\mathcal{T}$ are transient. The set $\mathcal{S}'$ is connected and all its states are persistent.

Proof. See Appendix A.2. □

The set $\mathcal{T}$ contains occupancy vectors with at least one agent who prefers an alternative with zero attractiveness. In the case in which the values of the intrinsic attractiveness parameters are positive for all alternatives, i.e. $a_l > 0 \; \forall l$, then the set $\mathcal{T}$ is empty and $\mathcal{S}'$ is equal to $\mathcal{S}$. Otherwise, assume that the alternatives with attractiveness strictly greater than zero are labeled by the first $L' \leq L$ integers.

In order to present the main result of the paper let us define the 1-step transition coefficient $T_{l \to m}$ using the definition of transition probabilities in (3.2)

Definition 3.1. The 1-step transition coefficient $T_{l \to m}$ between $\boldsymbol{n}, \boldsymbol{n} - \boldsymbol{\delta}_l + \boldsymbol{\delta}_m \in \mathcal{S}'$ reads

$$T_{l \to m}(\boldsymbol{n}) = \frac{P(\boldsymbol{n} - \boldsymbol{\delta}_l + \boldsymbol{\delta}_m | \boldsymbol{n})}{P(\boldsymbol{n} | \boldsymbol{n} - \boldsymbol{\delta}_l + \boldsymbol{\delta}_m)} = \frac{J(n_m, a_m, b_m)}{J(n_l - 1, a_l, b_l)} \frac{C(\boldsymbol{n} - \boldsymbol{\delta}_l + \boldsymbol{\delta}_m, \boldsymbol{a}, \boldsymbol{b})}{C(\boldsymbol{n}, \boldsymbol{a}, \boldsymbol{b})} \tag{3.4}$$

with

$$J(n, a, b) = \frac{a + n\, b}{n + 1} \quad .$$

Then a complete characterization of the "equilibrium" condition of the present model is provided by the following

Proposition 3.4. *The finite dimensional Markov chain described in Proposition 3.2 admits a unique stationary distribution $\pi(\boldsymbol{n}; \boldsymbol{a}, \boldsymbol{b})$.*

On $\mathcal{S}'$ the Markov chain is symmetric under time reversal and satisfies the detailed balance condition $\pi(\boldsymbol{n}') = T_{\boldsymbol{n} \to \boldsymbol{n}'} \pi(\boldsymbol{n})$ between two generic states $n, n' \in \mathcal{S}'$.

[2] This happens when none of the locations is empty.

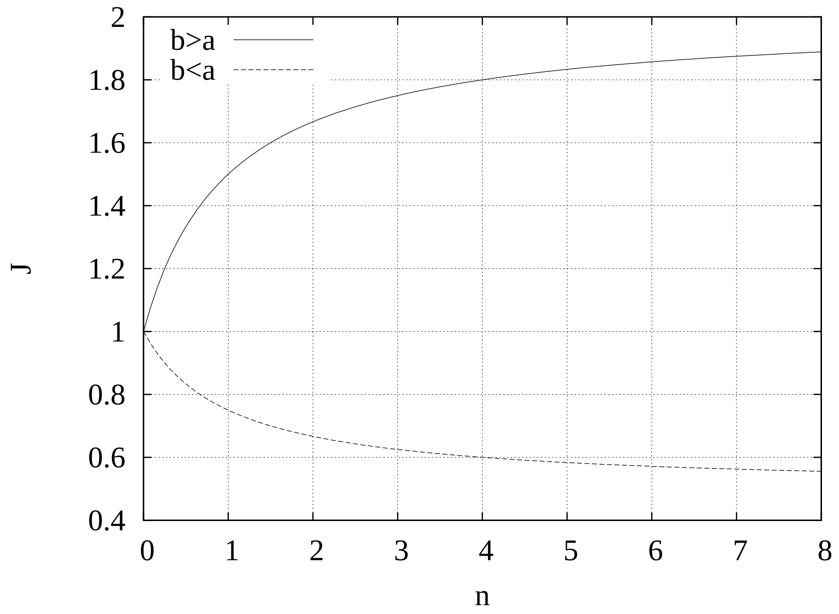

Figure 1: Two examples of the behavior of the "marginal attractiveness" parameter J as a function of the number of agents n for $b > a$ and $b < a$.

On $\mathcal{T} \subset \mathcal{S}$ the stationary distribution is zero: if $\boldsymbol{n} \in \mathcal{T}$ it is $\pi(\boldsymbol{n}) = 0$.
If $\boldsymbol{n} \in \mathcal{S}'$ the stationary distribution $\pi(\boldsymbol{n})$ reads

$$\pi(\boldsymbol{n}; \boldsymbol{a}, \boldsymbol{b}) = \frac{N! C(\boldsymbol{n}, \boldsymbol{a}, \boldsymbol{b})}{Z_N(\boldsymbol{a}, \boldsymbol{b})} \prod_{l=1}^{L'} \frac{1}{n_l!} \vartheta_{n_l}(a_l, b_l) \quad , \tag{3.5}$$

where

$$\vartheta_n(a, b) = b^n \frac{\Gamma(a/b + n)}{\Gamma(a/b)} = \begin{cases} \prod_{h=1}^{n} [a + b(h-1)] & n > 0 \\ 1 & n = 0 \end{cases} \tag{3.6}$$

and $Z_N(\boldsymbol{a}, \boldsymbol{b})$ is a normalization coefficient depending on the number of firms N and on the L-tuples $\boldsymbol{a}$ and $\boldsymbol{b}$.

Proof. See Appendix A.3. □

The above theorem contains one of the main results of our analysis and deserves some discussion. First, notice that all alternatives with zero intrinsic attractiveness, labeled by indices greater than L', disappear from expression (3.5). Second, the 1-step transition coefficient (3.4) between "near" states can be used to gain some insights into the behavior of the model. In equilibrium, the occupancy vector $\boldsymbol{n} - \boldsymbol{\delta}_l + \boldsymbol{\delta}_m$ is more probable than the occupancy vector $\boldsymbol{n}$ if the 1-step transition coefficient $T_{l \to m}(\boldsymbol{n})$ of an agent from l to m is greater than 1. The transition coefficient, in turn, depends on the ratio of the coefficients J of the two alternatives: the one that looses and the one that gains the generic agent under scrutiny. One can interpret this result by saying that, in our stochastic equilibrium, an agent is more likely to move from an alternative with a low J to an alternative with an high J. Thus, $J(n, a_l, b_l)$ can be thought as a measure of the "marginal" attractiveness of alternative l when it is preferred by n agents. It is immediate to check that J is a monotone function of n, increasing if $b > a$ and decreasing if $b < a$. Indeed $dJ/dn \sim b - a$. Then, comparing the values of a and b, it is possible to define two classes of alternatives. An alternative with $b > a$ is, in equilibrium, more attractive than other alternatives if it contains more firms. On the other hand, the attraction strength of an alternative with $b < a$ decreases when the number of agents choosing it increases, even if the externality parameter b is greater than zero. This seemingly counterintuitive conclusion derives from the fact that the stationary distribution of agents across alternatives depends on two effects: (i) the increase in the number of agents choosing a given alternative due to its ability to attract agents from the whole economy; (ii) the reduction in the number of agents picking a given alternative due to the random exit process. Our model postulates that these effects are both linear and the coefficient J captures their overall impact. In Figure 1 an example of the behavior of J as a function of n is reported, for the two cases $a > b$ and $a < b$.

The transition coefficient T contains also the ratio of the terms C computed in the final (numerator) and initial (denominator) state. After simplification, the ratio of the C's reduces to

$$\frac{C(\boldsymbol{n}-\boldsymbol{\delta}_l+\boldsymbol{\delta}_m,\boldsymbol{a},\boldsymbol{b})}{C(\boldsymbol{n},\boldsymbol{a},\boldsymbol{b})} = 1 + \frac{b_m - b_l}{N/(N-1)\,A + \boldsymbol{b}\cdot\boldsymbol{n}} \quad .$$

This term provides a correction to the ratio of factors J's that depends only on the difference of social externality strengths between any two alternatives and, for N sufficiently large, it is in general close to 1. In Proposition (3.4) an explicit expression for the normalization coefficient Z_N is not provided. A formal expression can be straightforwardly obtained by imposing a normalization condition $\sum_{\boldsymbol{n}\in\mathcal{S}} \pi(\boldsymbol{n}) = 1$ for (3.4). This procedure, however, is not, in general, very informative. One can obtain a more useful representation of the normalization coefficient by using the generating function of the stationary distribution.

Proposition 3.5. *Let $\boldsymbol{s} = (s_1, \dots, s_L)$ an L-tuple of real numbers. The generating function $\tilde{\pi}(\boldsymbol{s})$ of the stationary distribution $\pi(\boldsymbol{n})$ defined as*

$$\tilde{\pi}(\boldsymbol{s}) = \sum_{\substack{n_1,\dots,n_l=1 \\ \sum_l n_l = N}}^{+\infty} s_1^{n_1} \dots s_L^{n_L} \pi(n_1, \dots, n_L) \tag{3.7}$$

admits the following representation

$$\tilde{\pi}(\boldsymbol{s}) = \frac{1}{Z_N(\boldsymbol{a},\boldsymbol{b})} \left(\sum_{l=1}^{L} s_l \frac{d}{dx_l} \right)^{N-1} \sum_{l=1}^{L} (A + (N-1)b_l)\, s_l \frac{d}{dx_l} \prod_{l=1}^{L} (1 - x_l\, b_l)^{-a_l/b_l} \Bigg|_{\boldsymbol{x}=0} \quad , \tag{3.8}$$

where $\boldsymbol{x} = 0$ stands, with usual notation, for the set of conditions $x_1 = 0, \dots, x_L = 0$.

Proof. See Appendix A.4. □

As a first application of (3.8) we can obtain an expression for the normalization coefficient Z_N.

Proposition 3.6. *The normalization coefficient $Z_N(\boldsymbol{a},\boldsymbol{b})$ that appears in (3.5) admits the following representation*

$$Z_N(\boldsymbol{a},\boldsymbol{b}) = \left(\sum_{l=1}^{L} \frac{d}{dx_l} \right)^{N-1} \sum_{l=1}^{L} (A + (N-1)b_l) \frac{d}{dx_l} \prod_{l=1}^{L} (1 - x_l\, b_l)^{-a_l/b_l} \,\Big|_{\boldsymbol{x}=0} \quad . \tag{3.9}$$

Proof. From the definition of the generating function in (3.7) one has

$$\tilde{\pi}(1,\ldots,1) = 1$$

that reduces to

$$\tilde{\pi}(\mathbf{1}) = \frac{1}{Z_N(\boldsymbol{a},\boldsymbol{b})}\left(\sum_{l=1}^{L}\frac{d}{dx_l}\right)^{N-1}\sum_{l=1}^{L}(A+(N-1)b_l)\frac{d}{dx_l}\prod_{l=1}^{L}(1-x_l\,b_l)^{-a_l/b_l}\Bigg|_{\boldsymbol{x}=0}$$

so that (3.9) follows. □

Common externality coefficient

Our model allowed for different social externality coefficients b for different alternatives. However, as a first approximation, one might also think of the social externality effect as a force acting with a strength which does not depend from the specific alternative. In our notation this means assuming a constant b across all available alternatives. As showed above, this assumption is also suitable to describe cases in which social externalities are, to some extent, alternative-dependent but the size of the economy is large. In this case, only the alternative with the highest coefficient b's will be chosen by a relevant number of agents so that one can assume all other sectors as having $a = b = 0$, that is remove them from the dynamics.

Formally let us consider a situation in which we assume different intrinsic attractiveness a_l for each different alternative l. On the contrary the strength of the social externality is represented by a single parameter b, equal for all alternatives. Since one may have alternatives with zero intrinsic attractiveness we assume that the first $L' \leq L$ integers label the alternatives with strictly positive intrinsic attractiveness a. Then we have the following

Proposition 3.7. *If $b_l = b\ \forall l \in \{1,\ldots,L\}$ with constant $b > 0$, the stationary distribution defined in* (3.5) *reduces to*

$$\pi(\boldsymbol{n};\boldsymbol{a},b) = \frac{N!\Gamma(A/b)}{\Gamma(A/b+N)}\prod_{l=1}^{L'}\frac{1}{n_l!}\,\frac{\Gamma(a_l/b+n_l)}{\Gamma(a_l/b)}\quad . \tag{3.10}$$

Proof. See Appendix A.5. □

In this case alternatives do, in general, differ and are characterized by their specific attractiveness parameter a_l. In order to define a marginal distribution, one has to specify the parameter a of the alternative of interest.

Proposition 3.8. *The marginal distribution $\pi(n, a_l)$ of the number of agents choosing an alternative with intrinsic attractiveness a_l for the model in* (3.10) *reduces to the Polya distribution*

$$\pi(n;N,L,a_l,A,b) = \binom{N}{n}\frac{\Gamma(A/b)}{\Gamma(A/b+N)}\,\frac{\Gamma(a_l/b+n)}{\Gamma(a/b)}\,\frac{\Gamma((A-a_l)/b+N-n)}{\Gamma((A-a_l)/b)} \tag{3.11}$$

and the average occupancy of site $l \in \{1,\ldots,L\}$ with attractiveness a_l reads

$$< n_l >= N\,\frac{a_l}{A} \tag{3.12}$$

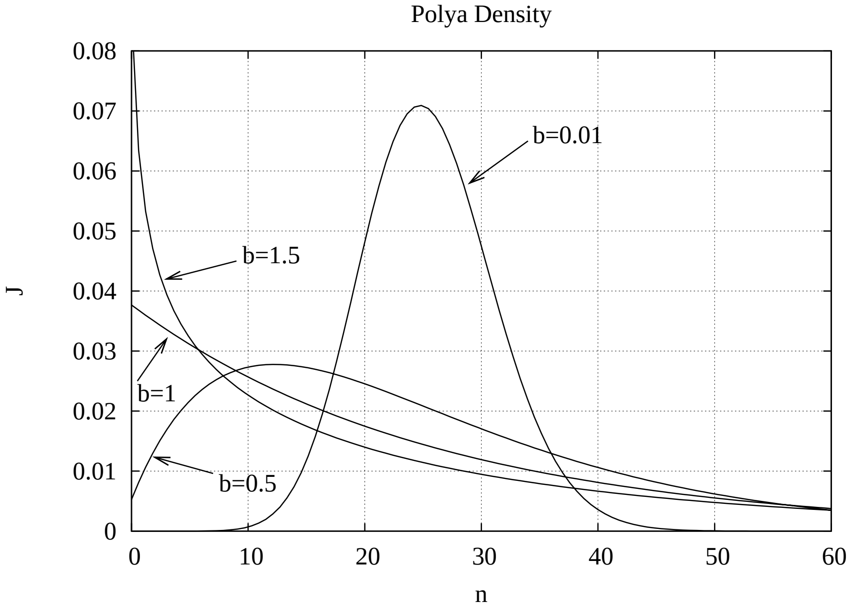

Figure 2: Polya marginal distributions (for different values of b). All distributions are computed for $N = 20000$, $L = 800$, and intrinsic attractiveness $a = 1$.

Proof. See Appendix A.5. □

The marginal distribution in (3.11) depends on the total number of agents N, the total number of alternatives L, the two global parameters $A = \sum_{j=1}^{L} a_j$ and b and the alternative-specific parameters a_l. Figure 2 reports the marginal distribution (3.11) for different values of the parameter b. As we observed before, an increase in the value of b induces an apparent change in the shape of the distribution and, in particular, an increase in the size of its support again hinting at more turbulent dynamics of choice.

No externality coefficient

Among the many specifications one can derive from (3.5) a natural benchmark case emerges assuming that there are no social externalities in agents' choices, that is assuming that $b_l = 0 \; \forall l$. In this case it is straightforward to prove the following

Proposition 3.9. *If* $b_l = 0 \; \forall l \in \{1, \ldots, L\}$ *the stationary distribution defined in* (3.5) *reduces to a Multinomial distribution*

$$\pi(\boldsymbol{n}; \boldsymbol{a}, 0) = \frac{N!}{A^N \prod_{l=1}^{L'} n_l!} \prod_{l=1}^{L'} a_l^{n_l} \quad . \tag{3.13}$$

The corresponding marginal distribution $\pi(n, a_l)$ *of the number of agents choosing an alternative with intrinsic attractiveness* a_l *reduces to the Binomial distribution*

$$\pi(n; N, L, a_l, A, 0) = \binom{N}{n} \left(\frac{a_l}{A}\right)^n \left(\frac{A - a_l}{A}\right)^{N-n} \quad . \tag{3.14}$$

No revision and large economy limits

In the present section we study the properties of the dynamic of our model when we neglect the possibility for the agents to revise their choices and we allow the population of agents to grow indefinitely. In order to do that we switch off the exit process and retain only the entry dynamics described in Assumption 3.1. This implies that the number of agents in the economy will increase linearly with time. Assuming that the process starts with no agents present in the economy, if $n_l(t)$ is the number of agents choosing the alternative l at time t, one has $\sum_l n_l(t) = t$. Let $\boldsymbol{n(t)} = (n_1(t), \ldots, n_L(t))$ be the occupancy vector at time t, the probability that the next agent chooses location l is

$$p_l(\boldsymbol{n(t)}) = \frac{a_l + b_l n_l(t)}{A + \boldsymbol{b}.\boldsymbol{n}(t)} \ , \tag{3.15}$$

with the same notation used in Proposition 3.1. This pure entry dynamics belongs to the family of Generalized Urns schemes discussed, for instance, in Dosi and Kaniovski (1994). In terms of the "fractional occupancy" $\boldsymbol{x}$, where $x_l(t) = n_l/t$, the previous probability defines what is typically called an "urn function"

$$q_l(\boldsymbol{x}, t) = \frac{a_l/t + b_l x_l}{A/t + \boldsymbol{b} \cdot \boldsymbol{x}} \ . \tag{3.16}$$

In this case, the urn function describes the probability that the new entrant agent select the alternative l, given the time t in which it enters the economy and the actual fractional occupancy $\boldsymbol{x}$. Notice that the urn function q_l depends on t both explicitly and implicitly - trough the dependence on the fractional occupancy $\boldsymbol{x}$.

Now let $\beta_l(\boldsymbol{x}, t)$ be a random variable which takes value one with probability $q_l(\boldsymbol{x}, t)$ and value zero otherwise. One can write

$$x_l(t+1) = x_l(t) + \frac{1}{t+1} \left(\beta_l(\boldsymbol{x(t)}, t) - x_l(t)\right) \ .$$

The expected value of variable $\beta_l(\boldsymbol{x}, t)$ is, by construction, equal to $q_l(\boldsymbol{x}, t)$, so that the previous equation can be rewritten as

$$x_l(t+1) - x_l(t) = \frac{1}{t+1} \left(q_l(\boldsymbol{x(t)}, t) - x_l(t)\right) + \frac{\epsilon_l(\boldsymbol{x(t)}, t)}{t+1} \ , \tag{3.17}$$

where $\epsilon_l(\boldsymbol{x(t)}, t) = \beta_l(\boldsymbol{q(t)}) - q_l(\boldsymbol{x(t)}, t)$ is a random variable with expected value equal to zero. In equation (3.17) the increment of the population share who prefers l is driven by two components: a deterministic one, proportional to the difference between the urn function q_l and the actual fraction of firms x_l, and a random term, captured by ϵ_l.

We will provide below a formal result necessary to analyze the limit of the dynamics described by (3.17) when the number of agents t becomes large. First, however, consider a simple heuristic analysis which, albeit incomplete, can be useful to understand what happens in the general case. Since the expected value of the second term of the right hand side in equation (3.17) is zero, one could say that on average, the element of the equation which actually drives the dynamics is the deterministic one. Indeed, considering the expected value conditional on the occupancy at the previous time step

$$\bar{x}_l(t) = \mathrm{E}\left[x_l(t) | \boldsymbol{x(t-1)}\right] \ ,$$

after some algebra one obtains

$$\bar{x}_l(t+1) - x_l(t) = \frac{1}{t+1}\frac{1}{\frac{A}{t} + \boldsymbol{b}\cdot\boldsymbol{x(t)}}\left[\frac{1}{t}\left(a_l - Ax_l(t)\right) + \sum_{j=1}^{L} x_j(t)x_l(t)(b_l - b_j)\right] , \tag{3.18}$$

where we substituted q_l with its expression in (3.16) and made use of the fact that $\mathrm{E}\left[\epsilon|\boldsymbol{x(t-1)}\right] = 0$. Such an expression can be tentatively used to derive some properties of the asymptotic behavior of the system. First of all, consider the case in which at least one b is different from zero. In this case, the first term inside the square brackets vanishes, with respect to the second term, proportionally to t^{-1}. The same applies to the first term of the denominator in front of the square brackets. In this case, retaining only the leading terms in the asymptotic expansion one has

$$\bar{x}_l(t+1) - x_l(t) \sim \frac{1}{t+1}\frac{1}{\boldsymbol{b}\cdot\boldsymbol{x(t)}}\sum_{j=1}^{L} x_j(t)x_l(t)(b_l - b_j) . \tag{3.19}$$

Notice that the coefficients a have completely disappeared from this expression and the asymptotic behavior seems completely driven by the coefficients b. In particular, if there exists an alternative l which possesses a social externality coefficient greater than any other alternative, that is $b_l > b_j, \forall j \neq l$, then, for this alternative, the right hand side of (3.19) is always positive, that is $\mathrm{E}[f_l(t+1)] > f_l(t)$. This means that the expected value of the fraction of agents in l at the next time step is always *higher* than the presently realized value. This seems to suggest that, with probability one, $f_l(t) \to 1$ when $t \to \infty$.

We will see below that this is actually the case. In order to derive more general conclusions, however, we need a few formal definitions. Consider a model with non-null social externality strength, $\boldsymbol{b} > 0$. Let $\boldsymbol{Q}(\boldsymbol{x})$ be the large t limit of the urn function, so that for each component $Q_l(\boldsymbol{x})$ one has

$$\lim_{t\to\infty} q_l(\boldsymbol{x}, t) = Q_l(\boldsymbol{x}) = \frac{b_l x_l}{\boldsymbol{b}\cdot\boldsymbol{x}} \tag{3.20}$$

and let $B(\boldsymbol{b}) \subseteq (S_{L-1})$ stand for the set of fixed points of (3.19), that is

$$B(\boldsymbol{b}) = \{\boldsymbol{x} \in S_{L-1} | \boldsymbol{Q}(\boldsymbol{x}) = \boldsymbol{x}\} . \tag{3.21}$$

Using (3.20), it follows that the set $B(\boldsymbol{b})$ contains all the points $\boldsymbol{x}$ which satisfy the following relation

$$x_l\left(b_l - \boldsymbol{x}\cdot\boldsymbol{b}\right) = 0 \quad \forall l \in 1, \dots, L .$$

This means that if two components of the vector $\boldsymbol{x} \in B$ are different from zero, they should be associated to alternatives with the same b. If we group equal elements of $\boldsymbol{b}$, the point in B takes the form $(0, \dots, x_k, \dots, x_{k+h}, \dots, 0)$, where the indices from k to $k+h$ are associated with vertex with the same social externalities coefficient. In other terms, the points in B are *linear combinations* of vertices with the same b. One has the following

Theorem 3.1. *Consider a model with non-null social externality strength, $\boldsymbol{b} > 0$. Then when t goes to infinity, the vector $\boldsymbol{x}(t)$ is almost surely inside the set $B(\boldsymbol{b})$, that is*

$$\lim_{t\to\infty} \mathrm{Prob}\left\{x_t \in B(\boldsymbol{b})\right\} = 1$$

Proof. See Appendix A.6. □

The basic meaning of the theorem is that fractional occupancies which are linear combinations of vertices with different values of b can *never* be observed when t is large. In other terms, when the number of agents in the economy diverges, only two types of distributions can possibly be observed: either all the agents choose the same alternative or they split choices across alternatives with the same coefficient b. Then, the set B contains at least L points, the vertices of the $L-1$ simplex. In general, the previous theorem does not provide any clue about what outcome, among the many possible, is actually selected. The following result is useful to further reducing the set of possible limit states

Theorem 3.2. *Consider a model with non null social externality strength,* $\boldsymbol{b} > 0$. *Without loss of generality, we can imagine to sort the components of* b *in such a way that* $b_1 \geq b_2 \geq \ldots \geq b_L$. *Suppose that the first* $K \leq L$ *components are equal and let* $C(\boldsymbol{b}) \subseteq S_{L-1}$ *be the set of fractional occupancies in which only the first* K *site are occupied*

$$C(\boldsymbol{b}) = \left\{ (x_1, \ldots, x_K, 0, \ldots, 0) \mid \sum_{j=1}^{K} x_j = 1 \right\}$$

then

$$\lim_{t \to \infty} \text{Prob}\left\{x_t \in C(\boldsymbol{b})\right\} = 1$$

Proof. See Appendix A.6. □

Theorem 3.2 tells us that only the alternatives with the *largest* social externality coefficients are populated in the limit. This finally proves our heuristic conclusion: if there exists an alternative whose b is larger than any other b, then, when the number of agents becomes large, the economy finds itself having all the agents choosing the same single alternative. On the other hand, if there are several alternatives which share the highest coefficient b, Theorem 3.2 predicts a constant positive (in probability) flows of agents moving from the alternative with lower b's toward the alternative with higher b's. Consequently, when t increases, agents increasingly concentrate among the latter alternatives and, in the limit, only these alternatives retain a positive fraction of agents. Theorem 3.2, however, does not give any hint on the way in which the population of agents is distributed across these possible alternatives.

The (partial) solution of the previous problem will be presented in the next Section. For the time being, in order to complete our analysis, let us analyze the case in which all coefficient b's are equal to zero, that is the economy lacks any social externality effect for any alternative. Following our heuristic approach and setting $\boldsymbol{b} = \boldsymbol{0}$ in (3.17) one has

$$\bar{x}_l(t+1) - x_l(t) = \frac{a_l - A x_l(t)}{t+1} \quad . \tag{3.22}$$

The right hand side of (3.22) becomes zero when

$$x_l = \frac{a_l}{\sum_{j=1}^{L} a_j} \ , \tag{3.23}$$

so that, as expected, each alternative contains, asymptotically, a number of agents proportional to its intrinsic attractiveness. In this case, indeed, the process retains no history: the choice of each agent is identical. At each time t, the distribution of occupancies follows a multinomial laws, with probabilities given by (3.23), so that the trivial result follows.

Recovering the Polya approach

We stressed that Theorem 3.2 does not give any hint on the way in which the population of agents is distributed across different alternatives when they are characterized by the same social externality parameter b. In the present section, taking a different approach, we show that it is possible to partially overcome this limitation. Indeed, the dynamical process described above admits an analogous representation in terms of a simple "entry" process in which agents choose their preferred alternative once for all according to a given probabilistic rule. In particular, the expression in (3.10) can be obtained via a Polya urn process.

Let us consider a urn containing balls of L different colors. The initial number of balls of color l is u_l. For each extraction, s balls of the extracted color are added. After N extractions, the probability of finding $\boldsymbol{n} = (n_1, \ldots, n_L)$ balls of the L colors is

$$\pi(\boldsymbol{n}, \boldsymbol{u}, s) = \frac{N!}{\prod_l n_l!} \frac{\prod_{j=0}^{L} u_j(u_j + s) \cdots (u_j + (n_j - 1)s)}{U(U + s) \cdots (U + (N - 1)s)} , \tag{3.24}$$

with $U = \sum_j u_j$ (for a derivation cf. for instance Johnson and Kotz (1977)). If we interpret the extraction of different colors as a the choice of different alternatives it is easy to show (substitute $u_j/s = a_j/b$) that equation (3.24) reduces to (3.10) (and the process is immediately extended also to non integer s and u's). Notice that the initial number of balls u_j (i.e. the initial "relevance" of the alternative) and the number of balls added at each extraction s (i.e. the strength of the social externality effect) enter in the definition of the "intrinsic benefits" a_j/b. Since the equilibrium distribution with N agents and L alternatives is equivalent to the distribution of Polya urns with L colors after N extractions, we can use theorems derived for the latter to derive the asymptotic properties of the former. In particular, applying a result in Polya (1931) and Johnson and Kotz (1977), one can conclude that the distribution of fractional occupancy span, in equilibrium, the entire simplex and follows generalized beta distribution.

4 Conclusions

In this paper we present a discrete choice model with social interactions. We describe a simple economy in which a population of heterogeneous agents choose among a set of several alternatives facing a stochastic utility function. The utility function depends on two terms capturing respectively the common, to all agents, characteristics of each alternatives and the ones that are idiosyncratic to each agent. The effects induced by social interactions among agents are captured simply assuming that the utility associated with a given alternative increases linearly with the number of times the alternative has been chosen in the past.

The essential novelty of our approach is the introduction in such a framework of a randomic revision mechanism of agents choices. This allows us to describe the dynamics of our economy as a Markov process with a finite number of states. We prove that such a process possesses a unique stationary distribution of agents across alternatives and we show how it looks like in some simple instantiations of the model. This distribution can be compared to empirical distributions to directly estimate the magnitude of the externalities associated with social interactions among agents. Moreover we are able to derive expressions for the transition probabilities between different states of the economy which can be used to perform comparative static exercises in order to explore the structural properties of the model.

There are many directions in which our analyses can be extended. Two, in particular, deserve to be mentioned. In the present version of the model, the agent allowed to revise his choice at each time step is randomly selected. It would be interesting to investigate what are the consequences of changing this hypothesis and assuming other mechanisms along the lines discussed in the introduction. Another important extension concerns the role of social interactions. For the sake of simplicity in this paper we assume that social interactions generates a positive linear externality, a sort of conformity effect, in individual choices. Exploring different implementations of the social term in the utility function assuming non-linearities or negative, instead of positive, effect of previous choices is surely worthwhile both at theoretical and empirical level.

References

Abramowitz, M. and I. Stegun (1964). *Handbook of Mathematical Functions.* Dover, New York.

Anderson, S., A. de Palma, and J.-F. Thisse (1992). *Discrete Choice Theory of Product Differentiation.* Cambridge, MIT Press.

Benabou, R. (1996). Equity and efficiency in human capital investment: the local connection. *Review of Economic Studies 63*, 237–264,.

Bottazzi, G., G. Dosi, G. Fagiolo, and A. Secchi (2007). Modeling industrial evolution in geographical space. *Journal of Economic Geography 7*, 651–672,.

Brian Arthur, W. (1989). Competing technologies, increasing returns and lock-in by historical events. *Economic Journal 99*, 116–131,.

Brian Arthur, W. (1990). Silicon valley locational clusters: When do increasing returns imply monopoly. *Mathematical Social Sciences 19*, 235–231.

Brian Arthur, W. (1994). *Increasing returns and path-dependency in economics.* University of Michigan Press, Ann Arbor.

Brian Arthur, W., Y. Ermoliev, and M. Kaniovski, Yu. (1986). Strong laws for a class of path-dependent stochastic process with applications. In V. Arkin, A. Shiraev, and R. Wets (Eds.), *Stochastic Optimization.* Springer-Verlag, Berlin.

Brock, W. and S. Durlauf (2001). Discrete choice with social interactions. *Review of Economic Studies 68*, 235–260,.

Dosi, G., Y. Ermoliev, and M. Kaniovski, Yu. (1994). Generalized urn schemes and technological dynamics. *Journal of Mathematical Economics 23*, 1–19.

Dosi, G. and Y. Kaniovski (1994). On 'badly behaved' dynamics. *Journal of Evolutionary Economics 4*, 93–123.

Durlauf, S. (1996). A theory of persistent income inequality. *Journal of Economic Growth 1*, 75–93,.

Feller, W. (1968). *An Introduction to Probability Theory* (Third ed.), Volume 1. Wiley and Sons, New York.

Fujita, M., P. Krugman, and A. Venables (1999). *The Spatial Economy: Cities, Regions, and International Trade.* The MIT Press, Cambridge.

Jaïbi, M. and T. ten Raa (1998). An asymptotic foundation for logit models. *Regional Science and Urban Economics 28*, 75–90.

Johnson, N. and S. Kotz (1977). *Urn Models and Their Applications.* New York, Wiley.

Luce, R. (1959). *Individual Choce Bahaviour.* New York, Wiley.

McFadden, D. (1984). Econometric analysis of qualitative response models. In Z. Griliches and M. e. Intrilligator (Eds.), *Handbook of Econometrics*, Volume 2. Elsevier, Amsterdam.

Montgomery, J. (1991). Social networks and labor-market outcomes: Toward an economic analyis. *American Economic Review 81*, 1408–1418,.

Pemantle, R. (1990). Nonconvergence to unstable points in urn models and stochastic approximations. *The Annals of Probability 18*, 698–712.

Polya, G. (1931). Sur quelques points de la théorie des probabilités. *Annals of Institute H. Poincaré 1*, 117–161.

Thurstone, L. (1927). A law of comparative judgment. *Psychological Review 34*, 273–286.

Topa, G. (2001). Social interactions, local spillovers and unemployment. *Review of Economic Studies 68*, 261–295,.

Tversky, A. (1972). Elimination by aspects: a theory of choice. *Psycological Review 79*, 281–299,.

Yellot, J. (1977). The relationship between luce's choice axiom, thurstone theory of comparative judgment, and the double exponential distribution. *Journal of Mathematical Psychology 15*, 109–146.

APPENDIX

A Proof of Propositions

A.1 Proof of Propositions 3.2

Proof. From Assumption 3.1 it is clear that the state of the system at time $t+1$ only depends on the state of the system at times t, and no memory is retained of the previous entry/exit events, so that the ensuing stochastic process possesses a Markovian nature. Since the number of agents N and of alternatives L is kept constant, the first part of the theorem immediately follows. Let us thus focus on the derivation of (3.2).

Assumption 3.1 postulates that at each time period, one and only one agent exits the economy and quits from the alternative previously chosen and, subsequently, only one agent chooses one of the L alternatives (including also the one in which exit has occurred).

Therefore, if the state of the economy at time t is $\boldsymbol{n} = (n_1, \dots, n_L)$, the state of the economy at time $t+1$ can be $\boldsymbol{n'}$ if either

1. there exist two alternatives, say l and m, $l \neq m$ such that $n'_{l,t+1} = n_{l,t} - 1$, $n'_{m,t+1} = n_{m,t} + 1$, and $n'_{h,t} = n_{h,t}$ for any h, $h \notin (l, m)$. In this case an agent is removed from alternative l and an agent chose alternative m; or

2. $\boldsymbol{n} = \boldsymbol{n'}$. In this case the entrant has chosen the same alternative of the exiting agent.

If the two occupancy vectors differ by more than one agent, i.e. $\boldsymbol{n'} \neq \boldsymbol{n} + \boldsymbol{\delta}_m - \boldsymbol{\delta}_l$ for all $l, m = \{1, \dots, L\}$, the transition probability is zero and the second row of (3.2) follows.

Given two indexes $l, m \in \{1, \ldots, L\}$, consider the couple of "near" states $\boldsymbol{n}$ and $\boldsymbol{n'} = \boldsymbol{n} + \boldsymbol{\delta}_m - \boldsymbol{\delta}_l$. The probability of transition between these two states can be written as

$$P(\boldsymbol{n'}|\boldsymbol{n}) = \Pr\{\boldsymbol{n} + \boldsymbol{\delta}_m - \boldsymbol{\delta}_l | \boldsymbol{n}\} = \Pr\{\text{agent rejects alternative } l\} \Pr\{\text{agent chooses alternative } m | \text{agent rejected alternative } l\} \tag{A.1}$$

where we drop the explicit mention of the parameters $\boldsymbol{a}$ and $\boldsymbol{b}$ in P. The probability of transition is expressed as a product of the probabilities of two events, denoted with $Pr\{\ldots\}$. This structure reflects the two-step nature of the exit/entry process, as described in Assumption 3.1. In particular, the probability of choosing j is conditional on the previous exit event associated with the rejection of alternative i. Let us now look more closely at these probabilities. Since the exiting agent is chosen at random from all incumbent agents it must be that

$$\Pr\{\text{agent rejects } i\} = \frac{n_l}{N} \ .$$

On the other hand, from Assumption 3.1, the probability of the entrant firm to locate in j can be written as

$$\Pr\{\text{agent chooses alternative } m | \text{agent rejected alternative } l\} = \frac{a_m + b_m\,(n_m - \delta_{m,l})}{H}$$

where H is a suitable normalization constant to be determined. Notice that the outcome of the exit event affects the subsequent entry event via the Kronecker term δ . The final transition probability can then be written as

$$P(\boldsymbol{n} + \boldsymbol{\delta}_m - \boldsymbol{\delta}_l | \boldsymbol{n}) = \frac{n_l}{N} \frac{a_m + b_m\,(n_m - \delta_{m,l})}{H} \ .$$

By imposing the normalization condition

$$\sum_{l,m=1}^{L} P(\boldsymbol{n} + \boldsymbol{\delta}_m - \boldsymbol{\delta}_l | \boldsymbol{n}) = 1$$

one obtains

$$H = \sum_{l=1}^{L} a_l + (1 - \frac{1}{N}) \sum_{l=1}^{L} b_l\, n_l \tag{A.2}$$

that proves the proposition. □

A.2 Proof of Proposition 3.3

Proof. Suppose that $a_1 = 0$ and that the system is in state $(1, \boldsymbol{n}_{L-1})$, where $\boldsymbol{n}_{L-1}$ stands for a vector of length $L-1$. If, at the next time step, alternative 1 looses an agent and the system jumps to a state of the type $(0, \boldsymbol{n'}_{L-1})$ it can never return, later, to the previous state $(1, \boldsymbol{n}_{L-1})$. Since the jump from state $(1, \boldsymbol{n}_{L-1})$ to state $(0, \boldsymbol{n'}_{L-1})$ has a finite probability, there is a finite probability that the state $(1, \boldsymbol{n}_{L-1})$ will never be reached again. That is, this state is transient.

Consider now a state $(n_1, \boldsymbol{n}_{L-1})$ in which there are $n_1 > 1$ agents choosing 1 with $a_1 = 0$. Starting from this state, the system has a positive probability of reaching a state of the type $(0, \boldsymbol{n'}_{L-1})$ in n_1 steps. At this point, the system can never return back to $(n_1, \boldsymbol{n}_{L-1})$. Then, this state is transient as well. It is easy to see that the previous reasoning can be repeated for all the alternatives l with $a_l = 0$. Therefore, all the states where one or more agents chooses an alternative with zero intrinsic attractiveness are transient.

Consider now the states in S'. An alternative with strictly positive a_l and positive b_l has a strictly positive probability of being chosen by an entering agent (see Assumption 3.1). Therefore, any state

$\boldsymbol{n} \in S'$, is reachable in a finite number of steps starting from any other state $\boldsymbol{n'} \in S'$, with a positive probability. The set S' is then connected and, consequently, made of persistent states (Feller, 1968, Theorem 3, p.392). ☐

A.3 Proof of Proposition 3.4

To build the proof it is useful to set two preliminary results. First, recall that a Markov chain which posses a stationary distribution is in general not required to satisfy the detailed balance condition. However, if one is able to find a distribution that satisfies all the detailed balance conditions arising between any possible pair of states of the system, then the chain is said to be reversible and the distribution is invariant (see, for instance, Feller (1968, p.414)).

The second result is summarized in the following

Lemma A.1. *The 1-step transition coefficients in* (3.4) *between states in $\mathcal{S}'$ commute, i.e. for two couples (i, j) and (h, k) if $\boldsymbol{n}, \boldsymbol{n} - \boldsymbol{\delta}_i + \boldsymbol{\delta}_j, \boldsymbol{n} - \boldsymbol{\delta}_h + \boldsymbol{\delta}_k \in S'$ it is*

$$T_{i\to j}(\boldsymbol{n})\, T_{h\to k}(\boldsymbol{n} - \boldsymbol{\delta}_i + \boldsymbol{\delta}_j) = T_{h\to k}(\boldsymbol{n})\, T_{i\to j}(\boldsymbol{n} - \boldsymbol{\delta}_h + \boldsymbol{\delta}_k) \tag{A.3}$$

Proof. Since the transition coefficient from an alternative to itself is 1, if $i = j$ or $h = k$ the Lemma is easily proved.

If $i \neq j$ and $h \neq k$, substituting in (A.3) the definition for T in (3.4) and after simplifying the C coefficients defined in (3.3), one obtains

$$\frac{J_j(n_j)}{J_i(n_i - 1)} \frac{J_k(n_k - \delta_{k,i} + \delta_{k,j})}{J_h(n_h - 1 - \delta_{h,i} + \delta_{h,j})} = \frac{J_k(n_k)}{J_h(n_h - 1)} \frac{J_j(n_j - \delta_{j,h} + \delta_{j,k})}{J_i(n_i - 1 - \delta_{i,h} + \delta_{i,k})} \quad ,$$

where the notation $J_k(n)$ for $J(n, a_k, b_k)$ is employed. One can directly check that for all possible cases the relation is satisfied noting that, due to the requirements $i \neq j$ and $h \neq k$, the values of the Kronecker delta's are not all independent. ☐

We are now able to undertake the proof of Proposition 3.4.

Proof. Proposition 3.3 states that the Markov chain possesses a single connected set of persistent states $\mathcal{S}'$ and, if there exists at least one alternative l with $a_l = 0$, also a set of transient states $\mathcal{T}$. Since the persistent states are all connected, the chain possesses a unique stationary distribution.

The stationary distribution will have probability 0 on all the states in $\mathcal{T}$.

Conversely, in order to compute the expression for the stationary distribution on the states of $\mathcal{S}'$ we make use of the first preliminary result above and we build the invariant density for our model using the detailed balance condition. The transition coefficient $T_{\boldsymbol{n}\to\boldsymbol{n}'}$ from an occupancy configuration $\boldsymbol{n}$ to any other occupancy $\boldsymbol{n}'$defined as

$$\pi(\boldsymbol{n}') = T_{\boldsymbol{n}\to\boldsymbol{n}'}\pi(\boldsymbol{n}) \tag{A.4}$$

can be computed using any suitable series of single-step "jumps" which go from $\boldsymbol{n}$ to $\boldsymbol{n}'$. Since (A.3) holds, as long as these jumps start from $\boldsymbol{n}$ and lead to $\boldsymbol{n}'$, the particular series of jumps one takes is irrelevant and the final transition coefficient reduces to the product of the coefficients T generated by the series of 1-step jumps. The factors $C(\boldsymbol{n}, \boldsymbol{a}, \boldsymbol{b})$ present in successive 1-step transition coefficients T cancel out, so that only the first and last ones are left. Moreover, at each jump in which the site l is involved, a term $J(n, a_l, b_l)$ is generated, with n equal to the number of agents choosing l at that time. Since $T_{l\to m} = T^{-1}_{m\to l}$ this term is at the numerator if a firm chooses l, and at the denominator if he rejects it. With these rules in mind, by applying recursively the definition of T in (3.4), one can see that

$$T_{\boldsymbol{n}\to\boldsymbol{n}'} = \frac{C(\boldsymbol{n}', \boldsymbol{a}, \boldsymbol{b})}{C(\boldsymbol{n}, \boldsymbol{a}, \boldsymbol{b})} \prod_{\substack{l=1 \\ \delta n_l \neq 0}}^{L} \Delta(|\delta n_l|, n_l, a_l, b_l)^{\delta n_l / |\delta n_l|} \tag{A.5}$$

with

$$\Delta(\delta n, n, a, b) = \prod_{h=n}^{n+\delta n-1} J(h, a, b) \quad \delta n > 0 \quad . \tag{A.6}$$

In principle, given the generic transition coefficient $T_{\boldsymbol{n}\to\boldsymbol{n}'}$ and using (A.4) it is possible to compute the probability distribution $\pi(\boldsymbol{n})$ for any occupancy vector $\boldsymbol{n}$, starting from a given occupancy $\boldsymbol{n}_0$. Since the number of states of the Markov chain is finite and the T terms are neither zero nor infinite, this procedure define a proper probability distribution for any $\boldsymbol{n}_0$.

In order to obtain $\pi(\boldsymbol{n})$ there exists, however, a simpler approach. Indeed, noting that, according to the definition in (A.6),

$$\frac{\Delta(n+h, 0, a, b)}{\Delta(n, 0, a, b)} = \Delta(h, n, a, b) \quad n > 0, h \geq 0 \quad , \tag{A.7}$$

even if the null vector $\mathbf{0}$ does not represent a proper occupancy vector, we can use $T_{\mathbf{0}\to\boldsymbol{n}}$ to obtain

$$\pi(\boldsymbol{n}) = \frac{N! C(\boldsymbol{n}, \boldsymbol{a}, \boldsymbol{b})}{Z_N(\boldsymbol{a}, \boldsymbol{b})} \prod_{l=1}^{L} \Delta(n_l, 0, a_l, b_l) \tag{A.8}$$

where the $Z_N(\boldsymbol{a}, \boldsymbol{b})$ represents a suitable and unknown normalization constant and $N!$ has been factored out to simplify following computations. Using (A.7) it is immediate to check that (A.8) satisfies (A.4) for any couple of states in S'. Finally, notice that

$$\Delta(n, 0, a, b) = \frac{\vartheta_n(a, b)}{n!}$$

so that (A.8) reduces to (3.5) and the Theorem is proved. □

A.4 Proof of Proposition 3.5

First, it is useful to introduce the generating function of the coefficients ϑ defined as

$$\tilde{\vartheta}(x; a, b) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \vartheta_n(a, b) \quad . \tag{A.9}$$

With this definition comes the formal property

$$\vartheta_n(a, b) = \frac{d^n}{dx^n} \tilde{\vartheta}(x, a, b)|_{x=0} \quad . \tag{A.10}$$

From the expression of ϑ in terms of Γ functions in (3.6) it follows that

$$\vartheta_n(a, b) \underset{n\to+\infty}{\sim} n!$$

so that (A.9) possesses a finite radius of convergence and (A.10) is meaningful.

In what follows we use the following property

Lemma A.2. *The generating function of the coefficient ϑ defined in* (A.9) *admits the representation*

$$\tilde{\vartheta}(x; a, b) = (1 - x\, b)^{-a/b} \quad . \tag{A.11}$$

Proof. Using (3.6) the definition (A.9) becomes

$$\tilde{\vartheta}(x; a, b) = \sum_{n=0}^{\infty} \frac{x^n\, b^n}{n!} \frac{\Gamma(a/b + n)}{\Gamma(a/b)}$$

which, as from definition 15.1.1 in Abramowitz and Stegun (1964, p.556), reduces to a hypergeometric function

$$\tilde{\vartheta}(x;a,b) = {}_2F_1(a/b;1,1,x\,b)$$

and equation 15.1.8 in Abramowitz and Stegun (1964, p. 556) proves the assertion. □

Using the previous lemma we can obtain the representation of the stationary distribution in (3.7).

Proof. Substituting the relation (A.10) in the definition of π in (3.5), from (3.7) one has

$$\tilde{\pi}(\boldsymbol{s}) = \frac{N!}{Z_N(\boldsymbol{a},\boldsymbol{b})} \sum_{\substack{n_1,\ldots,n_l=1 \\ \sum_l n_l=N}}^{+\infty} C(\boldsymbol{n},\boldsymbol{a},\boldsymbol{b}) \prod_{l=1}^{L} \frac{s_l^{n_l}}{n_l!}\frac{d^n}{dx_l^n}\tilde{\vartheta}(x_l,a,b)|_{x_l=0} \quad . \tag{A.12}$$

In this equation one can introduce the following substitution

$$C(\boldsymbol{n},\boldsymbol{a},\boldsymbol{b}) \to \left(A + \left(1-\frac{1}{N}\right)\sum_{l=1}^{L} b_l s_l \frac{d}{ds_l}\right)$$

and move this differential operator at the beginning of the expression to obtain

$$\tilde{\pi}(\boldsymbol{s}) = \frac{1}{Z_N(\boldsymbol{a},\boldsymbol{b})} \left(A + \left(1-\frac{1}{N}\right)\sum_{l=1}^{L} b_l s_l \frac{d}{ds_l}\right) \left(\sum_{\substack{n_1,\ldots,n_l=1 \\ \sum_l n_l=N}}^{+\infty} N! \prod_{l=1}^{L} \frac{s_l^{n_l}}{n_l!}\frac{d^n}{dx_l^n}\right) \prod_{l=1}^{L}\tilde{\vartheta}(x_l,a,b) \quad .$$

The third factor of the expression is a multinomial expansion. Once this expansion is collected one obtains

$$\tilde{\pi}(\boldsymbol{s}) = \frac{1}{Z_N(\boldsymbol{a},\boldsymbol{b})} \left(A + \left(1-\frac{1}{N}\right)\sum_{l=1}^{L} b_l s_l \frac{d}{ds_l}\right) \left(\sum_{l=1}^{L} s_l \frac{d}{dx_l}\right)^N \prod_{l=1}^{L} \tilde{\vartheta}(x,a,b)|_{x_l=0} \,.$$

Consider the two factor between parentheses. Expanding the derivatives with respect to $\boldsymbol{s}$ in the first factor and recollecting terms one has

$$\left(A + \left(1-\frac{1}{N}\right)\sum_{l=1}^{L} b_l s_l \frac{d}{ds_l}\right) \left(\sum_{l=1}^{L} s_l \frac{d}{dx_l}\right)^N = \left(\sum_{l=1}^{L} (A+(N-1)\,b_l)\,s_l\frac{d}{dx_l}\right)\left(\sum_{l=1}^{L} s_l \frac{d}{dx_l}\right)^{N-1}$$

which substituted in the previous expression gives (3.8) once the expression for $\tilde{\theta}$ in (A.11) is considered. □

A.5 Proof of Proposition 3.7 and 3.8

Lemma A.3. *Let $f_l(x)$ with $l \in (1,\ldots,L)$ be a collection of L real functions infinitely differentiable at $x=0$. Then the following applies*

$$(\sum_{l=1}^{L}\frac{d}{dx_l})^N \prod_{l=1}^{L} f_l(x_l)|_{\boldsymbol{x}=0} = (\frac{d}{dx})^N \prod_{l=1}^{L} f_l(x)|_{x=0}$$

for any integer N.

Proof. The statement is straightforward and can be checked by explicitly taking the left and right derivatives. □

Making use of the Lemma then we can go back to the

Proof of Proposition 3.7. Consider the expression for the normalization constant in (3.9). Under the assumption of constant b it becomes

$$Z_N(\boldsymbol{a}, b) = (A + (N-1)\, b) \left(\sum_{l=1}^{L} \frac{d}{dx_l} \right)^N \prod_{l=1}^{L} (1 - x_l\, b)^{-a_l/b} \;|_{\boldsymbol{x}=0}$$

and using Lemma A.3 it reduces to

$$Z_N(\boldsymbol{a}, b) = (A + (N-1)\, b) \left(\frac{d}{dx} \right)^N \; (1 - x\, b)^{-A/b}|_{x=0} \quad .$$

According to (A.10), the last part of the previous expression is the differential representation of a ϑ function and one has

$$Z_N(\boldsymbol{a}, b) = (A + (N-1)\, b)\; \vartheta_N(A, b)\; .$$

Substituting the expression above for the normalization constant in the definition of the stationary distribution (3.5), one gets

$$\pi(\boldsymbol{n}; \boldsymbol{a}, b) = \frac{N!}{\vartheta_N(A, b)} \prod_{l=1}^{L} \frac{1}{n_l!} \vartheta_{n_l}(a_l, b)$$

that reduces to (3.10) whenever the representation of the ϑ in terms of Γ functions provided by (3.6) is used.

□

Proof of Proposition 3.8. Using the expression for the normalization coefficient Z_N derived above, the generating function can be written

$$\tilde{\pi}(\boldsymbol{s}) = \frac{1}{\vartheta_N(A, b)} \left(\sum_{l=1}^{L} s_l \frac{d}{dx_l} \right)^N \prod_{l=1}^{L} (1 - x_l\, b)^{-a_l/b} \Bigg|_{\boldsymbol{x}=0} \quad . \tag{A.13}$$

Using the representation of the marginal distribution as a derivative of the generating function

$$\pi(n) = \frac{1}{n!} \frac{d^n}{ds_l^n}\; \tilde{\pi}(\boldsymbol{s}) \Bigg|_{\boldsymbol{s}=1}$$

and applying Lemma A.3 and (A.10), the expression in (A.13) becomes

$$\pi(n) = \binom{N}{n} \frac{\vartheta_n(a_l, b)\, \vartheta_{N-n}((A - a_l, b)}{\vartheta_N(A, b)} \quad .$$

Using the representation of the ϑ in terms of Γ functions provided in (3.6) one can see that the last expression is equivalent to (3.11).

The average number of agents choosing alternative $j \in \{1, \ldots, L\}$ can be computed as

$$< n_m >= \frac{d}{ds_m}\; \tilde{\pi}(\boldsymbol{s}) \Bigg|_{\boldsymbol{s}=1} \quad .$$

Using the expression in (A.13) one has

$$< n_m >= \frac{N}{\vartheta_N(A, b)} \frac{d}{dx_m} \left(\sum_{l=1}^{L} \frac{d}{dx_l} \right)^{N-1} \prod_{l=1}^{L} (1 - x_l\, b)^{-a_l/b} \Bigg|_{\boldsymbol{x}=0} \quad .$$

Taking the derivative with respect to x_m and using Lemma A.3 this reduces to

$$< n_m >= \frac{N\, a_j}{\vartheta_N(A,b)} \left(\frac{d}{dx}\right)^{N-1} \prod_{l=1}^{L} (1 - x_l\, b)^{-a_l/b - \delta_{m,l}} \Bigg|_{x=0}$$

and finally, with the help of (A.10), one has that

$$< n_m >= N\, a_m \frac{\vartheta_{N-1}(A+b,b)}{\vartheta_N(A,b)} = N\, \frac{a_m}{A} \quad .$$

□

A.6 Proof of Theorem 3.1 and 3.2

The proof of Theorem 3.1 is based on results obtained in Brian Arthur et al. (1986). Basically, we need to identify a Lyapunov function associated with the dynamics described in (3.17) in which the urn function is replaced with its large t limit. Moreover, some conditions related to the asymptotic behavior of the cumulated random effects must be fulfilled. The Lyapunov function is introduced in the following Lemma. Before, however, we need some formal definition.

Denote with $\overline{b} = \max_l\{b_l\}$ and $\underline{b} = \min_l\{b_l | b_l > 0\}$ the largest and smaller non-negative social externality coefficient b and with $\overline{a} = \max_l\{a_l\}$ the largest intrinsic attractiveness. One has the following

Lemma A.4. *The function $\nu(\boldsymbol{x}) = \overline{b} - \boldsymbol{b} \cdot \boldsymbol{x}$ possesses the following properties*

1. *ν is twice differentiable*
2. *$\nu(\boldsymbol{x}) \geq 0\,, \quad \forall \boldsymbol{x} \in S_{L-1}$*
3. *$< \boldsymbol{Q}(\boldsymbol{x}) - \boldsymbol{x}, \nabla\nu(\boldsymbol{x}) >$*

where $< \ldots >$ stands for the ordinary scalar product and ∇ indicates the gradient.

Proof. Point 1 and 2 are trivial. In order to prove point 3, rewrite the previous equation explicitly

$$\sum_{l=1}^{L} (Q_l(\boldsymbol{x}) - x_l)\, \partial_{x_l} \nu(\boldsymbol{x})$$

which, substituting the expression for $Q_l(\boldsymbol{x})$ in (3.20), after some algebra reads

$$-\frac{1}{\boldsymbol{b} \cdot \boldsymbol{x}} \left(\sum_{l=1}^{L} b_l^2\, x_l - (\boldsymbol{b} \cdot \boldsymbol{x})^2 \right) \quad .$$

The numerator of the previous expression is nothing but the variance of the values b's weighted with probabilities x. Then, the assertion is proved.

□

Proof of Theorem 3.1. Consider the difference $a_l(\boldsymbol{x}, t)$ between the urn function $q_l(\boldsymbol{x}, t)$ at a given time step t and the asymptotic limit $Q_l(\boldsymbol{x})$

$$a_l(\boldsymbol{x}, t) = q_l(\boldsymbol{x}, t) - Q_l(\boldsymbol{x}) = \frac{a_l\, \boldsymbol{b} \cdot \boldsymbol{x} + A\, b_l\, x_l}{(A + t\, \boldsymbol{b} \cdot \boldsymbol{x})\ \boldsymbol{b} \cdot \boldsymbol{x}} \ .$$

Then it is immediate to see that

$$0 \leq a_l(\boldsymbol{x}, t) \leq a(t) = \frac{(\overline{a} + A)\overline{b}}{(A + t\underline{b})\underline{b}} \quad \forall \boldsymbol{x} \in S_{L-1} \quad \text{and} \quad \forall l \in \{1, \ldots, L\}\ ,$$

so that one has

$$\sup_{\boldsymbol{x}} ||a_l(\boldsymbol{x}, t)|| \leq a(t)$$

and the series $\{a(t)/t\}$ admits a finite limit

$$\sum_{t=1}^{\infty} \frac{a(t)}{t} = (\overline{a} + A)\frac{\overline{b}}{\underline{b}} \left(\gamma_{\mathrm{E}} + \Psi \left(\frac{A + \underline{b}}{\underline{b}} \right) \frac{1}{A} \right) . \tag{A.14}$$

According to Theorem 3.1 in Brian Arthur et al. (1986), the finite limit of the summation in (A.14), the existence of the Lyapunov function $\nu(x)$ introduced in Lemma A.4, and the fact that the set $B(\boldsymbol{b})$ defined by (3.21) is made of a finite number of connected components is sufficient to prove the assertion. □

The proof of Theorem 3.2 is based on results obtained in Pemantle (1990). The basic intuition is to think to the deterministic part of (3.17), with q_l replaced by its limit Q_l, as a continuous time dynamical system. The points in $B(\boldsymbol{p})$ are fixed points of this system and can be classified as stable or instable. The following applies

Lemma A.5. *The point $\boldsymbol{x} \in B(\boldsymbol{p})$ is asymptotically stable or non hyperbolic only if all its non-zero components $x_l > 0$ are relative to the vertices with the largest value of social externality coefficient, $b_l = \overline{b}$. Otherwise, the point is asymptotically unstable.*

Proof. Consider the Jacobian matrix of the dynamical system

$$J_{l,j} = \frac{\partial}{\partial_j} \left(Q_l(\boldsymbol{x}) - x_l \right) = \delta_{j,l} \left(\frac{b_l}{\boldsymbol{b} \cdot \boldsymbol{x}} - 1 \right) - x_l \, \frac{b_l b_j}{\boldsymbol{b} \cdot \boldsymbol{x}} .$$

For each $\boldsymbol{x} \in B(\boldsymbol{p})$ there exits a $b^\star$ such that only alternatives with social externality coefficients equal to $b^\star$ posses a non zero firms share. Without loss of generality we can assume that these alternatives are the first $K \leq L$ alternatives, so that $\boldsymbol{x} = (x_1, \ldots, x_K, 0, \ldots, 0)$ and the Jacobian computed in that point reads

$$J_{l,j} = \begin{cases} -x_l \, \frac{b_j}{b^\star} & \text{if} \quad l \leq K \\ \delta_{l,j} \left(\frac{b_l}{b^\star} - 1 \right) & \text{if} \quad l > K \end{cases} .$$

The matrix posses a $(L-K) \times (L-K)$ lower-right diagonal block. Consequently $b_l/b^\star - 1$ ar all eigenvalues of the Jacobian. If there exists an l such that $b_l > b^\star$, the associated eigenvalue is positive. This proves the second part of the theorem. If $b^\star$ is the largest social externality coefficient, then all these eigenvalues are negative, but it remains to analyze the upper-left $K \times K$ block. This block is proportional to the tensorial product of the first K elements of the two vectors $\boldsymbol{x}$ and $\boldsymbol{b}$. Then, it possesses an eigenvalue $-\boldsymbol{x} \cdot \boldsymbol{b}/b^*$ with multiplicity one and eigenvalue 0 with multiplicity $L-1$. This proves the first part of the assertion. □

Proof of Theorem 3.2. It is immediate to verify that the process described in (3.17) satisfies all the requirements of Theorem 1 in Pemantle (1990). Then, according to that theorem, the probability to converge to (asymptotically) unstable fixed points in $B(\boldsymbol{b})$ is zero. Hence, since the probability to converge to $B(\boldsymbol{b})$ is 1, the convergence must be with probability one toward the non-unstable fixed points. Using Lemma A.5 the propositions follows.

□