Changing Identity: The Emergence of Social Groups[1]

Ulrich Horst
University of British Columbia, Canada
Humboldt University Berlin, Germany
horst@math.hu-berlin.de

Alan Kirman
GREQAM, EHESS, Université Aix-Marseille III, IUF, France
alan.KIRMAN@univmed.fr

Miriam Teschl
Robinson College, University of Cambridge, UK
mt367@cam.ac.uk

This version: September 2007

Abstract

The original Homo Economicus has progressed from an atomistic and self-interested individual to a socially embedded agent in modern economics. In particular, social interaction models suggest that the individual's own utility of undertaking an action may be influenced by the number of peers taking this same action. Hence, people gain by conforming to, or differentiating their behaviour from that of others. A number of papers have also suggested why people want to conform. In particular, Akerlof and Kranton (2000, 2002, 2005) suggest that people belong to certain groups and wish to adopt the corresponding social identity by behaving according to the behavioural prescriptions of these groups. In this paper, we present a social interaction model that is based on a different account of identity. The concept of identity used here is on a more personal level and suggests that people have desired self-images of themselves that they wish to attain at some time in the future. Hence, individuals aim to transform their current individual characteristics into those of their self-image. They try to achieve this by joining social groups and adopting the typical characteristics of these groups. However, groups will be modified over time by the people joining them. This may induce individuals to revise their previous choices and eventually to move on and to choose different groups. The model thus presents an endogeneous interaction structure and offers an account of endogenous group formation as well as an endogenous evolution of personal identity. We further study in what sense and under what conditions the dynamics at the individual and at the social level will reach an "equilibrium" and what the nature of such an equilibrium is.

Keywords: Economic agent, social interaction, conformity, personal identity, self-image, change

JEL Classification: D01, D63, Z13

[1]Alan Kirman would like to thank IAS Princeton for its hospitaltiy and the ACI "Complexité" for financial aid. We would also like to thank Kaushik Basu, Roland Benabou, Stéphane Luchini, Eric Maskin, and the members of the Princeton-IAS Economics and Psychology seminar and the Cornell economics seminar for their comments.

"A man is said to be the same person from childhood until he is advanced in years: yet though he is called the same he does not at any time possess the same properties; he is continually becoming a new person not only in his body but in his soul besides we find none of his manners or habits, his opinions, desires, pleasures, pains or fears, ever abiding the same in his particular self, some things grow in him while others perish."
Plato, Symposium 207D-208B

"To assume individuals as fixed or developing independently of economic activity means merely that we do not evaluate, in a normative analysis of economic activity, the way they got to be the way they are - and the way they change."
Herbert Gintis (1974, p. 415)

"...everything can be taken from a man but one thing: the last of the human freedoms – to choose one's attitude in any given set of circumstances, to choose one's own way."
Victor Frankl (1963, p. 104)

1 Introduction

In certain areas of economic theory, the standard concept of the economic agent as an independent, atomistic individual who, endowed with his preferences and budget constraints, considers prices as the sole source of information for his optimal choices, has over the last decades started to be replaced with the concept of a socially embedded individual (Manski 2000). This is an individual, who gains or loses in utility if he behaves as others do or adheres to social norms and rules, independently of his own private self-concerned preferences. Even more generally, the socially embedded individual gains or loses in utility with the number of peers undertaking the same action as he does. Taking account of particular "social preferences" of individuals as well as of the influences of other people's choices on individuals' own decisions offered economists a way of accounting for a number of social phenomena that have a direct impact on individual welfare, but also on a number of economic variables such as productivity and consumption. Economists have been able to explain for example, how educational investment and residential choice can lead to segregation into high-skilled and low-skilled residential areas (Bénabou 1993, 1996; Durlauf 1996), how social interaction can lead to different crime levels across time and space (Glaeser, Sacerdote, Scheinkman 1996), how social capital, defined, for example, as the average of a particular good chosen within a given group of people affects individual choice behaviour such as divorce (Becker and Murphy 2000) or even how the adherence to particular social norms can lead to the formation of "underclasses" (Oxoby 2004). "Non-market interactions" (Glaeser and Scheinkman 2001) have thus become an increasingly important area of research and study for economists.

A common characteristic of these models is that they depict situations based on the rather psychological assumption that "birds of a feather flock together", that is, people desire to conform to others or to given social standards of their particular neighbourhood or even that people choose to interact with those that are closest to them in terms of social distance (Akerlof 1997). The general idea therefore is that like is attracted to like and thus people may conform to the behaviour of others in order to be like them. In most economic models, this either presupposes a given interaction structure or a given set of groups and agents can choose where to situate themselves in the structure or to which groups they want to belong. One interesting result that these models often produce is that the gain in conformity can outweigh the satisfaction of (heterogeneous) individual preferences (Bernheim 1994) and it is this fact that can lead to particularly curious situations where individuals end up with a lower welfare than they would, if for example, they were to consider their own individual private preferences alone.[2]

More recently, Akerlof and Kranton (2000, 2002, 2005) have gone beyond social interaction or preferences for particular social goods and norms and have introduced the concept of "identity", in particular a person's social identity, into a standard model of behaviour. They have gone further since their concept of identity, provides a particular reason as to why people want to conform to others. The general idea is that individuals desire to be and to act in conformity with their social identity and that acting against it would cause a loss in identity, that is, anxiety or discomfort – either in themselves or in others – and this loss can trigger certain actions with the aim to restore the "sense of self" (2000, p. 715). They define identity as the social categories to which an individual belongs. These categories are linked with particular socially determined "self-images" that consist of specific characteristics as well as behavioural rules and prescriptions the individual should comply with if she participates in those categories. Akerlof and Kranton's models then show that individuals have an incentive to choose particular effort levels or actions such that their own individual characteristics fit with the socially determined "self-image" of the social categories to which they belong (or want to belong).

Akerlof and Kranton's models thus explain conformity through identity. However, one can again imagine situations where the "conformity effect" outweighs the satisfaction of individual preferences, that is, where actions that are carried out in accordance with one's social identity receive more weight than actions based on one's own individual preferences.[3] While these models

[2]That is to say that the social utility component of agents' preferences and their private incentives can "clash" (Durlauf 2002, p. 69) and individuals would be better off if an equilibrium prevails where private and social incentives go in the same direction (i.e. have the same sign). Similar issues are also discussed for example in Gordon et al. (2007, p.3) and Blume and Durlauf (2004).

[3]In Akerlof and Kranton (2002), they propose a utility function which is a convex combination of private individual preferences and social identity preferences with an exogenously given distribution of weights in favour

show that this is perfectly rational behaviour. However, one could also assume that there are certain circumstances in which actions are carried out in accordance to one's social identity, but they create some form of discomfort because one's own individual preferences have not been sufficiently satisfied. This is to say that there may be situations in which an individual feels under a social obligation or pressure to conform to his social identity, even though he would have liked to behave according to his own individual preferences. In such a case, one may assume that an agent may also want to leave the group and change his social identity in order to be able to choose more freely what he prefers doing.

Our paper will develop such an idea. It is situated in the context of identity and also of social interaction, in particular following the social interaction literature of e.g. Brock and Durlauf (2001, 2002), Glaeser and Scheinkman (2001) and Horst and Scheinkman (2006). It distinguishes itself from previous papers in two specific ways. First, we introduce the concept of personal identity. This is different from Akerlof and Kranton's concept of identity since it is not based on group memberships and similarities to others, but rather on the particular personal identity of a specific individual and the evolution of her characteristics over time (Horst et al. 2006; Kirman and Teschl 2004, 2006; Teschl 2007). We will assume that instead of acting in conformity to given social standards and thus of choosing to transform individual characteristics to bring them as close as possible to the characteristics of the "social" self-image, individuals have their own "private" self-image which they desire to attain by choosing to belong to particular social groups that help them achieving those desired characteristics. If a particular group does not help to achieve the person's self-image, this may cause a feeling of discrepancy between who the person currently is and who she wants to be, and to alleviate this feeling, the individual may want to change her group participation. Groups are therefore only a means to "realise oneself" and not purely an aim in themselves. In our model, an individual at time t will have particular (stochastic) preferences over her own characteristics, her self-image and her group-memberships (both described by a vector of characteristics), and the direction in which they will evolve in $t+1$. Personal identity is then the individualised process that explains the combination and interaction of these three aspects of a person over time.

The second aspect that we analyse in this paper concerns social interaction. People with different characteristics will join different groups (or types as we will call them) – but this will modify the characteristics of those groups over time. This change in group characteristics may push individuals to reconsider their previous choice of group and eventually to move on to choose a different group that helps them better to achieve their desired self-image. Hence, in contrast to

of social identity preferences. To justify this, Akerlof and Kranton state that "ethnographies suggest such a low value [for individual preferences]" (2002, p. 1175). If this were true in general, this would not be good news for the standard concept of the economic agent who usually simply maximises his own individual preferences.

previous models of the social interaction literature, we will present a model in which the interaction structure is endogenised. This means that we present an account of endogenous personal identity formation and an endogenous formation of social groups.

The next step is to see what we can say about the "society" that consists of people who care about their personal identity. That is, what social interaction models help to understand is the behavior of a population of actors (or of a "society") rather than that of a single actor. Hence, the general question that interests us in this paper is under what conditions the interactive stochastic network that consists of individuals whose personal identity evolves over time will be in equilibrium. An equilibrium will consist of a given number of social groups whose characteristics will no longer change, and with individuals who move between groups but where the distribution of individuals over groups does not change. Thus our equilibrium notion is not a static one in the sense that agents no longer move, but conveys the idea that the social groups stabilise and that the proportions of individuals belonging to each group remain constant. Under standard continuity conditions on the agents' behavior we prove the existence of such an aggregate equilibrium. In personal identity terms, this means that when personal identity is weak, that is, when individuals do not put "too much weight" on current states when revising their choices, and when preferences depend continuously on types, the aggregate equilibrium turns out to be unique and asymptotically stable. This means that the distribution of characteristics, self-images and groups settles down to a unique limit in the long run.

This paper is a theoretical contribution to the existing social interaction literature, a contribution however, that promotes a new behavioural foundation for individuals, which is consistent with psychological (and philosophical) research as we will explain in the following section. In the next section, we will present our concept of personal identity in more detai.l In Section 3 we develop the model. We discuss the results and conclude in Section 4.

2 The Psychology and Philosophy of Personal Identity

Identity has been introduced by Akerlof and Kranton (2000, 2002, 2005) as an extra element into a standard utility function. More specifically, it has been introduced in terms of an identity function, which depends on a person's own actions as well as on those of others, on her assigned social categories, and on how well her own given personal characteristics as well as actions match those of the assigned categories, which are represented by a given number of prescriptions. In their first paper (2000), they present a prototype model with two different social categories (green and red) both associated with a particular action. They assume, following an extensive review of socio-psychological literature, that each person has a number of internalised rules of behaviour and to act against them would cause anxiety. This may be the reason why a person belonging to the green category for example refrains from taking an action associated with the red category. However, even if the person's personal satisfaction from taking this action more than outweighs the psychological discomfort that carrying it out generates, she may cause anxiety in others, i.e. in both, the members of her own category as well as those of the other category. This anxiety constitutes a "loss in identity", which entails a drop in utility. To offset this loss, others may respond with actions, even if costly, that restores their damaged sense of identity. This in its turn will represent a loss to the person who violated the category prescriptions. Whether the person will or will not engage in this action is, in the end, the result of an arbitrage between the resultant advantages and disadvantages. A model of this type has been elaborated in more detail, for example in the context of education and schooling in Akerlof and Kranton (2002), where they explain the separation of pupils, who choose an effort level as well as a group membership, where the groups are the leading crowd, the nerdsänd the "burnouts". It turns out that pupils are better off the closer their own characteristics and behaviour are to the ideal self-image of the category to which they have chosen to belong.

These models thus depict identity as a form of social identification, where group membership and conformity with the group is decisive for a person's self-understanding. In addition to this, identity is psychologically regulated through internalised rules of the particular category or group to which the person belongs and the anxiety produced by violating them. Finally, these models also represent the idea that people have a universal need for belonging, which psychologists consider to be intrinsic to human nature (Baumeister and Leary 1995).

There are also a number of studies showing that self-understanding, and, linked to that, the relation between the "self" and others, is indeed influenced by the social background and culture a person lives in. However, cultures vary and hence different aspects of the "self" are made more salient in differing social contexts or are associated with different psychological processes such as cognition, emotion and motivation. For example, a (rather simplistic) distinction has been made between cultures focussing on an independent conception of the self (associated in particular with North-America) as opposed to those that promote an interdependent conception of the self (associated with East-Asian countries) (Markus and Kitayama 1991, Iyengar and Brockner 2001, Cross and Gore 2003). The essential aspect of the former view is the idea of the "self" as an autonomous, independent person, who is responsive to the social environment, but this social responsiveness is often only a means to express or assert the internal attributes of the "self" (Markus and Kitayama 1991, p. 226). For the independent self, to enter into, but also to leave social engagements freely is important and the person vigilantly protects his or her individualism and limits the influence of others (Cross and Gore 2003, p. 538). Indeed, while it is important for the "self" to be part of a larger collective unit, it is equally important for the person to be different and distinct from others. Hence, a thesis which has been advanced is that distinctiveness is a factor that underlies the selection of social groups in order to achieve a balance between similarity to others and uniqueness and individuality. In other words, a person aims to attain an "optimal distinctiveness" to secure a balanced self-concept in which individual characteristics coexist with being a member of particular social groups (Brewer 1991, 2003). If this optimal identity is threatened or challenged in any way, the individual will attempt to restore congruence or consistency between the self-concept and the group membership, either by adjusting the self-concept to fit better with the prototype of the social group to which the person belongs, or by choosing a group that is more congruent with the "self" (Brewer 2003, p. 484).

The latter idea is also reflected in psychological "self-discrepancy theory" (Higgins 1987). This theory suggests that a person possesses different *domains* of the self, namely the *actual* self, the *ideal* self and the *ought* self. These domains represent respectively the attributes someone possesses, would like to possess or should possess. They can be seen and evaluated from the *standpoint* of the person herself or from that of another person. Inconsistencies or discrepancies between these self-concepts, for instance between the actual self of the person and her ideal self, or between the actual self of the person and the ideal self such as another person ideally wants the person in question to be, will cause different amounts of emotional discomfort and uneasiness. The person will thus attempt to reduce these distances by changing the standpoint of evaluation, the domain of the self or will undertake some actions with the aim of restoring consistency between these selves.

The idea developed above is also that different self-conceptions (or self-images) serve as *regulatory references* with respect to which one tries to reduce certain distances or discrepancies by moving the current state as close to a desired end state as possible (Carver and Scheier 1990) or by distancing the current state as far as possible from an undesired state or self-image (Higgins et al. 1999). This implies that people have a rather extensive knowledge of themselves and their desires and that they have both the possibility and the intention to change over time (Cantor et al. 1986). It is assumed that people work on certain *life-tasks* based on the knowledge of their preferences, abilities, acknowledgement of social roles and other personal characteristics. These life-tasks, however, change over time as do their goals and achievements. By pursuing these different life-tasks, people develop different images of themselves in the future. That is, they are guided by the imagination of their own future possible selves, which are cognitive representations giving a personalised vision of their overall motives. These *possible selves* thus involve the potential for change of a person and represents *who* they could be and would like to be (Cantor et al. 1986).

In contrast, a person with an interdependent conception of the self has the tendency to see his or herself as part of a larger social relationship and is therefore more connected and less differentiated from others. Behaviour is contingent on, if not determined by what the person perceives to be the thoughts, feelings or actions of others. A person's basic motivation is to fit in with others and to fulfil social obligations (Markus and Kitayama 1991, p. 227). Individuals' statuses, roles, and positions and their related obligations and commitments constitute their self-understanding rather than their own individual characteristics, attributes or guiding self-concepts. Indeed, it has even been suggested that those with an interdependent self-conception are less subject to the need to maintain consistency or congruence as mentioned above, which is felt for example when social demands are opposed by individual desires or attitudes (Markus and Kitayama 1991, p. 240). This is because the interdependent conception of the self gives less weight to individual elements and characteristics, which means that in those circumstances, private feelings are generally regulated to fit with the requirement of a given situation. This suggestion is also in line with the idea that for the interdependent concept of the self, social information has more weight, whereas for the independent conception of the self, it is personal information (Iyengar and Brockner 2001). By social information is meant a person's knowledge of what others have said, chosen or done, whereas private information is a person's knowledge of what he or she has done or chosen in the past. Given that the goal of a person with an interdependent concept of self is to strive for interconnectedness and integration, what will be important is not how she or he behaved in the past, but how others behaved. A person with an independent conception however, tends to want to be consistent with his or her choices in the past and will try to maintain a coherent self-concept over time. In contrast, the interdependent self has a more malleable identity concept in that respect and will not feel so strongly this need for personal and intertemporal consistency (Iyengar and Brockner 2001, p. 18)

All this suggests that in human psychology, more is going on than expressing a preference to conform and to behave like others. It also suggests that identity is not related to social identification alone. Indeed, being member of particular groups and thus being associated with a specific social identity does not preclude one's choosing how much importance to attribute to the different aspects of those given groups (Sen 1999, 2006). And this attributed importance will depend on the varying degrees of interaction between a person's personal history and current characteristics, desired self-conceptions or self-images and the choice of groups, which together will contribute to a person's *personal identity*.

But what exactly then is personal identity? Philosophers have reflected on this question for centuries. It has attracted particular interest since two competing views of identity have emerged, namely numerical identity (being always one and the same over time) and qualitative identity (being qualitatively exactly the same as something else) (Ferret 1998). For a long time, philosophers maintained that qualitative identity is a precondition for numerical identity. Only because there is some criterion or some aspect of the person that remains qualitatively the same over time is it possible to say that we are dealing with one and the same, i.e. identical, person over time. But what is this criterion? It cannot be a person's body, because a body changes over time. It cannot be a person's memories either, because a person accumulates new memories and forgets others. It could be a person's DNA. But if by personal identity one understands more than a criterion that establishes the sameness and uniqueness of a person over time, but something that also incorporates information about *who* the person actually is, then DNA is not a satisfactory criterion. Hence, the requirement of qualitative identity is too restrictive in order to define personal identity. People do change throughout the course of their life and identity cannot be considered in any strict sense to be either "all or nothing", but it is rather a matter of degrees. Indeed, as Derek Parfit (1986) suggests, the question is less what identity *is*, but what *matters* for identity, that is, how much change can be admitted in order to still be able to say that person X at time t_1 and person Y at time t_2 are indeed the same person. This means that from one moment to the next, there must be enough psychological and physical *connectedness* between person X at t_1 and person Y at t_2 in order to say that X and Y are indeed one and the same person. Aggregating over several moments, the connectedness between each of them will constitute the *continuity* of a person over time, while at the same time admitting that between moment, say, t_2 and t_9, there need not be any "qualitative identity". All that matters for identity is that we can connect a person's changes over time and that change is not too radical from one moment to the next. However, this account of *continuity* will only tell us *who* the person is, if we allow not only for the possibility that the person changes (because of ageing, memory decay, experiences, etc.), but that the person also *chooses* to change, that is, actively intervenes in the process of her changes (Livet 2004). A person can to some degree choose *who* she wants to be, and at the same time maintain a certain degree of connectedness between the different episodes of life in order to guarantee her continuity over time. Personal identity, then, is not something given nor stable, but the process – at least to some extent self-chosen – that explains change and connectedness of a particular person over time.

This is what we aim to explore in our social interaction model, which we will base on a very general account of individual behaviour. The underlying idea of the model is that agents have particular (stochastic) preferences over the development of their personal states, which are characterised by individual characteristics, self-image and group (or *type*) participation. Individual characteristics and group membership can be considered as describing the *actual self* of the individual, whereas the self-image consists of a number of characteristics that the individual would like to possess and could thus be considered as an *ideal self* of the individual. However, these different selves are all part of the same person and can change over time. Indeed, the agent can for example choose to belong to a particular group, which will help him attaining certain characteristics that bring him closer to his self-image. The idea is that the agent knows that he will change by participating in specific groups. For example, an agent can join a tennis club because he knows that he will be able to take some tennis lessons and acquire those tennis skills that he has not had so far, but that he would have liked to have in his *ideal* image of himself.

The participation in groups will entail social interaction. What we assume is that several people with their different individual characteristics and self-images will join one and the same group – and this will contribute to the particular characteristics of the given social group. That is, a group is also the "aggregate" of its participants, and when people with a variety of individual characteristics join groups, they will, over time, contribute to a change in the "outlook" of those groups. However, this change in characteristics of social groups will affect the state of those individuals who have chosen these groups. What can happen for example is that groups have changed to such an extent that the original reason for joining that group is no longer valid for a particular person.[4] This may cause a discrepancy or uneasiness between the *actual self* of a person and his *ideal self* and in order to restore consistency between his different aspects, the person may want to change the group and chooses a different one that will help him better to achieve his desired self-image-characteristics. For example, suppose that more and more people join the above mentioned tennis club more in order to play tennis casually and to enjoy the company of their co-club-members in the evenings than to play serious tennis. This may upset the person who first has chosen the club in order to become a good tennis player and this will motivate him to look out for a different club where he will be able to find more competitive tennis partners.

In the next section, we thus present a global interaction model and more precisely a mean-field interaction model, where agents feel the impact of others through the empirical distribution of group characteristics, which in its turn, drives the evolution of groups. Generally, agents are endowed with a random utility function that depends on given groups, their current state and their realisation in the next period, as well as on a noise term. The randomness here reflects all the other factors not necessarily listed in the basic characteristics of an individual which may influence this individual's choices as well as identity. Thus, using probabilities reflects the idea that there is some "noise" in the system which represents the heterogeneity of the agents and which is not directly observable by the modeller. This is a typical assumption in social interaction models (e.g. Brock and Durlauf 2000). Other interpretations of this persistent randomness involve assuming fluctuations in the way agents choose, some bounded rationality on the part of the agents or even some process of learning that agents undergo through repeated choices (Blume 1995). Whatever the interpretation, they all fit with our model.

Choices will also entail some costs if the individual changes current individual characteristics too far from a given self-image or from social groups for example. This captures the idea that certain people may have a smaller or larger resistance to changes. It also avoids too radical changes

[4]This conflict has similarities to what was described by Fred Hirsch (1976), who observed that as more people obtain "status" through the accumulation of what he called "positional goods", the less status these goods offer.

in a person's identity. The maximisation of such a utility function will lead to a conditional probability that depends on the agent's current state given current groups. We then introduce an interaction condition that depend on the influence of the current individual state and on current groups. We use this to evaluate the *connectedness* of a person over time despite her changing states and hence to determine her personal identity as discussed above. Indeed, if there is little dependence on current individual state and current groups, then we will say that a person's *connectedness* between the different moments of her evolving life is low, that is, the personal identity of this person is, in this sense, relatively weak. Her current individual state and current groups have little influence on *who* the person will be tomorrow. In this sense, a person can make any random choice; it will not matter much to her and her overall life-story. She makes decisions essentially independently of her *personal information*, that is, her own decisions in the past and has therefore no particular incentive to maintain a feeling of consistency or congruence. Indeed, one could also say that at each moment, a new person or a new "self" emerges and all that we have is a sequence of successive selves, which are relatively independent or unconnected one from the other. The low dependency on current groups is similar to what Glaeser and Scheinkman (2001, p.3) call a *moderate social interaction (MSI)* condition. In relation to that, we will call a low dependency on the current individual state the *moderate individual interaction (MII)* condition. On the other hand, if these dependencies are strong, then personal identity will be stronger, which means that current individual state and current groups have a relatively stronger influence on *who* the person will be tomorrow. A person's identity is thus *continuous* over time, in the sense that the different moments that constitute this person's life are sufficiently connected in order to attribute them to one and the same person (or "self") even though the person is changing.

The advantage of our approach is that it allows for some remarkable personal changes to take place. One finds occasionally individuals who, for example, despite the highly constraining nature of the social or economic environment in which they live, manage to develop a life which is not at all consistent with their background. In our framework these may constitute low probability events but which, nevertheless, can happen. To take an example, consider the case of Cornelia Sorabji, a Parsee woman converted to Christianity, who was the first woman to obtain a law degree at Oxford and who went on to become a distinguished barrister and a militant human rights advocate. Her career and life, including her conscientiously chosen changes, can be thought of as a choice in our context. Indeed Amartaya Sen (2001, p. 331) says of her: "She chose her plural identities influenced by her background, but through her own decisions and priorities. In the last respect, she was not unique, despite the uniqueness of her chosen combination of identities." This sort of switch is possible in our model and it will have an important effect on the evolution of the individual's identity but also on the evolution of the social groups that make up society.

Following the social interaction literature, what we are interested in is what happens at the "macroscopic" level, given the particular individual behaviour at the "microscopic" level. More precisely, we want to define an appropriate notion of "equilibrium" for such a system. Then we establish under what conditions this interactive stochastic system will be in such an equilibrium. As it will turn out, when our MII and MSI conditions are satisfied, that is, when we place a bound on intrapersonal and interpersonal interaction, then the evolution of groups settles down asymptotically to a unique fixed point and the sequence of individual states will converge to a unique probability measure. It is this result which will allow us to say that when personal identity is weak (MII and MSI conditions are satisfied), society "stabilises" and will not undergo any further changes in the characteristics of the groups.

3 A Formal Model of Personal Identity with Social Interaction

Let us now be more specific about our formal model. We consider a model with an countably infinite set

$$\mathbb{A} = \{1, 2, \ldots\}$$

of interacting economic agents. Each agent is characterized by her *state*. An agent's state comprises a list of her *characteristics*, her *self-image* (or desired self) and her *social group*. The characteristics of the individual $a \in \mathbb{A}$ are represented by a vector in c^a in $\mathbb{R}^n$. Its entries may represent not only the standard preferences over goods but also many other dimensions such as the agents' attitude towards smoking (non-smoker, occasional smoker, smoker) or outdoor activities (athletic, unathletic), her athletic skills (professional tennis player, semi-professional tennis player, mediocre tennis player, beginner) her level of income (high, low, medium), or the amount of time she devotes to social activities for example. These are characteristics that can be changed over time. However, there may also be other characteristics which remain "unchangeable" such as race or a person's height for example. For convenience, we think of the set C of possible characteristics as being finite:

$$C = \{c^1, \ldots, c^m\}.$$

We assume that self-images will also be represented by n-dimensional vectors. We do not specify the origin of self-images. We simply assume that each person has at each moment of time particular desires or aspirations the person wishes to attain with greater or smaller intensity. These self-images do not necessarily contain only self-regarding characteristics. They may well contain components that are influenced by the social background and culture in which a person lives. The finite set of possible self images is denoted by

$$Z = \{z^1, \ldots, z^r\}.$$

At each point in time, the agent will have certain characteristics, a desired self-image and will have chosen a social group. These groups can be for instance peer groups, sport clubs, or political parties. A social group will be represented by a vector of the members' "typical", predominant or average characteristics.[5] These group-characteristics will be particularly appealing to certain agents. That is, a high profile tennis club, for example, may define itself in large part by a member's average income; on the other hand, a high-income dedicated tennis player is likely join a club with a disproportional high number or proportion of wealthy committed sportsmen. In order to capture the dependence of group characteristics on the average characteristics of the people that joint them, groups are represented as elements of the convex hull of the set of characteristics. More specifically we assume that the agents can choose from a finite set of groups and that the vector of groups belongs to the compact convex set

$$X \subset \mathbb{R}^n.$$

The agents' characteristics and self-images are not fixed but are modified over time either as the result of the selection of social groups, but they can also be modified by direct choice. For example, if a person decides to dress only in black, she can do so without needing to join a particular social group. Self-images can also change after a process of personal reflection and evaluation. Here, however, we do not specify in any further detail how and why self-images do change, but simply assume that they can do so. Social groups, on the other hand, evolve as a result of the choices of the individuals. The high profile tennis club, for instance, may lose some of its appeal to the more snobbish members over time, when it is joined by an increasing number of average income people. For each time $t \in \mathbb{N}$ we denote by

$$x_t := (x_t^1, \ldots, x_t^n) \in X, \quad c_t = (c_t^a)_{a=1}^\infty \quad \text{and} \quad z_t = (z_t^a)_{a=1}^\infty$$

the vector of possible groups at time t, the current configuration of characteristics and the current configuration of self-images, respectively. The *empirical distribution* associated to the configuration $(s_t^a)_{a=1}^\infty$ is denoted by

$$\varrho_t := \lim_{n \to \infty} \frac{1}{n} \sum_{a=1}^n \delta_{s_t^a}(\cdot)$$

provided the limit exists. Here $\delta_s(\cdot)$ denotes the Dirac measure that puts all mass on $s \in S$.

Remark 3.1 *Since the state space is finite the distribution ϱ_t can be identified with a vector $\varrho_t = (\varrho_t^{i,j,k})$ where $\varrho^{i,j,k}$ denotes the fraction of agents that choose group x_t^i and whose characteristic and self-image are given by c^j and z^k, respectively.*

[5] Our framework would be flexible enough to even introduce the members' self-image-characteristics as characteristics that define the social group. However, we do not explore this possibility at the moment.

The empirical distribution ϱ_t describes the overall distribution of groups, characteristics and self-images in the population at date $t \in \mathbb{N}$. The long run dynamics of aggregate behavior is thus described by the asymptotics of the sequence $\{\varrho_t\}$. What we shall eventually be interested in is whether the groups settle down and also whether the proportions of individuals choosing those groups stabilizes over time.

3.1 Utilities and Choices

As we said previously, at any point in time, an agent is specified by her characteristics, self-image and social group. Social groups, however, change over time as a result of the choices of the individuals. For technical reasons this is inconvenient. While the number of possible groups is fixed, the actual characteristics of those groups at each point may not be the same. To overcome this problem we "label" groups and call them "types". Individuals thus choose from or are characterized by the"types" (i.e. labelled groups) they choose, even if what those types represent is changing. For example, the Republican party is a "type" whose characteristics change but agents can still choose to vote Republican. Hence, we will label the groups by $1, 2, \dots, n$, and we assume that the agent $a \in \mathbb{A}$ chooses a *type* y_t^a from the set

$$Y := \{1, 2, \dots, n\}.$$

In what follows, when we refer to x_t, we mean the current configuration in terms of characteristics of "types" (groups), whereas y_t refers more generally to "types", that is labeled groups, whose characteristics can change over time.

An agent is then characterized by his *state* s_t^a,

$$s_t^a = (c_t^a, z_t^a, y_t^a).$$

a vector of characteristics, self-image, and a type. The state space will be denoted by

$$S = \{s = (c, z, y) : c \in C,\, z \in Z,\, y \in Y\}.$$

We assume that an agent's new state in the following period $t+1$ is given by the realization of a random variable whose distribution depends on the current state. More specifically, given s_t^a and her new state s_{t+1}^a the agent enjoys the (random) utility

$$u(x_t, s_t^a, s_{t+1}^a, \epsilon^a(s_{t+1}^a)).$$

Here $\epsilon^a(s_{t+1}^a)$ is a random variable whose realization my depend on the chosen state. If the realization is observable to the agent prior to making her choice, her (conditional) optimal action is given by[6]

$$\hat{s}^a_{t+1} = \arg\max_s u(x_t, s^a_t, s, \epsilon^a(s)).$$

Under mild restrictions on the utility function and the distribution of the random terms, the optimal choice is almost surely uniquely defined. In this case we denote the conditional distribution of her best response by

$$\pi_{x_t}(s^a_t; \cdot).$$

The transition probability π_x governs the one-step transition dynamics of individual characteristics and self images given the current vector of types x. It reflects the preferences an agent has over his future state, given the current characteristics of available types. Under the assumption that the taste shocks are independent and identically distributed across individuals, the agents act conditionally independently of each other, given their current states: while all the agents react to the prevailing composition of social groups, the transition to a new state is made independently of the current choice of others. As a result, the joint distribution of the new state $(s^a_{t+1})_{a=1}^{\infty}$ in the following period takes the product form

$$\Pi_{x_t}(s_t; \cdot) = \prod_{a=0}^{\infty} \pi_{x_t}(s^a_t; \cdot).$$

3.2 Examples of Utility Functions

Before we proceed with the dynamics of types, let us illustrate how our setting fits into the discrete choice framework associated with many models of social interaction. As it is customary in many models of social interactions, we assume that the noise term enters the utility function in an additive way so the utility function is of the form

$$U(x_t, s^a_t, s^a_{t+1}) + \epsilon^a(s^a_{t+1}). \tag{1}$$

We further assume that the random variables $\epsilon(s)$ $(s \in S)$ are independent and doubly exponentially distributed with parameter $\beta > 0$. That is,

$$\mathbb{P}[\epsilon(s) \leq b] = e^{-e^{\beta b}}.$$

Since $\epsilon(s)$ has a continuous distribution and the set of states is finite an agent's conditional optimal choice is uniquely determined:

$$\arg\max_{s \in S} \left\{ U(x_t, s^a_t, s^a_{t+1}) + \epsilon^a(s) \right\}.$$

[6] "If the realization is observable" means that once the observation is made, one has no longer to worry about expected utility. However, if the agent does not know what his or her "taste shock" or random perturbation will be tomorrow, it would be reasonable to maximise expected utility. If on the other hand the random term is observable, we have a deterministic optimization problem. For an outsider (or modeller) though, it looks as if the agent is making random choices when his or her shock is private information.

It has been shown by Brock & Durlauf (2001, 2002) that assuming the double exponential distribution implies that the probability of her choosing a specific state s^a is given by

$$\pi_{x_t}(s_t^a; s^a) = \frac{\exp\{\beta U(x_t, s_t^a, s^a)\}}{\sum_{\hat{s}\in S} \exp\{\beta U(x_t, s_t^a, \hat{s}^a)\}}. \tag{2}$$

Remark 3.2 *For the special case $\beta = 0$ any state is chosen with equal probability. When β tends to infinity, an agent will choose with increasing probability her best response given her current state. The quantity β may thus be viewed as parameterizing the sensitivity of choice. More precisely, it parameterizes both the dependence of an agents' choice on her present state and the strength of interaction, i.e. the dependence of her choice of the choices of others (through the types).*

Let us illustrate the general setting by some specific examples. We begin with an example where agents want to adopt for themselves, as much as possible, characteristics of some social group. However, there is a cost of movement that penalizes deviations from the current characteristic and some fixed self-image. This is the characterisation of a rather "conservative" behaviour, where agents do not want to change too much and prefer, other things being equal, to remain *who* they are.

Example 3.3 *Suppose that the agents' self image is a given fixed vector z^* so the choice variables are characteristics and social groups. Assume furthermore that the utility function given the current vector of types x_t and the current characteristic c_t^a takes the form*

$$U(x_t, c_t^a, s^a) = u\left(x_t^{y^a}, c^a\right) - J_1|c_t^a - c^a| - J_2|z^* - c^a| - J_3|x_t^{y_t^a} - c^a| + \epsilon(s^a). \tag{3}$$

Here $u(x_t^{y^a}, c^a)$ is an instantaneous utility associated with the current choice that is independent of the prevailing characteristic. The quantities J_1 and J_2 penalize deviations from current characteristics and the self image, while J_3 measures the dependence of an agent's utility of the difference between the chosen characteristic and the characteristics of the selected social group. The random shocks are independent and doubly exponentially distributed with $\beta = 1$ so the conditional joint distribution of labels and characteristics in the following period is then given by

$$\pi_{x_t}(s_t^a; s^a) = \frac{\exp\left\{u(c^a, x_t^{y^a}) - J_1|c_t^a - c^a| - J_2|z^* - c^a| - J_3|x_t^{y_t^a} - c^a|\right\}}{\sum_{\hat{s}^a\in S} \exp\left\{u(\hat{c}^a, x_t^{\hat{y}^a}) - J_1|c_t^a - \hat{c}^a| - J_2|z^* - \hat{c}^a| - J_3|x_t^{\hat{y}^a} - \hat{c}^a|\right\}}. \tag{4}$$

The next example captures a situation where an agent's cost of moving own characteristics closer to her desired self-image can be reduced by choosing an appropriate social group.

Example 3.4 *Suppose again that the self-image is fixed at some level z^* and consider the following modification of the utility function in (3):*

$$U(x_t, s_t^a, s^a) = u\left(x_t^{y^a}, c^a\right) - J_1|c_t^a + \alpha(x_t^{y^a} - c_t^a) - c| - J_2|z^* - c^a| + \epsilon(s^a). \tag{5}$$

As in the preceding example the parameter J_2 penalizes deviations of characteristics from the desired self-image. What distinguishes the current utility function from that in (3) is the fact that by choosing a social group an agent's old characteristics moves some distance α towards that group. This may reduce the cost of choosing a new characteristic in the vicinity of the agent's self-image z^*. In this case the transition probabilities are independent of the self-image and take the form

$$\pi_{x_t}(s_t^a; s^a) = \frac{\exp\left\{u(x_t^{y^a}, c^a) - J_1|c_t^a + \alpha(x_t^{y^a} - c_t^a) - c| - J_2|z^* - c^a|\right\}}{\sum_{\hat{s}^a \in S} \exp\left\{u(\hat{c}^a, x_t^{\hat{y}^a}) - J_1|c_t^a + \alpha(x_t^{\hat{y}^a} - c_t^a) - \hat{c}^a| - J_2|z^* - \hat{c}^a|\right\}}. \tag{6}$$

The general form of the utility function (1) includes situations where the agents' only choice variables can be types and where new characteristics and self-images could be implicitly defined in terms of the group choice. This is captured by a functional relation between the current characteristics (self-images) and chosen type and the new characteristics (self-images). More generally, we may allow for some randomness in this functional relation and consider a situation of the form

$$c_{t+1}^a = f\left(c_t^a, x_t^{y_t^a}, \theta_t^a\right) \quad \text{and} \quad z_{t+1}^a = g\left(z_t^a, x_t^{y_t^a}, \eta_t^a\right).$$

Here the realizations of the random variables θ_t^a and η_t^a are observable prior to making the decision. As an illustration we may think of a situation in which the characteristics and self-images move a random distance towards the group that they choose.

Example 3.5 *Consider the following modification of the utility function in (3) where there is no intrinsic value to an agents' choice:*

$$U(x_t, s_t^a, s^a) = -J\left|f\left(c^a, x_t^{y_t^a}, \theta_t^a\right) - g\left(z_t^a, x_t^{y_t^a}, \eta_t^a\right)\right| + \epsilon(y^a). \tag{7}$$

Here the agent uses the group characteristics to move her own characteristics closer to her desired self-image. The distance between her characteristics and self-image is penalized by a multiplier J and there is a random cost $\epsilon(y^a)$ associated with the choice of the label y^a. In this case the transition dynamics of the states take a more complex form even if we assume that the random variables θ^a and η^a are independent and identically distributed across time and agents.

3.3 Dynamics of Types

What we want to look at is the evolution of the types over time and how this is influenced by the movement of individuals between the types. For this we have to make assumptions about the way in which the characteristics of types are modified by the changes in the choices of the individuals. To obtain a first simple characterization of the dynamics of the sequence $\{x_t\}_{t\in\mathbb{N}}$, we assume that types are modified as a function of average behavior. In the first place we restrict ourselves to a simple linear updating rule for the dynamics of types. This leads us to make the following assumption.

Assumption 3.6 *There exist a Lipschitz continuous function $F = (f^1, \ldots, f^m)$ with constant L and a constant $\alpha \in (0, 1)$ such that*

$$x^i_{t+1} = \alpha x^i_t + (1-\alpha) f^i(\varrho_{t+1}) \quad \text{so that} \quad x_{t+1} = \alpha x_t + (1-\alpha) F(\varrho_{t+1}). \tag{8}$$

We illustrate Assumption 3.6 by the following example where types are simply points on the unit interval and the way in which they change is determined by the fractions of individuals choosing them.

Example 3.7 *(i) Suppose that $x^i_t \in [0,1]$ and that new types are convex combinations of the old type and the proportion of individuals choosing them. In this case*

$$f^i(\varrho_{t+1}) = \int \boldsymbol{1}_i(y)\varrho_{t+1}(dc, dy) = \lim_{n\to\infty} \frac{1}{n}\sum_{a=1}^{n} \boldsymbol{1}_i(y^a_{t+1})$$

We may view ϱ_{t+1} as a vector $(\varrho^{i,j,k}_{t+1})$, so

$$f^i(\varrho_{t+1}) = \sum_{j,k} \varrho^{i,j,k}_{t+1}.$$

As a result, the mapping F is Lipschitz continuous with constant 1.

(ii) The new types could also reflect the characteristics of those who selected them. In such a situation new types could be given as a convex combination of old types and the average characteristics of the members:

$$f^i(\varrho_{t+1}) = \frac{\int c\boldsymbol{1}_i(y)\varrho_{t+1}(dc, dy)}{\int \boldsymbol{1}_i(y)\varrho_{t+1}(dc, dy)}.$$

We have the following representation:

$$f^i(\varrho_{t+1}) = \frac{\sum_{j,k} c^j \varrho^{i,j,k}_{t+1}}{\sum_{j,k} \varrho^{i,j,k}_{t+1}}. \tag{9}$$

In this case Assumption 3.6 requires that an agent chooses each state with strictly positive probability. If each state is chosen with probability at least δ, then F is Lipschitz with constant $L(\delta)$. The Lipschitz constant is decreasing in δ.

Given the evolution of types $\{x_t\}$ the evolution of the aggregate behavior as described by the process $\{\varrho_t\}$ can almost surely be described by a deterministic recursive relation, due to the law of large numbers of independent random variables. In fact, all the agents that find themselves in the same state choose their new states - independently of others - with the same probabilities. Thus, there is a proportion $\pi_{x_t}(s_t^a; s_{t+1}^a)$ of agents with current state s_t^a that choose to be in state s_{t+1}^a in the next period. Since the fraction of agents in state s_t^a is $\varrho_t(s_t^a)$ the distribution of states at

$$\varrho_{t+1} = \int \pi_{x_t}(s^a; \cdot)\varrho_t(ds^a) := H(\varrho_t, x_t) \tag{10}$$

The *joint* dynamics of types and empirical distributions thus follows (almost surely) the deterministic recursive relation

$$\begin{pmatrix} \varrho_{t+1} \\ x_{t+1} \end{pmatrix} = \begin{pmatrix} H(x_t, \varrho_t) \\ \alpha x_t + (1-\alpha) F \circ H(x_t, \varrho_t) \end{pmatrix} =: G(x_t, \varrho_t). \tag{11}$$

3.4 Equilibrium States and the Long Run Dynamics of Aggregate Behaviour

So far we have introduced the concepts of individual and social change in the immediate future, i.e from one time-period to the next. To analyze what happens in the longer run, we need an *interaction condition* that puts restrictions on the agent's individual behaviour. More specifically, in order to guarantee an almost sure convergence of types and average actions, we need to impose a *weak interaction* condition on the agents' individual behavior. We express the weak interaction condition in terms of the following total variation norm:

$$\frac{1}{2}|\pi_x(s;\cdot) - \pi_x(\hat{s};\cdot)|_1 = \frac{1}{2}\sum_{\tilde{s}} |\pi_x(s;\tilde{s}) - \pi_x(\hat{s};\tilde{s})|.$$

Assumption 3.8 *(i) The individual transition probabilities depend continuously on both the social types and the current states. More precisely, there exist constants β_1 and β_2 such that uniformly in types and states we have that*

$$\frac{1}{2}\left|\pi_x(s;\cdot) - \pi_x(\hat{s};\cdot)\right|_1 \leq \beta_1 \quad \textit{and} \quad \frac{1}{2}\left|\pi_x(s;\cdot) - \pi_{\hat{x}}(s;\cdot)\right|_1 \leq \beta_2 \sum_{i=1}^{n} |x^i - \hat{x}^i|.$$

(ii) The constants β_1 and β_2 satisfy $\beta_1 + \beta_1 < 1$.

The coefficient β_1 measures the dependence of an agents' new type and characteristics on his current state. If β_1 is low, it represents a *moderate individual interaction (MII)* condition. In the special case where where agents choose their new states *independently* of their previous states,

$$\pi_x(s;\cdot) = \pi_x(\hat{s};\cdot),$$

and we can choose $\beta_1 = 0$. In a personal identity sense, this would mean that there is no *connectedness* between the different choices of the agent. The agent is nothing other than a succession of different selves. That is, agents choose randomly their new state, independently of their own personal history. If, on the other hand, the agents "remain *who* they are" in the sense that

$$\pi_x(s;s) = 1 \quad \text{and} \quad \pi_x(\hat{s};\hat{s}) = 0 \quad \text{for some } \hat{z} \neq z$$

we have that $\beta_1 = 1$ and Assumption 3.8 (ii) is violated. This means that agents have a very strong personal connectedness and that they stick to their previous decision and remain *who* they are[7]. However, even in the case where people *choose* to remain the same, this does not mean that agents do not, in fact, change. Indeed, the type to which they belong may well have been changed, but this has not motivated the agent to move on and to select a different type.

Example 3.9 *Consider again the utility function in (3). In this case β_1 is an increasing function of J_1 and $\beta_1 = 0$ means that $J_1 = 0$. For the utility function of Example 3.4 β_1 increases in both J_1 and α.*

The constant β_2 places a quantitative bound on the dependence of the new state on the current vector of types. If β_2 is low, it represents the *marginal social interaction (MSI)* condition. A low β_2 implies that the probability of choosing a new state depends little on current types. In other words, it means that the agents' preferences for future states do not change much if the characteristics of available types change a little. A high β_2 on the other hand implies that agents can react very strongly to changes in characteristics in available types[8].

[7]Note that β_1 is a bound placed on the *difference* between two different transition probabilities π_x. This means that in the case β_1 has a low value, the two transition probabilities π_x cannot be too different. In the case where β_1 has a higher value, the two transition probabilities may, but do not need to be, different,. Thus, in the latter case, there is a higher chance that an agent will end up in a particular state, given that he was in a specific state previously. That is, the agent may put "more weight" on his previous state.

[8]As in the case of β_1, β_2 places a bound on the *difference* between the two transition probabilities π_x and $\pi_{\hat{x}}$ if the distance between the two vectors of groups x and $\hat{x}$ is not too large. If the latter distance would be large, then even a very small β_2 would not impose any restrictions on the transition probabilities and we cannot say much about this situation. However, again, whereas a small β_2 implies that the two transition probabilities must be almost similar, a higher value of β_2 tells us that the two transition probabilities can be different (but do not need to be). Thus, in the latter case, there is a higher chance that agents will end in a particular state, given the previous characteristics of available groups. It is also clear that, if the distance between the two vectors of groups is very small, but β_2 rather high, the difference between the two transition probabilities must also be small.

Example 3.10 *Consider again the utility function in (3). In this case β_2 is an increasing function of J_3 and $\beta_2 = 0$ means that $J_3 = 0$. For the utility function of Example 3.4 β_3 increases in both J_1 and α and $\beta_3 = 0$ if and only if $\alpha = 0$.*

In terms of personal identity then, this means that if β_1 and β_2 are small, personal connectedness between two different moments is relatively low. Hence, personal identity is weak - a person's new state will depend little on her current state and on types. We have here again the typical *multiple self* conception where a person is nothing else than a succession of different selves.

3.4.1 Existence of Aggregate Equilibrium

In a model where agents' choices have a random component, the resultant ëxperimentsprevent agents' choices from converging pathwise to some steady state. An appropriate notion of equilibrium is not a particular state but rather a distribution of states that reflect the proportion of time the process spends in the various states. This notion of equilibrium is closely related to the notion of equilibrium on the aggregate level of the empirical distribution. In our model the agents choose their states conditionally independently of each other so the dynamics of social groups and empirical distributions follows a recursive deterministic dynamics. A fixed point (x^*, ϱ^*) of the mapping (11) may be viewed as an aggregate equilibrium. If the groups are initially located according to the vector x^* and the distribution of states is given by ϱ^*, then the system is in a steady state at the aggregate level. While the individual states fluctuate in a random manner, the social types do not move and the distribution of states is fixed. The existence of an aggregate equilibrium follows from Assumption 3.8 (i). It is easy to show that the continuity condition on the individual transition laws along with the continuity of F implies continuity of the map G. Since G maps a compact convex set into itself, it has a fixed point, due to Brower's theorem. We thus have show the following result.

Theorem 3.11 *Suppose that Assumption 3.6 and Assumption 3.8 (i) are satisfied. Then the map G has a fixed point so an aggregate equilibrium exists.*

3.4.2 Uniqueness of Aggregate Equilibrium

While the existence of an aggregate equilibrium follows immediately from mild continuity conditions on the agents' choice probabilities, a uniqueness and stability result requires much stronger assumptions. It has been established in the social interaction literature (Brock & Durlauf (2001, 2002), Horst & Scheinkman (2006, 2007)) that uniqueness of equilibrium is guaranteed if the interaction is sufficiently weak. Our Assumption 3.8 (ii) can be viewed as such a *weak interaction condition*. It requires both β_1 and β_2 to be small, that is that MII and MSI are satisfied. Thus, loosely speaking Assumption 3.8 says that agents do not put “too much weight” on current states when revising their choices, and that preferences depend continuously on types.

Our weak interaction condition of the agent’s individual behavior also constrains the macroscopic behavior of agents. The future distribution of agents over types is continuously dependent on the current distribution of agents and the current vector of types. That is, the weak interaction condition constrains the dependence of the future distribution of agents on the current distribution and the characteristics of the available types.

We are now ready to state our convergence result. Its proof is given in the appendix.

Theorem 3.12 *Suppose that Assumption 3.8 is satisfied and that $L\beta < 1$.*

a) The sequence $\{(x_t, \varrho_t)\}_{t\in\mathbb{N}}$ of types and empirical actions converges almost surely to the unique fixed point (x^*, ϱ^*) of the map G.

b) The stochastic kernel π_{x^*} has a unique stationary measure μ and the sequence of individual states converges in probability, i.e., for all states s,

$$\lim_{t\to\infty} \mathbb{P}[s_t^a = s] = \mu(s)$$

The first part of the theorem gives a sufficient (though not necessary) condition that guarantees that the types settle down to a unique limit x^* in the long run. Asymptotically, the agents new characteristics is thus chosen according to the transition probability

$$\pi_{x^*}(s_t^a; \cdot),$$

and in the long run, the full dynamics of the microscopic process $\{s_t\}_{t\in\mathbb{N}}$ is described by the stochastic kernel

$$\Pi_{x^*}(s_t : \cdot) = \prod_{a=1}^{\infty} \pi_{x^*}(s_t^a; \cdot).$$

Under our assumptions, the transition law π_{x^*} has a unique stationary distribution μ and, for any initial condition s_0, the distribution of individual states converges weakly to μ. In this sense the dynamics on the macroscopic level of types and average actions settle down to a unique deterministic limit (x^*, ϱ^*) whereas the dynamics on the microscopic level of individual behavior settles down to the probabilistic limit μ. In equilibrium therefore, the characteristics of types’ will not change anymore but each type will continue to be constantly chosen by a given fractions of agents. Agents themselves will continuously move from one group to the other according to given probabilities. Depending on the higher or lower entry-values of the probabilistic limit μ, agents will move more slowly or hop rather quickly from one type to the other. At each time, however, the individuals’ probabilities of being in one of the future states do not depend much on their current state and on the characteristics of available types. Hence, the agent's switching of types does not have any impact on his individual history because his different choices are disconnected from each other. During each step, the agent assimilates to a greater or lesser degree the characteristics of the respective group. Hence, in equilibrium, agents are multiple selves that randomly adapt to different social groups. Their personal identity though, i.e. their connected and related change over time, has lost much of its significance.

4 Discussion and Conclusion

We have presented a dynamic model of social interaction that is based on an account of individual behaviour which is enriched by taking into account the idea that people are concerned about the evolution of their personal identity. In particular we allow people to influence that evolution. They choose to participate in particular groups, not solely in order to conform to a certain set of rules or actions of their social environment, but because they know that the participation at those groups will have an effect on their individual characteristics. People do not belong to fixed groups but rather use the latter as a means to an end. They can learn particular abilities or acquire characteristics that they would like to possess according their desired self-image, that is, "who" they would like to be at some point in the future. Social interaction comes through the selection of social groups. However, what is important in our model is that the choices made by individuals as to which group to join in turn determine the evolution of those groups. This is because the groups' characteristics will be modified by the people joining them. This in its turn will induce individuals to revise their evaluations of groups and possibly to choose an alternative to their present choice.

Our aim has been to define a notion of what might be thought of as an equilibrium of this dynamic process. Then we established the conditions under which the interactive dynamics of individual behaviour and social groups converge to a unique and stable equilibrium. As has been found in other social interaction models, we show that convergence to such an equilibrium is guaranteed if interaction is sufficiently weak. In our dynamic framework, this requires the satisfaction of two conditions – the *marginal social interaction (MSI)* condition and the *marginal individual interaction (MII)* condition. This means that if the agents' choices of their new states are sufficiently independent of current groups and current individual states, a unique stable equilibrium exists. Translated back into a personal identity context, this means that people's choices are relatively *unconnected* from one moment to the next, that is, they are sufficiently independent of people's personal history and social backgrounds. Paradoxically, perhaps, this can be thought of as saying that the personal identity of individuals is relatively weak. An equilibrium under these circumstances means that on the macroscopic level, there will have emerged a fixed number of social groups whose evolution of characteristics has been stabilised. On the microscopic level, individuals will continue to move between groups. However, the distribution of individuals between groups will not change in the long run.

Our results then, in a sense, contrast with models of segregation in which agents, endowed with explicit preferences for particular states, can choose where to situate themselves on a given network or interaction structure (e.g. Bénabou 1993, Akerlof 1997). If agents in our model had strong preferences for particular states, that is, if, the MII or MSI condition would be violated, then we may not have convergence to a unique equilibrium independently of initial conditions, but we may have a number of different possible equilibria. Consequently, any segregated or stabilised society may *not* be the result of the particular preferences of individuals (e.g. for their education, the place where they live or even the racial composition of their neighbourhood, such as in Schelling (1971)), but could also be because choices are made on a random basis without any specific rationale. However, even in our framework, we may still obtain convergence with very particular preferences. If all agents would choose their current state with probability one, then nothing ever changes and the system would obviously be in equilibrium but one in which there is no social mobility. Hence, the observation of any particular formation or segregation of groups does not necessarily reveal much about the underlying behaviour of individuals. In our particular case, it may be that if individuals did care more about the evolution of their personal identity, they would be able to live in a more dynamic society where they and groups continue to change and where they would supposedly have a larger set of possibilities to achieve any desired self-images and the kind of life they would like to live.

This consideration is clearly related to the freedom of choice and opportunity literature (e.g. Gravel 1994, Sen 1985, Pattanaik and Xu 1998, Sugden 1998 etc.) and the question would then be to what extent a person can achieve what he desires. Indeed, an interesting extension of the proposed model would be to see under which configuration of society, individuals would achieve their highest welfare and/or had their greatest freedom of choice to become "who" they want to be. This would also be well in line with what in Ball (2003) refers to as "counterfactual history" (p. 201). In Ball's introductory paper into "The Physical Modelling of Human Social Systems" (2003), a framework that is close to that of our own model, he states that this kind of approach is helpful in order to lay out a range of different "counterfactual" scenarios, to stimulate the reflection on particular issues concerning the interaction between personal desires and preferences and membership of particular social groups. Indeed, one of the aims in this paper has also been to show that social interaction does not always consist in conformity or in aligning behaviour to particular reference groups as previous social interaction papers have often assumed, but that individuals can also be seen as "autonomous" and "independent" beings, despite being members of a particular social structure in which they participate and with which they interact.

5 References

Akerlof George A., 1997, "Social Distance and Social Decision", *Econometrica* 65(5), pp. 1005-1027.

Akerlof George A., Kranton Rachel E., 2000, "Economics and Identity", *Quaterly Journal of Economics* vol 65 n 3, pp. 715-753.

Akerlof George A., Kranton R. E., 2002, "Identity and Schooling: Some Lessons for the Economics of Education", *Journal of Economic Literature* 40 (4), pp. 1167-1201.

Akerlof George A., Kranton Rachel E., 2005, "Identity and the Economics of Organisation", *Journal of Economic Perspectives* 19(1), pp. 9-32.

Ball Philip, 2003, "The Physical Modelling of Human Social Systems", *Complexus* 1, pp. 190-206.

Baumeister Roy F., Leary Mark R., 1995, "The Need to Belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation", *Psychological Bulletin* 117, pp. 497-529.

Becker Gary S., Murphy Kevin M., 2000, *Social Economics: Market Behavior in a Social Environment*, Havard University Press.

Bernheim Douglas B., 1994, "A Theory of Conformity", *Journal of Political Economy* 102(5), pp. 841-877.

Blume Lawrence E., 1995, "The Statistical Mechanics of Best-Response Strategy Revision", *Games and Economic Behavior* 11, pp. 111-145.

Blume Lawrence E., Durlauf Steven N., 2004, "The Interactions-Based Approach to Socioeconomic Behavior", in Durlauf Steven N., Young Peyton (eds.), *Social Dynamics*, The MIT Press, pp. 15-44.

Brewer Marilynn, 1991, "The Social Self: On Being the Same and Different at the Same Time", *Personality and Social Psychology Bulletin* 17(5), pp. 475-482.

Brewer Marilynn, 2003, "Optimal Distinctiveness, Social Identity, and the Self", in Leary Mark, Tangney June P. (eds.), "Handbook of Self and Identity", The Guilford Press, pp. 480-491.

Brock William A., Durlauf Steven N., 2000, "Interaction-Based Models", Technical Working Paper 258, retrieved at: http://www.nber.org/papers/T0258.

Brock William A., Durlauf Steven N., 2001, "Discrete Choice with Social Interactions", *Review of Economic Studies* 68, pp. 235-260.

Brock William A., Durlauf Steven N., 2002, “A Multinomial Choice Model of Neighborhood Effects,” *American Economic Review* 92, pp. 298-303.

Cantor Nancy, Markus Hazel, Niedenthal Paula, Nurius Paula, 1986, “On Motivation and the Self-Concept”, in Sorrentino Richard M., Higgins E. Tory (eds.),*Handbook of Motivation and Cognition* vol.1, Guilford Press.

Carver Charles S., Scheier Michael F., 1990, “Principles of Self-Regulation: Action and Emotion”. in Sorrentino Richard M., Higgins E. Tory (eds.),*Handbook of Motivation and Cognition* vol.2, Guilford Press.

Cross Susan E., Gore Jonathan S., 2003, “Cultural Models of the Self”, in Leary Mark, Tangney June P. (eds.), “Handbook of Self and Identity”, The Guilford Press, pp. 536-564.

Durlauf Steven N., 1996, “A Theory of Persistent Income Inequality”, *Journal of Economic Growth* 1, pp. 75-93.

Durlauf Steven N., 2001, “A Framework for the Study of Individual Behavior and Social Interactions”, *Sociological Methodology* 31(1), pp. 47-87.

Ferret Stéphane, 1998, *L'idéntité*, Flammarion.

Frankl Victor, 1963, *Man's Search for Meaning: An Introduction to Logotherapy*, Washington Square Press.

Gintis Herbert, 1974, “Welfare Criteria with Endogenous Preferences: The Economics of Education”, *International Journal of Economics* 15(2), pp. 415-430.

Glaeser Edward L., Scheinkman José A., 2001, “Non-Market Interaction”, HIER Discussion Paper No. 1914.

Gordon Mirta B., Nadal Jean-Pierre, Phan Denis, Semeshenko Viktoriya, 2007, “Discrete Choices under Social Influence: Generic Properties”, Working paper, retrieved at http://halshs.archives-ouvertes.fr/docs/00/13/54/05/PDF/GoNaPhSe07_Customers GeneralProperties.pdf

Gravel Nicolas, 1994, “Can a Ranking of Opportunity Sets Attach an Intrinsic Importance to Freedom of Choice?”, *The American Economic Review* 84(2), pp. 454-458.

Higgins E. Tory, 1987, “Self-Discrepancy: A Theory Relating Self and Affect”, *Psychological Review* 94(3), pp. 219-340.

Higgins E. Tory, Grant Heidi, Shah James, 1999, "Self-Regulation and Quality of Life: Emotional and Non-Emotional Life Experiences", in Kahneman Daniel, Diener Ed, Schwarz Norbert (eds.), *Well-Being: The Foundations of Hedonic Psychology*, Russell Sage Foundation.

Hirsch Fred, 1976, *Social Limits to Growth*, Routledge.

Horst Ulrich, 2000, *Asymptotics of locally and globally interacting Markov chains arising in microstructure models of financial markets*, Shaker-Verlag.

Horst Ulrich, Kirman Alan, Teschl Miriam, 2006, "Changing Identity: The Emergence of Social Groups", Greqam Working Paper No. 2006-51.

Horst Ulrich, Scheinkman José, 2006, "Equilibria in systems of social interactions", *Journal of Economic Theory* 130, pp. 44-77.

Horst Ulrich, Scheinkman José, 2007, "A limit theorem for systems of social interaction", Working Paper.

Iyengar Sheena, Brockner Joel, 2001, "Cultural Differences in Self and the Impact of Personal and Social Influences", in Wosinska Wilhelmina, Cialdini Robert B., Barrett Daniel W., Reykowski Janusz (eds.), *The Practice of Social Influence in Multiple Cultures*, Lawrence Erlbaum Associates, pp. 13-32.

Kirman Alan, Teschl Miriam, 2004, "On the Emergence of Economic Identity", *Revue de Philosophie Economique* 9, pp. 59-86.

Kirman Alan, Teschl Miriam, 2006, "Searching for Identity in the Capability Space", *Journal of Economic Methodology* 13(3), pp. 299-325.

Livet Pierre, 2004, "La pluralité cohérente des notions de l'identité personnelle", *Revue de Philosophie Economique* 9, pp. 29-58.

Manski Charles F., 2000, "Economic Analysis of Social Interactions", *Journal of Economic Perspectives* 14(3), pp. 115-136.

Markus Hazel R., Kitayama Shinobu, 1991, "Culture and the Self: Implications for Cognition, Emotion and Motivation", *Psychological Review* 98(2), pp. 224-253.

Oxoby Robert J., 2004, "Cognitive Dissonance, Status and Growth of the Underclass", *Economic Journal* 114(498), pp. 727-749.

Parfit Derek, 1986, *Reasons and Persons*, Oxford University Press.

Pattanaik Prashanta, Xu Yongsheng, 1998, "On Preference and Freedom",*Theory and Decision* 44, pp. 173198.

Plato, *The Symposium*, Penguin Books.

Schelling Thomas, 1971, "Dyanmic Models of Segregation", *Journal of Mathematical Sociology* 1, pp. 143-186.

Sen, Amartya. 1985, *Commodities and Capabilities*, North-Holland.

Sen Amartya, 1999, *Reason Before Identity: The Romanes Lecture*, Oxford University Press.

Sen Amartya, 2001, "Other People", *Proceedings of the British Academy* 111 , pp.319-335

Sen Amartyam, 2006, *Identity and Violence: The Illusion of Destiny*, Allen Lane.

Sugden Robert, 1998, "The Metric of Opportunity", *Economics and Philosophy* 14, pp. 307-337.

Teschl Miriam, 2007, "Personal Identity Decisions for Examining Change", Working Paper, Robinson College, Cambridge, UK.

A Proof of Theorem 3.12

Our weak interaction condition guarantees that the maps $\varrho \mapsto H(x,\varrho)$ and $x \to H(x,\varrho)$ are contractions with respect to $|\cdot|_1$ with constants β_1 and β_2 uniformly in $x \in X$ and $\varrho \in U$, respectively[9]. More precisely, under Assumption 3.8,

$$|H(x,\varrho) - H(x,\hat{\varrho})|_1 \leq \beta_1 |\varrho - \hat{\varrho}|_1 \quad \text{and} \quad |H(x,\varrho) - H(\hat{x},\varrho)|_1 \leq \beta_2 |x - \hat{x}|_1.$$

This means that the future distribution of agents over types is continuously dependent on the current distribution of agents and the current vector of types. That is, the weak interaction condition restraints the dependence of the future distribution of agents on the current distribution and the characteristics of available types. Stated differently, if we had two almost similar but still different current distributions of agents and characteristics of available types, the future distribution of agents in both “worlds” would be almost the same.

Proof of Theorem 3.12:

a) In order to guarantee a convergence of the empirical distributions, we need to impose a growth condition on F. To this end, we equip $X \times U$ with the norm

$$|(x,\varrho)^t| := \max\{|x|_1, |\varrho|_1\}.$$

and denote by $|DF(\varrho)|_1$ the column-sum-norm of the derivative $DF(\varrho)$ of the function F introduced in (8). With $L := \sup_\varrho |DF(\varrho)|_1$, the map $\varrho \mapsto F(\varrho)$ is Lipschitz continuous with constant L:

$$|F(\varrho) - F(\hat{\varrho})|_1 \leq L|\varrho - \hat{\varrho}|_1.$$

In particular, the maps $x \mapsto \alpha x + (1-\alpha) F \circ H(x,\varrho)$ and $\varrho \mapsto \alpha x + (1-\alpha) F \circ H(x,\varrho)$ are Lipschitz with constant $\alpha + (1-\alpha)L\beta_2$ and $(1-\alpha)L\beta_1$, respectively. As a result, the map

$$(x,\varrho) \mapsto \alpha x + (1-\alpha) F \circ H(x,\varrho)$$

and the map G defined in (11) are Lipschitz continuous with respective constants

$$\alpha + (1-\alpha)L\beta \quad \text{and} \quad \max\{\alpha + (1-\alpha)L\beta, \beta\}.$$

In particular, G is a contraction if $\beta L < 1$. Since G maps the compact convex set $X \times U$ continuously into itself, and so the assertion follows from standard fixed-point arguments.

[9] For details we refer the reader to Horst (2000).

b) Under Assumption 10, the stochastic kernel π_{x^*} has a unique stationary measures μ. Since the sequence $\{x_n\}_{n \in \mathbb{N}}$ converges almost surely to x^*, we can apply similar arguments as in Horst (2000) to prove convergence is distribution of individual states.

$\square$