

Changing Identity: The Emergence of Social Groups

Ulrich Horst

University of British Columbia, Canada

Alan Kirman

Richard B Fisher Member, IAS, Princeton,
GREQAM, EHESS, Université Aix-Marseille III, IUF, France

Miriam Teschl

GREQAM, France
Robinson College, University of Cambridge, UK

Preliminary and incomplete version May 2006

Abstract

Homo Economicus has progressed from an atomistic and self-interested individual in standard economics to a socially embedded agent in modern economics who is endowed with a particular social identity or with specific preferences for the latter. While this vision makes the economic agent, more realistic, its representation by adding entries in an agent's utility function poses problems. The distinction between the agent's own preferences and those that society imposes on the agent becomes both blurred, and difficult to make. In order to separate these two facets of preferences, we explore the idea of an agent's personal identity of which his social identity is one aspect. To consider personal identity means endowing the economic agent with the capacity to change, in particular, in the social space. This greatly complicates the problem of making welfare judgements since it is difficult to define a welfare improvement in the future. We next present a model in which the evolution of peoples' identities is stochastic. This can explain the appearance of individuals who seem to defy social convention. Our model shows how personal identity changes. We can find a social equilibrium only if personal identity matters little. But this suggests, paradoxically, that for individuals to have well-established identities, society must be one which is not in equilibrium, but one which will continue to change and in which both the social groups and the individuals are constantly being modified.

Keywords: Economic agent, social preference, social identity,
personal identity, concept of change, social equilibrium

JEL Classification: D01, D63, Z13

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)

1 Introduction

The typical Homo Economicus of modern economic theory has, for a long time, been conceived of as an atomistic individual, unconnected to and unconcerned about other people and their behaviour and who is represented by a strictly private, i.e. personal and purely self-concerned preference ordering. More recently, however, the idea that Homo Economicus is a social being and thus not independent of other people's behaviour or of his cultural and social background, has become a serious consideration in economics. Different sorts of social preferences, i.e. preferences that had in one sense or the other a relation to the agent's membership in a social community have been attributed to him. He is now considered as caring about other people or about his social status in a given society. This status is seen as depending on his class-membership or social background, and he is now seen as having developed a preference for norm and rule following which lead him to take into account notions of fairness and to adopt pro-social behaviour. Yet, while this development has been a necessary step towards a more realistic depiction of the economic agent, it suffers from the drawback that in many models the individual has been fitted with an exogenously given *social identity* that simply shapes and influences the agent's behaviour without his being directly involved in any way. Thus, endowing the economic agent with social preferences has the disadvantage that it induces the agent to act - under certain circumstances - against his "personal preferences" without leaving him any choice. Our aim in this paper is to present a model that accounts for endogenous identity formation and which allows for the individual behaving as a member of a social community, but also gives him a choice as to which community to belong to.

Underlying our model is the observation that as the economic agent has become socially delineated in economics, he has lost his *sovereignty* - yet the latter is a fundamental assumption of modern economics. That is, he became social but did not simultaneously develop the capacity to reflect upon and evaluate how acceptable to him were social influences on his behaviour. Thus, he is not sovereign with respect to the evaluations that he can make. However, he is also not sovereign in the sense of preferring whatever *he* thinks proper for himself as an individual. The social preferences his choices reflect might be in direct conflict with what *he* prefers as an individual. This is conceptually inconvenient and troublesome for at least two reasons. First, even if an individual has a social identity, which influences his behaviour, he is not totally defined by it. Indeed, the individual has the capacity to change and thus to shape and to influence his social identity. That is, an individual is a specific human being with a *personal* identity that is represented by his desire to be and to become *who* he wants to be. Social identity is just one aspect of this individual's personal identity. Having a social identity does not necessarily preclude an individual from choosing another social identity. To do this he must be able to think and to reflect on whether his social and cultural background or his position within society is the most adapted to his personal endeavours. Thus, in order to restore the economic agent's sovereignty, we must develop a model of behaviour that depicts an individual who is embedded in society but is still capable of *choosing* to change, as the result of his own critical reflection. Second, a given social identity that influences and shapes an individual's behaviour will make welfare evaluations more problematical if not impossible. Once again, if we do not have a way of distinguishing between an individual's personal preference and a preference that his social identity has led him to have, we can never be sure if the individual's preferences or the society's preferences as implanted *in the individual* will be satisfied. Thus, making sensible welfare judgements would seem to involve making a distinction between an individual's chosen or premeditated social preference and an individual's unquestioned social preference that he *assumed* for his social identity as a result of his place in society. Without this it seems difficult to say whether an individual is really better off or not and this undermines our capacity to make any judgements about social welfare.

In this paper, we present a model that attempts to give a direct response to the concern about the agents' role in shaping his social identity. Indeed, instead of presenting an individual as having an exogenously given social identity or some presupposed social preferences, our model incorporates endogenous social identity formation and an evolution of the latter over time both an an individual as well as a social level. Individuals, represented by choice probabilities that depend on personal characteristics, will choose to belong to specific social groups ("types"), which have certain characteristics. The choice of the social group will modify the characteristics of the individuals, for example an individual may try to adapt his characteristics to those of his group. However, when several individuals choose a particular type this will in turn, modify the type's

characteristics over time, and this change may motivate individuals to revise their ideas and to move on and choose yet again, different types. The model shows that this society that consists of individuals and types will, under certain conditions, reach a steady-state equilibrium with a fixed number of types with established characteristics. Moreover the individuals continues to change types but chooses each type with a given probability which no longer changes over time. This is true, if the individuals' choice of types only weakly depends on their current personal and type-characteristics and the overall availability of types in the society. This means that a society will be composed of a given set of social identities if the personal identity of individuals over time does not matter too much, i.e. if individuals' are sufficiently malleable to move from one social group to another. This result is both intuitive and counterintuitive. It indicates that social groups crystallise more strongly and are more stable the weaker the individual's personal identity over time. This means that agents adapt to a new social identity each time they move from one group to the other. This is in contradiction to other models in economics, which argue that society stabilises the more the individual agents are defined and characterised by rigid preferences. Indeed, this is the problem of exogenously given social identities, which by the logic of the individual's choice behaviour will be chosen if the characteristics of these identities match with the individual's preferences. However, the latter are influenced by that selfsame social identity. Our model breaks this circularity by introducing a dynamic of change of individual as well as of social group characteristics. There is no reason to assume that individuals and groups always remain the same over time. Individuals as well as society change over time and this dynamic of change needs to be captured by economic models if we wish to give a realistic account of the evolution of economic identity. Thus, sensible welfare evaluation needs to make a distinction between an individual's chosen or premeditated social preference and an individual's unquestioned social preference that he *adapted* to his social identity if this analysis will be able to say whether the individual reached his highest possible welfare level or not.

In the next section (2.), we will present the core arguments behind our model. In order to do this, we proceed by introducing three different concepts of identity, namely the *what*-, *where*-, and *who*-identity of the economic agent (Kirman and Teschl 2004) and these will provide the basis of our model together with a *measure of continuity* of an agent over time. Whereas the *what*-identity refers to the representation of the standard economic agent, the *where*-identity models of the economic agent have been developed over recent years and incorporate either directly or indirectly the idea that the economic agent is a social being who occupies a place in a social space. We will review some of the models that account for a variety of individuals' social preferences and point out that while these models paint a more realistic picture of the economic agent, they do not, consider the individual's and society's capacity for change and thus endogenous identity formation. Finally we will introduce the idea of the *who*-identity by considering psychological and

philosophical accounts of personal identity. Armed with this information, we will lay down the groundwork for the model that will follow in the subsequent section (3.). The last section (4.) will conclude and outline some of the possible avenues for future research on the basis of endogenous identity formation.

2 The What, Where and Who of Personal Identity

The standard economic agent, in particular as he is formally represented in general equilibrium models, is typically characterised as exclusively concerned about her personal and self-interested well-being. In addition to this hypothesis about his motivational, this Homo Economicus was considered to be an atomistic and socially isolated being, and this is derived from the underlying assumption that the agent's action did not have any influence on other people's behaviour and vice versa. If he did have social interactions they were beyond the control of the individual and taken as given. Homo Economicus is represented by a stable preference ordering over consumption bundles or consumption paths, and exogenously given budget-constraints¹. The agent then simply chooses the best attainable bundle. This characterisation explained *what* the economic agent is, namely a rational utility maximiser, who changed her behaviour only in order to adapt her choices to a modified price or income level that affected her budget-constraints. We call this representation of the individual and the characterisation of behaviour the *what*-identity of the economic agent. It is a very simplified description of the economic agent, especially insofar as it relies on the rather audacious claim that an individual has a preference ordering over goods and actions for the present as well as for all future periods. Given this *once-and-for-all* preference ordering, she then *reacts*, like an automaton, to changes in constraints and no revaluation or changes in preferences are thought to occur. However, underlying the given preference ordering is the idea that the individual was capable of specifying what is best for herself. That implies a consciously reflecting individual, commonly considered to be a rational and *sovereign* consumer, who is able to evaluate different choices and to order them. This parsimonious model has been constantly amended and adapted to different choice situations that go far beyond pure market behaviour.

Indeed, in the last few decades, more and more economists rallied around the idea that “man is a social animal and his choices are not rigidly bound to his own preferences only. An act of choice for this social animal is, in a fundamental sense, always a social act” (Sen 1973, pp. 252-3). This was a general recognition that economics could not abstract from the social environment in which production, exchange and distribution would take place. The respective social background was recognised to have specific influences on the individual and the individual was thought to affect,

¹In equilibrium, constraints are endogenous. However, exogenous refers here to the idea that the agent herself does not have any influence on her constraints (price-taker assumption).

even involuntarily, other people's behaviour and beliefs through his choices. Individual action was not any longer externality-free. The economic agent thus left his conceptual isolation and his preferences were no longer solely and exclusively self-concerned. The economic agent was thought of as caring about other people, about social and behavioural norms and rules, about fairness, his status, his personal characteristics and his distance to others in terms of these characteristics in a social space. Thus, information about *where* the economic agent is and in which social context he lives has been introduced into economic reasoning. We call this the *where*-identity of the economic agent.

In most cases, economic agent's *where*-identity has been incorporated into a *what*-identity model. That is, the mechanism of choice, namely constrained utility maximisation, has been maintained and different aspects of an individual's social preferences and his dependency on other people's actions have been introduced as arguments into the utility function. A social interaction model such as Akerlof's (1997) *Social Distance and Social Decision* for example, presupposes that individuals occupy different locations in a social space that they partly inherited and partly acquired through their choices. They thus have, from the outset a given *where*-identity. Individuals maximise their utility by choosing an action that affects their intrinsic preferences as well as their social position, i.e. *where* they want to be, by "interacting" with other individuals who equally occupy different social locations. Assuming that the benefits of interaction increase when interacting with people who are close to the individuals' position and that people form static expectations about the future position of their partners, the model shows that choices to interact with those who are socially close might outweigh choices according to individuals' intrinsic preferences. This eventually leads, according to standard criteria, to non-optimal economic outcomes, as individuals do not attain the welfare maximising equilibrium they would achieve if they focused solely on their individual preferences. "Interacting" with those who are socially close might also have negative consequences. However, this does not preclude social interaction with them. Indeed, as Akerlof states: "The jealousy and envy of friends, relatives, and neighbors [...] result in the same incentives to keep close to one's origins as [...] positive benefits" (p. 1011). This also explains class stability or the existence of subcultures.

The conceptual idea of social distance is also kept in Akerlof and Kranton's series of paper on *Economics and Identity* (2000, 2002, 2005). They introduce an individual's *identity*-function into a standard utility function. This identity function represents an individual's social identity and epitomises the psychological idea that people act in order to guard or to achieve a *sense of self*. Akerlof and Kranton characterise identity as the social categories to which an individual belongs. This means that the individual is already provided with an explicit *where*-identity. These categories have behavioural prescriptions and each individual has a particular distance from these prescriptions depending on her own characteristics. The individual will choose effort levels or

actions that fit and enhance the salience of her identity and thus sense of self. This means that the individual attempts to minimise the distance between her own characteristics and a socially determined self-image to which she adheres, in order to attain her highest possible utility. Akerlof and Kranton indicate that the choice of identity might be one of the most important decisions, but note that “identity ‘choice’, however, is very often limited” (2000, p. 726). That is, social categories are given and it is this, unexplained, feature that, provides society with a stable structure. Indeed, the choice of a category to which one does not naturally belong might violate other people’s sense of self and incite these people to undertake punitive actions that are aimed at restoring their social self-image derived from their internalised identity-rules. The expectancy of punitive actions might thus prevent individuals from choosing categories outside their own.

Other models, such as Becker and Murphy (2000) for example, consider how change is possible if a person has a particular place within the social structure. Their models shows that it is very difficult for individuals to increase the distance from their social group. Becker and Murphy introduce, as well as different goods *social capital* into an individual’s utility function. Social capital represents social influences on utility and is supposed to be complementary to certain goods in the utility function. This complementarity is a key element in explaining “that social forces have a tyrannical power over individual behaviour” (p. 9). Assuming that the stock of social capital is the average of a specific good chosen by all members of the same group, the complementarity of the social capital with this specific good makes it clear that a change in an individual’s idiosyncratic variable will not particularly modify this individual’s behaviour concerning this good as long as the complementary social capital did not change either. Equally, when a variable that is common to all members of a given group changes, a single individual action will not alter social capital. Only the sum of different individual choices will affect social capital through the mechanism of social conformity and social interaction.

Conformity is, to a large extent, built into the above models, often because the individual has from the outset a specified *where*-identity. Other models, however, show that individuals end up with a specific *where*-identity influenced by their specific personal and social preference structure and thus without presupposing any particular place in society yet. Bernheim (1994) for example introduces the idea that individuals care about social status or esteem in addition to their intrinsic preferences for standard consumption. This specific preference will, under certain conditions, give rise to conformity. Status should be conferred by a person’s motives or virtues. However, because individual proclivities cannot be observed directly, individuals signal to other members of a society their motives by choosing those actions that they think will convey a positive image. Conformity emerges if individuals put a sufficiently high weight on social status. They are then willing to suppress their individual preferences in order to adhere to a single homogenous norm of behaviour because they know that even small departures of the social norm will damage

their status.

Where they are has thus been influenced by their individual preferences for social esteem. Whereas this result only considers the case where all individuals value status and esteem equally, the model can also be adapted to the case where individuals care about the opinion of different population subgroups. In that case, several norms arise to which individuals adhere and thus lay the basis for subcultures. Norm following in general will be more manifest in those societies, where esteem and social status are more important. Bernheim notes that there is therefore scope for public policies: “If we assume that the government can, through publicity, make one value [...] more focal than another, then policy makers may be able to affect real decisions without traditional forms of economic intervention” (p. 865). The *where*-identity thus turns economists into social engineers even without economic tools. This, however, can also have unexpected consequences as Oxoby (2004) shows in a model that explains *Cognitive Dissonance, Status and Growth of the Underclass*.

The underlying idea is that individuals experience some psychological uneasiness (or cognitive dissonance) if they are unable to achieve what a social norm or status imposes (e.g. poor people cannot achieve the status associated with high consumption). That is, status-seeking individuals are concerned as to how they stand in relation to the average population regarding a specific norm. If they fare much worse than others and if the relative value of their status in overall utility is high, they suffer psychological pain in terms of envy or disappointment. In order to reduce it, they engage in (costly) dissonance reduction by changing the social norm to which they adhere and thus modifying what they think to be status worthy. The underclass is thus created: poor people abandon the norm of attaining high consumption by reducing their effort and favouring the status derived from leading a leisurely and relaxed life (which brings them to the edge of criminality and welfare dependency). This model thus shows that making some norms more focal might indeed have negative consequences for the social welfare.

Where-identity models thus recognise the importance of individual as well as social characteristics and consider the interaction between people as a major ingredient of behaviour. It also becomes clear that including social aspects in the standard *what*-identity model complicates welfare assessment on an individual as well as social level. More specifically, it opens the perspective that not only might individual behaviour impair social welfare (as traditional economics would readily accept), but also that social behaviour may harm individual welfare. The latter aspect is, however, a controversial claim, since *where*-identity models still start from the assumption that it is the individual who has these social preferences. And if he does not reach the social optimal position, it is the externality of other people’s behaviour that constrains in most cases the economic agent’s welfare. The very existence of well defined preferences, has, however, not been questioned. Indeed, Akerlof and Kranton (2000) do not presume, for example, that people are

aware of their motivations and argue that “Whether or not [the individual] consciously realizes she is maximizing a utility function [that contains an identity function], she does so nevertheless” (p. 719, FN 6). Thus, one cannot even claim that reflection about the nature and origin of a preferences happened at the moment of forming one’s utility function. This is certainly contrary to the idea of Homo Economicus being a *sovereign* individual. To restore the idea of sovereignty, it seems to be important to know if an agent’s preferences are explicitly this individual’s preferences given his social identity or if these are preferences that this individual is *made to have* because of his social identity. Assuming that the individual is not aware of motivations, as Akerlof and Kranton do, excludes any such distinction.

This calls for what Amartya Sen describes as *Reason Before Identity* (1999). His critique of the unthinking use of social identity can readily be extended to the adherence of any social norm and rule or “myopic” status-seeking that influences human behaviour. “In any explanatory exercise,” he writes, “note has to be taken of local knowledge, regional norms, and particular perceptions and values that are common in a specific community” (2004, p. 91). However, this should not preclude any choice and reasoning about identity. Indeed, even if cultural beliefs and attitudes influence a person’s behaviour and the nature of reasoning, it will never determine them fully. In fact, culture and respectively communities or social groups do not have one single set of defined attitudes, beliefs and norms. There is enough internal variation and difference such that the individual has not to adhere to a unique set of norms but can choose between different ones, even within the same group. A person, furthermore, can always be seen as belonging to several groups and social categories and it is up to the individual to recognise her multiple membership and to decide if she wants to adhere to them, how much weight to put on them, or even if she wants to leave particular groups. A person’s identity is thus, as Sen would say, a matter of choice. Yet, as we will argue, this identity does not only depend on choice of the person’s membership to different groups. Identity is also defined through those characteristics that make this choice possible. That is, choice depends on the individual’s own characteristics, preferences and desires and the reflective relationship she engages with her surrounding according to *who* she wants to be. We will call this aspect the *who*-identity of the economic agent. This *who*-identity is inspired by the idea that an individual has not only a social identity, but first and foremost a personal identity that the individual shapes and creates throughout her life. This means that a person acknowledges the fact of her own change over time. It also means that there is a reciprocal relationship between the circumstances that influence individual decisions and individual choices that will change those circumstances in order to achieve *who* one wants to be.

Personal identity and a person’s *self* are fundamental concepts in psychology and widely discussed in philosophy. Indeed, there is a long tradition in psychology assuming that a person has different facets of the self or different self-images. Psychologists such as James (1890), Cooley

(1902), Freud (1923), Mead (1934), etc. famously distinguished between aspects such as the “spiritual self” and the “social self”, or the “ego-id-superego” indicating that a person with personal aims and goals occupies not only a place in a social space, but also in moral and spiritual “spaces”, which affect a person’s aspirations as well as images of herself. Of course, occupying these different spaces not only guides a person’s action but also involves the risk that the respective aspirations enter in conflict with each other, i.e. that the various images of the self are at a certain *distance* from each other and are not coherent. This causes psychological discomfort (or cognitive dissonance), which again reflect back on people’s actions.

Modern versions of these “multiple selves” visions of the human being recall the idea of *self-discrepancy* (Higgins 1987) and indicate that a person possesses different domains of the self, namely the *actual* self, the *ideal* self and the *ought* self. These selves represent respectively the attributes someone possesses, would like to possess or should possess, either seen and evaluated from the point of view 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 but 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 some actions that restore coherence between these selves. Hence, different selves or self-images serve as *regulatory references*, with the intention 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 about themselves and have the possibility as well as the intention to change over time (Cantor et al. 1986). They work on certain *life-tasks* based on the knowledge of their preferences, abilities, acknowledgement of social roles and other personal characteristics. These life-tasks change over time and so will 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 their imagination of their own future possible selves, which are cognitive representations giving a personalised vision of their overall motives. These *possible selves* thus represent the potential of change of a person and visualises *who* they could be and would like to be.

However, while psychologists explore the cognitive representations and interactions of different selves that contribute to a person’s sense of identity and thus motivate this person to change his actions or beliefs, philosophers have concentrated on the nature and essence of personal identity especially on the problem that a person changes over time, but his identity is claimed to be the same. The philosophical question of personal identity can be split in different, partially

overlapping questions². It considers, for example, *what* it is to be a person. It furthermore asks what it takes for this person to persist. Philosophers were thus looking for an identity-criterion according to which a person at one time can be considered to be the same person at another time. Personal identity is also seen as representing and characterising what makes a person *who* he is. There is always something unique about a person. This it is often claimed, stems, at least in part, from the fact that he has a capacity of choice and can, at least marginally, influence the course of his life.

Thus, the philosophical account of personal identity attempts to provide models which reconcile persistence and change of a person over time, including the possibility of the person changing himself³ There are several difficulties to overcome. First, if the person's identity were defined in

²See e.g. Olson 2002, Oksenberg Rorty 1976. For overviews on different philosophical accounts of personal identity see Ferret 1998, Glover 1988, Noonan 2003.

³The questions of what it is for something to be a person, as well as of her persistence through time have sometimes been matched together. However, whereas fingerprints or a genetical code can be considered as criteria that establish if a person at one moment of time is identical with a person at a later moment of time, it can hardly be claimed that fingerprints or the genetical code are what makes a person a person. Thus, in order to be able to say something specific about a person over time, there must be some way of recognising this person as the selfsame person over time.

Evidently, a person has a body and some specific psychological or mental states (such as self-images) and memories that will change over time. But it is because they change that it seems to be difficult to consider them to be the criteria that help to recognise a person as the same person. One might thus say that a person is recognisable as the same person because she has a specific personality or character. But while saying that personality is an intrinsic attribute of what it is to be a person, it can hardly be considered as a criterion that stands for the persistence of a person through time, other than by saying that personality is some unchanging substance that remains the same throughout time and changes of that person. This, however, would imply that as soon as the personality of a person even marginally changes, this person ceases to exist and another person comes into existence. It would also contradict the idea that a person is able to change herself: for, as soon as the person (defined by her personality) changes her personality, she will become discontinuous with herself. One solution to these problems is to say that personal identity is only about numerical identity and not about qualitative identity. This means that a person is numerically the same if she remains one and only one individual over time. A person would be qualitatively the same if she would share at each moment the exactly same properties (except location in space and time). While in the first case that person remains the same even though she qualitatively changes, this is not assured in the second case. However, the numerical identity of a person still has to be defined in terms of some criterion. If A's memories are, say, transferred onto B's brain and vice versa, the question is if A is numerically identical with her memory contents or with her body. Furthermore, if this criterion can be changed, the question is how much change is allowed to say that a person is still the same person. If Derek Parfit is Derek Parfit because of his memories, but then undergoes an operation that successively replaces his memories by Greta Garbo's memories, when does Parfit cease to be himself and be Garbo instead? In this case it seems that personal identity is a question of degree and there will never be a precise answer. The line where Parfit ceases to exist can be drawn arbitrarily and may be a matter of convention.

This is far from a new preoccupation in economics, since when Sidgwick (1907 pp. 418-419) was discussing utilitarianism he observed, "If the Utilitarian has to answer the question, "why should I sacrifice my own happiness

terms of an identity-criterion that remains qualitatively the same throughout time, the person would, strictly speaking, lose his identity as soon as this criterion changes. We thus have to drop this strict identity relation and allow the criterion to change. However, if the criterion changes, we need at least some more or less enduring or suitably connected successive representations of that criterion in order to be able to talk of the same individual through time (Parfit 1984). Yet, we need more than simple connectedness if we wish to account for the individual's possibility to change himself. One way of doing this is to decouple personal identity into at least two identity criteria or functions one of which remains the same but is nevertheless responsible for the change of the other. (Livet 2004).

Applying the psychology and philosophy of personal identity to the economic agent by specifying *who* he is thus means not only explaining say *what* the economic agent is and *where* he is situated in a social and moral space, but also endowing the economic agent with a capacity for change. Thus introducing the *who*-identity of an economic agent means accounting for continuity and change of this agent through time.

for the greater happiness of another?" it must surely be admissible to ask the Egoist, "Why should I sacrifice a present pleasure for one in the future? Why should I concern myself about my own future feelings any more than about the feelings of other persons?" It undoubtedly seems to Common Sense paradoxical to ask for a reason why one should seek one's own happiness on the whole; but I do not see how the demand can be repudiated as absurd by those who adopt views of the extreme empirical school of psychologists, although those views are commonly supposed to have a close affinity with Egoistic Hedonism. Grant that the Ego is merely a system of coherent phenomena, that the permanent identical "I" is not a fact but a fiction, as Hume and his followers maintain; why, then, should one part of the series of feelings into which the Ego is resolved be concerned with another part of the same series, any more than with other series?"

Derek Parfit (1984) was led to the conclusion that the question about personal identity is not so much what identity *is*, but what *matters* in identity. The idea is that there must be enough psychological connectedness between two successive moments of time in order to say that A is identical with B (or survives in B if B's body is different from that of A). An overlapping chain of psychologically connected time-slices of some human carrier guarantees continuity of a person even if the psychological features of that person at the beginning of the chain are not identical anymore with those at the end of the chain. A different solution to the above questions is to say that one criterion of identity cannot be responsible for both aspects of personal identity, i.e. persistence and change (Livet 2004). Instead, the interplay of four different identity-functions (a person's body, her memories, preferences and social status), each in itself too weak to assure identity of the person, contribute together to the "functioning" and persistence of personal identity over time. Moreover, whereas the body and memory assure that the individual is, in some concrete sense, the same over time, preferences and social status guarantee the possibility of change, including of the person changing herself. If the consequences of some actions that were carried out on the basis of the person's social status enters in conflict with her personal preferences, the person will revise the priorities of her social status from the point of view of her personal preferences and vice versa.

While one identity-function is in the process of being changed, the other remains constant and it is this that guarantees persistence. Personal identity over time then is the overlapping chain of changing and persisting identity-functions.

The pure *what*-identity of the economic agent specifies that the individual is a rational utility maximiser with a given preference ordering. If we thus take this unchanging preference ordering as the criterion of personal identity, we trivially assure this individual's continuity but exclude any change other than by external constraints. However, if we think that an individual is not only determined by *what* he is, we have to introduce the possibility for the individual to choose to change. Thus, to account for an economic agent's identity, we need more than an unchanging preference ordering. If this ordering alone would represent his identity, then as soon as it changes, he would lose the basis of his identity.

The *where*-identity of the economic agent accounted for the fact that the individual has not only self-interested preferences, but also social preferences. On the one hand, these preferences are influenced by the position in the social space that he occupies, on the other hand the individual has preferences about the place in the social space that he wants to occupy. We will consider that place as being an identification with one of a number of possible social groups each with its own characteristics. A standard argument is that, although the individual is guided by two different kinds of preferences, they can both be included in the same utility function. But this simply suggests that there is a trade-off between the two and a more critical reflection about social influences in terms of *reason before identity* is excluded. Nor, is the possibility of a considered change of identity envisaged. The *who*-identity of the economic agent is influenced by the idea that underlying the individual is a dynamic of change. In the *who*-identity model of the economic agent that we will present in the next section, preferences are grounded in the agent's characteristics and because his characteristics change over time, so will his preferences. These preferences are represented as choice probabilities (p_{x_t}), which thus stand for *what* the economic agent is at each moment of time. Hence, we call choice probabilities the *what-criterion* of the economic agent's identity. Furthermore, the economic agent is living and acting in a social world, thus occupying a place in a social space. However, he will choose *where* this place should be by selecting his participation in certain social groups according to his preferences. Social groups themselves have specific characteristics and it is in fact, not the group that the economic agent is choosing, but the characteristics of the group. One vision which is discussed by Akerlof and Kranton, is that an individual can use the characteristics of the group as a means to reduce the distance between the agent's current characteristics and those of some self-image that the agent wants to realise. The characteristics of the social groups (or *types* as we call them) are the *where-criterion* of the economic agent's identity (x_t). The choice of social groups will thus change the economic agent since he will acquire new characteristics. However, the relationship is reciprocal: the agent's characteristics that he brings to the social group will affect the characteristics of the social group. Moreover, supposing that there are several individuals with different characteristics who choose the same social group, this social group will be changed by the choices of these individuals. This

change in characteristics though, will motivate the agent, once again, to reevaluate his previous choice of that social group. Hence the individual will make a new choice whereby he decides on the basis of his preferences either to stay in the group or “to move on” and to choose to participate at a different group. Thus, the idea of a dynamic change of the individual through time is matched with the idea of a dynamic change in the social structure. The probability of having certain characteristics and types and ending up with new characteristics in new types is the *who-criterion* of the economic agent’s identity and it represents the idea of change through time, based on both, the individual’ choices and the social change (π_{x_t}). It symbolizes the possibility to change oneself through one’s own choices. Evidently, given the stochastic nature of the agent’s preferences, the *who-criterion* gives a picture of the possible selves that the agent may become given his current characteristics and group-memberships. However, while the *who-criterion* indicates the possible change of the agent from one time-period to the next, or, stated differently, his preference for the future, it does not yet tell us anything about the different connections between the agent’s changes over time. For this, we need to introduce a *measure of continuity* (β) of the agent’s personal identity. This measure consists of the sum of two different measures (β_1 and β_2), one of which concerns the agent’s connectedness to his past, the other his responsiveness to varying social opportunities. They can also be seen as a measure of personal connectedness and of social connectedness of the agent over time respectively. The former (β_1) measures the agent’s dependence on his new group-memberships (or types) and characteristics on his current characteristics and group-memberships (or types). The latter (β_2) measures the dependence of the agent’s future characteristics and types on currently available groups in a given society. A high connectedness to the past means that the agent’s new characteristics and types depend strongly on his current characteristics and types. That indicates that *where* the agent is coming from and *who* he was in the past influences his future changes. The agent can thus be traced through time because one state of characteristics and types leads to the other. A high responsiveness to varying social opportunities means that the agent’s new characteristics and types depend strongly on which social groups are currently available. In this case, the agent can be followed through time according to the changes in his social environment to which he strongly reacts. Thus, eventually, personal identity is represented by the *who-criterion* that can be decoupled into two criteria, the *what* and the *where-criterion* that in turn constitute the economic agent’s change over time, and by his *measure of continuity* that consists of his personal and social connectedness over time. In this sense, continuity and change contribute together to give an account of *who* the agent is.

What are the consequences and results of these individual and social dynamics? In what sense and under what conditions does the structure of the society stabilise? Our model shows that society will reach a stationary state if two ”weak interaction” conditions are satisfied. Society

stabilises if the path-dependency of an individual's successive states is very weak (small β_1). and if the agent's choice of future state does not depend too much on changes in the types available. (small β_2). In a stationary state the characteristics of the social groups and the proportions of individuals choosing them converge. This means that on a macroscopic level, the society no longer evolves. On a microscopic level though, individuals will continue moving from one group to the other and thus changing their characteristics. In the limit, individuals will move from one group to the other with the same probabilities. One might almost say that the agent's preference for a future state exists almost independently of the agent himself.

The result is surprising and stands in contrast to many models developed previously *where-identity* models. Indeed, in those models, society stabilise because individuals have a well defined identity in terms of their preferences over personal and social outcomes, which are unchanging. Furthermore, society stabilises even when individuals are very sensitive to social influences. In our model however society stabilises because individuals are relatively flexible, that is they are not too sensitive to *who* they are not to changes in the available groups. It can happen, as we will see in the next section, that a situation of multiple long run outcomes can occur, even if a certain degree of connectedness, or continuity, of a person is allowed. Such outcomes are dependent on initial conditions. Such difficulties arise, insofar as individuals have the capacity to choose and to influence the course of their life and in so doing modify the social structure that allows them to become *who* they want to be. In that sense, the most interesting society to live in is the one which is not stable.

3 An example

Why does our model involve stochastic rather than deterministic choices? It is clear that this is important for evaluating the evolution of the society or economy in which the agents live but we need a behavioral justification for this assumption. In the first place, individual choices will be conditioned by many factors which cannot all be incorporated into the model and the idea of 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.

An advantage of this approach is that it allows for some remarkable personal changes to take place. One finds occasionally individuals who 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 constitute low probability events but which, inevitably, happen from time to time. 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 went on to become a distinguished barrister and a militant human rights advocate. Her career

would be thought of as a choice in our context albeit an unlikely one. Indeed Amartya Sen (2001) 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.

We will observe this sort of switch and it will have an important effect on the evolution of the individual but, also, on the evolution of the social groups that make up society.

4 Searching for Identity

We now present the basic characteristics of our model. Firstly we specify the set of possible characteristics that an individual agent can have. We consider the characteristics of any individual as being represented by a vector in \mathbb{R}^m , though we might want, in general, to consider some more elaborate space. In particular, we think of the possible vectors of characteristics as being finite in number. Furthermore social types are also represented by vectors in the same characteristics space. One could think of each "social type" as having a "representative" whose characteristics correspond to the image of that type. At each point in time, the agent will have certain characteristics and will have chosen a type. The choice of the type constitutes the answer to the where criterion for the individual. The types themselves will, in our model, change depending on who has chosen them. Just as the image of the Republican party is modified as the people who declare themselves to be members or to vote Republican change. Furthermore the agents own characteristics will be modified over time, in part, for example, as the result of his choice of social type.

More formally: Let $C = \{c^1, \dots, c^n\}$ be a finite set of *characteristics* and $X \subset \mathbb{R}^m$ be a compact convex set of *types*, $C \subset X$. We denote by

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

the vector of possible types and the configuration of characteristics at date $t \in \mathbb{N}$, respectively. As we have said, this vector represents the agent's *where-criterion* and stands for the different *where-identities* from which the agent can choose.

We thus have three reference-points: what the person is and does, what the person can choose to become (if she converts her choices of specific characteristics into functionings), and what the person would value to be and to do. Sen referred to the first aspect as a persons functionings, to the third aspect as a persons capabilities, and to the second aspect as the instrumental freedoms a society offers its individuals to foster capabilities. In our own interpretation, we see the first aspect as an objective evaluation of a persons functionings, the second aspect as the capabilities, i.e. the freedoms to be choose to select those characteristics that would bring the person closer to what she would value to be and to do, and the third aspect as those functionings a person

would (subjectively) like to become. More precisely, the conditional distribution of agent a 's type in period t takes the form

$$p_{x_t}(c_t^a; \cdot)$$

for a transition probability p from $C \times X$ to X where $p_{x_t}(c_t^a; \cdot)$ is concentrated on the set $X_t := \{x_t^1, \dots, x_t^m\}$. This choice probability depends on who the agent is now, that is, the *what-criterion* of the agent's identity and represents the "preferences" of the individual over social types.

Now, at each point in time the social types evolve as a result of the choices of the individuals. For technical reasons this is inconvenient since the set of types at each point is not the same. To avoid this we can think of individuals as choosing from m "labels", even if what those labels represent is changing. Thus, the Republican party's characteristics change but agents can still choose to vote Republican.

Hence, we will label the types by $1, 2, \dots, m$, and rather than choosing types, we assume that the agents choose a label $y_t^a \in Y := \{1, 2, \dots, m\}$ with probability

$$\hat{p}_{x_t}(c_t^a; y_{t+1}^a) = p_{x_t}(c_t^a; x_t^{y_{t+1}^a}).$$

After all the agents have selected their respective types or labels, nature chooses their new characteristics stochastically. Thus the evolution of the agents is stochastic and not deterministic. The choice of characteristics is made according to the probability measure

$$\tilde{p}_{x_t}(c_t^a, y_{t+1}^a; \cdot)$$

where \tilde{p} is a stochastic kernel from $C \times Y$ to C . Thus we can write down the probabilities of the new characteristics given the old characteristics and vector of types. In particular, the conditional distribution of the random variables c_{t+1}^a given c_t^a and also the current vector of types x_t is given by

$$\hat{\pi}_{x_t}(c_t^a; c_{t+1}^a) = \sum_{y=1}^m p_{x_t}(c_t^a; y) \tilde{p}_{x_t}(c_t^a, y; c_{t+1}^a).$$

The distribution of an agent's new characteristics and type depends on his current one. Specifically, the joint conditional distribution of the random pair $z_{t+1}^a = (c_{t+1}^a, y_{t+1}^a)$ takes the form

$$\pi_{x_t}(z_t^a; z_{t+1}^a) = p_{x_t}(z_t^a; y_{t+1}^a) \tilde{p}_{x_t}(z_t^a, y_{t+1}^a; c_{t+1}^a).$$

The transition probability π from $C \times Y \times X$ to $C \times Y$ governs the dynamics of individual characteristics and types. This captures what we call the *who-criterion* of an agent's identity. It represents the idea of how an agent changes and the agent's preference for change. An agent changes both because his own characteristics, his *what-criterion* and - as we will see shortly - the characteristics of his group, his *where-criterion*, will be modified over time. Because of this,

an individual always has the possibility to revise his choices and can consequently choose either to stay with the same type or to “move on” and to identify with a different type. Notice that, if we want to obtain specific results concerning the evolution of types we will have to put a lot of structure on the transition probabilities and then to justify that structure in economic terms. However, our overall aim here is to see under what general conditions the system has some sort of stability at the aggregate and individual level. It is nevertheless useful at this point to introduce some examples of how the probabilities of individual moves might be generated.

Example 4.1 *Let us first consider a model of group choice where the utility function given the current vector (x_t, c_t, z_t) takes the form*

$$U(x_t, c_t, z_t, c, z, y) = u(c, z, x_t^y) - J_1|c_t - c| - J_2|z_t - z| + \epsilon(c, z, y). \quad (1)$$

Here y denotes the group label; $u(c, z, x_t^y)$ is an instantaneous utility associated with the current choice and J_1 and J_2 specify the dependence of an agent’s utility from its current characteristics and self-images. The random shocks $\epsilon(c, z, y)$ are independent and doubly exponentially distributed with parameter $\beta > 0$:

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

They are observable to the agent prior to making an action. Upon observation of the realizations the agent chooses the vector (c, z, y) that maximizes her utility. As a result, the conditional joint distribution of types (labels), characteristics and self-images in the following period given (x_t, c_t, z_t) is

$$\pi_{x_t}(c_t, z_t; c, z, y) = \frac{\exp(\beta u(c, z, x_t^y) - J_1|c_t - c| - J_2|z_t - z|)}{\sum_{\hat{c}, \hat{z}, \hat{y}} \exp(\beta(u(c, z, x_t^{\hat{y}}) - J_1|c_t - \hat{c}| - J_2|z_t - \hat{z}|))} \quad (2)$$

Example 4.2 *Let us now assume that the agents only choice variables are the types and self images. Once a type is chosen, nature picks the new characteristics with a distribution μ that depends on the chosen type and self image. Specifically, let us assume that the agents expected utility from choosing the label y at time t is given by*

$$V(x_t, c_t, z_t, z, y) = \int_C U(x_t, c_t, z_t, c, z, y) \mu(y, y; dc_{t+1})$$

where U is the utility function (1). The random shocks depend on the choice variables (z, y) . If we retain the assumption that their realizations are observable prior to taking an action, the linear structure of the utilities implies that the choice probabilities take the form (2) if we replace the quantities $u(c, z, x_t^y)$ and $J_1|c_t - c|$ by their expected values with respect to μ .

Before proceeding it is worth noting that in what follows we will confine ourselves to a situation in which there are only social identities or groups and individual characteristics. Thus agents at

each point choose social identities and move their characteristics and but the self image is no longer explicit.

To illustrate our basic ideas we ran some simulations in which agents choose their social types with probabilities that depend on the distance of the types from their own characteristics. They move a distance α towards the group that they choose,

$$c_{t+1}^a = c_t^a + \alpha(x_t^{y_{t+1}^a} - c_t^a),$$

and they choose the groups with probabilities that are a linear function of the distances from the social types. See figures 1,2,3 and 4. Of course, to be strictly consistent with our model, agents' characteristics should become the point in characteristics space nearest to the one which is chosen by the rule. Groups' characteristics are then modified as a linear combination of the previous type and the average of those who identify with the group.

In the figures what is varied is the initial "seeding" of the agents' choices and the presence, or not, of an individual with more weight than the others and moreover, one who is not influenced by her social group's characteristics. The presence of such an influential and stubborn individual can, as can be seen lead to the persistence of differences between social groups as in figure 4.

Now we return to our model and assume for the moment, as in the simulations, that the agents act conditionally independently of each other, given the individual characteristics and the current vector of types. This means that for any given sequence of types $\{x_t\}_{t \in \mathbb{N}}$ the dynamics of the process $\{z_t\}_{t \in \mathbb{N}}$ is specified in terms of the product kernel

$$\Pi_{x_t}(z_t; \cdot) := \prod_{a=1}^{\infty} \pi_{x_t}(z_t^a; \cdot).$$

In order to study the asymptotics of both individual and aggregate behavior, we denote by \mathcal{M} the class of all probability measures on $C \times Y$ and by

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

the *empirical distribution* associated to the configuration z_t provided the limit exists in the weak sense, i.e., for any bounded function $f : C \times Y \rightarrow \mathbb{R}$, it exists

$$\int f d\varrho_t := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{a=1}^n f(z_t^a).$$

Remark 4.3 *Since C and Y are finite sets, ϱ_t may be viewed as a vector $\varrho_t = (\varrho_t^{i,j})$ ($i = 1, 2, \dots, n$, $j = 1, 2, \dots, m$) in \mathbb{R}^{nm} where $\varrho^{i,j}$ denotes the fraction of agents that choose type x_t^i and characteristic c^j .*

The long run dynamics of aggregate behavior is described by the asymptotics of the sequence $\{\varrho_t\}$. What we are interested in is whether the types settle down and also whether the proportions of individuals choosing those types stabilizes over time.

By the law of large numbers of independent random variables,

$$\varrho_{t+1} = \int \pi_{x_t}(z_t^a; z^a) \varrho_t(dz^a) := H(\varrho_t, x_t) \quad \Pi_{x_t}(z_t; \cdot)\text{-a.s.} \quad (3)$$

Thus, for any given sequence $\{x_t\}_{t \in \mathbb{N}}$ the evolution of the process $\{\varrho_t\}_{t \in \mathbb{N}}$ can almost surely be described by a deterministic recursive dynamics.

Now 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 4.4 *There exist smooth functions $F : U \rightarrow X$, $F = (f^1, \dots, f^m)$, defined on the compact convex U of all probability distributions on states and a constant $\alpha \in (0, 1)$ such that*

$$x_{t+1}^i = \alpha x_t^i + (1 - \alpha) f^i(\varrho_{t+1}) \quad \text{and so} \quad x_{t+1} = \alpha x_t + (1 - \alpha) F(\varrho_{t+1}). \quad (4)$$

We illustrate Assumption 4.4 by means of 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 4.5 (i) *Suppose that $x_t^i \in [0, 1]$ and that new types are convex combinations of the old type and the proportion of individuals choosing them. In this case*

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

In view of Remark 4.3 we may view ϱ_{t+1} as a vector $(\varrho_{t+1}^{i,j})$, and so

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

(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:*

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

In view of Remark 4.3 we have the following representation:

$$f^i(\varrho_{t+1}) = \frac{\sum_{j=1}^n c^j \varrho_{t+1}^{i,j}}{\sum_{j=1}^n \varrho_{t+1}^{i,j}}. \quad (5)$$

In this case Assumption 4.4 requires that an agent chooses each state with strictly positive probability.

In view of (3) the joint dynamics of types and empirical distributions follows almost surely the deterministic dynamics

$$\begin{pmatrix} \varrho_{t+1} \\ x_{t+1} \end{pmatrix} = \begin{pmatrix} H(x_t, \varrho_t) \\ \alpha x_t + (1 - \alpha) F \circ H(x_t, \varrho_t) \end{pmatrix} =: G(x_t, \varrho_t) \quad (6)$$

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 analyse what happens over time, we also need a measure of the *continuity* of the changes in the agent's choices. This will enable us to put constraints on the connectedness between the different time-periods of an agent. These constraints consist of two parts, one that refers to the personal connectedness, i.e. the dependence of the current state in terms of characteristics and types of an individual on previous states; the other refers to the social connectedness of an agent over time, i.e. the dependence of the current state on previously available types. The imposition of a continuity condition on an agent's changes over time can be seen as 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.

To this end, we denote by $|\cdot|_1$ the l_1 -norm and recall that

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

defines the total variation norm of the signed measure $\pi_x(z; \cdot) - \pi_x(\hat{z}; \cdot)$. In terms of the total variation norm we can now introduce our weak interaction condition.

Assumption 4.6 *There exist constants β_1 and β_2 such that $\beta := \beta_1 + \beta_2 < 1$ and*

$$\frac{1}{2} |\pi_x(z; \cdot) - \pi_x(\hat{z}; \cdot)|_1 \leq \beta_1$$

uniformly in all the possible types and states, and

$$\frac{1}{2} |\pi_x(z; \cdot) - \pi_{\hat{x}}(z; \cdot)|_1 \leq \beta_2 |x - \hat{x}|_1.$$

The coefficient β_1 measures the dependence of an agents' new type and characteristics on his current state $z_t^a = (c_t^a, y_t^a)$. In the special case where where agents choose their new states *independently* of their previous states,

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

and we can choose $\beta_1 = 0$. This means that the agents have no personal connectedness and that they are non-continuous with themselves. That is, they do not care *where* they came from and *what* they choose is independent of their own past history. If, on the other hand, the agents “remain *who* they are” in the sense that

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

we have that $\beta_1 = 1$ and our weak interaction condition is violated. This means that agents have a strong personal connectedness and that they stick to their previous decision and remain *where* they are. In the case of a high β_1 value, agents are continuous in terms of their personal connectedness and can be traced through time because one state of characteristics and types (can) depend on their previous one⁴. 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 has been changed, but this has not motivated the agent to move on and to select a different type.

The constant β_2 , on the other hand, places a quantitative bound on the dependence of z_{t+1}^a on the current vector of types x_t . A low β_2 implies that the agents' preferences for future states given their current states do not change much if the available types change a little. This means that agents do not react strongly to changes in available types, or, stated differently, that their preferences depend continuously on types. In that sense, the agents' social connectedness is very low. A high β_2 implies that agents can react very strongly to changes in available types and underlines the fact that the agents' social connectedness is strong. In this case, agents are continuous in terms of their social connectedness and can be traced through time because their states in terms of characteristics and types (can) depend on previously available types⁵.

⁴Note 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 can be different, but do not need to be. Thus, in the latter case, there is a higher chance that an agent will end in a particular state, given that he was in a particular previous state.

⁵As 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 types x and \hat{x} is not too large. If the latter distance would be large, then even a very small β_2 is unable to 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

Thus, the *measure of continuity* tells us that if β_1 is high and β_2 is low, personal identity is guaranteed through the agent's personal connectedness between his different states, and if β_1 is small and β_2 is high, personal identity is guaranteed through the agent's social connectedness to different states. That is, whereas in the first case we have a strong path-dependency between the different states of an agent over time, in the latter case, the individual's path over time depends on the changes in his social environment. In that sense, we might say that the stronger sense of personal identity is stemming from the idea of personal connectedness. Social connectedness bridges the different time-periods of a person only through changes in available types.

However, if both values are as small as the *weak interaction condition* requires, the agent is continuous neither in terms of a personal connectedness nor in terms of a social connectedness. This means that an agent's preference for a future state is not only independent of his current state, but also of the choice between different available types.⁶ A low measure of continuity (or a weak interaction condition of the agent's individual behaviour) implies that the different states of an agent's journey through time are not very connected to each other. In that sense, an agent's personal identity is not continuous. Consequently, we might say that an agent's preference really becomes, in the true sense of the word, a probability with which the agent happens to wind up in his future state and it loses any meaning in terms of a preference of a continuous agent over time. Thus, loosely speaking Assumption 4.6 says that agents do not put too much weight on current states and types when revising their choices, and that preferences depend continuously on types.

We illustrate this by means of the following example.

Example 4.7 *Let $c > 0$, and suppose there are only two characteristics, $C = \{-c, c\}$, and two types, $x_t = (x_t^{-1}, x_t^{+1}) \in [-1, 1]^2$. Thus, $Y = \{-1, +1\}$. Assume moreover that the individual transition probabilities do not depend on current types and that an agent a with $c_t^a = \pm c$ chooses the states $(+c, +1)$ and $(-c, -1)$ with probabilities*

$$\pi_{x_t} [+c; +c, +1] = \frac{1}{2} - \frac{1 + x_t^{+1}}{8} = \frac{3 + x_t^{+1}}{8} \quad \text{and} \quad \pi_{x_t} [-c; -c, -1] = \frac{1}{2} - \frac{1 - x_t^{-1}}{8} = \frac{3 - x_t^{-1}}{8}$$

need to be. Thus, in the latter case, there is a higher chance that agents will end in a particular state, given that previously there were some particular types available. Most evidently, if the distance between the two vectors of types is very small, but β_2 rather high, the difference between the two transition probabilities need also to be small.

⁶To visualise what this means imagine that in your current state, you can be either Kate Winslet or Juliette Binoche. In your future state you participate in a football club. A low value of β_1 thus means that you participate at the football club almost independently of whether you have been Kate Winslet or Juliette Binoche. Now imagine that you are Kate Winslet and that there exists a replica of Earth called Mars on which you can choose from almost similar available types as on Earth. A low value of β_2 thus means that you as Kate Winslet will choose to participate in the football club, almost independently of whether you live on Earth or on Mars. Thus, your preference to participate in the football club is almost independent of whether you are Kate Winslet or Juliette Binoche and of whether you live on Earth or on Mars. Surely, in this situation, personal identity does not matter much.

respectively, and any of the remaining three states with equal probability, i.e.

$$\pi_{x_t} [+c; -c, +1] = \pi_{x_t} [+c; +c, -1] = \pi_{x_t} [+c; -c, -1] = \frac{5 - x_t^{+1}}{24}.$$

and

$$\pi_{x_t} [-c; -c, +1] = \pi_{x_t} [-c; +c, -1] = \pi_{x_t} [-c; +c, +1] = \frac{5 + x_t^{-1}}{24}.$$

Thus, for an agent with a positive characteristic the most likely characteristic in the following period is $+c$ and the probability is increasing in x^{+1} . Likewise, an agent with a negative characteristic selects $-c$ with high probability, any of the states is selected at least with probability $\frac{1}{6}$ and at most with probability $\frac{1}{2}$,

$$\pi_{x_t}[c_t^a; z_{t+1}^a] \in \left[\frac{1}{6}, \frac{1}{2} \right],$$

independently of current states and types. Moreover,

$$\pi_x [+c; +c, +1] - \pi_x [-c; +c, +1] = \frac{3 + x^{+1}}{8} - \frac{5 + x^{-1}}{24} = \frac{4 + 3x^{+1} - x^{-1}}{24} \leq \frac{1}{3}$$

because $x^{\pm 1} \in [-1, 1]$. By analogy, $\pi_x [-c; -c, -1] - \pi_x [+c; -c, -1] \leq \frac{1}{3}$ while

$$|\pi_x [+c; \pm c, \mp 1] - \pi_x [-c; \pm c, \mp 1]| \leq \left| \frac{5 - x^{+1}}{24} - \frac{5 + x^{-1}}{24} \right| = \left| \frac{x^{+1} + x^{-1}}{24} \right| \leq \frac{1}{12}.$$

Thus,

$$\sum_z |\pi_x(+c; z) - \pi_x(-c; z)| \leq \frac{2}{3} + \frac{2}{12} = \frac{5}{6},$$

and so $\beta_1 = \frac{5}{12}$. With regards to the dependence of individual transition laws on x note that

$$|\pi_x [+c; +c, +1] - \pi_{\hat{x}} [+c; +c, +1]| \leq \frac{1}{8} |x^{+1} - \hat{x}^{+1}| \leq \frac{1}{8} |x - \hat{x}|_1$$

and

$$|\pi_x [+c; \pm c, \mp 1] - \pi_{\hat{x}} [+c; \pm c, \mp 1]| \leq \frac{1}{24} |x - \hat{x}|_1$$

so that

$$\sum_z |\pi_x(\pm c; z) - \pi_{\hat{x}}(\pm c; z)| \leq \left(\frac{1}{8} + \frac{3}{24} \right) |x - \hat{x}|_1 = \frac{1}{4} |x - \hat{x}|_1.$$

Thus, $\beta_2 = \frac{1}{4}$ and $\beta = \beta_1 + \beta_2 = \frac{5}{12} + \frac{1}{4} = \frac{2}{3}$.

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 available types. Stated differently, if we had two almost similar but still

different current distributions of agents and available types, the future distribution of agents in both “worlds” would be almost the same.

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

Theorem 4.8 *Suppose that Assumption 4.6 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 z ,*

$$\lim_{t \rightarrow \infty} \mathbb{P}[z_t^a = z] = \mu(z)$$

The first part of the theorem states that in the long run types settle down to a unique limit x^* . Asymptotically, the agents new characteristics is thus chosen according to the transition probability

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

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

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

Under Assumption 3, the transition law π_{x^*} has a unique stationary distribution μ and, for any initial condition z_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 be changed 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 available types. Hence, the agent’s switching of types does not have any impact on his individual history because his different stops 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 left with a social identity but one that changes according to when the agents will be where. Their personal identity though, i.e. their continuous and related change over time, has lost almost all significance.

To see this result, let us return to the situations considered in Example 4.5.

Example 4.9 In Example 4.5 we only had two possible types: -1 and $+1$.

(i) If types are modified solely according to the proportion of agents choosing them,

$$f^{+1}(\varrho_t) = \varrho_{t+1}^{+1,+1} + \varrho_t^{+1,-1} \quad \text{and} \quad f^{-1}(\varrho_t) = \varrho_{t+1}^{-1,-1} + \varrho_t^{-1,+1}.$$

Thus, $L = 1$ when $c \leq 1$ and the assumptions of Theorem 4.8 are satisfied for any $\beta < 1$.

(ii) If types are modified according to the average characteristic of the agents joining, the analysis becomes more involved. In this case

$$f^{+1}(\varrho_t) = c \frac{\varrho_t^{+1,+1} - \varrho_t^{+1,-1}}{\varrho_t^{+1,+1} + \varrho_t^{+1,-1}} \quad \text{and} \quad f^{-1}(\varrho_t) = -c \frac{\varrho_t^{-1,-1} - \varrho_t^{-1,+1}}{\varrho_t^{-1,-1} + \varrho_t^{-1,+1}}.$$

Thus,

$$\nabla f^{+1}(\varrho_t) = c \left(\frac{\varrho_t^{+1,-1}}{(\varrho_t^{+1,+1} + \varrho_t^{+1,-1})^2}, -\frac{\varrho_t^{+1,+1}}{(\varrho_t^{+1,+1} + \varrho_t^{+1,-1})^2}, 0, 0 \right),$$

and

$$\nabla f^{-1}(\varrho_t) = c \left(0, 0, -\frac{\varrho_t^{-1,+1}}{(\varrho_t^{-1,-1} + \varrho_t^{-1,+1})^2}, \frac{\varrho_t^{-1,-1}}{(\varrho_t^{-1,-1} + \varrho_t^{-1,+1})^2} \right),$$

If the choice probabilities are as in Example 4.7, then each agent selects each state with probability at least $\frac{1}{6}$, independently of individual characteristics and types. Thus, at least $\frac{1}{3}$ of the agents choose any of the types, $\varrho_t^{+1,+1} + \varrho_t^{+1,-1} \geq \frac{1}{3}$ and $\varrho_t^{-1,-1} + \varrho_t^{-1,+1} \geq \frac{1}{3}$. Moreover,

$$\frac{\varrho_t^{+1,\pm 1}}{(\varrho_t^{+1,+1} + \varrho_t^{+1,-1})^2} \leq \frac{1}{\varrho_t^{+1,+1} + \varrho_t^{+1,-1}} \quad \text{and} \quad \frac{\varrho_t^{-1,\pm 1}}{(\varrho_t^{-1,-1} + \varrho_t^{-1,+1})^2} \leq \frac{1}{\varrho_t^{-1,-1} + \varrho_t^{-1,+1}}$$

so that $L \leq 3c$. Since $\beta = \frac{2}{3}$, the assumptions of Theorem 4.8 are satisfied for $c \in (0, \frac{1}{2})$.

What would happen if our *weak interaction condition* of an agent's individual behaviour that led to the equilibrium is violated? Let us assume, for example, that an agent's preference does not continuously depend on available types, i.e. that his preference changes if types do. In that case, the individual's personal identity is largely determined by his social connectedness. In this situation we would be able to find multiple long run outcomes, and which of these will occur will, however, depend on the initial conditions.

Example 4.10 Suppose that there are two characteristics, $C = \{-1, +1\}$, two types, $Y = \{-1, +1\}$ and that $x_t^{-1} \in [-1, 0]$ and $x_t^{+1} \in [0, 1]$. Let us assume that agents choose their new characteristics in reaction to current characteristics and types and let $\varrho_t = (\varrho_t^{-1}, \varrho_t^{+1})$ be the empirical distribution of characteristics at time t . Types are updated in reaction to the proportion of people choosing the corresponding characteristics:

$$x_{t+1}^{\pm 1} = \alpha x_t^{\pm 1} \pm (1 - \alpha) \varrho_{t+1}^{\pm 1}.$$

Loosely speaking, the evolution of types reflect the dynamics of average characteristics. Let us further assume that an agents a sticks to his characteristics c_t^a if the majority of agents choose c_t^a . Otherwise he selects a characteristic with equal probability. More precisely,

$$\pi_{x_t}(1; 1) = \begin{cases} 1 & \text{if } x_t^{-1} + x_t^{+1} > 0 \\ 0.5 & \text{otherwise} \end{cases} \quad \text{and} \quad \pi_{x_t}(-1; -1) = \begin{cases} 1 & \text{if } x_t^{-1} + x_t^{+1} < 0 \\ 0.5 & \text{otherwise.} \end{cases}$$

Thus, the individual transition probabilities do not depend continuously on the current types. If $x_t^{-1} + x_t^{+1} > 0$, the law of large numbers for independent random variables yields

$$\varrho_{t+1}^{+1} = \varrho_t^{+1} + \frac{1}{2}\varrho_t^{-1} \quad \text{while} \quad \varrho_{t+1}^{-1} = \frac{1}{2}\varrho_t^{-1} \quad \text{and} \quad x_{t+1}^{-1} + x_{t+1}^{+1} > 0.$$

As a result, “initial conditions matter.” The long run evolution of empirical characteristics depends on the starting point (x_0^{+1}, x_0^{-1}) of the sequence of types. For instance, if $x_0^{-1} + x_0^{+1} > 0$, then

$$\varrho_1^{+1} = \varrho_0^{+1} + \frac{1}{2}\varrho_0^{-1} \quad \text{while} \quad \varrho_1^{-1} = \frac{1}{2}\varrho_0^{-1} \quad \text{and} \quad x_1^{-1} + x_1^{+1} > 0,$$

so that

$$\varrho_1^{+1} = \varrho_0^{+1} + \frac{1}{2}\varrho_0^{-1} + \frac{1}{4}\varrho_0^{-1} \quad \text{while} \quad \varrho_2^{-1} = \frac{1}{4}\varrho_0^{-1} \quad \text{and} \quad x_2^{-1} + x_2^{+1} > 0.$$

Thus, an induction argument shows that $\varrho_t^{-1} \rightarrow 0$ and $\varrho_t^{+1} \rightarrow 1$ as $t \rightarrow \infty$. By analogy, $\varrho_t^{-1} \rightarrow 1$ and $\varrho_t^{+1} \rightarrow 0$ as $t \rightarrow \infty$ if $x_0^{-1} + x_0^{+1} < 0$. Overall,

$$\lim_{t \rightarrow \infty} \varrho_t^{+1} = \begin{cases} 1 & \text{if } x_0^{-1} + x_0^{+1} < 0 \\ 0 & \text{if } x_0^{-1} + x_0^{+1} > 0 \\ 0.5 & \text{if } x_0^{-1} + x_0^{+1} = 0. \end{cases}$$

In this case, we can interpret the economic agent as being endowed with a specific set of preferences, whereby he prefers to stick to his previous decision if a majority of people choose in the same way as he does. The agent thus feels comfortable with himself as long as his decision is consistent with a majority of other people. If his choice is not consistent with that of the majority, the individual starts doubting his previous decision and becomes indifferent or undecided in terms of his preference in relation to his future behavior. This corresponds to a preference for conformism. The latter implies a social connectedness of the individual through time in terms of current available types. Thus, depending on which current types exist, the individual will behave differently. To put it in another way, his future state depends on current available types. This situation indeed resembles to some extent to the previously considered “social identity models”. However, the model at hand shows that a society can have different outlooks, i.e. equilibria, depending on the initial conditions. These initial conditions not only refer to the actual available groups, but also to the specific preference the individuals adopt. However, most certainly, there

is no reason to assume that individuals do exactly have this form of preference for conformity. As soon as this preference would change, an equilibrium cannot be guaranteed. In fact, we might imagine individuals' preferences to change if their connectedness through time not only depends on available groups, as in this case, but also on their past states and thus personal history.

Finally, we give an example to illustrate the fact that individuals do not necessarily settle down if $\beta_1 = 1$ even if the empirical distribution of characteristics is constant. This is a very simple example that shows that if a person's evolution is path dependent, they will have different preferences for, or probabilities of choosing different characteristics at different times.

Example 4.11 *Let us assume that there are only two characteristics, $C = \{-1, +1\}$ and that the agents do not react to types. If the individual transition laws take the form*

$$\pi(\pm 1; \mp 1) = 1,$$

then $\beta_1 = 1$. Let us also assume that $\varrho_0^{-1} = \varrho_1^{-1} = \frac{1}{2}$. In this case, the empirical distribution of agents' characteristic does not change through time, but the probability of an agent choosing a certain characteristic does not converge because, for any $t \in \mathbb{N}$ any all agents,

$$\mathbb{P}[c_{2t+1}^a = -1 | c_0^a = +1] = 1 \quad \text{while} \quad \mathbb{P}[c_{2t}^a = +1 | c_0^a = +1] = 0.$$

5 Discussion

Our model differs from previous models of identity insofar as it considers not only the aspects of people's social identity and thus of adding *where* they are to *what* they are, but goes further and tries to explain *who* they are in successive time periods. The idea is to add a further dimension to personal identity of which social identity is one particular aspect. Accounting for a person's *who*-identity focuses specifically on considerations about the change and the continuity of a person and of society over time.

There are two major differences between our model and social-identity-models. Here society stabilises neither because people are already situated within certain social groups and thus act according to their norms and rules, nor because people develop a preference for specific norms and rules that then become embedded in specific social groups. Quite the opposite, society stabilises when people's preferences as to which groups they wish to identify with are relatively independent of *where* they are coming from and of the groups that are available at any point in time. Indeed, the smaller the degree of connectedness between successive states of one and the same person and the less responsive the individual is to available groups, i.e. the less important the continuity of personal identity, the more preferences are stochastic but stable, in the sense, that their probabilities of choosing their "types" settle down over time. We might also say that in

long run equilibrium preferences exist independently of the person, i.e. each person's probabilities of choosing the different groups will be the same. . In equilibrium, all that matters is that agents find themselves in different social groups at different moments of time. Thus, instead of having a situation where social identity is an aspect of personal identity, personal identity dissolves itself in social identity. Even if the individual continues moving in equilibrium, at each time the way in which he changes is essentially disconnected from the various stages of his journey up to that point. Once he has chosen a group the person adapts to a greater or lesser degree to the specific characteristics of that social group before moving on to the next. Indeed, social identity also predominates in equilibrium in the sense that it is the group-membership that changes the characteristics of the agent, but the agent no longer affects the characteristics of the group. However, again, this is what happens in equilibrium and the equilibrium itself is only attained if individuals' personal identity matters little. We might say that the price to pay for a stable society is to rid individuals of their personal concerns.

However, we have seen that if agents react strongly to variations in the various types available, then, given their preference for social conformism, several long run outcomes of the society are possible, depending on the initial conditions of that society. This is a special case which resembles a number of social-identity-models considered previously in the literature. However, the example shows clearly that it is the social connectedness of agents over time which becomes central in this case. Yet, this is a rather weak ingredient of the overall vision of personal identity that we wish to present, because it reflects the simple idea that people change as a result of external modifications. It is thus a special case for, given the overall dynamics of the general model, we have no reason to assume that individuals will not modify themselves modify their preferences over time.

More importantly, if personal identity matters in terms of people's preferences that express a dependence on *where* they are coming from and *who* they want to be, i.e. if we take account of the personal connectedness of individuals, society might not become stable at all. This might sound counterintuitive, but it is not if we look at the dynamics of the model that relies on the simple idea that social groups affect people's characteristics, but people's characteristics also affect the outlook of social groups. If there is indeed an element of independence between social and individual preferences and people develop a particular preference for a social group because they want to *adopt* its characteristics and not because people want to *adapt* their characteristics to that of the social group, then we can conceive of a situation in which there is continuous change and movement of individuals. Each time they have the impression that they would need to start *adapting* to a social group instead of freely *adopting* certain characteristics of that group in order to fit to the group, they may decide to move on.

That is, people, by joining groups, are bringing their different characteristics to them and thus modify the outlook of those groups over time. This change in characteristics of the group

might make people reluctant to continue staying in a particular group and motivates them to look out for different groups whose characteristics correspond more to those with which they feel less dissonance. Thus, the model shows why there can be situations in which a person's embeddedness in a social context or a person's holding of a preference for a particular place in society does not imply that she loses her capacity for change and that society's dynamics will settle to a steady state. In fact, it seems rather intuitive to think that only under very particular circumstances, will society's structure stabilize in this way and individuals have to live with only a limited number of choices of *who* they can be. However, again, these circumstances occur not because people care too much about *who* they are, but too little. In that specific case we might say that it is society's preferences that replace those of individuals.

It is clear that for the moment, the model so far is only a first rudimentary approach to the problem. Several issues are not yet taken enough into account. First of all, individuals do not have a well specified rationality. They are simply endowed with choice probabilities whose origins we did not specify. However, even in this very general framework we were able to obtain some idea of what conditions might lead to individual choices and society's structure stabilising. If we wish to obtain more specific results, it is clear that people's transition probabilities, or what we called an individual's *who-criterion*, need more detailed specification. Thus, the question is why people with particular characteristics and group-memberships end up with those characteristics and memberships.

The extension of the model that we intend to work on is the problem of "self image". We might imagine that people have particular self-images to which they would like to correspond and such images, could equally be described in terms of particular characteristics. Then people would choose those groups whose characteristics are most similar to those of their self-image. Given that group-characteristics change over time, this creates a distance between their self-image and the perceived characteristics of the group to which they have chosen to belong. To reduce this distance (or internally felt discrepancy or cognitive dissonance), individuals continue to move on and choose a different group that is closer to their self-images. However, it is also clear that, in certain circumstances, it might also be the self-image that changes. Indeed, one might conceive that a change in the perceived characteristics of the social group introduces new characteristics (or information) that induce the individual to reconsider her choices and aspirations. Thus, additional information might contribute to a different perspective of *who* one is and *who* one wants to be. The person then might either continue to change her characteristics in order to achieve her self-image or change her self-image and continue to be *who* (and *where*) she is. In that case, conditions must be found that differentiate between people *adapting* to a given situation by reducing their aspirations or *adopting* a new situation and thereby invigorating their desires. Whereas *adaptation* means to conform to given possibilities, *adoption* would imply creative involvement with new opportunities.

That this difference matters has long since been pointed out by Amartya Sen, who consistently raised awareness about those unfavourable social and economic circumstances to which people adapt and live through with “cheerful endurance” (Sen 1984, p. 309).

The same observation was made by George Bernard Shaw when he said, ”The reasonable man adapts himself to the world: the unreasonable one persists in trying to adapt the world to himself. Therefore all progress depends on the unreasonable man”. *Man and Superman (1903) Maxims: Reason*

Another more realistic feature that we might wish to include in the model is to remove the assumption that individuals move freely from one group to the other. It is obvious that in the real world, people might be barred from entry to and obstructed from exit of social groups and this fact indeed acts as a brake on any change. However, while this is true regarding particular social groups, classes and access to different social facilities and is certainly interesting to analyse, our model does not need to be thought in terms of specific existing groups only. Indeed, the implicit set of what we call “types” is larger than that of existing social groups. “Types” can also be seen as objects, actions and even norms that are describable in terms of characteristics and whose feature it is that they are shared by different people. There are obvious if rather trivial cases such as for example that of specific sport-shoes which have been used first for the purpose for which they were intended, but because pop-stars and other people started using them, their characteristics changed over time and thus also the people who wear them. In that sense, our model is really intended as a way to think about continuity and change of a person choosing particular groups, objects and actions to create her personal identity over time.

Another aspect that is worth considering is that the model does not differentiate between the different ways in which the characteristic of social groups can change. Indeed, it might be claimed that certain social groups attract from the outset quite conformist individuals whose characteristics do not differ much from that of the group. In such cases (e.g. religious groups or political parties), the potential for, and rapidity of ,change of this group is much less compared to others, which attract a quite diverse range of people with varying characteristics (e.g. tennis club or university). That is, the model, by presuming one specific mechanism of change (e.g. the new type is a convex combination of old type and average characteristics of members) is silent about what determines the speed of change and thus eventually admits any degree of change. However, an interesting aspect of group change (and consequently also of personal change) over time is not only how characteristics change, but also what sort of characteristics contributes to change. Indeed, the idea that different people bring different characteristics to the group and thus participate in its change is quite convincing. However, depending on the characteristics that are brought to the group, the speed of change can be radically modified. If David Beckham joins Real Madrid, he definitely brings characteristics to the group that contribute to its continued

fame and success as a football club. When, on the other hand, Michael Gorbatshev came to power in the communist regime in the Soviet Union, his particular characteristics contributed to its complete disintegration. Thus, whereas in some cases, many different characteristics of many different people are needed to change the overall outlook of particular groups, sometimes a few characteristics of a single person suffice to trigger an almost complete change of specific groups.

While these are certainly limitations, we think that the substantive contribution of the model is to provide a different view of the social embeddedness of an individual. Indeed, whereas social-identity models assume strong individual preferences for conformism and social status that contribute to the stability of society, we envisage an endogenous social identity formation on the basis of personal identity considerations, i.e. a continuous and connected change of personal and social characteristics over time with a strong feed-back between the two. Under these assumptions, we see that society only settles down if individuals are sufficiently malleable and rather than pursuing an individualistic course, adapt, at least to some extent, at each point in time to the characteristics society imposes. This, on the one hand, is indeed in sharp contrast to social-identity models, which assume that individuals who have strong individual preferences depending on their social group. On the other hand, in many cases, social-identity models do not consider any dynamics and are thus silent about the development of a person's preference and identity over time. In that sense, those models might be considered as special cases of our model once equilibrium has been attained. However, by taking this route one could argue that one is not making models of *sovereign* individuals, but of *social fools*.

6 Conclusion

The main argument of this paper is that we need a more complete model of economic identity. Recently, the characterisation of the economic agent has been extended from an atomistic and strictly self-concerned individual to a socially embedded individual. However, while this is a necessary step to increase the realism of the concept of an economic agent, it suffered from the drawback that the individual has become to some extent over-socialised and this in two ways. Either the individual is being endowed right from the start with a given social identity that specifies and influences the agent's preferences. Or the individual expresses a preference for a particular social position, status or norm with the effect that this preference might become so important that it overrules her intrinsic or private preferences. It is certainly an empirical fact that a person's social environment has influences on preferences or that a person adheres to specific rules despite of underlying and maybe even opposite personal preferences. However, these situations should not be considered as universal. The person, even if socially embedded, still maintains the capacity to evaluate social influences and to decide about the observance of specific

norms and rules or about the importance to achieve a specific social position in a given society. This capacity does not simply imply that the person trades her social preferences (for norms, status etc.) for her private ones and will choose one or the other according to its lower cost or higher utility. A person might rather feel a discrepancy between her own personal preferences that she is attempting to satisfy and the social necessities to which she thinks she must adhere. This discrepancy may not be resolved by simply acting either according to her preferences or to the social expectations, but only by changing either her personal preferences or by attempting to change her social environment. Indeed, if the person decided to act according to her personal preferences, she would be constantly faced with a violation of social norms or with contradictions with her social environment. If she adhered to the standards of her social environment, she would overlook her private preferences. In that sense, either one or the other will be put on one side to some extent and her internal discrepancy cannot be claimed to be resolved. Standard welfare economics does not make any difference between these two situations. If the person acted such as to satisfy her personal preferences or her social preferences, in any case it is assumed that she did satisfy *her* preferences and her new welfare level expresses exactly this satisfaction. However, one may doubt whether both forms of satisfaction transmit the same level or aspect of benefit. Indeed, it might be claimed that the satisfaction of personal preferences “feels” different from the satisfaction of or adherence to social rules. If, on the other hand, the person is trying to appease her discrepancy between personal and social preferences, she might look out for possibilities of change - either of herself or of her environment, that would rid her of her internal conflict and transmit a level of satisfaction that is compatible with both, her personal and her social preferences. Of course, in this case one must have to differentiate between a situation in which a person changes and thus *adapts* her personal preferences to her social environment because particular social circumstances did not leave her enough opportunities to do otherwise and a situation in which a person freely *adopts* a given set of characteristics that allow her to fully deploy her personal identity and *who* she wants to be. These are indeed important aspects, which should be included in any consideration of a socially integrated economic agent but which so far have not received much attention in the economic literature.

7 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”, *Quarterly 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.
- 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.
- 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.
- Cooley Charles H., 1964 (1902), *Human Nature and the Social Order*, Schocken Books.
- Ferret Stéphane, 1998, *L'identité*, Flammarion.
- Freud Sigmund, 1961 (1923), "The Ego and the Id", In: Strachey J. (ed.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud* (Vol. 19, pp. 3-66), Hogarth Press.
- Gintis Herbert, 1974, "Welfare Criteria with Endogenous Preferences: The Economics of Education", *International Journal of Economics* 15(2), pp. 415-430.
- Glover Jonathan, 1988, *I: The Philosophy and Psychology of Personal Identity*, Penguin Books.
- 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.
- Horst Ulrich, 2000, *Asymptotics of locally and globally interacting Markov chains arising in microstructure models of financial markets*, Shaker-Verlag.
- James William, 1948 (1890), *Psychology*, World.

- Kirman Alan, Teschl Miriam, 2004, "On the Emergence of Economic Identity", *Revue de Philosophie Economique* 9, pp. 59-86.
- Livet Pierre, 2004, "La pluralité cohérente des notions de l'identité personnelle", *Revue de Philosophie Economique* 9, pp. 29-58.
- Mead George H., 1934, *Mind, Self, Society*, University of Chicago Press.
- Noonan Harold W., 2003 (1989), *Personal Identity*, Routledge.
- Oksenberg Rorty Amelie (ed.), 1976, *The Identity of Persons*, University of California Press.
- Olson Eric, 2002, "Personal Identity", Entry in the on-line Stanford Encyclopedia of Philosophy, <http://plato.stanford.edu/entries/identity-personal/>
- Oxoby Robert J., 2004, "Cognitive Dissonance, Status and Growth of the Underclass", *Economic Journal* 114(498), pp. 727-749.
- Parfit Derek, 1984, *Reasons and Persons*, Oxford University Press.
- Sen Amartya, 1973, "Behaviour and the Concept of Preference", *Economica* 40(159), pp. 241-259.
- Sen Amartya, 1984, *Resources, Values and Development*, Basil Blackwell.
- Sen Amartya, 1999, *Reason Before Identity: The Romanes Lecture*, Oxford University Press.
- Sen Amartya, 2001, "Other People", *Proceedings of the British Academy* 111 , 319-335
- Sen Amartya, 2004, "Social Identity", *Revue de Philosophie Economique* 9, pp. 7-28.

A Proof of Theorem 4.8

Our weak interaction condition guarantees that the maps $\varrho \mapsto H(x, \varrho)$ and $x \mapsto H(x, \varrho)$ are contractions with respect to $|\cdot|_1$ with constants β_1 and β_2 uniformly in $x \in X$ and $\varrho \in U$, respectively⁷. More precisely, under Assumption 4.6,

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

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 restrains the dependence of the future distribution of agents on the current distribution and the available types. Stated differently, if we had two almost similar but still different current distributions of agents and available types, the future distribution of agents in both “worlds” would be almost the same.

Proof of Theorem 4.8:

- 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 (4). 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 (6) 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.

⁷For details we refer the reader to Horst (2000).

b) Under Assumption 3, 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.

□

Appendix: Figures 1-4

$\alpha = 0.10$
tendency to change = distance dependent
iteration #000

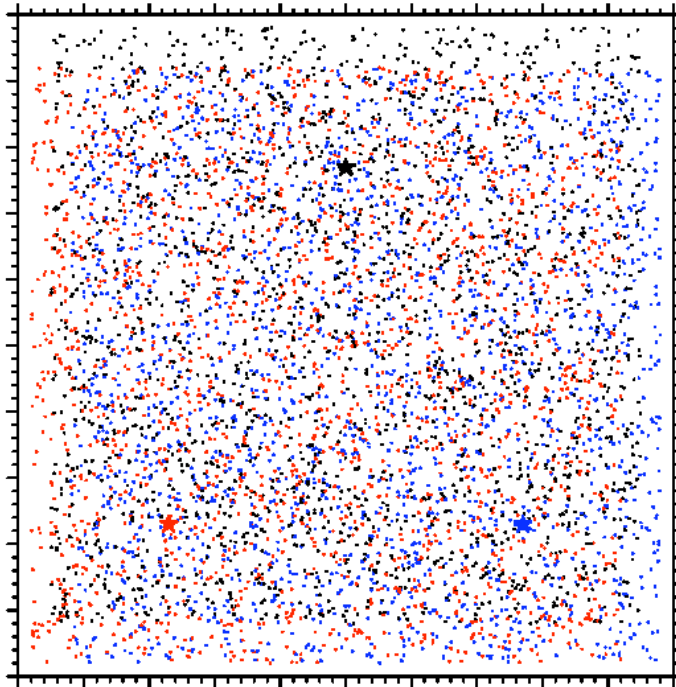


Figure 1a

$\alpha = 0.10$
tendency to change = distance dependent
iteration #014

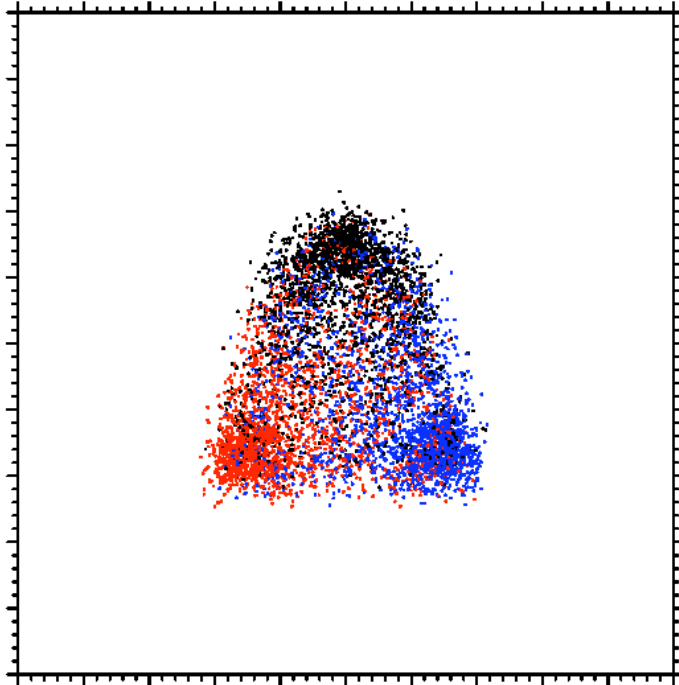


Figure 1b

$\alpha = 0.10$
tendency to change = distance dependent
iteration #029

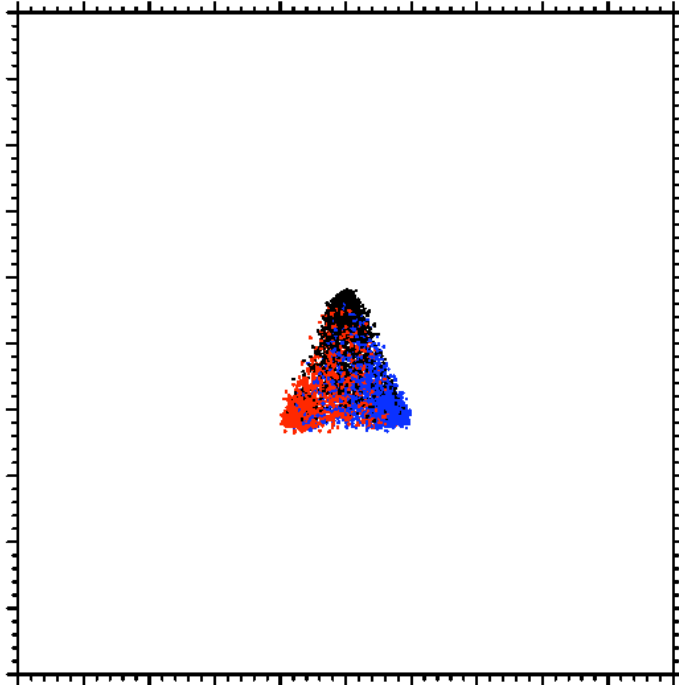


Figure 1c

$\alpha = 0.10$
tendency to change = distance dependent
(strong initial clustering)
iteration = start

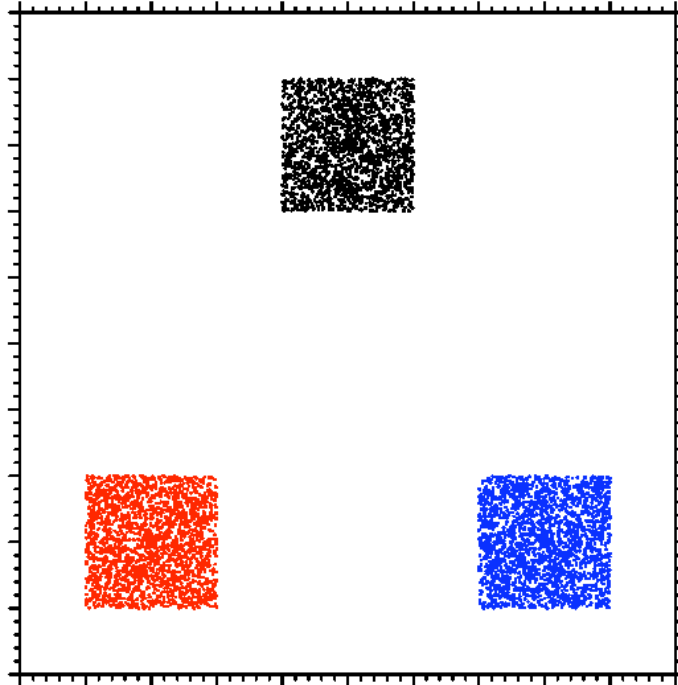


Figure 2a

$\alpha = 0.10$
tendency to change = distance dependent
(strong initial clustering)
iteration #014

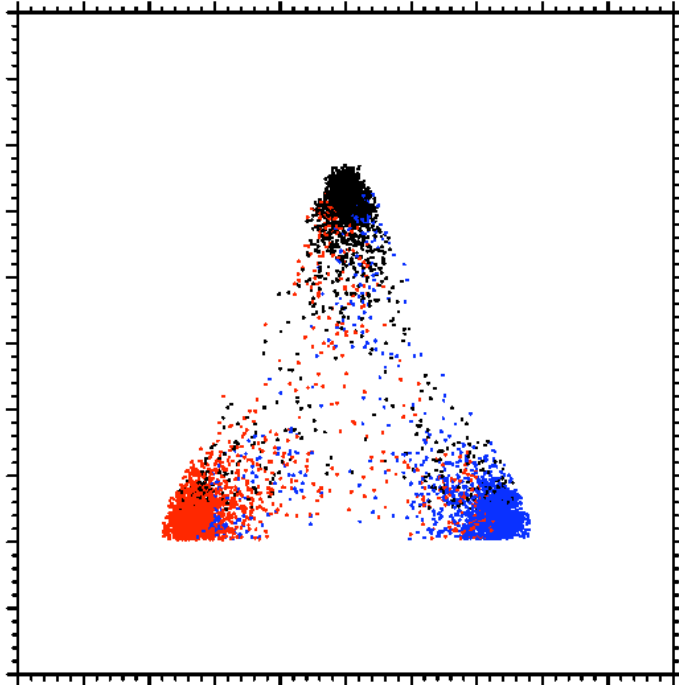


Figure 2b

$\alpha = 0.10$
tendency to change = distance dependent
(strong initial clustering)
iteration #029

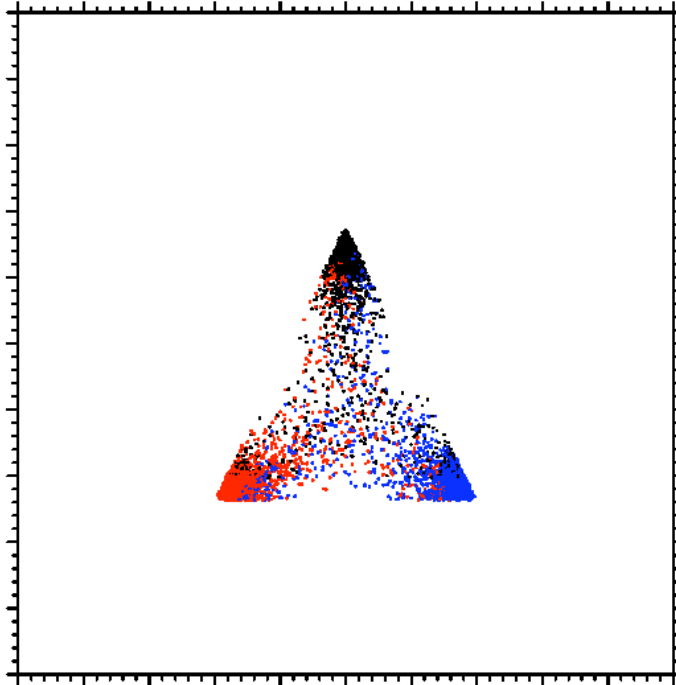


Figure 2c

$\alpha = 0.01$
tendency to change = distance dependent
(strong initial clustering)

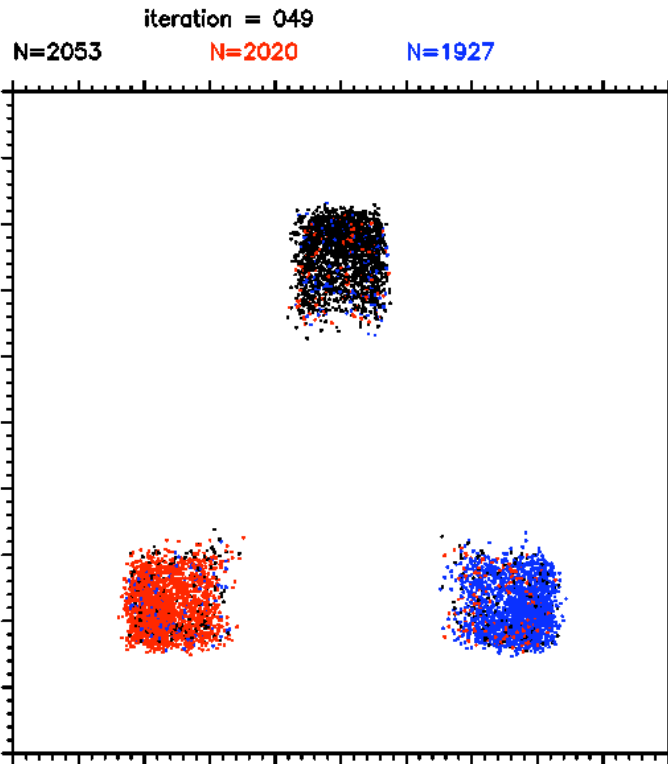


Figure 3a

$\alpha = 0.01$
tendency to change = distance dependent
(strong initial clustering)

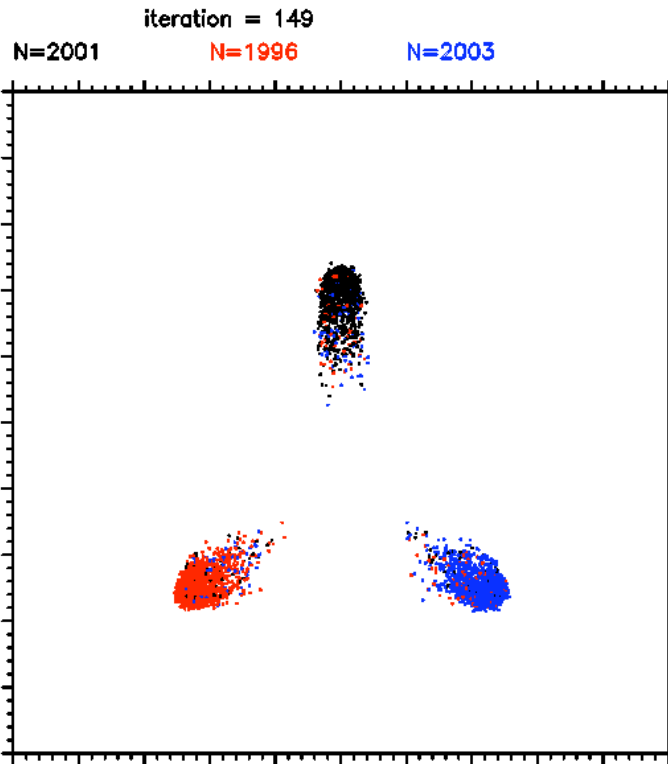


Figure 3b

$\alpha = 0.01$
tendency to change = distance dependent
(strong initial clustering)

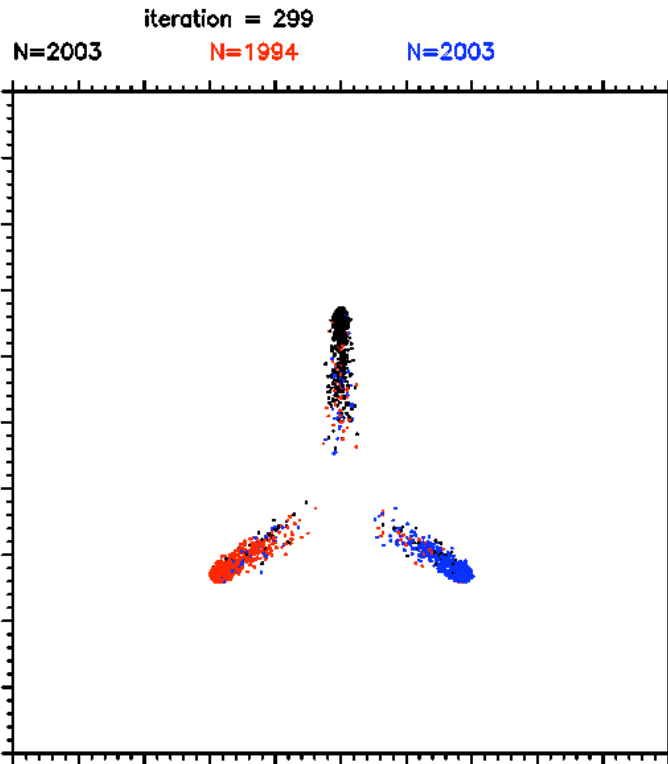


Figure 3c

$\alpha = 0.01$

tendency to change = distance dependent
(strong initial clustering)
(strong individual LEADER)

iteration = 099

N=1892

N=2083

N=2025

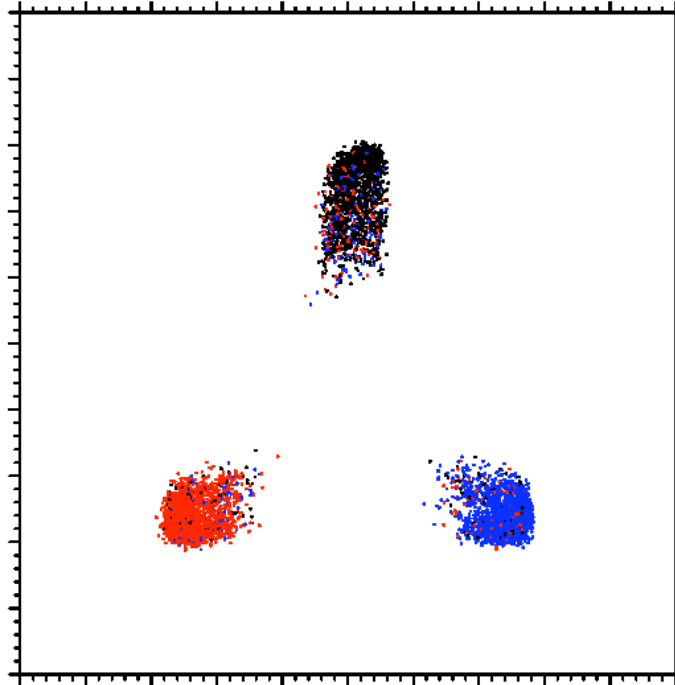


Figure 4a

$\alpha = 0.01$

tendency to change = distance dependent
(strong initial clustering)
(strong individual LEADER)

iteration = 499

N=818

N=2479

N=2703

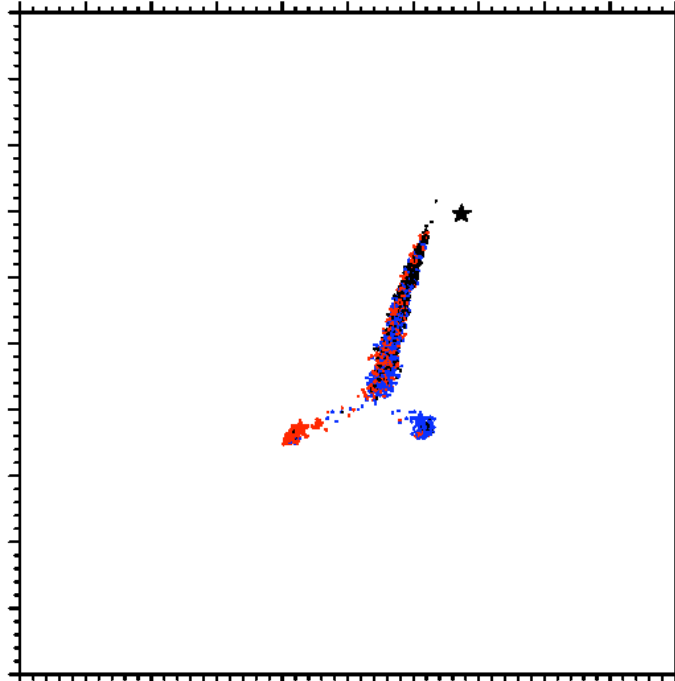


Figure 4b

$\alpha = 0.01$

tendency to change = distance dependent
(strong initial clustering)
(strong individual LEADER)

iteration = 999

N=1

N=2086

N=3913

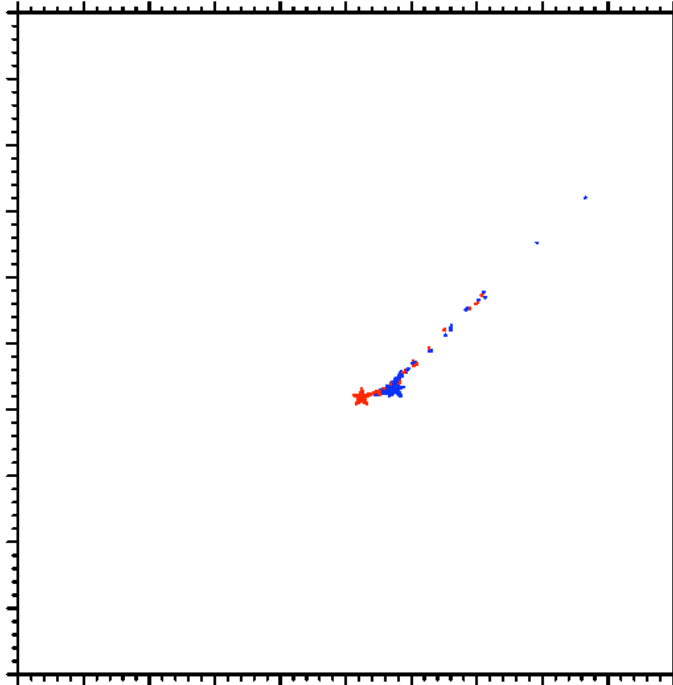


Figure 4c